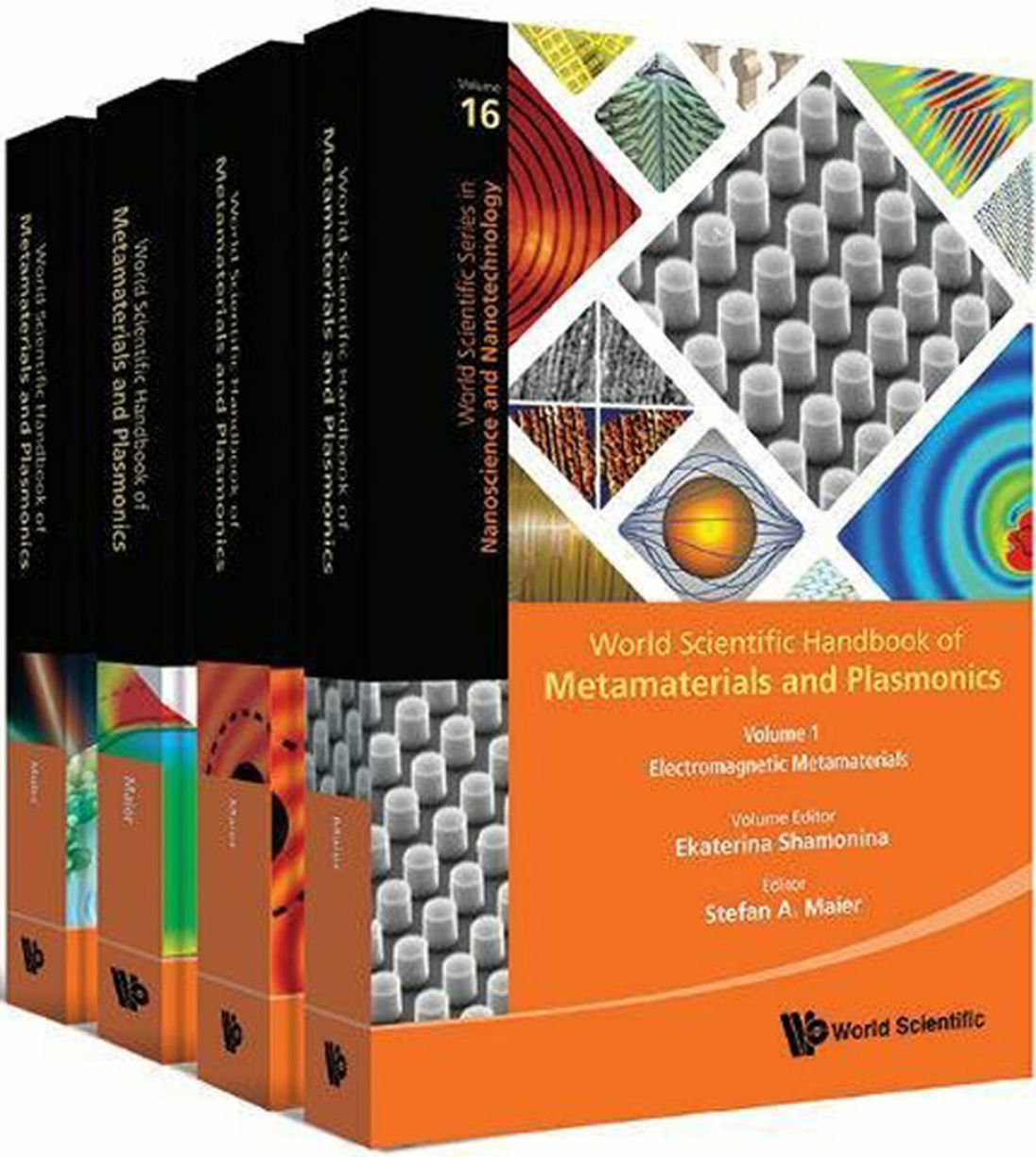

Volume 16
World Scientific Series in Nanoscience and Nanotechnology
World Scientific Handbook of Metamaterials and Plasmonics
World Scientific Handbook of Metamaterials and Plasmonics
World Scientific Handbook of Metamaterials and Plasmonics
World Scientific Handbook of Metamaterials and Plasmonics
World Scientific Handbook of Metamaterials and Plasmonics
Volume 1
Electromagnetic Metamaterials
Volume Editor
Ekaterina Shamonina
Editor
Stefan A. Maier
World Scientific

World Scientific Handbook of

Metamaterials and Plasmonics

Volume 1

Electromagnetic Metamaterials

World Scientific Series in Nanoscience and Nanotechnology*

ISSN: 2301-301X

The Series aims to cover the new and evolving fields that cover nanoscience and nanotechnology. Each volume will cover completely a subfield, which will span materials, applications, and devices.

Published

Vol. 16 *World Scientific Handbook of Metamaterials and Plasmonics (In 4 Volumes)*
Volume 1: Electromagnetic Metamaterials
Volume 2: Elastic, Acoustic, and Seismic Metamaterials
Volume 3: Active Nanoplasmonics and Metamaterials
Volume 4: Recent Progress in the Field of Nanoplasmonics
edited by Stefan A Maier (Imperial College London, UK)

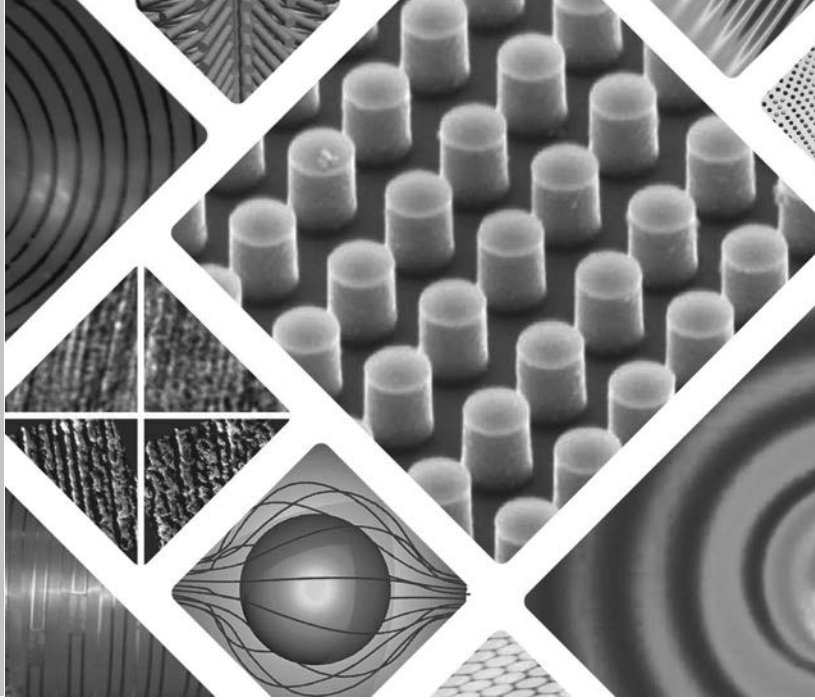
Vol. 15 *Molecular Electronics: An Introduction to Theory and Experiment Second Edition*
by Juan Carlos Cuevas (Universidad Autónoma de Madrid, Spain) and Elke Scheer (Universität Konstanz, Germany)

Vol. 14 *Synthesis and Applications of Optically Active Nanomaterials*
by Feng Bai (Henan University, China) and Hongyou Fan (Sandia National Laboratories, USA)

For further details, please visit: http://www.worldscientific.com/series/wssnn

(Continued at the end of the book)

Volume
16
World Scientific Series in
Nanoscience and Nanotechnology



World Scientific Handbook of
Metamaterials and Plasmonics

Volume 1
Electromagnetic Metamaterials

Volume Editor
Ekaterina Shamonina
University of Oxford, UK

Editor
Stefan A. Maier
Imperial College London, UK

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Cataloging-in-Publication Data
Names: Maier, Stefan A., editor.
Title: World Scientific handbook of metamaterials and plasmonics / edited by:
Stefan Maier (Imperial College London, UK).
Description: Singapore ; Hackensack, NJ : World Scientific, [2017] |
Series: World Scientific series in nanoscience and nanotechnology, ISSN 2301-301X ;
volume 16 | Includes bibliographical references and index.
Identifiers: LCCN 2017029450 | ISBN 9789813227613 (set) | ISBN 9789813227620
(v. 1 ; hardcover ; alk. paper) | ISBN 9813227621 (v. 1 ; hardcover ; alk. paper) |
ISBN 9789813227637 (v. 2 ; hardcover ; alk. paper) | ISBN 981322763X
(v. 2 ; hardcover ; alk. paper) | ISBN 9789813227644 (v. 3 ; hardcover ; alk. paper) |
ISBN 9813227648 (v. 3 ; hardcover ; alk. paper) | ISBN 9789813227651
(v. 4 ; hardcover ; alk. paper) | ISBN 9813227656 (v. 4 ; hardcover ; alk. paper)
Subjects: LCSH: Metamaterials. | Plasmons (Physics)
Classification: LCC TK7871.15.M48 W67 2017 | DDC 621.3028/4--dc23
LC record available at https://lccn.loc.gov/2017029450

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Where applicable, figures in this publication will appear in color for the online version.

Desk Editor: Rhaimie Wahap

Typeset by Stallion Press
Email: enquiries@stallionpress.com

Printed in Singapore

Preface by Main Editor

It is our pleasure to present you this *Handbook of Metamaterials and Plasmonics*, charting the tremendous progress that has occurred in this exciting area of research over the last years. What continues to fascinate me about the field above all is its interdisciplinary broadness — we have arrived at a stage where metamaterials make an impact on many arrays of science where control over waves is a prominent ingredient — be they electromagnetic, acoustic, elastic, or even seismic! In these four volumes, we hence attempt to set out the richness of the field, taking metamaterials in the widest sense as artificial media with sub-wavelength structure for control over wave propagation.

Volume 1 focuses on the fundamentals of electromagnetic metamaterials in all their richness, including metasurfaces and hyperbolic metamaterials. Volume 2 widens the picture to include elastic, acoustic, and seismic systems, whereas Volume 3 presents nonlinear and active photonic metamaterials. Finally, Volume 4 includes recent progress in the field of nanoplasmonics, used extensively for the tailoring of the unit cell response of photonic metamaterials.

In its totality, we hope that this handbook will be useful for a wide spectrum of readers, from students to active researchers in industry, as well as teachers of advanced courses on wave propagation. I want to thank the volume editors Ekaterina Shamon-

ina, Richard Craster, Sébastien Guenneau, Ortwin Hess and Javier Aizpurua, and all the authors for their excellent and sustained work to put these four volumes together.

Stefan A. Maier
Imperial College London, UK
May 2017

Preface by Volume Editor

It is not a coincidence that this volume on electromagnetic metamaterials is the first among the four volumes in the *Handbook of Metamaterials and Plasmonics*. Historically, the concept of metamaterials was first introduced and verified on electromagnetic metamaterials, rapidly covering a wide frequency range from radio frequencies to the visible. It was electromagnetic metamaterials that were first shown to exhibit intriguing phenomena at odds with our everyday experience and physical intuition. It was electromagnetic metamaterials for which the perfect lens, the concept of negative refraction, transformation optics and invisibility cloaks were first proposed creating an unprecedented stir among physicists and engineers. In the early days of the subject, at the very beginning of the 21st century, the terms "metamaterials" and "electromagnetic metamaterials", were used as synonyms. After a few years of extensive research at a breathtaking pace, a variety of unusual electromagnetic responses were looked at, examined and understood, and applications in a wide frequency range of electromagnetic waves were proposed. It turned out to be possible to design and fabricate structures that could produce electromagnetic responses practically at will, responses not available from natural materials.

This volume offers the readers 12 chapters from pioneers of the research in electromagnetic metamaterials, summarising the

progress, including historical surveys of fundamental discoveries, treatises on theory, design and fabrication, an overview of phenomena and applications and also an outlook and vision on the future. The versatility of the research field is reflected in the diversity of the topics and types of structures and phenomena presented by the distinguished authors of this volume. A wide frequency range is covered, from low RF to optical metamaterials. A large range of constituent materials are presented, from the classical metal-dielectric variety, to all-dielectric structures, semiconductor superlattices, and of course the newly fashionable graphene. A variety of geometries are described, from bulk 3D structures to 2D metasurfaces, one-dimensional waveguides and zero-dimensional devices made of a small number of individual elements. This volume will serve as a reference book and as a source of inspiration for experts and newcomers alike. The structure of the volume is as follows.

In Chapter 1, Ari Sihvola elaborates on the idea that is at the very heart of the concept of metamaterials, how a macroscopic structure comprising small subwavelength unit cells may acquire qualitatively new properties not present in its constituent components and exhibit dispersion behaviour that is not directly predictable from the dispersion of the constituent materials. The chapter demonstrates the power of mixing rules in explaining a rich variety of unusual macroscopic phenomena from microwaves to optics.

In Chapter 2, Mário Silveirinha presents an overview of recent developments in the characterisation of electromagnetic and quantum metamaterials using effective medium methods. It is highlighted that both electromagnetic and electronic systems can be homogenised in a unified manner based on the introduction of an effective Hamiltonian operator to describe the time evolution of the macroscopic initial states as well as the stationary states of the relevant system and to determine quadratic forms related to energy density and energy transport. Examples of homogenisation of electromagnetic metamaterials, of semiconductors and of graphene superlattices are considered.

In Chapter 3, Igor Smolyaninov provides a review of a class of metamaterials exhibiting hyperbolic behaviour. The chapter demonstrates that the study of this important class of metamaterials, originally introduced to overcome the diffraction limit in optical imaging, has extended far beyond its original goal, coming up with a large number of novel phenomena resulting from the broadband singular behaviour of the density of photonic states. Examples given include enhanced quantum-electrodynamic effects, new stealth technology, thermal hyperconductivity, high Tc superconductivity, gravitation theory analogues, and photonic hypercrystals.

Chapter 4 by Francisco Medina, Francisco Mesa, Raúl Rodríguez-Berral and Carlos Molero is dedicated to the topic of extraordinary transmission in metamaterials. The authors provide a state-of-the-art review of recent advances in analytical and circuit modelling simulation techniques. A variety of transmission based structures are discussed from 1D arrays of slits and 2D arrays of holes in metal plates to hollow metal waveguides, metal gratings with finite-thickness dielectric slabs and fishnets and stacked structures.

Chapter 5 by Karim Achouri and Christophe Caloz is dealing with synthesis and realisations of electromagnetic metasurfaces, 2D arrays of subwavelength unit cells, engineered to provide desired reflection and transmission to incident electromagnetic waves. Advantages over bulk 3D metamaterials include ease of fabrication and smaller loss due to reduced size, while still providing great flexibility and functionality. The authors provide a comprehensive review of design rules enabling practical implementation of scattering particles required to provide the desired electromagnetic response. A metasurface synthesis method based on transverse susceptibility tensors is applied both to reciprocal and non-reciprocal electromagnetic transformations, generalised refraction, polarisation rotation and orbital angular momentum multiplexing.

Chapter 6 by Sergei Tretyakov, Viktar Asadchy and Ana Díaz-Rubio offers a general survey of linear bianisotropic metasurfaces reviewing homogenisation models and methods to synthesise compact, easy to manufacture, and low-cost metasurfaces with required

functionalities. Providing numerous examples including thin lenses, frequency-selective surfaces, mantle cloaks or polarisation converters, the authors employ the generalised Huygens' principle and discuss how to identify suitable unit cell topologies that would enable the desired electromagnetic response of a metasurface.

Chapter 7 by Francesco Monticone and Andrea Alù provides a comprehensive study of fundamentals of scattering covering general concepts of scattering effects, power conservation and causality. A review of different techniques enabling control of the scattering processes "at the extreme", from drastic suppression to enhancement of the scattering cross-section, is provided and illustrated by a variety of examples including cloaking and invisibility, superscattering, light trapping, energy harvesting, directional scattering, biochemical sensing, and enhanced light-matter interaction at the micro- and nanoscale. Fundamental limits of scatterers employing active, nonlocal, nonreciprocal, and nonlinear materials are discussed.

Chapter 8 by Alexander Krasnok, Roman Savelev, Denis Baranov and Pavel Belov is introducing all-dielectric nanophotonics, a rapidly developing branch of metamaterials studying light interaction with high-refractive-index dielectric nanoparticles with optically-induced electric and magnetic Mie resonances. The chapter describes optical properties of high-index dielectric nanoparticles, methods of their fabrication and recent advances in practical applications, including metasurfaces, nanoantennas for quantum source emission engineering, Fano resonances in all-dielectric nanoclusters, surface-enhanced spectroscopy and sensing.

Chapter 9 by Ilya Shadrivov and Dragomir Neshev reviews tunable metamaterials, i.e. metamaterials with properties that are variable under external influence such as control voltage, temperature or magnetic field. The tuning mechanisms are classified into three groups, tuning by changing the structural geometry via mechanical deformation of either constituent elements or their mutual arrangements, tuning by changing the properties of the constituent materials e.g. their conductivity and tuning by changing the surrounding environment. Different approaches to fabrication are presented which are

shown to differ at microwaves, at terahertz and at optical frequencies, and potential applications and future developments are discussed.

Chapter 10 by Allan Boardman, Alessandro Alberucci, Gaetano Assanto, Yuriy Rapoport, Vladimir Grimalsky, Vasyl Ivchenko and Eugen Tkachenko provides an overview of metamaterials capable of carrying spatial solitons, beams of electromagnetic energy that rely upon balancing diffraction and nonlinearity in order to retain their shape. Theoretical models for spatial solitons are developed both for double-negative and hyperbolic metamaterials. Examples are considered illustrating nonlinear plasmonic effects including a nonlinear field concentrator ('electromagnetic black hole') and nonlinear plasma waves.

In Chapter 11, Richard Syms, Ian Young and Laszlo Solymar report on their advances in developing an application of RF metamaterials–metamaterial catheter receivers for high-resolution internal magnetic resonance imaging for early detection of bile duct cancer, which outperform 'traditional' detection techniques by employing magnetoinductive waves propagating by virtue of coupling between metamaterial elements, enabling low-loss guiding of RF signals resulting in an improved resolution achievable in MRI. The chapter provides an overview of magnetoinductive waves, describes the advances in the design of thin-film magnetoinductive cables and reports on experiments demonstrating high-resolution imaging combined with an MRI safe design and endoscopic compatibility.

In Chapter 12, Jordi Naqui, Ali Karami Horestani, Christophe Fumeaux and Ferran Martín report on their work on metamaterial-inspired microwave sensors, designed on the basis of symmetry-related electromagnetic properties of transmission lines loaded with resonant metamaterial elements. A novel sensing principle based on the disruption of symmetry is described and prototype devices demonstrating the working principles of sensors using resonance-based and frequency-splitting sensors are presented, aiming at potential applications including contactless linear and angular displacement and velocity sensors, alignment sensors and permittivity sensors.

I would like to thank all the authors of this volume for sharing with the readers their knowledge, wisdom and passion in the dynamic and exciting field of electromagnetic metamaterials. I am certain that our readers will find this volume to be an excellent reference book, a source for inspiration, and an enjoyable read.

Ekaterina Shamonina
Oxford
May 2017

Contents

CHAPTER 1

Electromagnetic Metamaterials: Homogenization and Effective Properties of Mixtures

ARI SIHVOLA

Aalto University, Finland
ari.sihvola@aalto.fi

This chapter presents homogenization principles for metamaterials and dielectric mixtures. The focus is on the principles with which homogenization brings forth emergence: such qualitatively new properties in the effective characterization that are not present in the constituent components that make the medium. The classical Maxwell Garnett mixing result is paralleled with the deterministic scattering problem of a composite sphere, and the resulting equivalence is exploited in translating known mixing results to scattering and absorption properties of simple scatterers and vice versa. Emphasis is given to the way mixing can lead to situations where the composite obeys an unexpected dispersion behavior that is not directly predictable from the knowledge of the dispersion of the component materials. The characteristics of the localized resonances and their dependence on geometrical and structural parameters in plasmonic nanoparticles are analyzed. Theoretical bounds for the effective permittivity of mixtures are discussed as well as situations in which these bounds can allow enhanced polarization and percolation effects.

Finally, anisotropic and in particular hyperbolic media are modelled with a special application to radially anisotropic (RA) particles which may display surprising macroscopic effects like, for example, anomalous absorption.

1.1. Introduction: Electromagnetic metamaterials

The age of the present wave of metamaterials research is equal to that of the ongoing century. Its evolution is dynamic, new ideas and concepts are presented, the pace of development is so fast that terminology and language have not yet reached stability. Even the very definition of the metamaterial concept is fluid: the varying emphases can for instance be followed in the changing page edits of Wikipedia during the past 15 years.[1–3] Essential in metamaterials is their character of displaying emergent properties that result from the complex interaction of their "ordinary-material" building blocks. The present-day definition[4] stresses especially the ordered structure in which the base materials are assembled to make a metamaterial whole.

The idea of metamaterials has diffused into a wide range of disciplines in applied physics and beyond. Being coined by Rodger M. Walser in 2000,[5,6] this term has since secured a solid place in various fields. Metamaterials literature contains contributions from people in electrical and microwave engineering, optics and photonics, microelectronics, materials science and nanotechnology. Interest extends to acoustic, mechanical, elastic and thermal metamaterials. Materials should have properties on demand, they should be reconfigurable and even programmable. Also, the meta-idea has expanded from plain materials into more general meaning: terms like metasystems, metastructures and metafunctions are commonly in use. Metatronics[7] emphasizes the possibilities of combining the modularity and simplicity of circuit electronics thinking into optical nanostructures through the use of complex material response. Another characteristic feature of modern metamaterial engineering is the multifunctionality of structures, and the design often takes inspiration from biology or can be even biomimetic. These methodologies

lead to new constructions like, for example, mimumes–microfibrous multifunctional metamaterials.[8]

It is, however, in electromagnetics where this paradigm had its beginnings.[2] And in fact there exists a very long history of research on complex electromagnetic material effects, artificial materials, periodic structures and bianisotropics, etc. with associated rich collection of documented results.[9] This also means that much of the results from "classical" electromagnetics literature can be reused in analyzing properties of modern metamaterials. In particular, this is the case for studies of macroscopically homogeneous and effective media, hybrid materials, mixtures[10] and composites.[11]

In the present chapter, the power of mixing rules and homogenization principles in the analysis of the macrosocpic properties of metamaterials will be shown. Classical mixing rules may seem straightforward "recipes" to produce a composite, and here the metaphor hits the point: like in cooking, the end result is nonlinear and emergent. The taste of the final dish cannot be reduced to those of its ingredients.

1.2. Homogenization and mixing laws

In the following, let us look at dielectric homogenization with the aim of computing the effective permittivity ε_{eff} of a heterogeneous structure[a] as a function of the structural and material parameters of the components. The analysis is quasi-static, which means that the effective description is valid when the size of the heterogeneities is sufficiently small compared to the wavelength of the electromagnetic excitation. For microwave applications, snow is a good example of homogenizable medium, but for waves within the optical range, the size of the inclusions has to be below several tens of nanometers.

One of the problems in homogenization is that heterogenous media are very seldom well-defined ordered lattices. Sometimes the structure is fully random, but also in many engineering composites

[a]In this chapter, the permittivities denoted by ε are dimensionless and relative: $\varepsilon = \varepsilon_{\text{abs}}/\varepsilon_0$ with ε_0 the free-space permittivity.

where regularity is the objective, it is not perfect. Hence, the structural description of the system has to be given with a set of much fewer number of parameters than there are degrees of freedom in the system. Consequently, many of the resulting mixing formulas are approximative.

1.2.1. *Maxwell Garnett formula and effective polarizability*

The problem of a single scatterer, exposed to an exciting field, is deterministic. A homogeneous and isotropic dielectric sphere with permittivity ε in uniform static field creates a dipolar perturbation in its surroundings. The important parameter measuring the strength of this dipole field is the polarizability. For a sphere, the normalized polarizability reads[b]

$$\alpha = 3\frac{\varepsilon - 1}{\varepsilon + 2}. \tag{1.1}$$

Let us increase scatterer complexity from a homogeneous sphere and continue with another well-defined deterministic problem. Consider an isotropic core–shell structure: the spherical core has radius a and relative permittivity ε_2, and it is surrounded by a spherical coating with radius b and permittivity ε_1. The main response of a small scatterer is its dipole moment, which is the product of the exciting field and the polarizability of the particle. For the composite sphere, the normalized polarizability is [10, Section 4.3]

$$\alpha_{\mathrm{k}} = 3\frac{(\varepsilon_1 - 1)(\varepsilon_2 + 2\varepsilon_1) + g(2\varepsilon_1 + 1)(\varepsilon_2 - \varepsilon_1)}{(\varepsilon_1 + 2)(\varepsilon_2 + 2\varepsilon_1) + 2g(\varepsilon_1 - 1)(\varepsilon_2 - \varepsilon_1)}, \tag{1.2}$$

where $g = (a/b)^3$ is equal to the volume fraction of the core in the scatterer.

The formula (1.2) projects the three internal parameters of the particle into a single number, the polarizability. This quantity, α_{k}, is uniquely determined by the internal structure, but not vice versa.

[b]The normalization is the following: $\alpha_{\mathrm{abs}} = \varepsilon_0 V \alpha$ with V the volume of the scatterer.

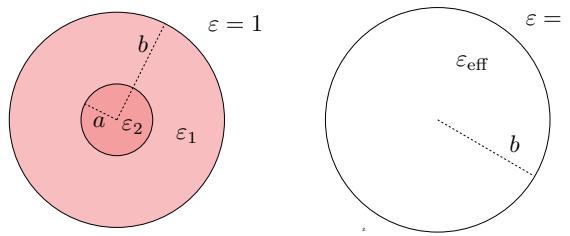


Fig. 1.1. Core–shell sphere and its external equivalent: homogeneous effective dielectric sphere.

There exist an infinite number of possible composite spheres with the same polarizability.

This brings forth the question what is the simplest one of all the possibilities. An obvious answer is the homogeneous sphere. The equivalence is displayed in Fig. 1.1. Let us forget the internal structure of the composite sphere and look for a homogeneous sphere that externally looks exactly like the original layered sphere in a locally uniform static field.[c] The effective permittivity of the homogenized sphere is (cf. equation (1.1))

$$\varepsilon_{\text{eff}} = 1 + \frac{\alpha_{\text{k}}}{1 - \alpha_{\text{k}}/3}, \tag{1.3}$$

where α_{k} comes from (1.2).

The effective permittivity ε_{eff} can be solved as a function of ε_1, ε_2, and g from equations (1.2) and (1.3), leading to

$$\varepsilon_{\text{eff}} = \varepsilon_1 + 3\varepsilon_1 g \frac{\varepsilon_2 - \varepsilon_1}{\varepsilon_2 + 2\varepsilon_1 - g(\varepsilon_2 - \varepsilon_1)}. \tag{1.4}$$

A fascinating observation[13] is that this formula (1.2) is exactly the same as the classical Maxwell Garnett (MG) mixing rule[14] which appeared in the literature in the year 1904. The MG formula predicts the effective permittivity of a mixture where inclusions of permittivity ε_{i} occupy a volume fraction p in environment permittivity ε_{e} (see

[c]This principle of homogenization can also be called *internal homogenization*.[12]

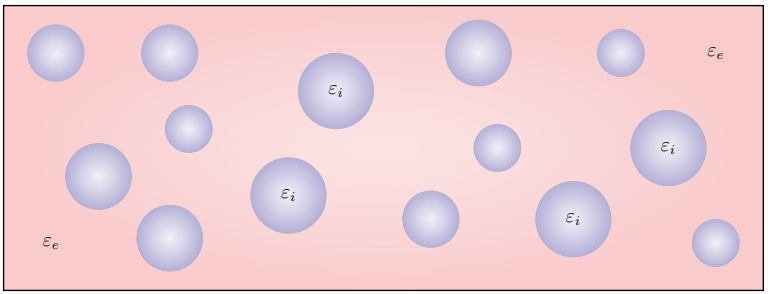


Fig. 1.2. A two-phase mixture where randomly positioned spherical inclusions of permittivity ε_i occupy a volume fraction p in environment with permittivity ε_e.

Fig. 1.2), and

$$\varepsilon_1 = \varepsilon_\mathrm{e}, \quad \varepsilon_2 = \varepsilon_\mathrm{i}, \quad g = p. \tag{1.5}$$

This means that the permittivity that results from the effective-polarizability condition of a composite sphere is exactly the same as the MG prediction when the core is treated as the inclusion and the surrounding shell as the host medium.

This observation has far-reaching consequences. Since there is an exact correspondence between the homogenization of the layered sphere on one hand and MG mixing rule on the other, we can transfer all results for MG mixing known in the vast literature over the past hundred years into the domain of dielectric behavior of core–shell structures and vice versa. In later sections, let us make use of this possibility.

1.2.2. *Bruggeman mixing formula*

On the other hand, it is also worth noting that since there is no exact solution for the effective permittivity of a random sample of dielectric mixture, several different homogenization formulas coexist in the literature along with MG. Another widely used one is the so-called Bruggeman formalism which predicts the effective permittivity according to the formula[15]

$$(1-p)\frac{\varepsilon_\mathrm{e} - \varepsilon_\mathrm{eff}}{\varepsilon_\mathrm{e} + 2\varepsilon_\mathrm{eff}} + p\frac{\varepsilon_\mathrm{i} - \varepsilon_\mathrm{eff}}{\varepsilon_\mathrm{i} + 2\varepsilon_\mathrm{eff}} = 0. \tag{1.6}$$

Here, ε_{eff} is not explicitly given, but the form Eq. (1.6) emphasizes the symmetry between the two phases of the mixture, ε_{i} and ε_{e}. The form also stresses the self-complementary character of Bruggeman mixing formula, in other words an interchange of the host and guest ($\varepsilon_{\text{i}} \leftrightarrow \varepsilon_{\text{e}}, p \to 1-p$ in Eq. (1.6)) does not change the value of ε_{eff}. In this respect, calling the components "host" and "guest" does not appreciate the full equality between the two phases.

One attempt to cover the variety of homogenization rules involving mixtures with spherical inclusions is the following[16]

$$\frac{\varepsilon_{\text{eff}} - \varepsilon_{\text{e}}}{\varepsilon_{\text{eff}} + 2\varepsilon_{\text{e}} + \nu(\varepsilon_{\text{eff}} - \varepsilon_{\text{e}})} = p\frac{\varepsilon_{\text{i}} - \varepsilon_{\text{e}}}{\varepsilon_{\text{i}} + 2\varepsilon_{\text{e}} + \nu(\varepsilon_{\text{eff}} - \varepsilon_{\text{e}})}. \tag{1.7}$$

A dimensionless parameter ν chooses the "character" of the mixing rule: $\nu = 0$ returns the MG rule, $\nu = 2$ gives the Bruggeman formula, and $\nu = 3$ gives another choice, the so-called coherent potential approximation.[17]

These families of mixing rules have been also generalized to cover microgeometries other than spherical. Since a homogeneous ellipsoid in a static field can be solved exactly, mixtures with ellipsoidal inclusions possess concise effective permittivity formulas. They, however, need to be treated with additional care due to the fact that their geometrical non-symmetry makes their orientation distribution affect the result. Furthermore, the effective properties of multiphase mixtures can also be analyzed using mixing rules in which the polarization contributions of all phases are included.[10]

1.3. Homogenization is not dispersion-neutral

Temporal dispersion—in other words the variation of the permittivity with frequency (or wavelength) of the excitation field—is often very strongly and subtly affected by mixing. Even if common dispersion models are comparatively simple functions of frequency on one hand, and on the other, basic mixing rules (like the MG formula Eq. (1.4)) have a straightforward appearance, the behavior of the effective permittivity may display a rather complicated behavior as function of frequency.

1.3.1. *Redshifting plasmonic peaks*

Dispersive effects are prominent in plasmonic metamaterial applications. The field of plasmonics deals with special effects that take place in particles with negative permittivity, often in the nanoscale level. In particular, localized plasmons in nanoparticles can be very effectively studied with the concept of polarizabilities and mixing rules.

1.3.1.1. *Multipole resonances*

The polarizability formula Eq. (1.1) shows a singularity for permittivity value $\varepsilon = -2$, meaning that even particles that are small compared to the wavelength can experience a "resonance" at a frequency where the permittivity attains this value, and dielectric losses are sufficiently small.

The plasmons, despite being confined into a deeply subwavelength particle, can be very efficient radiators. A common measure for scattering is the so-called scattering efficiency Q_{sca}, which is the scattering cross-section divided by the geometrical cross-section of the particle.[18] With the same normalization are defined the absorption (Q_{abs}) and extinction (Q_{ext}) efficiencies, extinction being the sum of absorption and scattering. For an isotropic and homogeneous sphere, the cross-sections can be computed from the so-called Lorenz–Mie coefficients involving spherical Bessel and Hankel functions.[19,20]

For a loss-less scatterer, absorption cross-section should vanish, and scattering equals extinction. Figure 1.3 shows a contour plot of the scattering efficiency of a small loss-less negative-permittivity sphere, as function of relative permittivity between -3 and -1 and the size parameter. Conspicuous in this map are bright and narrow lines; they are strong resonances of the electric multipoles. For very small spheres, these happen at

$$\varepsilon = -1 - \frac{1}{n}, \quad n = 1, 2, 3, \ldots, \tag{1.8}$$

where $n = 1$ corresponds to the electric dipole, $n = 2$ to the electric quadrupole and so on. However, for spheres of finite size, this

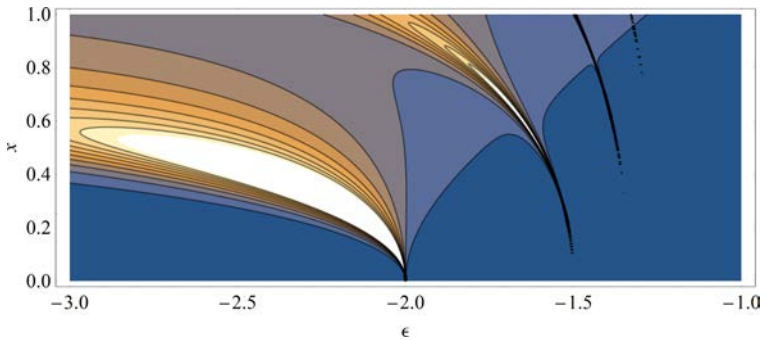


Fig. 1.3. Contour plot of the scattering efficiency of a dielectric, non-magnetic sphere in the plasmonic regime, as function of permittivity ε and size parameter $x = 2\pi a/\lambda$. Loss-less case: scattering and extinction efficiencies are the same. The resolution of the computations cannot distinguish the evermore narrower multipolar peaks that curve to the horizontal axis and accumulate finally at $\varepsilon = -1$.

relation is not exact. Numerically, it can be seen that for the electric dipole resonance of a non-magnetic sphere, the dependence on size parameter x is approximately

$$\varepsilon_{\text{dipole}} = -2 - \frac{12}{5} \cdot x^2, \tag{1.9}$$

where the size parameter is $x = 2\pi a/\lambda$, as a function of the radius a of the sphere and the wavelength λ. The error in this prediction Eq. (1.9) is less than 1% up to size parameter 0.4. In Fig. 1.3, this size dependence displays itself as the curving of the scattering peaks to the left, in other words to more negative permittivity values. And since the permittivity dispersion of plasmonic materials, and materials in general, shows increase with frequency, this phenomenon means that the plasmonic resonances are redshifted (moving to longer wavelengths) when the size increases.

Likewise, the asymptotic dependence of the electric quadrupole resonance can be written as:

$$\varepsilon_{\text{quadrupole}} = -\frac{3}{2} - \frac{5}{14} \cdot x^2 \tag{1.10}$$

Figure 1.3 reveals another interesting phenomenon in the frequency dependence of scattering by small spheres. At exactly $\varepsilon = -2$,

there is a vertical isocontour line reaching up to values of $x \approx 0.5$. This means independence of scattering efficiency on the size for small scatterers for this special plasmonic permittivity value. An insensitivity of the scattering on scatterer size represents an anomalous behavior of small scatterers which deviates strongly from the ordinary Rayleigh scattering[d] with the size dependence of fourth power ($\propto x^4$).

1.3.1.2. *Accuracy of quasistatic prediction*

Even if the results in Fig. 1.3 are calculated using the full-wave Lorenz–Mie machinery, it is worth noting that the polarizability formula Eq. (1.1) captures the scattering, absorption and extinction behavior for spheres of size much smaller than wavelength. Let us examine the accuracy of the quasistatic predictions by invoking realistic plasmonic materials nanoparticles.

A common material used in plasmonics is silver, which has negative permittivity over the optical range. Its complex permittivity can be fitted[22] to experimental data[23] into Drude dispersion formula

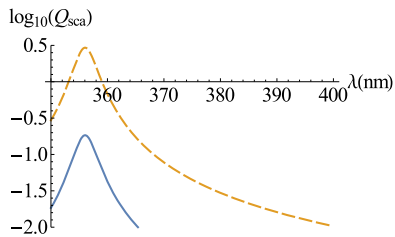
$$\varepsilon_{\text{Silver}}(\lambda) = \varepsilon'_{\text{Silver}} - \mathrm{j}\,\varepsilon''_{\text{Silver}} = \varepsilon_\infty - \frac{(\lambda/\lambda_p)^2}{1 - \mathrm{j}\lambda/\lambda_d}, \tag{1.11}$$

where $\varepsilon_\infty = 5.5, \lambda_p = 130\,\text{nm}$, and $\lambda_d = 30\,\mu\text{m}$. This is a fairly good model for the permittivity of silver in the range $320\,\text{nm} < \lambda < 700\,\text{nm}$.[e]

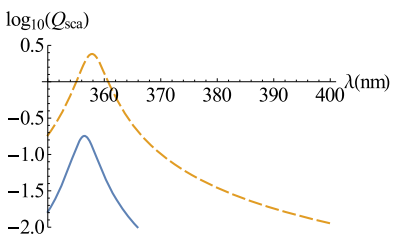
Figures 1.4 and 1.5 illustrate the accuracy of the static analysis. The scattering and absorption cross-sections are compared with the full-wave solution. The asymptotic expansions of the scattering and

[d]The blue color of sky is a result of Rayleigh scattering, due to the fact that the elastic scattering by atmospheric gases is caused by particles very small compared with the wavelength. Hence, the blue part of the light spectrum is scattered much more strongly. As a side note, the term "Rayleigh scattering" is anachronistic: John William Strutt published his analysis on skylight properties in 1871.[21] However, at that time he was not yet Lord Rayleigh. He rose into peerage and became 3rd Baron Rayleigh later, in 1873, at the death of his father.

[e]Here the time-harmonic convention $\exp(\mathrm{j}\omega t)$ is used.

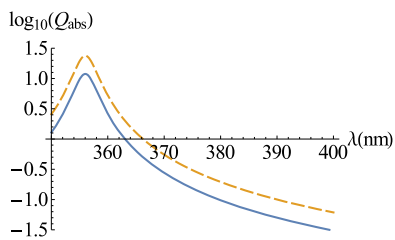


(a) Quasi-static prediction

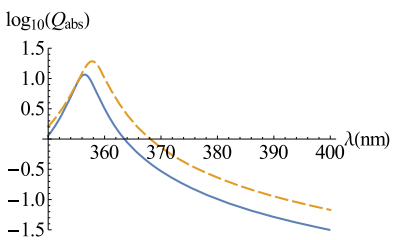


(b) Full Mie analysis

Fig. 1.4. Comparison of the scattering efficiency of a silver nanosphere (with diameter 10 nm (solid blue line), 20 nm (dashed orange line)), calculated from the polarizability approximation (quasistatic) and full-wave solution.



(a) Quasi-static prediction



(b) Full Mie analysis

Fig. 1.5. The same as in Fig. 1.4, for the absorption efficiency $Q_{\rm abs}$.

absorption efficiencies give the following connection to the polarizability:

$$Q_{\rm sca} \approx \frac{8}{27}|\alpha|^2 x^4, \quad Q_{\rm abs} \approx \frac{4}{3}\alpha'' x, \tag{1.12}$$

where α'' is the (negative of)[f] the imaginary part of the polarizability. The static approach explains very well qualitatively the scattering absorption behavior for 10 nm and 20 nm diameter silver spheres. The only feature that the static computation ignores is the slight redshift that can be seen in the Mie solution in Figs. 1.4 and 1.5

[f]For the time-harmonic notation $\exp(\mathrm{j}\omega t)$, the imaginary part of the polarizability is negative for lossy (dissipative) media.

for the 20 nm sphere. Note also the dominance of absorption over scattering in this small-scatterer limit.

1.3.1.3. *Non-spherical particles*

Although the polarizability-based static treatment does not predict the redshifting due to the size dependence, it can astonishingly well explain the effects of shape and structure on plasmonic peaks.

A natural direction to examine the effect of morphology on plasmonic resonances is to allow the shape of the nanoparticle to be ellipsoidal. The advantage is the availability of a rather simple closed-form solution for the ellipsoid in a static field. The geometry of an ellipsoid is defined by its three semi-axes a_x, a_y and a_z. From these, the so-called depolarization factors can be calculated: the factor N_x (in the a_x-direction) is

$$N_x = \frac{a_x a_y a_z}{2} \int_0^\infty \frac{ds}{(s + a_x^2)\sqrt{(s + a_x^2)(s + a_y^2)(s + a_z^2)}}. \tag{1.13}$$

The other depolarization factor N_y (N_z), results by interchange of a_y and a_x (a_z and a_x) in the above integral.

The three depolarization factors for any ellipsoid satisfy

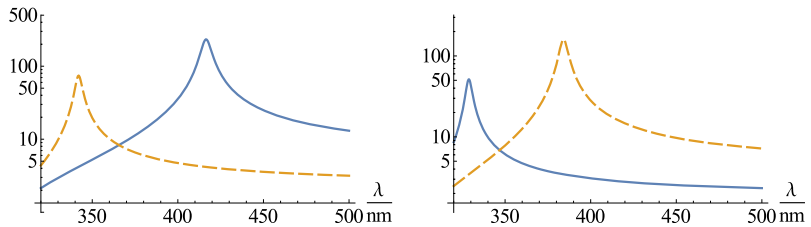
$$N_x + N_y + N_z = 1. \tag{1.14}$$

A sphere has three equal depolarization factors of $1/3$. The other two special cases are a disc (depolarization factors $1, 0, 0$) and a needle $(0, 1/2, 1/2)$.

The three components of the polarizability dyadic of an isotropic ellipsoid with permittivity ε are then ($i = x, y, z$)

$$\alpha_i = \frac{\varepsilon - 1}{1 + N_i(\varepsilon - 1)}. \tag{1.15}$$

This relation, a generalization of Eq. (1.4), reveals that the localized plasmon resonance is very strongly affected by the shape of the nanoparticle: the singularity (zero of the denominator in Eq. (1.15))



(a) Prolate spheroid. (b) Oblate spheroid.

Fig. 1.6. The magnitude of the polarizability components of prolate and oblate spheroids (ellipsoids of revolution, with axis ratios 2:1:1 and 1:2:2). Solid blue line: polarizability along the axis of revolution, dashed orange line: along the two transversal axes. Silver model according to equation (1.11) is used.

takes place at

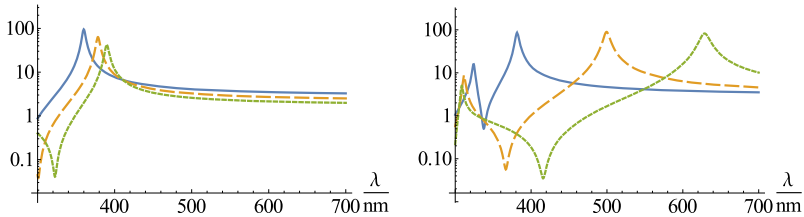
$$\varepsilon_{\mathrm{lp}} = 1 - \frac{1}{N_i} \tag{1.16}$$

and since N_i can vary between 0 and 1, a nanoparticle with any negative permittivity value can resonate, given a suitable shape. In particular, when N_i decreases from the sphere value 1/3, the value $\varepsilon_{\mathrm{pl}}$ decreases (more negative values than -2), meaning a redshift of the plasmonic peak. This happens for an elongated structure (the depolarization factor is zero for a needle when the electric field is oriented along the axis). Indeed, this effect of redshifting is known for nanorods.[24]

In Fig. 1.6, the polarizability components of two silver nanoparticles of ellipsoidal shape are displayed. Two special cases of spheroids (ellipsoids of revolution) are given: a prolate (the axis of revolution is longer than the two transversal axes) and an oblate spheroid (axis of revolution is shorter than the two others). For the prolate (needle-like) particle, there is a strong redshift for the plasmon resonance for axial excitation, and a blueshift for the transversal ones; in the case of oblate spheroid the situation is the opposite.

1.3.1.4. *Core–shell structures*

Another much studied and interesting nanoparticle geometry a core–shell structure. There a spherical core is surrounded by a concentric



(a) Plasmonic core, dielectric shell (b) Dielectric core, plasmonic shell.

Fig. 1.7. The magnitude of the polarizability components of complementary core–shell composite spheres. The plasmonic material is silver (dispersion model according to Eq. (1.11)), and the dielectric is silica (dispersionless permittivity in the optical region 2.25). The fractional volume of silver is g. Solid blue line: $g = 0.9$, dashed orange line: $g = 0.5$, dotted green line: $g = 0.3$.

spherical shell of different material than the core. The polarizability of such a two-layer particle obeys Eq. (1.2), with ε_1 and ε_2 being the permittivities of the shell and the core, respectively, and $0 < g < 1$ is the fractional volume that the core takes from the whole scatterer.

It is fascinating that plasmonic core–shell particles behave qualitatively very differently depending on whether the core or shell is of negative permittivity. This effect is shown in Fig. 1.7. When a plasmonic core is surrounded by a dielectric shell (here taken as silica with permittivity 2.25), the effect is again a redshift. The strength of the shift is dependent on the permittivity of the dielectric shell, and its relative thickness; the shift increases with both. This fact allows tunability of the localized plasmon resonance properties. Experimental research has shown[25] that, by controlling the annealing conditions in the growth process of a bulk nanoplasmonic, Bi_2O_3–Ag eutectic-based metamaterial, the spectral position, width and intensity of the resonance can be modified.

However, for the complementary case (dielectric core surrounded by a plasmonic shell), two resonances appear, due to the fact that there are now two separate surfaces of the metallic phase.[g] The magnitude of these effects, visible in Fig. 1.7, can be very drastic: again

[g]Sometimes this phenomenon of excitation of two resonances in the core–shell structure is called *mode hybridization*.[26]

very strong redshifts are to be expected if the plasmonic layer is very thin. This has also been observed experimentally for hollow gold nanospheres.[27]

Note that these morphological redshifting effects are much stronger than the shift in Fig. 1.4 which was due to the increasing size of the scatterer. In order to have further degrees of freedom in engineering the spectrum of nanoparticles, these effects can also be combined. Hence, an ellipsoidal core–shell inclusion displays an even stronger redshift than a solid ellipsoid or spherical core–shell particle, when the electric field is along the long axis. Also the size of these more complex inclusions affects the position, linewidth and strength of the plasmonic peaks. Furthermore, other types of shaping the inclusion geometry can have a strong effect on the resonances: for example, a transformation of the sphere towards a cube with rounded corners, computational models predict redshifting of the dipole resonance.

1.3.2. *Translation of dispersion results into mixtures*

Homogenization affects dispersion. For example, a mixture where Drude-type particles are embedded in neutral dielectric matrix displays a resonating response. Assuming an ordered lattice of inclusions (and hence no percolation), isolated conducting inclusions do not make the continuum conducting, and the effective permittivity becomes asymptotically real valued for low frequencies. However, a strong response appears at a frequency determined by the parameters of the components, and especially the shape of the inclusions.

In more detail, for a mixture with spherical inclusions, it follows that if the inclusions follow the Drude dispersion

$$\varepsilon(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2 - \mathrm{j}\omega\gamma} \tag{1.17}$$

and the host dielectric has permittivity of ε_d, the mixture is Lorentzian

$$\varepsilon_{\mathrm{eff}}(\omega) = \varepsilon_{\infty,\mathrm{eff}} + \frac{\omega_{p,\mathrm{eff}}^2}{\omega_{0,\mathrm{eff}}^2 - \omega^2 + \mathrm{j}\omega\gamma_{\mathrm{eff}}}, \tag{1.18}$$

with the effective high-frequency permittivity, plasma frequency, the resonance frequency and damping amplitude as

$$\varepsilon_{\infty,\mathrm{eff}} = \varepsilon_d + 3p\varepsilon_d \frac{\varepsilon_\infty - \varepsilon_d}{\varepsilon_\infty + 2\varepsilon_d - p(\varepsilon_\infty - \varepsilon_d)}, \tag{1.19}$$

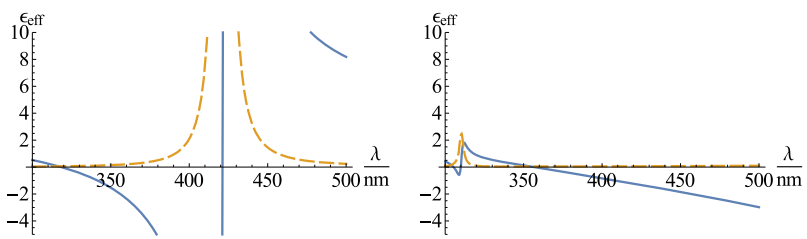
$$\omega_{p,\mathrm{eff}} = \sqrt{p}\,\frac{3\varepsilon_d}{(1-p)\varepsilon_\infty + (2+p)\varepsilon_d}\,\omega_p, \tag{1.20}$$

$$\omega_{0,\mathrm{eff}} = \sqrt{\frac{1-p}{(1-p)\varepsilon_\infty + (2+p)\varepsilon_d}}\,\omega_p, \tag{1.21}$$

$$\gamma_{\mathrm{eff}} = \gamma. \tag{1.22}$$

However, an intuitive understanding of the effects appearing in mixing can be gathered from the equivalence between the effective permittivity of a core–shell particle and the homogenization of a mixture *à la* MG (cf. Section 1.2). This opens up an interesting possibility to transfer the composite nanoparticle results into the domain of plasmonic mixtures.

For example, the qualitative difference between the spectra of the complementary core–shell particles in Fig. 1.7 means that also the mixture of silver spheres in dielectric behaves very differently from the Swiss cheese type of mixture where spherical holes occupy a given volume fraction in silver matrix. This is illustrated in Fig. 1.8 where a 50–50 mixture of silver and air is homogenized in two settings.



(a) Silver spheres in air. (b) Air holes in silver matrix.

Fig. 1.8. The effective permittivity (solid blue line: real part, dashed orange line: imaginary part) of a mixture of silver and air. The volume fraction of silver is 0.5. The model Eq. (1.11) is used. Note the drastic difference of the curves (raisin pudding in (a) and Swiss cheese in (b)) when plotted in the same scale.

The case of silver inclusions in air displays a very strong Lorentzian resonance, according to equations Eq. (1.19)–(1.22), whereas the complementary holey structure obeys a much smoother dispersion. For example, to achieve an $\varepsilon = -1$ composite with low losses and broader bandwidth, the Swiss-cheese structure is hence much more desirable than the complementary "raisin-pudding" design. In Ref.,[22] this phenomenon is exploited in the conceptual design of a broadband subwavelength lens.

Another example of the effect of mixing on the frequency dependence of permittivity is homogenization of materials with relaxation-type dispersion. Materials with permanent dipole moments, like water, display this kind of dispersion, often termed as Debye dispersion

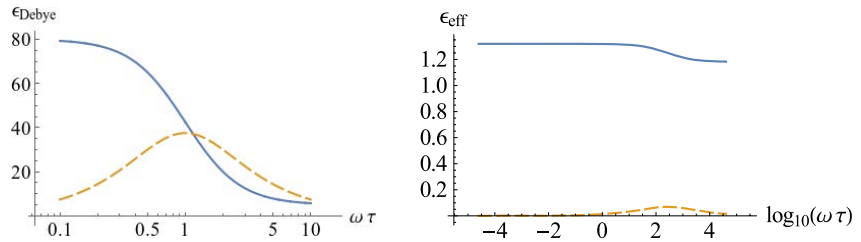
$$\varepsilon(\omega) = \varepsilon + \frac{\varepsilon_s - \varepsilon_\infty}{1 + \mathrm{j}\omega\tau} \tag{1.23}$$

with ε_s and as ε_∞ the low- and high-frequency permittivities, respectively, and τ the relaxation frequency. For Debye medium, there is a strong decrease of the real part of the permittivity around the relaxation frequency $\omega_{\mathrm{rel}} = 1/\tau$ where losses are also strong; however the peak in the imaginary part is not as sharp as for Lorentz-dispersion around the resonance.

For a mixture where Debye-medium particles are embedded into dispersionless host medium, also the homogenized continuum follows exactly the Debye dispersion law, but the model parameters change. In particular, the relaxation frequency takes place at a different region:

$$\omega_{\mathrm{rel,eff}} = \omega_{\mathrm{rel}} \frac{(1-p)\varepsilon_s + (2+p)\varepsilon_e}{(1-p)\varepsilon_\infty + (2+p)\varepsilon_e} \tag{1.24}$$

with ε_e as the host medium permittivity and p the volume fraction of the Debye phase. The frequency shift be very considerable: for example, water at room temperature has values $\varepsilon_s \approx 80.1$, $\varepsilon_\infty \approx 4.9$, and $\tau \approx 10.1$ ps, and this means that the absorption peak of a cloud or rain is shifted from 16 GHz to 200 GHz! Figure 1.9 illustrates this effect.



(a) Bulk Debye medium. (b) Debye spheres in air.

Fig. 1.9. Debye dispersion for water and water-in-air mixture: real (solid blue) and imaginary (dashed orange) part of the effective permittivity. Water parameters: $\varepsilon_s = 80.1$, $\varepsilon_\infty = 4.9$. In (b), the volume fraction of water is 0.1.

If the inclusions are ellipsoidal, an even more drastic relaxation shift can happen. For a mixture with ellipsoids (permittivity ε_i and volume fraction p) with depolarization factor N in the direction of the electric field, the effective permittivity component reads[h]

$$\varepsilon_\mathrm{eff} = \varepsilon_\mathrm{e} + p\varepsilon_\mathrm{e}\frac{\varepsilon_\mathrm{i} - \varepsilon_\mathrm{e}}{\varepsilon_\mathrm{e} + (1-p)N(\varepsilon_\mathrm{i} - \varepsilon_\mathrm{e})}. \tag{1.25}$$

Then the relaxation frequency becomes

$$\omega_\mathrm{rel,eff} = \omega_\mathrm{rel}\frac{(1-p)N(\varepsilon_s - \varepsilon_e) + \varepsilon_e}{(1-p)N(\varepsilon_\infty - \varepsilon_e) + \varepsilon_e} \tag{1.26}$$

and if $N \to 1$, the shift increases from the sphere case $N = 1/3$. This corresponds to flat discs with the field excitation normal to the face of the disc, like in the case of falling flattened raindrops in a vertically polarized radio beam.

However, if the ellipsoids in the mixture are not aligned equally but are oriented with a certain distribution, the effective permittivity dyadic has to be computed by weighing the polarizabilities with the orientation distribution. If all ellipsoids are randomly oriented, there is no special direction and the mixture is isotropic. Then the effective

[h]Note that the mixture is anisotropic in this case.

permittivity follows the formula

$$\varepsilon_{\text{eff}} = \varepsilon_{\text{e}} + \varepsilon_{\text{e}} \frac{\frac{p}{3} \sum_{j=x,y,z} \frac{\varepsilon_{\text{i}} - \varepsilon_{\text{e}}}{\varepsilon_{\text{e}} + N_j(\varepsilon_{\text{i}} - \varepsilon_{\text{e}})}}{1 - \frac{p}{3} \sum_{j=x,y,z} \frac{N_j(\varepsilon_{\text{i}} - \varepsilon_{\text{e}})}{\varepsilon_{\text{e}} + N_j(\varepsilon_{\text{i}} - \varepsilon_{\text{e}})}}. \tag{1.27}$$

Finally, the correspondence between the internal and external homogenization principles allows transferring these relaxation-shift results into the analysis of dielectric behavior of composite particles where a Debye-dispersive core is surrounded by a dielectric layer.

1.4. Homogenization is not simple arithmetic

The previous section emphasized the marriage of dispersion and mixing and presented examples where offsprings were created that display a new type of dispersive behavior. Aside from the transformation of frequency dependence in homogenization, other unexpected behavior can also result from the mixing process produces in the effective permittivity as function of the structural parameters.

To compute the macroscopic mass density of a mixture is very easy: it is the volume averages of the component densities. The macroscopic permittivity, on the contrary, is no simple arithmetic average of the component permittivities. Rather, it can be and often is a very nonlinear[i] function of the fractional volume of the inclusions. It is even possible to construct a composite whose effective permittivity is higher than any of the component permittivities.[28] This phenomenon has been known for a long time in other fields of engineering and materials science: eutectics[29] may allow an alloy to have superior properties over its ingredients. For example, a mixture of lead and tin has lower melting temperature than pure lead or

[i]Nonlinearity refers here to the functional dependence of the effective permittivity on the primary parameters (volume fractions, component permittivities), not to a nonlinearity in the sense that the permittivity depends on the amplitude of the exciting electric field.

pure tin. The "whole being greater than the sum of its parts" is in harmony with the metamaterials ethos.

1.4.1. *Polarization limits*

Mixtures exist whose macroscopic properties are fairly predictable and non-surprising. This is the case especially for composites that consist of isotropic and loss-less phases with permittivities (real-valued scalars) close to each other. Even the morphology of the microstructure does not play a large role in this case. However, when the contrast increases, predictions start to differ.

A random dielectric mixture does not have a unique effective permittivity. Taking samples with a fixed amount of inclusion particles, each sample is different even if the fractional composition is the same. Hence, the effective permittivities are also different. Bounds, however, exist between which all values have to lie. By Monte Carlo simulations, one can create virtual random samples with distribution of inclusions in host space. In the improbable case that all inclusions would cluster as parallel plates through the continuum, the medium would be extremely anisotropic, with two distinct values for the two perpendicular field excitations

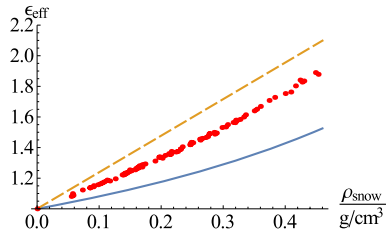
$$\varepsilon_{\mathrm{eff,max}} = \frac{\varepsilon_{\mathrm{i}}\varepsilon_{\mathrm{e}}}{p\varepsilon_{\mathrm{e}} + (1-p)\varepsilon_{\mathrm{i}}}, \quad \varepsilon_{\mathrm{eff,min}} = p\varepsilon_{\mathrm{i}} + (1-p)\varepsilon_{\mathrm{e}}. \tag{1.28}$$

These so-called Wiener bounds,[30,31] come naturally from the ellipsoidal mixing rule Eq. (1.25) with depolarization factor choices 1 and 0, respectively.

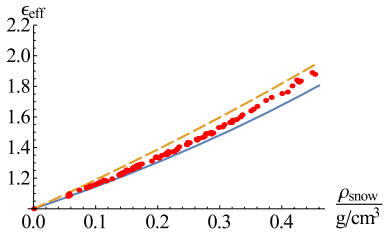
Wiener bounds do not restrict much, and sharper bounds were presented by Hashin and Shtrikman (HS).[32] In fact, the lower HS limit is exactly the MG rule Eq. (1.4), and the upper limit is the complementary MG formula where the roles of the inclusion and the host are reversed:[j]

$$\varepsilon_{\mathrm{HS,max}} = \varepsilon_{\mathrm{eff,MG}}(\varepsilon_{\mathrm{i}}, \varepsilon_{\mathrm{e}}, p), \quad \varepsilon_{\mathrm{HS,min}} = \varepsilon_{\mathrm{eff,MG}}(\varepsilon_{\mathrm{e}}, \varepsilon_{\mathrm{i}}, 1-p). \tag{1.29}$$

[j]This is valid if the inclusion permittivity is larger than that of the environment ($\varepsilon_{\mathrm{i}} > \varepsilon_{\mathrm{e}}$). For the "Swiss cheese" case ($\varepsilon_{\mathrm{i}} < \varepsilon_{\mathrm{e}}$), the limits are interchanged.



(a) Measurements within Wiener bounds.



(b) Measurements within HS bounds.

Fig. 1.10. Bounds for a dielectric mixture of air and ice, modelling dry snow. Permittivity as function of snow density ρ. The volume fraction of ice is $p = \rho/\rho_{\text{ice}}$ where the ice density is $\rho_{\text{ice}} = 0.917\,\text{g/cm}^3$. The permittivity if ice is 3.19. Also shown measurement data by Mätzler[36] (red points).

HS bounds are not the strictest ones. Milton,[33] Felderhof[34] and Helsing[35] have published even sharper limits.

As an example of the applicability of the bounds, let us focus on snow. Dry snow is a mixture of ice grains (permittivity around 3.19 at microwave frequencies) in air matrix. Wiener and HS limits for snow permittivity as function of snow density are shown in Fig. 1.10 with the field measurements data by Christian Mätzler.[36] These data are gathered using a resonator instrument operating near 1 GHz. However, due to the fact that ice is fairly dispersionless in UHF and microwave wavelengths, these dry snow permittivity values apply within a broad frequency range, from 1 MHz up to 10 GHz.

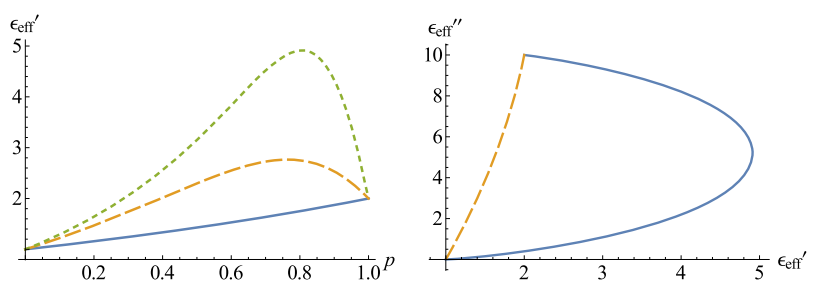
The curves in Fig. 1.10 seem to be fairly loose limits for the experimental data, which is quite astonishing considering the fact that snow is rather coarse-grained and the length scale of the microstructure is not very much smaller than the dimension of the volume over which the measuring instrument averages (maximum dimension tens of centimeters).[36] These nature-generated snow samples represent most probable states of the snow distribution, and therefore there should be no need to care for very improbable system states, like, for example, that all snow grains would suddenly be clustered in a very ordered structure.

To take an analogy to thermodynamics, the distributions of molecules in a sample of gas (their positions and velocities) are

always very sharply peaked. Average values of the sample parameters are practically indistinguishable from the most probable value. For example, to describe the state of a glass of water, around 10^{24}–10^{25} coordinates are required.[37] In the macroscopic specification of thermodynamic systems, the scales to be considered are temporal: one needs to compare the macroscopic measurement slowness to the rapidity of atomic movements. Compared to the case of dielectric mixtures, these scales are extremely distant from each other. For the example of the snow permittivity, the scales of sample and the "microscopic" level are much closer to each other, allowing for larger deviations from the expected values.

The snow permittivity behaves in a quite "dull" manner in Fig. 1.10, monotonically and predictably increasing towards the permittivity of ice. However, if the dielectric response becomes more complicated, the behavior becomes more metamaterial-like, and bounds and limits can be overcome. This happens when strong losses are present or any cross-coupling effects are allowed, like magneto-electric response.

Figure 1.11 illustrates the way losses in the component materials may cause a special enhancement in the real part of the effective permittivity of a mixture. The curves show a substantial polarization



(a) Real part of effective permittivity for increasing losses.

(b) HS limits for the case $\varepsilon_{\mathrm{i}} = 2 - \mathrm{j}\,10$ and $\varepsilon_{\mathrm{e}} = 1$.

Fig. 1.11. (a) MG prediction for the real part of the effective permittivity when the losses of the inclusion phase increase $\varepsilon_{\mathrm{i}} = 2 - \mathrm{j}\,0$ (solid blue), $\varepsilon_{\mathrm{i}} = 2 - \mathrm{j}\,5$ (dashed orange), and $\varepsilon_{\mathrm{i}} = 2 - \mathrm{j}\,10$ (dotted green). The environment is air $\varepsilon_{\mathrm{e}} = 1$. (b) The HS limits in the complex $\varepsilon_{\mathrm{eff}}$ plane.

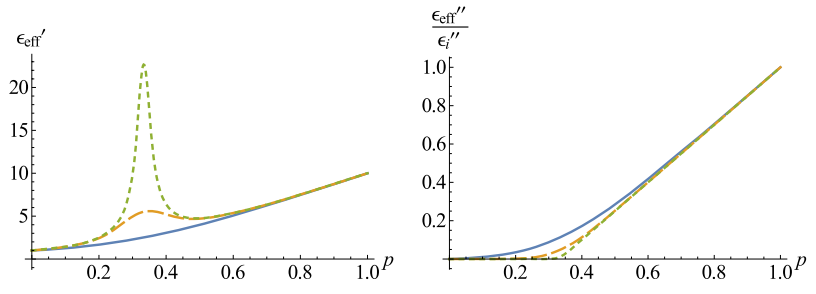
enhancement effect, in other words the real part of the effective permittivity $\varepsilon'_{\text{eff}}$ of a mixture can be much larger than the real part of the permittivity of either of the components, when sufficient losses are introduced into the inclusion phase. If other mixing formulas are applied, the prediction for the enhancement magnitude changes, but it is still there.[28] Figure 1.11 also shows that the HS bounds (MG and its complement), in the complex ε_{eff} plane allow quite a wide range of permissible values, and all the other mixing rules predict ε_{eff} values that are located in this restricted domain.

It is important to note that the enhancement effect highlighted in Fig. 1.11 is by no means any modelling artefact. It is observed in many real-world constellations, like in colloidal and biological cell-suspensions.[38] In connection with insulating membranes that separate electrolyte solutions, the phenomenon is also called Maxwell–Wagner effect,[39] and in the language of bioelectromagnetics, it is known as alpha dispersion, taking place at low frequencies (10 Hz to a few kHz).

1.4.2. *Percolation behavior*

For mixtures where strongly contrasting permittivity phases are interacting with each other, percolation[40] is also often present. Percolation is a phenomenon where a minute variation in the structure or geometrical parameters of a mixture causes a notable and abrupt change in macroscopic properties. In the context of dielectric mixtures, this means a very non-linear functional dependence of the effective permittivity on the volume fractions of the phases.

Percolation is conspicuous for mixtures with high losses. Figure 1.12 shows the prediction of the Bruggeman mixing formula of a composite where lossy spherical inclusions are embedded in free space. Indeed, the macroscopic permittivity curve is no longer smooth, rather it has a nearly discontinuous character when the inclusion losses increase. The imaginary part $\varepsilon''_{\text{eff}}$ is practically zero until the so-called percolation threshold, here happening at $p = 1/3$. There it linearly starts to increase towards the the value ε''_{i}. The real part $\varepsilon'_{\text{eff}}$, on the other hand experiences a very strong enhancement



(a) $\varepsilon'_{\text{eff}}$. (b) $\varepsilon''_{\text{eff}}$.

Fig. 1.12. Bruggeman prediction for the real part of the effective permittivity when the losses of the inclusion phase increase $\varepsilon_{\text{i}} = 10 - \text{j}\,1$ (solid blue), $\varepsilon_{\text{i}} = 10 - \text{j}\,100$ (dashed orange), and $\varepsilon_{\text{i}} = 10 - \text{j}\,2000$ (dotted green). The environment is air $\varepsilon_{\text{e}} = 1$.

around the threshold, decreasing again when the inclusion volume fraction increases.

It is worth noting that predictions of various mixing rules for the details in percolation behavior vary, although the phenomenon itself is there. The unified mixing rule Eq. (1.7) shows this clearly: assuming a very high inclusion-to-background permittivity ratio, this formula distills down to

$$\varepsilon_{\text{eff}} = \varepsilon_{\text{e}} \frac{1 + p(2 - \nu)}{1 - p(1 + \nu)}, \tag{1.30}$$

which breaks down when $p = 1/(1 + \nu)$. This fraction can be interpreted as percolation threshold, and for the Bruggeman case ($\nu = 2$), the threshold is $1/3$ of which agrees with the example in Fig. 1.12. Bruggeman mixing rule by its symmetric character is widely used and applicable to random mixtures, as opposed to ordered lattices which are more close to the domain of MG principle. The percolation threshold for MG ($\nu = 0$) is pushed up to $p = 1$.

1.4.3. *Magnetoelectric coupling*

Among interesting complex electromagnetic material responses is magnetoelectric coupling. Electric excitation creates magnetic response and vice versa. If such a medium is isotropic (the response

is independent on the vector direction of the excitation field), we talk about *bi-isotropic* materials.[41] The constitutive relations of bi-isotropic media are compactly written by a matrix that transmits the relation between the electric and magnetic fields $(\mathbf{E}, \mathbf{H})$ and flux densities $(\mathbf{D}, \mathbf{B})$:

$$\begin{pmatrix} \mathbf{D}/\sqrt{\varepsilon_0} \\ \mathbf{B}/\sqrt{\mu_0} \end{pmatrix} = \begin{pmatrix} \varepsilon & \chi - \mathrm{j}\kappa \\ \chi + \mathrm{j}\kappa & \mu \end{pmatrix} \begin{pmatrix} \sqrt{\varepsilon_0}\,\mathbf{E} \\ \sqrt{\mu_0}\,\mathbf{H} \end{pmatrix} = \mathcal{M} \begin{pmatrix} \sqrt{\varepsilon_0}\,\mathbf{E} \\ \sqrt{\mu_0}\,\mathbf{H} \end{pmatrix}, \quad (1.31)$$

where the fields and displacements have been normalized with the free-space permittivity and permeability (ε_0, μ_0) to carry the same units (square root of energy density $\sqrt{\mathrm{Ws/m^3}}$). Hence, the material matrix $\mathcal{M}$ becomes dimensionless. Cross-coupling is quantified by the chirality parameter κ and the non-reciprocity parameter χ. Chirality parameter appears in handed media, and κ is often termed *Pasteur* parameter.[41] The non-reciprocity parameter χ goes often under the name *Tellegen* parameter.[42]

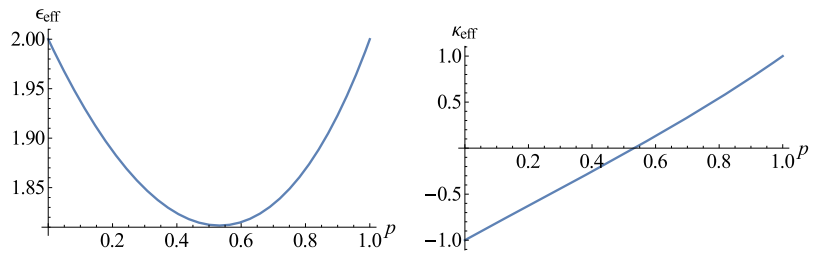
For mixtures involving bi-isotropic media, the mixing principles can be taken from the simple non-magnetoelectric domain by replacing permittivities with the material matrices that include the four parameters.[43] For example, the bi-isotropic MG mixing formula reads

$$\mathcal{M}_{\mathrm{eff}} = \mathcal{M}_e + 3p\mathcal{M}_e \cdot [\mathcal{M}_i + 2\mathcal{M}_e - p(\mathcal{M}_i - \mathcal{M}_e)]^{-1} \cdot (\mathcal{M}_i - \mathcal{M}_e), \quad (1.32)$$

with the bi-isotropic components of the environment included in $\mathcal{M}_e$, those of the inclusion in $\mathcal{M}_i$, and again p being the volume fraction of the inclusion phase. All parameters will be coupled through the matrix inversion operation in the formula, but the effective parameters can be written as explicit functions of the primary parameters.

As an example, consider a mixture of two chiral phases which are otherwise identical but are structurally mirror images of each other. This means that their permittivities and permeabilities are the same ($\varepsilon_i = \varepsilon_e$, $\mu_i = \mu_e$), and the chirality parameters are opposite numbers $\kappa_i = -\kappa_e$.

Figure 1.13 shows the effective permittivity and chirality parameters for the case $\varepsilon_i = \varepsilon_e = \mu_i = \mu_e = 2$, $\kappa_i = -\kappa_e = 1$. Even



(a) Effective permittivity. (b) Effective chirality.

Fig. 1.13. Effective parameters of chiral-in-chiral mixture where otherwise identical but mirror-image phases are mixed. MG prediction, the parameters are $\varepsilon_i = \varepsilon_e = 2$, $\mu_i = \mu_e = 2$, $\kappa_i = -\kappa_e = 1$.

if both phases have the same permittivity, the effective permittivity is lower than that of the components (the same is valid for the effective permeability). Furthermore, the chirality parameter changes monotonously from the environment value to that of the inclusion, although it is not symmetric: for the $p = 0.5$ mixture, the chirality, although quite weak, does not completely vanish.

1.4.4. *Hyperbolic media and anomalous losses*

Anisotropy is one of the interesting directions of complex response in materials and mixtures. In particular, when anisotropy is connected with plasmonics, the resulting material effects can be particularly fascinating in the dielectric response.

1.4.4.1. *Cartesian and spherical uniaxial anisotropy*

Structural alignment leads to anisotropy in the macroscopic dielectric response, and the permittivity has to be quantified by more parameters than a single scalar. In the simplest form, anisotropy is uniaxial. For uniaxial media, the permittivity dyadic has two components: axial (ε_r) and transversal (ε_t) permittivity. Two composite structures that behave macroscopically in this manner are shown in Fig. 1.14 where the axial direction is normal to the plates in the stacked-plates structure and along cylinder axes in the aligned-needles structure.

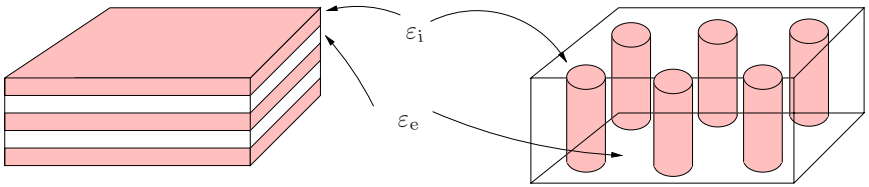


Fig. 1.14. Uniaxial two-phase composite structures: alternating stacked plates and aligned circular cylinders.

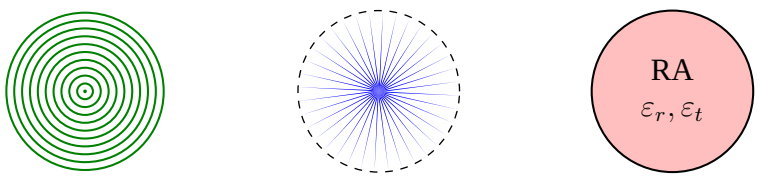


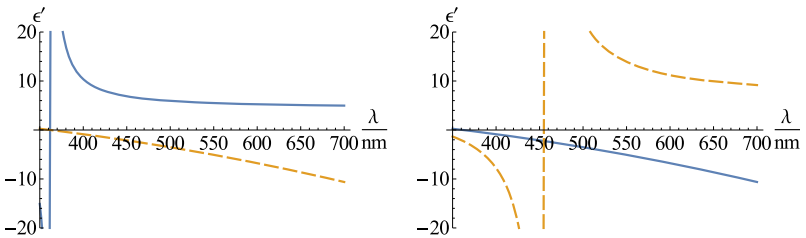
Fig. 1.15. Bulbic (onion-like) and porcupic (radially spiked) composite spheres are effectively RA.

Instead of the Cartesian setting of Fig. 1.14, the uniaxial anisotropy can also be radial and spherically symmetric. This is the case shown in Fig. 1.15 which displays inclusions that are called RA spheres. The two types of RA sphere constructions are a *bulbic* sphere (onion-like layers) and a *porcupic* sphere (porcupine spikes).[44,k]

Homogenization of bulbic and porcupic uniaxial structures is straightforward once the inhomogeneities are truly subwavelength: all four components are special cases of the aligned ellipsoid mixing formula Eq. (1.25). For the stacked-plates/bulbic structure, the axial permittivity corresponds to depolarization factor $N = 1$ and the transversal one to $N = 0$. On the other hand, the aligned needles/porcupic design leads to $N = 0$ for axial component and $N = 1/2$ in the transversal direction.

The cases $N = 1$ and $N = 0$ give the Wiener lower and upper bounds Eq. (1.28), which consequently means that for the bulbic

[k]An even more general spherical anisotropy results by allowing the tangential anisotropy to have different permittivity components in the two perpendicular angular directions. Such a sphere is called *systropic*.[45]



(a) Layered structure (bulbic). (b) Needle-type structure (porcupic).

Fig. 1.16. Silver–silica composite with uniaxially anisotropic structure. Volume fractions of both phases are equal. Solid blue line: axial; dashed orange line: transversal component of the real part of the permittivity. The model equation (1.11) is used for silver, and the permittivity of silica is assumed 2.25.

structure, the axial permittivity is smaller than the transversal one ($\varepsilon_r < \varepsilon_t$, negative uniaxiality), and vice versa for the porpupic case ($\varepsilon_r > \varepsilon_t$, positive uniaxiality).

However, when one of the phases in bulbic or porcupic composites is plasmonic, the axial and transversal effective permittivity components may also change sign. In the particular case that these two eigenvalues of the permittivity dyadic have different signs, the medium can be called *indefinite* or *hyperbolic*.[46] This effect is illustrated in Fig. 1.16 where the components of the (real part of the) permittivity dyadic are shown as a function of wavelength for a silver–silica composite (both phases occupy 50% of the volume). It can be seen that in both designs, a broadband hyperbolic behavior appears. The difference in the parameters is that in the bulbic case the transversal permittivity is negative, while in the porcupic constellation, it is the axial permittivity component which is negative.

1.4.4.2. *Radially anisotropic sphere*

RA spheres have been shown to display particularly interesting properties[47,48] due their cloaking potential, and anomalous absorption cross-section. Again, these properties can be studied based on purely quasi-static analysis.

The RA sphere is complicated from inside: it is inhomogeneous (the eigenaxes of the permittivity are not constant) and anisotropic,

but looking from outside, it is completely spherically symmetric and effectively isotropic. Hence, its polarizability is a scalar.[1]

A quasi-static analysis gives the polarizability of an RA sphere as follows:

$$\alpha_{\mathrm{RA}} = 3\frac{\varepsilon_r\sqrt{1+8\varepsilon_t/\varepsilon_r}-\varepsilon_r-2}{\varepsilon_r\sqrt{1+8\varepsilon_t/\varepsilon_r}-\varepsilon_r+4}. \tag{1.33}$$

This degenerates into Eq. (1.1) in the isotropic case $\varepsilon_r = \varepsilon_t$. For $\varepsilon_r \neq \varepsilon_t$, several interesting observations can be drawn.

First of all, α_{RA} vanishes when

$$\varepsilon_t = \frac{\varepsilon_r+1}{2\varepsilon_r}, \quad \varepsilon_r > 0 \quad \text{or} \quad \varepsilon_r < -2 \tag{1.34}$$

meaning that a material sphere can be invisible for a static electric field. In the positive-permittivity side, this makes sense intuitively: a combination of ε_r and ε_t values, one larger, one smaller than $+1$, can compensate each others' effect and make the sphere disappear. The internal homogenized sphere has $\varepsilon_{\mathrm{eff}} = 1$. What is not obvious is that the polarizability can also vanish in the region where ε_r is negative but ε_t positive (region of indefinite anisotropy). In that case, there is a very strong field concentration in the center of the scatterer, even if the sphere does not cause any perturbation to the external field.

On the other hand, the singularity of α_{RA} gives us the plasmonic dipole resonances. It is interesting that there are two types of singularities. The condition for singularity is

$$\varepsilon_t = \frac{2-\varepsilon_r}{\varepsilon_r}, \quad \varepsilon_r < 0 \quad \text{or} \quad \varepsilon_r > +4. \tag{1.35}$$

The first type of resonance is a generalization of the corresponding electrostatic $\varepsilon = -2$ resonance in the isotropic case and corresponds to $\varepsilon_r < 0$ in Eq. (1.35), in other words, in the fully plasmonic ("double-negative") region. The other branch of singularity appears

[1]Radial anisotropy has been generalized to ellipsoidal inclusions.[44] There the "radial" direction has to be understood in the sense of the corresponding dimension in the ellipsoidal coordinate system.

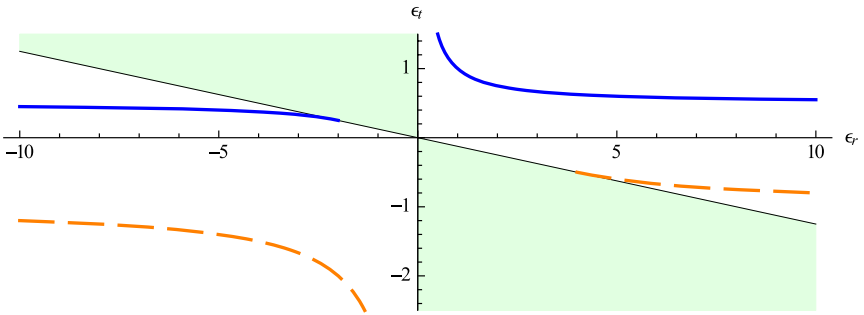
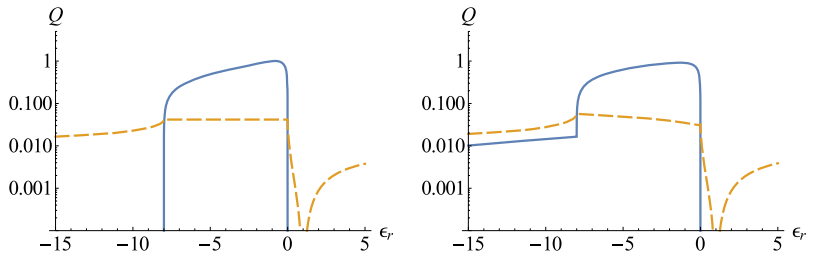


Fig. 1.17. Polarizability characteristics of an RA sphere in the $\varepsilon_r - \varepsilon_t$ plane. Solid blue lines are invisibility contours, dashed orange lines mark the singularity conditions. Light-green shaded area is the anomalous absorption region.

in the indefinite quadrant $\varepsilon_r > 0, \varepsilon_t < 0$. The field distribution is different for these two cases: for the first one, it is the ordinary localized surface plasmon, whereas for the indefinite-region singularity, the field energy is concentrated into the (center) of the RA sphere.

Of particular interest is the domain where $\varepsilon_t/\varepsilon_r < -1/8$. Then the square root in Eq. (1.33) becomes complex. In other words, the RA sphere contains an imaginary part even if its permittivity components are real. This phenomenon has been termed *anomalous absorption* and has been studied in detail in Ref. 48. The analysis requires a regularization of the field behavior in the center of the sphere and concludes that the focusing effect of the RA sphere makes it possible to enhance by an arbitrary amount the effect of intrinsic losses. Figure 1.17 charts the various RA regions in the complex $\varepsilon_r - \varepsilon_t$ plane.

The phenomenon of anomalous absorption by RA spheres can also be studied for the dynamic excitation. In Fig. 1.18, the absorption efficiency of an RA sphere with size parameter $x = 0.5$ is illustrated. The prediction of quasi-static analysis (coming from the imaginary part of Eq. (1.33), with Eq. (1.12)) is compared against the full-wave Mie scattering computation. Even if the sphere is not deeply subwavelength, it can be seen that the scattering cross-section is very well explained by the static analysis, and also the anomalous



(a) Quasistatic prediction. (b) Full Mie analysis.

Fig. 1.18. The absorption (solid blue) and scattering (dashed orange) efficiencies of an RA sphere of size parameter $x = 0.5$. The tangential permittivity component is $\varepsilon_t = 1$ and the radial component ε_r varies. Comparison of the quasi-static prediction and the full anisotropic Mie analysis.

absorption in the region $-8 < \varepsilon_r/\varepsilon_t < 0$. However, the dynamic analysis also reveals additional anomalous absorption, although small in magnitude. This happens for hyperbolic anisotropies with more negative values of ε_r than in the light-green region of Fig. 1.17.

1.5. Conclusions: Metamaterials as Vessels of Emergence

Analysis of metamaterials means transformation of small-scale structure into a picture on a higher level. New qualities appear on the macroscopic description, but at the cost of sacrificing all structural details and information about the microstructure. Yet, the price may be worth paying since there may be no other way into these new desired properties. Metamaterials are vessels of emergence.

The focus in this chapter has been the use of classical mixing principles in understanding the emergence in metamaterials. The discussion centered on the diversity of the possibilities how homogenization may create new effective dielectric responses in mixtures and composites. The mixtures consist of elements that are made of ordinary material and whose dielectric responses follow known dielectric functions, like Drude or Debye models. Mixing laws, even though fairly simple equations of combining the permittivities of the components and very few structural parameters, lead to effective behavior that

is qualitatively different from those of the component phases. Here, there is a certain analogy to the world of cellular automata[49] where a simple rule can generate very complex behavior which cannot be directly seen from the initial conditions. There is no shortcut: the only way to know the end state is to computationally go through the evolution, step by step.

What is the domain of metamaterials? Sometimes metamaterials are defined very inclusively: the definition may embrace, for example, assemblies that could be called metamaterial-inspired structures, like antennas in the shape resembling a split-ring. Leaving such single-element designs aside, mainstream metamaterials are such macroscopic continua in which microstructure details produce a collective effect. Then also, homogenization is truly in the core of analysis and understanding of metamaterials. The inclusions and structure that cause the interesting and desired effective behavior is deeply sub-wavelength in spatial scale.[m] This is the domain of mixing principles which are based on quasi-static analyses and solutions of generalized Laplace equation. The power of mixing rules in explaining many macroscopic phenomena contained in the effective permittivity functions was demonstrated by many examples from microwaves to optics in the present chapter. The examples (Sections 1.3.1 and 1.4.4) also shed light on the essential question of how much smaller must the scatterer size (or characteristic length of the microstructure) be compared to the wavelength of the electromagnetic excitation.

Electromagnetic homogenization has a long history. As Christian Brosseau[50,51] has noted, it can be seen as a play in five scenes. The beginnings up to J.C. Maxwell Garnett in the first decade of 20th century lead to the Bruggeman effective-medium homogenization from 1930s. The bounding principles by Hashin, Shtrikman and others in 1960s guided towards comprehension of percolation. Finally, the arrival of powerful computers revolutioned computational electromagnetics and gave unprecedented possibilities

[m]This applies of course to electromagnetic and acoustic metamaterials; in connection with thermal metamaterials one cannot talk about wavelength.

to simulate random mixtures and test predictions of effective-medium theories.

And the show is going on. After these five scenes, we are presently living through and acting in the sixth one: homogenization of metamaterials. Effective-medium theories and mixing principles reveal the richness and variety of unpredicted properties that are possible through arranging building blocks and tailoring their interaction to design novel metamaterial structures.

References

1. Sihvola, A. (2003). Electromagnetic emergence in metamaterials. In eds. S. Zouhdi, A. Sihvola, and M. Arsalane, *Advances in Electromagnetics of Complex Media and Metamaterials*, vol. 89, pp. 1–17 (Kluwer Academic Publishers, Dordrecht).
2. Sihvola, A. (2007). Metamaterials in electromagnetics, *Metamaterials* **1**(1), pp. 2–11.
3. Sihvola, A. (2009). Metamaterials: A personal view, *Radioengineering* **18**(2: Part I: Special Issue on Artificial EM Materials and Metamaterials), pp. 90–94.
4. https://en.wikipedia.org/wiki/Metamaterial, (accessed 27 August 2015).
5. Walser, R. M. (2000). Metamaterials: What Are They? What Are They Good For? Talk at American Physical Society Meeting, Minneapolis, pp. 20–24.
6. Walser, R. M. (2003). Electromagnetic metamaterials. In eds. A. Lakhtakia, W. S. Weiglhofer, and I. J. Hodgkinson, *Complex Mediums II: Beyond Linear Isotropic Dielectrics*, vol. 4467, pp. 1–15. SPIE.
7. Engheta, N. (2008). Metactronics: Optical circuits and information processing in nanoworld. In *Proc. of META'08, NATO Advanced Research Workshop*, p. 533.
8. Lakhtakia, A. (2015). From bioinspired multifunctionality to mimumes, *Bioinspir. Biomim. Nan.* **4**, pp. 168–173.
9. Shamonina, E. and Solymar, L. (2007). Metamaterials: How the subject started, *Metamaterials* **1**(1), pp. 12–18.
10. Sihvola, A. (1999). *Electromagnetic Mixing Formulas and Applications*. vol. 47, *Electromagnetic Waves Series*, (IEE Publishing, London).
11. Milton, G. W. (2002). *The Theory of Composites*. (Cambridge University Press, Cambridge).
12. Chettiar, U. K. and Engheta, N. (2012). Internal homogenization: Effective permittivity of a coated sphere, *Opt. Exp.* **20**(21), pp. 22976–22986.
13. Sihvola, A., Kettunen, H. and Wallén, H. (2013). Mixtures and composite particles: Correspondence of effective description. In *Proc. EMTS 2013, URSI*

International Symposium on Electromagnetic Theory, pp. 908–911, pp. 21–24 May, Hiroshima, Japan.
14. Garnett, J. C. M. (1904). Colours in metal glasses and metal films, *Trans. R. Soc.* **203**, pp. 385–420.
15. Bruggeman, D. A. G. (1935). Berechnung verschiedener physikalischer Konstanten von heterogenen Substanzen. I. Dielektrizitätskonstanten und Leitfähigkeiten der Mischkörper aus isotropen Substanzen, *Ann. Phys.* **24**, pp. 636–679.
16. Sihvola, A. (1989). Self-consistency aspects of dielectric mixing theories, *IEEE Trans. Geosci. Remote Sens.* **27**(4), pp. 403–415.
17. Elliott, R. J., Krumhansl, J. A. and Leath, P. L. (1974). The theory and properties of randomly disordered crystals and related physical systems, *Rev. Mod. Phys.* **46**, pp. 465–543.
18. Bohren, C. F. and Huffman, D. R. (1983). *Absorption and Scattering of Light by Small Particles* (Wiley, New York).
19. Lorenz, L. (1890). Lysbevægelser i og uden for en af plane Lysbølger belyst Kugle, *Kongelige Danske Videnskabernes Selskabs Skrifter.* **6**, pp. 2–62.
20. Mie, G. (1908). Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen, *Ann. Phys.* **25**, pp. 377–445.
21. Strutt, J. W. (1871). On the light from the sky, its polarization and colour, *Philosophical Magazine* **41**, pp. 107–120, pp. 274–279.
22. Wallén, H., Kettunen, H. and Sihvola, A. (2009). Composite near-field superlens design using mixing formulas and simulations, *Metamaterials* **3**(3–4), pp. 129–139.
23. Johnson, P. B. and Christy, R. W. (1972). Optical constants of the noble metals, *Phys. Rev. B.* **6**, pp. 4370–4379.
24. Smythe, E. J., Cubukcu, E. and Capasso, F. (2007). Optical properties of surface plasmon resonances of coupled metallic nanorods, *Opt. Exp.* **15**(12), pp. 7439–7447.
25. Sadecka, K., Toudert, J., Surma, H. B. and Pawlak, D. A. (2015). Temperature and atmosphere tunability of the nanoplasmonic resonance of a volumetric eutectic-based Bi_2O_3-Ag metamaterials, *Opt. Exp.* **23**(15), p. 241792.
26. Sonnefraud, Y., Verellen, N., Sobhani, H., Vandenbosch, G. A., Moshchalkov, V. V. Dorpe, P. V., Nordlander, P. and Maier, S. A. (2010). Experimental realization of subradiant, superradiant, and Fano resonances in ring/disk plasmonic nanocavities, *ACS Nano* **4**(3), pp. 1664–1670.
27. Schwartzberg, A. M., Olson, T. Y., Talley, C. E. and Zhang, J. Z. (2006). Synthesis, characterization, and tunable optical properties of hollow gold nanospheres, *J. Phys. Chem. B* **110**(40), pp. 19935–19944.
28. Sihvola, A. (2002). How strict are theoretical bounds for dielectric properties of mixtures?, *IEEE Trans. Geosci. Remote Sens.* **40**(4), pp. 880–886.
29. Pawlak, D. A., Turczynski, S., Gajc, M., Kolodziejak, K., Diduszko, R., Rozniatowski, K., Smalc, J. and Vendik, I. (2010). Metamaterials: How far are we from making metamaterials by self-organization? The microstructure of

highly anisotropic particles with an SRR-like geometry, *Adv. Funct. Mater.* **20**(7), pp. 1116–1124.
30. Wiener, O. (1904). Lamellare doppelbrechung, *Physikalische Zeitschrift* **5**(12), pp. 332–338.
31. Wiener, O. (1910). Zur theorie der refraktionskonstanten, *Berichte Über die Verhandlungen der Königlich-Sächsischen Gesellschaft der Wissenshaften zu Leipzig* (62), pp. 256–277.
32. Hashin, Z. and Shtrikman, S. (1962). A variational approach to the theory of the effective magnetic permeability of multiphase materials, *J. Appl. Phys.* **33**(10), pp. 3125–3131.
33. Milton, G. W. (1981). Bounds on the complex permittivity of a two component composite material, *J. Appl. Phys.* **52**(8), pp. 5286–5293.
34. Felderhof, B. (1984). Bounds for the complex dielectric constant of a two-phase composite, *Phys. A.* **126**(3), pp. 430–442.
35. Helsing, J. (1993). Bounds to the conductivity of some two-component composites, *J. Appl. Phys.* **73**(3), pp. 1240–1245.
36. Mätzler, C. (1996). Microwave permittivity of dry snow, *IEEE Trans. Geosci. Remote Sens.* **34**, pp. 573–581.
37. Callen, H. B. (1960). *Thermodynamics.* (John Wiley & Sons, New York).
38. Schwan, H. P., Schwarz, G., Maczuk, J. and Pauly, H. (1962). On the low-frequency dielectric dispersion of colloidal particles in electrolyte solution, *J. Phys. Chem.* **66**(12), pp. 2626–2635.
39. Wagner, K. W. (1914). Erklärung der dielektrischen Nachwirkungsvorgänge auf Grund Maxwellscher Vorstellungen, *Archiv für Elektrotechnik* **2**(9), pp. 371–387.
40. Stauffer, D. and Aharony, A. (1994). *Introduction to Percolation Theory.* (Taylor & Francis, London, U.K.).
41. Lindell, I. V., Sihvola, A. H., Tretyakov, S. A. and Viitanen, A. J. (1994). *Electromagnetic Waves in Chiral and Bi-Isotropic Media* (Artech House, Norwood, Massachusetts).
42. Tellegen, B. D. H. (1948). The gyrator, a new electric network element, *Philips Res. Rep.* **3**(2), pp. 81–101.
43. Sihvola, A. H. (1992). Bi-isotropic mixtures, *IEEE Trans. Antennas Propag.* **40**(2), 188–197.
44. Rimpiläinen, T., Wallén, H. and Sihvola, A. (2015). Radial anisotropy in spheroidal scatterers, *IEEE Trans. Antennas Propag.* **63**(7), pp. 3127–3135.
45. Rimpiläinen, T., Pitkonen, M., Wallén, H., Kettunen, H. and Sihvola, A. (2014). General systropy in spherical scatterers, *IEEE Trans. Antennas Propag.* **62**(1), pp. 327–333.
46. Poddubny, A., Iorsh, I., Belov, P. and Kivshar, Y. (2013). Hyperbolic metamaterials, *Nat. Photon.* **7**(12), pp. 948–957.
47. Kettunen, H., Wallén, H. and Sihvola, A. (2013). Cloaking and magnifying using radial anisotropy, *J. Appl. Phys.* 114(4), p. 044110.

48. Wallén, H., Kettunen, H. and Sihvola, A. (2015). Anomalous absorption, plasmonic resonances, and invisibility of radially anisotropic spheres, *Radio Science* **50**(1), pp. 18–28.
49. Wolfram, S. (2002). *A New Kind of Science*. (Wolfram Media, Champaign, Illinois).
50. Brosseau, C. (2006). Modelling and simulation of dielectric heterostructures: A physical survey from an historical perspective, *J. Phys. D: Appl. Phys.* **39**, pp. 1277–1294.
51. Sihvola, A. (2013). Homogenization principles and effect of mixing on dielectric behavior, *Photon. Nanostruct.* **11**(4), pp. 364–373.

CHAPTER 2

Effective Medium Theory of Electromagnetic and Quantum Metamaterials

MÁRIO G. SILVEIRINHA

University of Lisbon, Portugal
mario.silveirinha@co.it.pt

Here, we present an overview of recent developments in the characterization of electromagnetic and quantum metamaterials using effective medium methods. It is highlighted that both electromagnetic and electronic systems can be homogenized in a unified manner based on the introduction of an effective Hamiltonian operator that describes the time evolution of the macroscopic initial states as well as the stationary states of the relevant system. Furthermore, it is shown that in some circumstances quadratic forms of the fields, such as the energy, can be exactly determined using the effective medium theory.

2.1. Introduction

The wave propagation in periodic or in random systems can be rather challenging to study as an exact treatment of problem is generally

unfeasible and a numerical analysis with no approximations requires large-scale time-consuming computations. In addition, brute force numerical calculations provide very limited insights of the physical mechanisms that determine the wave propagation. Fortunately, it is often possible to describe "low-energy" wave phenomena by resorting to an effective medium description that regards the structure of interest as a continuum.

For example, natural materials are formed by a collection of atoms or molecules arranged — in case of crystalline structures — in a periodic lattice. Even within a classical framework, wherein the atoms are regarded as electric or magnetic dipoles, the propagation of light in a natural material can be immensely complex: each atom or molecule scatters and absorbs light, and the propagation problem must be solved self-consistently. Auspiciously, provided the light wavelength is much larger than the lattice constant, i.e. larger than the characteristic spacing between the atoms, the complex light–matter interactions can be homogenized and the material can be regarded as a continuum described by a certain effective permittivity and permeability.[1–4] Indeed, natural materials may often be regarded as continuous media without any granularity for light wavelengths as short as some tens of nanometers.

As a second example, consider the propagation of electron waves under the influence of an ionic lattice in the context of one-body Schrödinger equation, e.g. the propagation of electrons in a semiconductor material. The electron wave function is scattered by the electric potential created by the ionic lattice, and hence the time evolution of a given initial electronic state is challenging to characterize. Fortunately, for low-energy phenomena the ionic lattice can be effectively homogenized in such a manner that its effect on the wave propagation can be described by an effective mass.[5] In other words, the medium may be regarded as a continuum provided the electron mass is suitably redefined to take into account the influence of the scattering centers.

The objective of this chapter is to present a unified overview of the recent developments in the research of wave propagation in

complex media using effective medium methods, for both light and electron waves. The contents of this chapter are largely based on the ideas originally introduced in Refs. 6 and 7, and which were further developed in subsequent works.

2.2. Microscopic Theory

We consider a generic physical system whose dynamics are characterized by a one-body Schrödinger-type equation of the form:

$$\hat{H}\psi = i\hbar\frac{\partial}{\partial t}\psi. \tag{2.1}$$

Here, $\hbar$ is the reduced Planck constant, $\hat{H}$ is the Hamiltonian operator that determines the time evolution of the system and ψ is the wave function that describes the state of the system. In general, ψ is a multicomponent vector (a spinor). Evidently, this type of formulation is suitable to characterize the propagation of electron waves in a bulk semiconductor or in a semiconductor superlattice. Interestingly, the propagation of light can also be described using a similar formulation. Indeed, Maxwell's equations can be written in a compact form as[7]

$$\begin{pmatrix} 0 & i\nabla\times\mathbf{1}_{3\times3} \\ -i\nabla\times\mathbf{1}_{3\times3} & 0 \end{pmatrix}\cdot\mathbf{f} = i\frac{\partial\mathbf{g}}{\partial t}, \tag{2.2}$$

where $\mathbf{f} = (\mathbf{e} \quad \mathbf{h})^T$ is a six-element vector with components determined by the microscopic electric and magnetic fields and $\mathbf{g} = (\mathbf{d} \quad \mathbf{b})^T$ is a six-element vector with components determined by the electric displacement and the magnetic induction fields. In electromagnetic metamaterials, the $\mathbf{g}$ and $\mathbf{f}$ fields are related by a space-dependent material matrix $\mathbf{M} = \mathbf{M}(\mathbf{r})$ through the constitutive relation $\mathbf{g} = \mathbf{M}\cdot\mathbf{f}$. In conventional isotropic media, the material matrix is simply:

$$\mathbf{M} = \begin{pmatrix} \varepsilon\mathbf{1}_{3\times3} & 0 \\ 0 & \mu\mathbf{1}_{3\times3} \end{pmatrix}. \tag{2.3}$$

Hence, by introducing the operator $\hat{H}$ given by

$$\hat{H} = \hbar \begin{pmatrix} 0 & i\nabla \times \mathbf{1}_{3\times 3} \\ -i\nabla \times \mathbf{1}_{3\times 3} & 0 \end{pmatrix} \cdot \mathbf{M}^{-1} \tag{2.4}$$

and identifying the state vector with the $\mathbf{g}$ field, $\psi = \mathbf{g}$, Maxwell's equations can be expressed as in Eq. (2.1). It should be noted that in the electromagnetic case, $\hat{H}$ is unrelated to the energy of the system and should be simply regarded as an operator that describes the time evolution of the classical electromagnetic field. Moreover, in the previous discussion it is implicit that the relevant materials are non-dispersive, i.e. the permittivity ε and the permeability μ are frequency independent. Yet, the previous ideas can be generalized to dispersive media. Indeed, for dispersive materials the multiplication operator $\mathbf{M}^{-1}$ should be replaced by a suitable time convolution operator. In that case, the action of $\hat{H}$ on ψ at a generic time instant t depends not only on ψ at the same time instant, but also on the values of the state vector in the past $t' < t$. Thus, for dispersive media Eq. (2.1) should be understood as a generalized Schrödinger-type equation.

In the following subsections, we further elaborate on some physical systems to which the theory applies.

2.2.1. *Electromagnetic metamaterials*

The first case of interest corresponds to that of electromagnetic metamaterials, i.e. mesoscopic structures formed by dielectric or metallic inclusions arranged in a periodic lattice.[8] The electromagnetic response of metamaterials is mainly determined by the geometry of its constituents, rather than directly by the chemical composition. The light propagation in these structures is described by Maxwell's equations, which, as previously discussed, can be recast in the Schrödinger form with the time evolution operator given in Eq. (2.4). An illustrative metamaterial unit cell is represented in

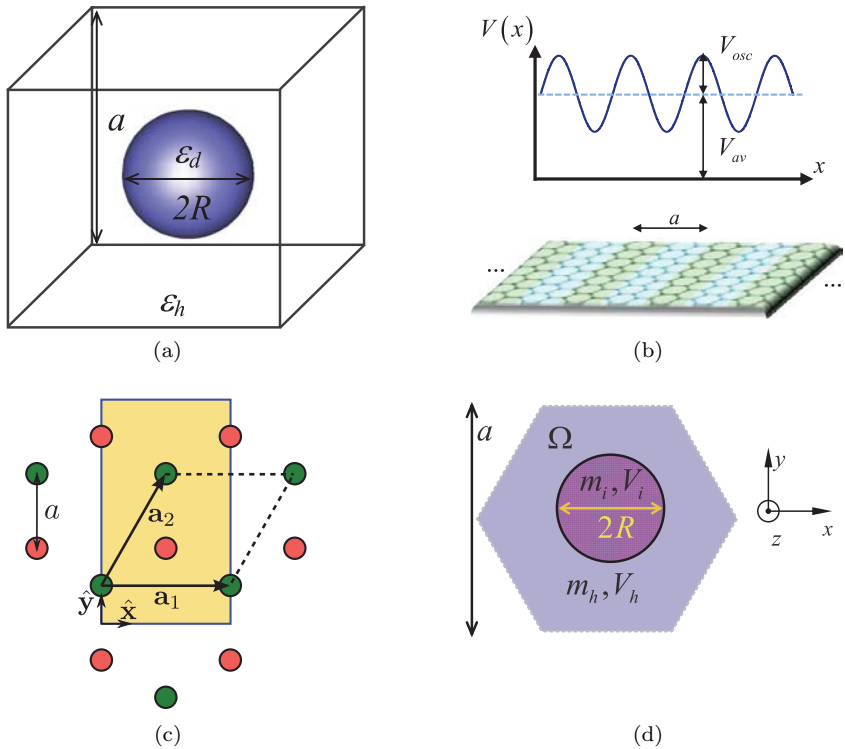


Fig. 2.1. Sketch of the geometries of several metamaterial and semiconductor superlattices. (a) Unit cell of an electromagnetic metamaterial formed by a cubic array of spherical dielectric inclusions embedded in a host medium. (b) Graphene superlattice created by a periodic electrostatic potential. (c) Nanopatterned 2D-electron gas. (d) Unit cell of a semiconductor superlattice formed by two different materials.

Fig. 2.1(a) and corresponds to a spherical dielectric inclusion embedded in a metallic region (the host material). The 3D metamaterial is formed by the periodic repetition of the unit cell. The problem of homogenization of metamaterials, the extraction of effective medium parameters, and the limitations inherent to an effective medium description have been extensively discussed in the literature.[9–26]

2.2.2. *Semiconductor superlattices*

In a groundbreaking study, Esaki and Tsu suggested in 1969 that new quantum effects can be observed at an intermediate physical scale if a mono-crystalline semiconductor is either periodically doped or if the composition of a semiconductor alloy is periodically varied.[27] Such structures are known as semiconductor superlattices and can be regarded as the precursors and the semiconductor counterparts of modern electromagnetic metamaterials. The original proposal of Esaki and Tsu has set the stage for the engineering of the electron transport,[27–29] and led to the development of novel electronic materials with ultrahigh mobilities, wherein the electrons can experience a near zero effective mass, and other breakthroughs.[30–32]

The time evolution of the electron wave function in a semiconductor superlattice may be characterized by a Hamiltonian of the form:

$$\hat{H} = -\frac{\hbar^2}{2}\nabla \cdot \left(\frac{1}{m}\nabla\right) + V, \tag{2.5}$$

where V is a static electric potential and m is the electron (effective) mass in the pertinent semiconductor. Usually, V and m are periodic functions of the spatial coordinates. For example, the above Hamiltonian may describe the physics of a 2D electron gas (e.g. a semiconductor quantum well) modulated by an electrostatic potential created by nanopatterned scattering centers (Fig. 2.1(c)).[33] A different possibility is to periodically change the material composition (Fig. 2.1(d)). In this case, the superlattice has the hexagonal symmetry and is formed by two different semiconductors.

2.2.3. *Graphene superlattices*

Graphene is a one-atom thick material discovered at the turn of the 21st century.[34,35] Remarkably, graphene has a relativistic-type electronic spectrum such that the relation between energy and momentum is linear.[36] This unusual property enables mimicking quantum relativistic effects in a condensed-matter platform and creates many exciting opportunities in nanoelectronics.[36–38] Interestingly, it is possible to tailor the electronic transport in graphene

using the superlattice concept.[39–45] This can be done with patterned gates that impress a space-dependent electric potential on a graphene sheet (Fig. 2.1(b)). Other solutions take advantage of the electric potential induced by crystalline substrates such as boron nitride, or create a desired electric potential through the controlled deposition of adatoms. The low-energy electronics in graphene near the K point is described by a 2D massless Dirac Hamiltonian[35]

$$\hat{H} = -i\hbar v_F(\boldsymbol{\sigma} \cdot \nabla) + V, \tag{2.6}$$

where $v_F \sim 10^6 m/s$ is the Fermi velocity, $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_x, \boldsymbol{\sigma}_y)$, $\boldsymbol{\sigma}_x, \boldsymbol{\sigma}_y$ are the Pauli matrices and V is the periodic electrostatic potential. The wave function ψ in graphene is a pseudo-spinor, and hence has two components.[35] From a physical point of view, each component of the pseudo-spinor is associated with a specific trigonal sublattice of graphene.

2.3. Effective Medium Theory

The goal of an effective medium theory is to provide an approximate and simplified description of the wave propagation in some complex system. Due to the spatially inhomogeneous nature of the structure, the wave function (in case of electronic systems) or the electromagnetic fields (in case of light waves) can vary wildly in the characteristic length scale determined by the "granularities". In this Chapter, we restrict our attention to periodic structures, and hence the characteristic length scale is defined by the lattice period. Usually, one is interested in "low-energy" phenomena for which the wave packet envelope varies slowly in space. An effective medium theory aims to describe the dynamics of the wave packet envelope. These ideas will be made more precise in the following subsections.

2.3.1. *Spatial averaging and the envelope function*

The envelope function is intuitively the slowly varying in space of part of the state vector ψ. It is defined here as:

$$\Psi(\mathbf{r}, t) \equiv \{\psi(\mathbf{r}, t)\}_{\text{av}}, \tag{2.7}$$

where $\{\,\}_{\rm av}$ is a linear operator that performs a spatial averaging. The averaging operator is completely determined by the response to plane waves, determined by the function $F(\mathbf{k})$ such that

$$\{e^{i\mathbf{k}\cdot\mathbf{r}}\}_{\rm av} = F(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}}. \tag{2.8}$$

Thus, the action of the averaging operator on a generic plane wave with wavevector $\mathbf{k}$ yields another plane wave with the same wavevector, but with a different amplitude determined by $F(\mathbf{k})$. Because of the linearity of the operator $\{\,\}_{\rm av}$, its action on a generic function is determined by Fourier theory and is given by a spatial convolution. For example, it is possible to write the envelope function as

$$\Psi(\mathbf{r},t) = \int d^N\mathbf{r}'\, f(\mathbf{r}')\psi(\mathbf{r}-\mathbf{r}',t), \tag{2.9}$$

where N is the space dimension (e.g. $N=3$ for a 3D metamaterial, and $N=2$ for graphene). The weight function f is the inverse Fourier transform of F so that

$$f(\mathbf{r}) = \frac{1}{(2\pi)^N}\int d^N\mathbf{k}\, F(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}}. \tag{2.10}$$

Related ideas have been developed by Russakov in the context of macroscopic electromagnetism.[46] In this Chapter, it is assumed that the averaging operator corresponds to an ideal low-pass spatial filter such that

$$F(\mathbf{k}) = \begin{cases} 1, & \mathbf{k}\in \text{B.Z.} \\ 0, & \text{otherwise.} \end{cases} \tag{2.11}$$

In the vast majority of the cases of interest, the set B.Z. stands for the first Brillouin zone of the periodic lattice, but sometimes other choices can be relevant. Unless something different is explicitly stated, it will always be assumed that B.Z. is the first Brillouin zone.

With these definitions, the envelope function $\Psi(\mathbf{r},t)$ has no relevant spatial fluctuations on the scale of a unit cell, i.e. the microscopic fluctuations are filtered out by the averaging operator. Hence, we will also refer to $\Psi(\mathbf{r},t)$ as the macroscopic state vector. In general, we

say that a given state vector ψ is macroscopic when it stays invariant under the operation of spatial averaging:

$$\psi(\mathbf{r}) = \{\psi(\mathbf{r})\}_{\text{av}} \quad \text{(macroscopic state vector)}. \tag{2.12}$$

Importantly, a macroscopic state cannot be more localized in space than the characteristic period of the material.

2.3.2. *The effective Hamiltonian*

The effective Hamiltonian is the operator that describes the time evolution of the envelope function. Specifically, suppose that the initial state vector is macroscopic, so that $\psi_{t=0} = \Psi_{t=0}$. In general, the time evolution of an initial macroscopic state does not yield a macroscopic state at a later time instant, i.e. $\psi(\mathbf{r}, t) \neq \Psi(\mathbf{r}, t)$ for $t > 0$. We define the effective Hamiltonian $\hat{H}_{ef}$ in such a manner that $\Psi(\mathbf{r}, t)$ calculated using $\hat{H}_{ef}$ is coincident with the spatially averaged microscopic state vector $\{\psi(\mathbf{r}, t)\}_{\text{av}}$, $\psi(\mathbf{r}, t)$ being determined by the microscopic Hamiltonian $\hat{H}$.[7,47] These ideas are illustrated in the diagram of Fig. 2.2.

The time evolution of the macroscopic state vector is determined by a generalized Schrödinger equation:

$$(\hat{H}_{\text{ef}}\Psi)(\mathbf{r}, t) = i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r}, t). \tag{2.13}$$

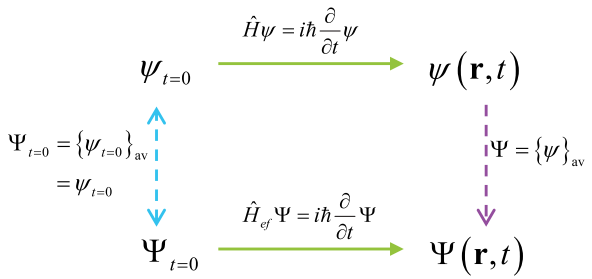


Fig. 2.2. Schematic relation between the time evolutions determined by the macroscopic and microscopic Hamiltonians: for an initial macroscopic state the effective medium formulation ensures that $\Psi = \{\psi\}_{\text{av}}$ for $t > 0$. Reprinted with permission from Ref. 47.

From the definition of the effective Hamiltonian, it is clear that it must ensure that

$$\{\hat{H}\psi\}_{\text{av}} = \hat{H}_{ef}\Psi. \tag{2.14}$$

It was shown in Ref. 7 that the action of the effective Hamiltonian on the wave function can be written as a convolution in space and in time:

$$(\hat{H}_{\text{ef}}\Psi)(\mathbf{r},t) = \int d^N\mathbf{r}' \int_0^t dt'\, \mathbf{h}_{\text{ef}}\left(\mathbf{r}-\mathbf{r}', t-t'\right)\cdot\Psi(\mathbf{r}',t'). \tag{2.15}$$

In general, the kernel $\mathbf{h}_{\text{ef}}(\mathbf{r},t)$ is represented by a square matrix $[\mathbf{h}_{\sigma,\sigma'}]$ because Ψ is typically a multicomponent vector. In the electronic case, the dimension of $\mathbf{h}$ is determined by the number S of spin or pseudospin degrees of freedom ($\sigma = 1,\ldots,S$), whereas in the light case the dimension of $\mathbf{h}_{\text{ef}}$ is $S = 6$. Equation (2.15) implies that the effective Hamiltonian depends on the past history ($0 < t' < t$) of the macroscopic state vector, rather than just on the instantaneous value of Ψ.

It is convenient to introduce the Fourier transform of $\mathbf{h}_{\text{ef}}(\mathbf{r},t)$ defined as

$$\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k},\omega) = \int d^N\mathbf{r} \int_0^{+\infty} dt\, \mathbf{h}_{\text{ef}}(\mathbf{r},t)e^{i\omega t}e^{-i\mathbf{k}\cdot\mathbf{r}}. \tag{2.16}$$

The Fourier transform is bilateral in space and unilateral in time. The unilateral Fourier transform in time can also be regarded as a Laplace transform. In the Fourier domain, the action of the effective Hamiltonian reduces to a simple multiplication:

$$(\hat{H}_{\text{ef}}\Psi)(\mathbf{k},\omega) = \boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k},\omega)\cdot\Psi(\mathbf{k},\omega). \tag{2.17}$$

In the above, $\Psi(\mathbf{k},\omega)$ is the Fourier transform of the macroscopic state vector,

$$\Psi(\mathbf{k},\omega) = \int d^N\mathbf{r} \int_0^{+\infty} dt\, \Psi(\mathbf{r},t)e^{i\omega t}e^{-i\mathbf{k}\cdot\mathbf{r}}, \tag{2.18}$$

and $(\hat{H}_{\text{ef}}\Psi)(\mathbf{k},\omega)$ is defined similarly. The convergence of the Fourier transforms is ensured for $\text{Im}\{\omega\} > 0$. The function $\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k},\omega)$

completely determines the effective Hamiltonian. Because of the properties of the spatial averaging operator, it is possible to enforce that

$$\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k},\omega) = 0, \quad \text{when } \mathbf{k} \notin \text{B.Z.} \tag{2.19}$$

This property ensures that the effective Hamiltonian is a smoothened version of the microscopic Hamiltonian. In the next subsections, it is explained how $\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k},\omega)$ can be calculated for $\mathbf{k} \in$ B.Z.

2.3.3. *Calculation of $\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k},\omega)$ with a time domain approach*

Let us consider an initial macroscopic state of the form $\psi_{t=0} \sim e^{i\mathbf{k}\cdot\mathbf{r}}\mathbf{u}_l$ with $\mathbf{k} \in$ B.Z. Here, $(\mathbf{u}_l)$ represents a basis of unit vectors that generates the S-dimensional vector space wherein ψ is defined. Because of the periodicity of the system, the microscopic time evolution of this initial state yields a state vector $\psi(\mathbf{r},t)$ with the Bloch property. Specifically, $\psi(\mathbf{r},t)e^{-i\mathbf{k}\cdot\mathbf{r}}$ is a periodic function in space for any fixed t. For the same reason, $\hat{H}\psi$ has also the Bloch property. Importantly, the operation of spatial averaging only retains spatial harmonics with wavevector inside the B.Z., and hence it follows that the dependence of $\{\psi\}_{\text{av}}$ and $\{\hat{H}\psi\}_{\text{av}}$ on the spatial coordinates is of the form $e^{i\mathbf{k}\cdot\mathbf{r}}$ for any time instant. In other words, within the effective medium approach the time evolution of a plane wave-type initial state yields another plane wave-type state. Moreover, it is possible to write

$$\{\psi\}_{\text{av}}(\mathbf{r},t) = \psi_{\text{av}}(t)e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{2.20a}$$

$$\{\hat{H}\psi\}_{\text{av}}(\mathbf{r},t) = (\hat{H}\psi)_{\text{av}}(t)e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{2.20b}$$

with

$$\psi_{\text{av}}(t) = \frac{1}{V_{\text{cell}}}\int_{\Omega} d^N\mathbf{r}\ \psi(\mathbf{r},t)e^{-i\mathbf{k}\cdot\mathbf{r}}, \tag{2.21a}$$

$$(\hat{H}\psi)_{\text{av}}(t) = \frac{1}{V_{\text{cell}}}\int_{\Omega} d^N\mathbf{r}\ \hat{H}\psi(\mathbf{r},t)e^{-i\mathbf{k}\cdot\mathbf{r}}, \tag{2.21b}$$

where Ω represents the unit cell and V_{cell} is the respective volume.

Taking now into account that $\Psi = \{\psi\}_{\text{av}}$ and $\hat{H}_{\text{ef}}\Psi = \{\hat{H}\psi\}_{\text{av}}$, and substituting Eq. (2.20) into Eq. (2.15), it is seen that

$$(\hat{H}\psi)_{\text{av}}(\omega) = \boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k},\omega)\cdot\psi_{\text{av}}(\omega). \tag{2.22}$$

In the above, $\psi_{\text{av}}(\omega)$ and $(\hat{H}\psi)_{\text{av}}(\omega)$ stand for the unilateral Fourier (Laplace) transforms of the functions in Eq. (2.21). Hence, if we denote $\psi^{(l)}$, $l = 1,\ldots,S$ as the microscopic state vector determined by the time evolution of the initial state $\psi^{(l)}_{t=0} = i/\hbar\, e^{i\mathbf{k}\cdot\mathbf{r}}\mathbf{u}_l$ (the proportionality constant was fixed as $i/\hbar$ for convenience), it follows from the previous analysis that the effective Hamiltonian is given by

$$\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k},\omega) = [(\hat{H}\psi^{(1)})_{\text{av}} \quad \cdots \quad (\hat{H}\psi^{(S)})_{\text{av}}]\cdot[\psi^{(1)}_{\text{av}} \quad \cdots \quad \psi^{(S)}_{\text{av}}]^{-1}. \tag{2.23}$$

Thus, $\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k},\omega)$ can be written as the product of two matrices, whose columns are determined by the vectors $\psi^{(l)}_{\text{av}}(\omega)$ and $(\hat{H}\psi^{(l)})_{\text{av}}(\omega)$.

In summary, for a given $\mathbf{k} \in$ B.Z. the effective Hamiltonian can be found by solving S microscopic time evolution problems associated with initial states of the form $\psi^{(l)}_{t=0} = i/\hbar\, e^{i\mathbf{k}\cdot\mathbf{r}}\mathbf{u}_l$. The effective Hamiltonian is written in terms of the Fourier transforms in time of the functions (2.21).

2.3.4. *Calculation of $\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k},\omega)$ with a frequency domain approach*

The effective Hamiltonian may also be determined based on frequency domain calculations. To prove this, we note that $\psi_{\text{av}}(\omega)$ and $(\hat{H}\psi)_{\text{av}}(\omega)$ can be written explicitly as

$$\psi_{\text{av}}(\omega) = \frac{1}{V_{\text{cell}}}\int_{\Omega} d^N\mathbf{r}\ \psi(\mathbf{r},\omega)e^{-i\mathbf{k}\cdot\mathbf{r}}, \tag{2.24a}$$

$$(\hat{H}\psi)_{\text{av}}(\omega) = \frac{1}{V_{\text{cell}}}\int_{\Omega} d^N\mathbf{r}\ \hat{H}\psi(\mathbf{r},\omega)e^{-i\mathbf{k}\cdot\mathbf{r}}, \tag{2.24b}$$

where $\psi(\mathbf{r},\omega)$ is the unilateral Fourier transform of $\psi(\mathbf{r},t)$. Applying the unilateral Fourier (Laplace) transform to both members of the microscopic Schrödinger equation (2.1) and using the property

$\partial_t \psi(\mathbf{r},t) \leftrightarrow -i\omega\psi(\mathbf{r},\omega) - \psi_{t=0}(\mathbf{r})$, it follows that

$$[\hat{H} - \hbar\omega] \cdot \psi(\mathbf{r},\omega) = -i\hbar\psi_{t=0}(\mathbf{r}). \tag{2.25}$$

Hence, $\psi^{(l)}(\mathbf{r},\omega)$ can be directly found by solving the above equation for $-i\hbar\psi^{(l)}_{t=0} = e^{i\mathbf{k}\cdot\mathbf{r}}\mathbf{u}_l$, with $l = 1,\ldots,S$. Once $\psi^{(l)}(\mathbf{r},\omega)$ is known, one can determine $\psi^{(l)}_{\text{av}}$ and $(\hat{H}\psi^{(l)})_{\text{av}}$ using Eq. (2.24) and finally obtain the effective Hamiltonian from Eq. (2.23).

It is interesting to note that for $-i\hbar\psi^{(l)}_{t=0} = e^{i\mathbf{k}\cdot\mathbf{r}}\mathbf{u}_l$, Eq. (2.25) implies that $(\hat{H}\psi^{(l)})_{\text{av}} - \hbar\omega\psi^{(l)}_{\text{av}} = \mathbf{u}_l$. Substituting this result into Eq. (2.23), one may also write the effective Hamiltonian as

$$\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k},\omega) = \hbar\omega + [\psi^{(1)}_{\text{av}} \quad \cdots \quad \psi^{(S)}_{\text{av}}]^{-1}. \tag{2.26}$$

2.3.5. *The electromagnetic case*

In the case of light waves, using the time evolution operator (2.4) and $\psi = \mathbf{g} = (\mathbf{d} \quad \mathbf{b})^T$, it is possible to rewrite Eq. (2.25) as

$$\begin{pmatrix} 0 & i\nabla\times\mathbf{1}_{3\times3} \\ -i\nabla\times\mathbf{1}_{3\times3} & 0 \end{pmatrix} \cdot \mathbf{f}(\mathbf{r}) - \omega\mathbf{g}(\mathbf{r}) = -i\mathbf{g}_{t=0}(\mathbf{r}), \tag{2.27}$$

with $\mathbf{f} = (\mathbf{e} \quad \mathbf{h})^T$. If we decompose the six-vector $\mathbf{g}_{t=0}$ as $\mathbf{g}_{t=0} = (\mathbf{j}_e \quad \mathbf{j}_m)^T$, the above system can be spelled out as

$$\begin{aligned} \nabla\times\mathbf{e} &= +i\omega\mathbf{b} - \mathbf{j}_m \\ \nabla\times\mathbf{h} &= -i\omega\mathbf{d} + \mathbf{j}_e \end{aligned}. \tag{2.28}$$

These correspond to the standard microscopic Maxwell's equations in the frequency domain with fictitious electric-type and magnetic-type sources, $\mathbf{j}_e$ and $\mathbf{j}_m$, respectively. Clearly, in the homogenization problem the sources have a plane-wave spatial dependence $e^{i\mathbf{k}\cdot\mathbf{r}}$. Thus, the effective response of a composite medium can be calculated by exciting the medium with fictitious macroscopic sources. This idea is the essence of the "source-driven" homogenization method originally introduced in Ref. 6. Next, we prove that the effective response obtained using the approach of section 3.4 is coincident with what is

obtained using the theory of Refs. 6, 17–21. Related effective medium formalisms have also been presented in Refs. 22–26.

To this end, define $\mathbf{G}_{\rm av} = (\mathbf{D}_{\rm av} \quad \mathbf{B}_{\rm av})^T$ and $\mathbf{F}_{\rm av} = (\mathbf{E}_{\rm av} \quad \mathbf{H}_{\rm av})^T$ such that (compare with Eq. 2.24a):

$$\mathbf{G}_{\rm av} = \frac{1}{V_{\rm cell}} \int_\Omega d^N\mathbf{r}\ \mathbf{g}(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}}, \tag{2.29a}$$

$$\mathbf{F}_{\rm av} = \frac{1}{V_{\rm cell}} \int_\Omega d^N\mathbf{r}\ \mathbf{f}(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}}. \tag{2.29b}$$

Furthermore, let us introduce the effective material matrix $\mathbf{M}_{\rm ef}(\mathbf{k},\omega)$ such that for arbitrary macroscopic sources $\mathbf{j}_e$ and $\mathbf{j}_m$ one has

$$\mathbf{G}_{\rm av} = \mathbf{M}_{\rm ef}(\mathbf{k},\omega) \cdot \mathbf{F}_{\rm av}. \tag{2.30}$$

Then, it can be shown from Eqs. (2.4) and (2.22) that

$$\boldsymbol{\mathcal{H}}_{\rm ef}(\mathbf{k},\omega) = \hbar \begin{pmatrix} 0 & -\mathbf{k}\times\mathbf{1}_{3\times3} \\ \mathbf{k}\times\mathbf{1}_{3\times3} & 0 \end{pmatrix} \cdot \mathbf{M}_{\rm ef}^{-1}(\mathbf{k},\omega). \tag{2.31}$$

This proves that the effective Hamiltonian can be written in terms of the effective material matrix $\mathbf{M}_{\rm ef}(\mathbf{k},\omega)$ calculated with the source-driven homogenization.[6] In particular, the time evolution of the macroscopic electromagnetic fields can be determined from Eq. (2.13), which is equivalent to the macroscopic Maxwell's equations:

$$\begin{pmatrix} 0 & i\nabla\times\mathbf{1}_{3\times3} \\ -i\nabla\times\mathbf{1}_{3\times3} & 0 \end{pmatrix} \cdot \mathbf{F}(\mathbf{r},t) = i\frac{\partial}{\partial t}\mathbf{G}(\mathbf{r},t), \tag{2.32}$$

where $\mathbf{G} = \{\mathbf{g}\}_{\rm av}$ and $\mathbf{F} = \{\mathbf{f}\}_{\rm av}$ are the macroscopic electromagnetic fields. Consistent with the conventional theory of spatially dispersive materials,[48] the fields $\mathbf{G}$ and $\mathbf{F}$ are related by a space–time convolution whose kernel is determined by the inverse Fourier transform of $\mathbf{M}_{\rm ef}(\mathbf{k},\omega)$. It can be shown that for reciprocal structures the effective material matrix satisfies:

$$\mathbf{M}_{\rm ef}(\mathbf{k},\omega) = \mathbf{U}\cdot\mathbf{M}_{\rm ef}^T(-\mathbf{k},\omega)\cdot\mathbf{U}, \quad \text{with } \mathbf{U} = \begin{pmatrix} \mathbf{1}_{3\times3} & 0 \\ 0 & -\mathbf{1}_{3\times3} \end{pmatrix}. \tag{2.33}$$

In most electromagnetic metamaterials, the inclusions are either dielectric or metallic particles, and thus do not have an intrinsic magnetic response ($\mu = \mu_0$). In this case, it is evident that independent of the excitation one has $\mathbf{B}_{\text{av}} = \mu_0 \mathbf{H}_{\text{av}}$. This result together with the reciprocity constraint (2.33) implies that for metamaterials formed by non-magnetic particles, the material matrix is of the form

$$\mathbf{M}_{\text{ef}}(\mathbf{k}, \omega) = \begin{pmatrix} \overline{\varepsilon}_{\text{ef}}(\mathbf{k}, \omega) & 0 \\ 0 & \mu_0 \mathbf{1}_{3\times 3} \end{pmatrix}. \tag{2.34}$$

Thus, the effective response of metal-dielectric metamaterials is completely characterized by a non-local dielectric function $\overline{\varepsilon}_{\text{ef}}(\mathbf{k}, \omega)$, consistent with Ref. 6.

In summary, it was proven that the time evolution of macroscopic electromagnetic field states characterized by a certain $\mathbf{g}_{t=0}$ is rigorously described by the operator $\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k}, \omega)$ given by Eq. (2.31). The effective Hamiltonian is written in terms of an effective material matrix $\mathbf{M}_{\text{ef}}(\mathbf{k}, \omega)$, which is exactly coincident with that originally introduced in Ref. 6, based on the idea of source-driven homogenization.

2.3.6. *Stationary states*

A key property of the effective Hamiltonian is that its energy spectrum coincides with that of the microscopic Hamiltonian.[7] The energy spectrum of the macroscopic Hamiltonian is determined by the non-trivial solutions of the stationary Schrödinger equation

$$[\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k}, \omega)|_{\omega = E/\hbar} - E] \cdot \Psi = 0, \tag{2.35}$$

where E stands for the energy of a certain stationary state. Likewise, in the electromagnetic case the photonic band structure calculated with the effective Hamiltonian is coincident with the exact band structure obtained using a microscopic theory.[6]

The enunciated result can be understood noting that in a time evolution problem (with no source excitation) the state vector can be written as a superposition of eigenmodes. The eigenmodes have a time variation of the form $e^{-i\omega_n t}$, $\omega_n = E_n/\hbar$ being the relevant

eigenfrequencies. Importantly, since the macroscopic and microscopic state vectors are related by the spatial-averaging operation ($\Psi = \{\psi\}_{\rm av}$), both Ψ and ψ have the same-type of time oscillations. In other words, the averaging affects only the space coordinates, while the time coordinate is not averaged in any manner. As a consequence, the spectrum of the microscopic and macroscopic Hamiltonians must be the same (Eq. (2.35)).

Strictly speaking, it is possible that some special microscopic states are not predicted by the effective medium Hamiltonian. We refer to such states as "dark states".[7] A dark state corresponds to a microscopic stationary state ψ that has a zero projection into the subspace of macroscopic states, i.e. a state such that $\Psi = \{\psi\}_{\rm av} = 0$. Dark states appear only in degenerate singular cases, and for very specific forms of the microscopic Hamiltonian. Thus, typically the band structures of the microscopic and effective Hamiltonians are indeed the same.

2.4. Applications

Next, we present several examples that illustrate the application of the ideas developed in Section 2.3 to both electromagnetic metamaterials and quantum structures.

2.4.1. *Homogenization of electromagnetic metamaterials*

The interest in modern electromagnetic metamaterials was sparked by a seminal study by J. Pendry, who showed that a composite material with a simultaneously negative permittivity and permeability can make a perfect lens.[49] Among many other proposals, it was suggested that such a doubly negative metamaterial can be realized based on dielectric spherical particles embedded in a metallic host material[20,50,51] (Fig. 2.1(a)). The following discussion is focused on the metamaterial homogenization.

As mentioned in Section 2.3.5, a metamaterial formed by dielectrics and metals is completely characterized by a non-local

dielectric function of the form $\overline{\varepsilon}_{\rm ef}(\mathbf{k},\omega)$. Such a description is rather powerful but still quite complex. Indeed, both the frequency and the wavevector are independent parameters, and thus the non-local dielectric function depends on four independent variables. It is desirable to further simplify the theory. This can be done based on the hypothesis that the material response is local. Specifically, if one assumes that the homogenized medium can be characterized by a local permittivity, $\overline{\varepsilon}_L(\omega)$, and by a local permeability, $\overline{\mu}_L(\omega)$, then it can be shown that the local parameters are linked to the non-local dielectric function as[6,21]

$$\frac{1}{\varepsilon_0}\overline{\varepsilon}_{\rm ef}(\mathbf{k},\omega) = \overline{\varepsilon}_L(\omega) + \frac{c\mathbf{k}}{\omega} \times (\overline{\mu}_L^{-1}(\omega) - \mathbf{1}) \times \frac{c\mathbf{k}}{\omega}. \tag{2.36}$$

Note that the local parameters are independent of the wavevector. The above relation implies that the local permittivity satisfies

$$\overline{\varepsilon}_L(\omega) = \frac{1}{\varepsilon_0}\overline{\varepsilon}_{\rm ef}(\mathbf{k}=0,\omega). \tag{2.37}$$

On the other hand, the local permeability can be written in terms of the second-order derivatives of the non-local dielectric function with respect to the wavevector. For example, assuming that $\overline{\mu}_L(\omega)$ is a diagonal tensor it is easy to check that

$$\mu_{L,zz}(\omega) = \frac{1}{1 - \left(\frac{\omega}{c}\right)^2 \frac{1}{2\varepsilon_0} \left.\frac{\partial^2 \varepsilon_{{\rm ef},yy}}{\partial k_x^2}\right|_{\mathbf{k}=0}}. \tag{2.38}$$

The formulas for the other components of the permeability can be obtained by considering permutations of the indices x, y, and z. In practice, $\overline{\varepsilon}_{\rm ef}(\mathbf{k},\omega)$ is calculated using computational methods. Several solutions have been reported in the literature: an integral equation based approach,[6] a finite-difference frequency domain method,[19] and a time domain scheme.[20] The derivatives with respect to the wavevector are calculated using finite differences.[6,19,20]

Figure 2.3 depicts the extracted local effective parameters for a metamaterial with unit cell as in Fig. 2.1(a).[20] For long wavelengths, the metamaterial has an isotropic response, and hence both the permittivity and permeability are scalars. The radius of the

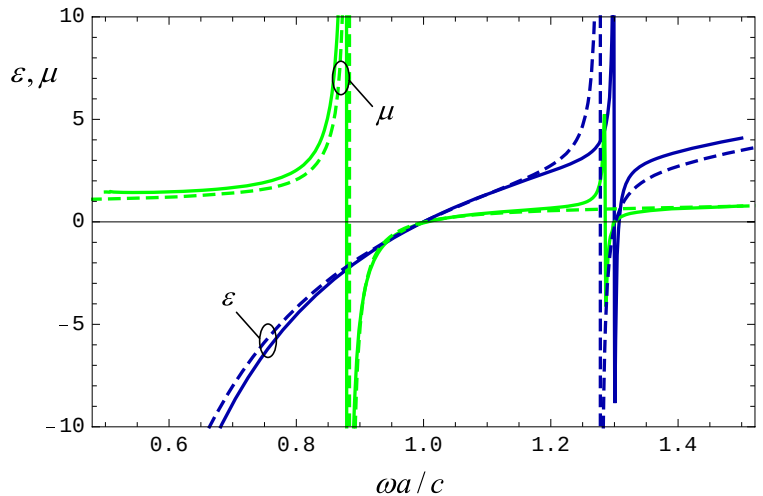


Fig. 2.3. Local effective permittivity ε and local effective permeability μ as a function of frequency for a simple cubic array of dielectric spheres with $\varepsilon_d = 73.1$ and $R = 0.4a$ embedded in a metallic host. Solid lines: Effective medium theory; Dashed lines: Lewin's formulas.[52] Adapted from Ref. 20.

dielectric spheres is $R = 0.4a$, a being the lattice constant. The spheres have permittivity $\varepsilon_d = 73.1$ and are embedded in a metallic material with a permittivity modelled by a Drude-type dispersion, $\varepsilon_h = \varepsilon_\infty(1 - \omega_p^2/\omega(\omega + i\Gamma))$. In the simulations, it was assumed that the normalized plasma frequency is $\omega_p a/c = 1.0$, that $\varepsilon_\infty = 3.6$, and for simplicity the metallic loss was neglected $\Gamma = 0$. The parameters of the metamaterial were tuned to ensure that $\varepsilon_L \approx \mu_L$ (matched index material) over a broad range of frequencies close to $\omega = \omega_p$.

Figure 2.3 also shows the local effective parameters predicted by a mixing formula proposed by Lewin.[52] As seen, there is an overall excellent agreement between the full wave homogenization results and Lewin's formula, especially near $\omega = \omega_p$. Interestingly, in this design the conditions $\varepsilon = \mu = -1$ are observed simultaneously, and thus this metamaterial can mimic to some extent a Veselago–Pendry's lens, as further discussed in Ref. 51. Many other examples of the application of the effective medium theory can be found in Refs. 6, 19, 20.

2.4.2. *Zincblende semiconductors*

The electronic structure of semiconductors is typically studied using perturbation schemes, usually known as $k \cdot p$ methods.[53–62] The $k \cdot p$ theory is rooted on the knowledge of the electronic band structure at highly symmetric points of the Brillouin zone determined by some eigenstates u_{n0}. These eigenstates are used as a basis to construct the wave function for a generic wavevector. Hence, the electron wave function is usually described by a multicomponent vector, whose elements are the coefficients of the expansion of the wave function in the basis u_{n0}. The $k \cdot p$ theory can be applied to bulk materials as well as to semiconductor heterostructures.

In contrast, G. Bastard developed the concept of the envelope-function approximation in the analysis of semiconductor heterostructures during the 1980s.[55,61,62] The idea of an envelope-function is closely related to the pseudopotential method used in condensed matter physics.[5] Interestingly, there is a profound connection between the effective medium methods developed in Section 2.3 and Bastard's theory.[47,63] Specifically, it was rigorously shown from "first principles" that when the effective medium approach is applied to bulk semiconductors with a zincblende structure — using as a starting point Kane's eight-band $k \cdot p$ theory[56] — the corresponding effective Hamiltonian is characterized by an energy-dependent effective mass and an effective potential.[47] Moreover, the obtained effective Hamiltonian is equivalent in the long-wavelength limit to that used in Bastard's theory.[55] Thus, the formalism of Section 2.3 recovers Bastard's theory and shows that the corresponding effective Hamiltonian describes rigorously the time evolution of the macroscopic electronic states in bulk semiconductors. Moreover, it puts into perspective that the envelope function approximation used in semiconductor physics is intrinsically related to the effective medium methods used in the context of macroscopic electrodynamics.[47,59]

The zincblende lattice corresponds to a face-centered cubic lattice with two atoms per unit cell, and is characteristic of binary III–V compounds such as GaAs, GaSb, InSb, and II–VI compounds such as HgTe and CdTe.[53,55] The electronic structure of these materials

is formed by a conduction band, a light-hole band, a heavy-hole band, and a split-off band.[53,57] The wave function associated with the conduction band has the same symmetry as a monopole (S-type symmetry), while the wave functions associated with the remaining (valence) bands have the same symmetry as a dipole (P-type symmetry). The split-off band usually lies well below the remaining bands, and often its effects on the wave propagation can be neglected. The zincblende compounds can be characterized by the effective Hamiltonian[47]:

$$\boldsymbol{\mathcal{H}}_{\mathrm{ef}}(\mathbf{k}, E) = E_c + \frac{\hbar^2}{2m_0}k^2\left[1 + \frac{\varepsilon_P}{3}\left(\frac{-2}{\tilde{E}_v} + \frac{-1}{\tilde{E}_v - \Delta}\right)\right]. \tag{2.39}$$

Here, $E = \hbar\omega$ is the electron energy, $\tilde{E}_v = E_v - E + \hbar^2k^2/(2m_0)$, m_0 is the free electron rest mass, Δ is the spin-orbit split-off energy, $\varepsilon_P = 2P^2m_0/\hbar^2$ is Kane's energy, and E_c and E_v determine the energy levels at the edges of the conduction and light-hole bands, respectively. The split-off energy Δ determines the energy offset between the split-off band and the other two valence bands.

For long wavelengths, the term $\hbar^2k^2/(2m_0)$ in the definition of $\tilde{E}_v$ can be neglected. Within this approximation, the effective Hamiltonian can be written as

$$\hat{H}_{\mathrm{ef}}(E) = -\frac{\hbar^2}{2}\nabla\cdot\left(\frac{1}{m_{\mathrm{ef}}}\nabla\right) + V_{\mathrm{ef}}, \tag{2.40}$$

with the effective potential given by

$$V_{\mathrm{ef}}(E) = E_c, \tag{2.41}$$

and the dispersive effective mass $m_{\mathrm{ef}} = m_{\mathrm{ef}}(E)$ defined as

$$\frac{1}{m_{\mathrm{ef}}} = \frac{1}{m_0} + v_P^2\left(\frac{2}{E - E_v} + \frac{1}{E - E_v + \Delta}\right), \tag{2.42}$$

where $v_P = \sqrt{\varepsilon_P/(3m_0)}$ is Kane's velocity. Thus, the electron wave propagation in the bulk semiconductor compound is completely characterized by an energy-dependent effective mass and by an effective potential. The effective mass formula (2.42) is well known in the context of Bastard's envelope function approximation[55] (p. 88). It

is important to highlight that the effective parameter "dispersive mass" is distinct from the usual effective mass $m^* = \hbar^2[\partial^2 E/\partial k^2]^{-1}$ obtained from the curvature of the energy diagram.

In case of narrow-gap semiconductors, Δ is typically a few times larger than the band gap energy $|E_g| = |E_c - E_v|$. In these conditions, for energies in the band gap or in vicinity of the band gap only the first term in brackets in Eq. (2.42) is important. This yields the linear dispersive mass approximation:

$$m_{\text{ef}} \approx \frac{E - E_v}{2v_P^2}. \tag{2.43}$$

There is an interesting analogy between the propagation of electron waves in a zincblende semiconductor compound and the propagation of light waves in an isotropic material. Indeed, by comparing the stationary Schrödinger equation associated with the Hamiltonian (2.40) with the Helmholtz equation, $\nabla \cdot [\mu^{-1}\nabla E_z] + (\omega^2/c^2)\varepsilon E_z = 0$ that describes the propagation of transverse electric (TE) electromagnetic waves (with $\mathbf{E} = E_z\hat{\mathbf{z}}$ and $\partial/\partial z = 0$) in a material characterized by the parameters ε and μ, it is possible to establish the following correspondence[47,63,64]:

$$\begin{aligned} E - V_{\text{ef}} &\leftrightarrow \varepsilon \\ m_{\text{ef}} &\leftrightarrow \mu. \end{aligned} \tag{2.44}$$

Thus, $E - V_{\text{ef}}$ may be regarded as the semiconductor dual of the electric permittivity, whereas m_{ef} may be regarded as the semiconductor dual of the magnetic permeability. This analogy can be useful to establish parallelisms between phenomena in electromagnetic metamaterials and in semiconductor superlattices. For example, in Ref. 63, it was shown how such a correspondence can be used to design novel semiconductor materials with extreme anisotropy, such that the effective mass is zero along some preferred direction of motion and infinite for perpendicular directions. Furthermore, based on these ideas it was theoretically suggested that the perfect lens concept and a light tunneling phenomenon in metamaterials have semiconductor analogues.[64,65]

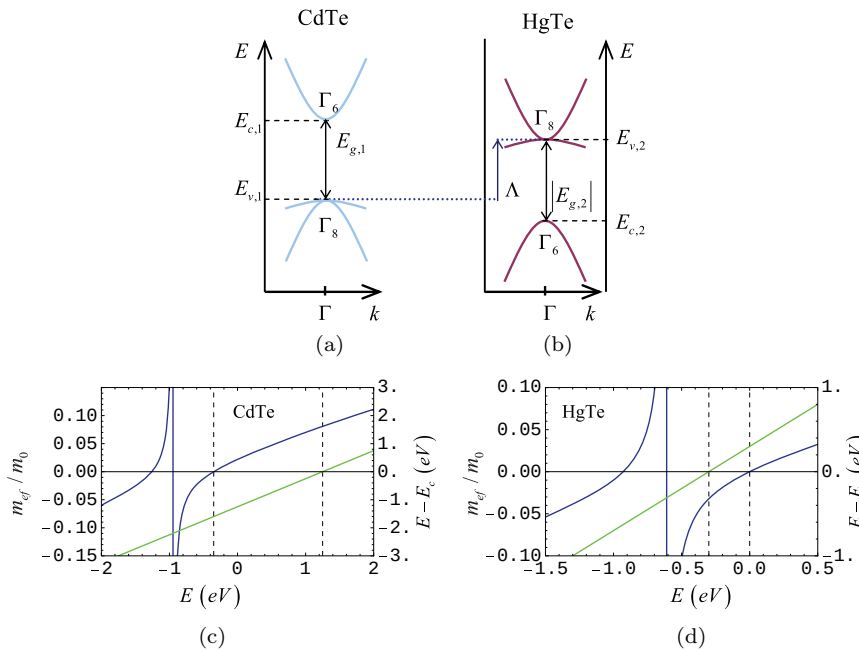


Fig. 2.4. (a) and (b) Sketch of the electronic band diagrams of CdTe and HgTe. (c) and (d) Effective parameters $m_{\rm ef}$ and $E - V_{\rm ef}$ as a function of the normalized energy for CdTe with a regular band structure and for HgTe with an inverted band structure. The energy of the valence band edge of HgTe is arbitrarily taken equal to zero and the valence band offset is $\Lambda = 0.35\,eV$. Blue lines: $m_{\rm ef}/m_0$; Green lines: $E - E_c$ in $[eV]$. The dashed vertical gridlines indicate the edges of the light-hole valence and conduction bands, and delimit the relevant bandgaps. Adapted from Refs. 47 and 63.

It is interesting to further discuss the properties of the effective parameters of mercury–cadmium–telluride (HgCdTe) compounds. The electronic band structures of cadmium–telluride (CdTe) and of mercury–telluride (HgTe) (a group II–VI degenerate semiconductor) are sketched in Figs. 2.4(a) and 2.4(b), respectively. Following the usual convention, the conduction band with S-type symmetry is denoted by Γ_6, whereas the light-hole and heavy-hole valence bands with P-type symmetry are denoted by Γ_8. The heavy-hole band is the nearly flat valence band. As seen in Fig. 2.4(a), the binary compound CdTe has a regular band structure with the conduction band

lying above the valence bands. Interestingly, even though unusual, it is possible that the order of the conduction band (with S-type symmetry) and of the valence bands (with P-type symmetry) is interchanged. An example of a material with this remarkable property is HgTe, which, as illustrated in Fig. 2.4(b), has an "inverted" band structure. Therefore, HgTe has a negative band gap energy ($E_g = E_c - E_v < 0$).[47,63,64]

Figures 2.4(c) and 2.4(d) show the effective parameters m_{ef} and $E - V_{\text{ef}}$ as a function of the electron energy for CdTe and HgTe, respectively. The stationary energy states occur at the energy intervals for which m_{ef} and $E - V_{\text{ef}}$ have the same sign, and the band gaps occur in the range wherein m_{ef} and $E - V_{\text{ef}}$ have different signs. The band gap that separates the conduction band from the light-hole valence band is delimited by the vertical dashed gridlines in Fig. 2.4.

The plots in Fig. 2.4 show that in the band gap a semiconductor with a regular band structure ($E_g > 0$, e.g. CdTe) behaves, from the point of view of the wave propagation, as a material with $\varepsilon < 0$ and $\mu > 0$ (*epsilon* negative –ENG– material), whereas a semiconductor with an inverted band structure ($E_g < 0$, e.g. HgTe) behaves as a material with $\varepsilon > 0$ and $\mu < 0$ (*mu* negative –MNG– material). Due to these remarkable properties, the electron wave propagation in HgTe–CdTe superlattices exhibits remarkable analogies with the propagation of light in ENG–MNG metamaterials.[63–65] In general, it is estimated that the ternary semiconductor alloy $Hg_{1-x}Cd_xTe$ exhibits a regular band structure for $x > 0.17$ and an inverted band structure for $x < 0.17$, $0 \leq x \leq 1$ being the mole fraction of cadmium.

2.4.3. *Homogenization of semiconductor superlattices*

The effective Hamiltonian (2.40) enables a macroscopic description of the electron wave propagation in a bulk semiconductor, in the same manner as the permittivity and the permeability characterize the light propagation in a natural material. Similar to electromagnetic metamaterials, it is possible to tailor the transport properties of electrons by mixing different semiconductors in a periodic structure. To illustrate this, we consider the superlattice formed by two material

phases arranged in a hexagonal lattice (Fig. 2.1(d)). It is assumed that the circular scattering centers (the inclusions) have dispersive mass and effective potential m_i, V_i, whereas the background material is characterized by the parameters m_h, V_h. Furthermore, it is supposed that the electrons can move only in the *xoy* plane so that the structure is intrinsically 2D. For example, the structure may correspond to a quantum well. In what follows, we assume that the host region is the ternary alloy $Hg_{0.75}Cd_{0.25}Te$ and that the scattering centers are made of HgTe. Moreover, for simplicity, the effective parameters m, V are taken to be the same as in a bulk semiconductor.

The effective Hamiltonian of the superlattice can be found using the general approach outlined in Section 2.3.4. In particular, to obtain $\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k}, E)$ one needs to solve Eq. (2.25) with respect to ψ with $\hat{H}$ given by Eq. (2.5) where $m = m(\mathbf{r}, E)$ and $V = V(\mathbf{r})$ are periodic functions of the spatial coordinates. Because $\hat{H}$ is a scalar operator ($S = 1$), one needs to solve a single microscopic problem with $\psi_{t=0} \sim e^{i\mathbf{k}\cdot\mathbf{r}}$. The effective Hamiltonian is given by $\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k}, E) = (\hat{H}\psi)_{\text{av}}/\psi_{\text{av}}$ with $(\hat{H}\psi)_{\text{av}}$ and ψ_{av} defined as in Eq. (2.24). A finite difference frequency domain discretization of Eq. (2.25) is reported in Ref. 66.

It is desirable to further simplify the effective medium description. For low-energy phenomena, this can be done relying on a Taylor expansion of the effective Hamiltonian with respect to the wavevector[47]:

$$\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k}, E) \approx V_{\text{ef}}(E) + \frac{\hbar^2}{2}\mathbf{k} \cdot \bar{M}_{\text{ef}}^{-1} \cdot \mathbf{k}, \tag{2.45}$$

where $V_{\text{ef}}(E) = \boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k} = 0, E)$ is the energy-dependent effective potential of the superlattice and $\bar{M}_{\text{ef}}$ is the energy-dependent effective mass tensor of the superlattice, with its inverse determined by

$$\bar{M}_{\text{ef}}^{-1} = \frac{1}{\hbar^2}\left[\frac{\partial^2 \boldsymbol{\mathcal{H}}_{\text{ef}}}{\partial k_i \partial k_j}\bigg|_{\mathbf{k}=0}\right]. \tag{2.46}$$

Note that similar to the effective permeability in electromagnetic metamaterials (Eq. (2.38)), the superlattice effective mass is written in terms of the second-order derivatives of the effective Hamiltonian with respect to the wavevector.

For the particular superlattice under study, it is clear that by symmetry the effective mass tensor is a scalar. Hence, within the validity of Eq. (2.45) the effective Hamiltonian in the space domain is such that

$$\hat{H}_{\rm ef}(\mathbf{r}, E) = -\frac{\hbar^2}{2}\nabla \cdot \left[\frac{1}{M_{\rm ef}(E)}\nabla\right] + V_{\rm ef}(E). \tag{2.47}$$

Therefore, the electron wave propagation in the superlattice is completely characterized by the effective parameters $V_{\rm ef}(E)$ and $M_{\rm ef}(E)$. The order of the operators in the term $\nabla \cdot [M_{\rm ef}^{-1}\nabla]$ is consistent with the generalized Ben Daniel-Duke boundary conditions.[53,55,57] Indeed, Eq. (2.47) implies that Ψ and $M_{\rm ef}^{-1}(\partial/\partial n)\Psi$ are continuous at an abrupt interface between two materials, which ensures the continuity of the normal component of the probability current at the interface.

Interestingly, it is possible to write approximate analytical formulas for $V_{\rm ef}(E)$ and $M_{\rm ef}(E)$. Indeed, by exploiting the correspondences $E - V \leftrightarrow \varepsilon$ and $m \leftrightarrow \mu$, and the analogies between electromagnetic metamaterials and semiconductor superlattices one finds that[67]

$$V_{\rm ef} = V_h(1 - f_V) + V_i f_V, \tag{2.48}$$

$$M_{\rm ef} = m_h \frac{(1 - f_V)m_h + (1 + f_V)m_i}{(1 + f_V)m_h + (1 - f_V)m_i}, \tag{2.49}$$

where f_V is the volume fraction of the HgTe inclusions. The mixing formula for the effective mass is nothing more than the semiconductor counterpart of the classical Clausius–Mossotti formula well known in electromagnetism.[67]

To illustrate the application of these ideas, next we discuss the design of a zero-gap semiconductor superlattice with linearly dispersing bands. This design is motivated by the fact that heuristically one may expect that by combining a material with a positive band gap (e.g. $Hg_{0.75}Cd_{0.25}Te$) with a material with a negative band gap (e.g. HgTe) it may be possible to design a material with a zero gap, such that the electrons experience a zero effective mass and an ultrahigh mobility.

It can be theoretically shown that the condition to have a zero gap is $M_{\rm ef}(E = V_{\rm ef}) = 0$.[67] Using the analytical formulas (2.48)–(2.49), and assuming that the dispersive mass of the relevant materials is described by the linear mass approximation (2.43), it can be shown that the zero-gap regime occurs when the volume fraction of the HgTe circular scattering centers satisfies[67]

$$f_{V0} = \frac{E_{v,h} + E_{v,i} - 2V_h}{E_{v,h} - E_{v,i} - 2(V_h - V_i)}. \tag{2.50}$$

For the particular design considered here, the critical volume fraction is $f_{V0} = 0.247$. Figure 2.5(a) shows the electronic band structures of three different semiconductor superlattices characterized by certain f_V/f_{V0} indicated in the insets. The superlattice period is

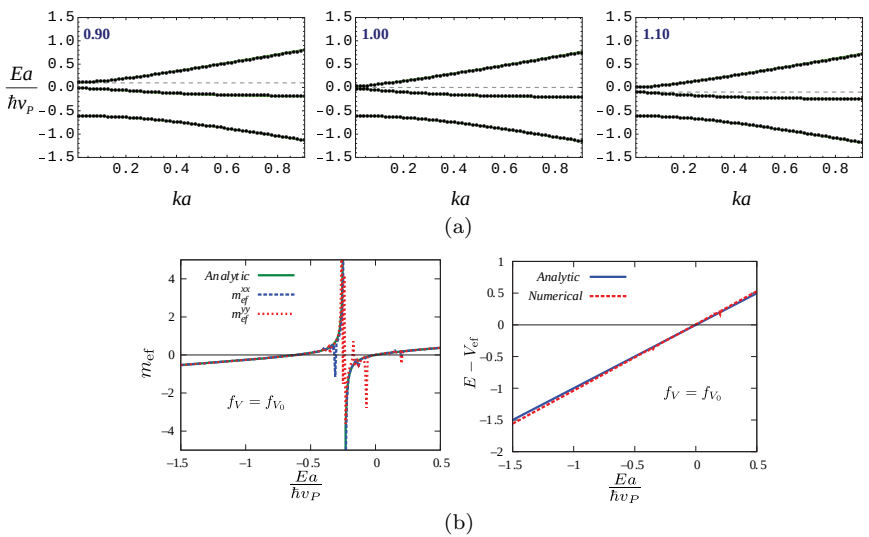


Fig. 2.5. (a) Electronic band structures of $Hg_{0.75}Cd_{0.25}Te$–HgTe superlattices for different values of the normalized volume fraction f_V/f_{V0} of the HgTe inclusions shown in the insets. Solid lines: effective medium theory; Discrete points: "exact" band structure obtained with the microscopic Hamiltonian. The HgTe scattering centers have circular shape and are arranged in a triangular lattice with period $a = 12a_s$. The dashed horizontal grid line represents the edge of the hybridized conduction (S-type) band. (b) Effective parameters of the superlattice with $f_V = f_{V0}$ as a function of the electron energy. Adapted from Refs. 66 and 67.

$a = 12a_s$, $a_s = 0.65\,\text{nm}$ being the atomic lattice constant. The solid lines were obtained using the analytical effective parameters (2.48)–(2.49), whereas the discrete points were found by calculating the energy spectrum of the periodic microscopic Hamiltonian (Eq. (2.5)) with the plane wave method. The effective medium formulas predict almost exactly the electronic stationary states for $ka < 1.0$, and only fail near the edges of the Brillouin zone (not shown). Moreover, consistent with the previous discussion, when $f_V/f_{V0} = 1$, the band gap that separates the conduction and valence bands closes and the energy dispersion is linear. The edge of the conduction band (with S-type symmetry) corresponds to the energy level wherein $E = V_{\text{ef}}$ and is marked with a dashed horizontal line in the figures. Thus, for structures with $f_V/f_{V0} < 1$ (e.g. $f_V/f_{V0} = 0.9$) the superlattice is characterized by a regular band structure, whereas for $f_V/f_{V0} > 1$ (e.g. $f_V/f_{V0} = 1.1$) it is characterized by an inverted band structure. The topological transition wherein the conduction and valence bands interchange positions takes place at $f_V/f_{V0} = 1$.

Figure 2.5(b) compares the analytic effective parameters (Eqs. (2.48)–(2.49)) with the effective parameters determined based on the rigorous numerical calculation of the Taylor expansion of $\boldsymbol{\mathcal{H}}_{\text{ef}}$ (Eq. (2.45)). Apart from some spurious resonances in the dispersive mass, the numerically calculated $V_{\text{ef}}(E)$ and $M_{\text{ef}}(E)$ agree extremely well with the analytical formulas. The spurious resonances are related to the excitation of heavy-hole type states, which are not predicted by the effective medium theory because they are dark states.[66]

The zero-gap regime can have important consequences in the electron transport. In particular, due to the zero-mass property the superlattice may have a giant non-linear response when excited by an external electromagnetic field.[67] The non-linear effects are particularly strong up to terahertz frequencies and when the chemical potential is near the tip of the Dirac-type point.[67]

2.4.4. *Homogenization of graphene superlattices*

The electron transport in graphene is determined by the energy dispersion near the Dirac points K and K' and is intrinsically isotropic:

the electron velocity is $v_F \approx 10^6\, m/s$ independent of the direction of propagation. Importantly, it has been theoretically predicted that a superlattice created by a 1D-electrostatic periodic potential (Fig. 2.1(b)) can be used to collimate an electron beam with virtually no spatial spreading or diffraction,[39–41] such that the electrons can only propagate along a preferred direction of space. This remarkable phenomenon can be conveniently modelled using an effective medium approach.[7]

The effective Hamiltonian that describes the propagation of Dirac fermions associated with the K point can be found by applying the general ideas described in Section 2.3.4 to the 2D massless Dirac Hamiltonian (2.6), being the microscopic electric potential $V = V(x)$ a periodic function of x. Note that the Dirac Hamiltonian (2.6) already provides an effective medium description of the electron propagation in graphene, but here we consider a second level of homogenization such that the superlattice itself can be regarded as a continuum. Because the Dirac fermions are described by a pseudospinor, the effective Hamiltonian $\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k}, E)$ is a 2×2 matrix. The computation of $\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k}, E)$ involves solving Eq. (2.25) for two independent initial macroscopic states ($S = 2$). In general, this requires the discretization of the relevant partial differential equation using finite differences.[7]

Similar to Sections 2.4.1 and 2.4.3, to obtain some local effective parameters it is useful to expand the effective Hamiltonian in a Taylor series:

$$\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k}, E) \approx \boldsymbol{\mathcal{H}}_{\text{ef}}(0, E) + k_x \frac{\partial \boldsymbol{\mathcal{H}}_{\text{ef}}}{\partial k_x}(0, E) + k_y \frac{\partial \boldsymbol{\mathcal{H}}_{\text{ef}}}{\partial k_y}(0, E). \quad (2.51)$$

It was found in Ref. 7 that for 1D graphene superlattices the Taylor expansion can be rewritten as $\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k}, E) \approx \hbar v_F \boldsymbol{\sigma}_{\text{ef}}(E) \cdot \mathbf{k} + V_{\text{ef}}(E)$, where $V_{\text{ef}}(E)$ is an energy-dependent effective potential (a scalar) and $\boldsymbol{\sigma}_{\text{ef}}$ is the tensor[7]:

$$\boldsymbol{\sigma}_{\text{ef}} \approx v_{xx} \boldsymbol{\sigma}_x \hat{\mathbf{x}} + v_{yy} \boldsymbol{\sigma}_y \hat{\mathbf{y}}. \quad (2.52)$$

Here, v_{ii} are some scalars weakly dependent on E and $\boldsymbol{\sigma}_i$ are the Pauli matrices. Moreover, a detailed analysis shows that

$$V_{\text{ef}}(E) \approx V_{\text{av}} - \alpha E, \quad \text{with } \alpha = v_{xx} - 1, \tag{2.53}$$

where V_{av} is the mean value of the microscopic potential $V = V(x)$.[7]

From these results, it follows that the time evolution of the macroscopic wave function in the spatial domain is determined by a modified Dirac equation[68]:

$$\left[-i\hbar v_F \boldsymbol{\sigma}_x \frac{\partial}{\partial x} - i\hbar v_F \chi \boldsymbol{\sigma}_y \frac{\partial}{\partial y} + V_{av}\right] \cdot \Psi = i\hbar \frac{\partial}{\partial t}\Psi, \quad t > 0, \tag{2.54}$$

where $\chi = v_{yy}/v_{xx}$ is by definition the *anisotropy ratio* of the superlattice. Thus, the superlattice can be regarded as a continuum described by the effective parameters χ and V_{av}. Comparing Eq. (2.54) with the original Dirac equation (2.6), it is seen that the main effect of the fluctuating electrostatic potential is to tailor the electron velocity in the direction perpendicular to the stratification, so that it becomes $v_F|\chi|$. Thus, the anisotropy ratio controls the velocity of the electrons along the y-direction. The velocity of the electrons along the x-direction is unaffected by the fluctuating potential due to the Klein tunneling effect.[7]

Figure 2.6 depicts the anisotropy ratio χ as a function of the amplitude of the oscillating part of the microscopic potential, $V_{\text{osc}} = V_1$ (see Fig. 2.1(b)). Two cases are considered: (i) a Krönig–Penney type potential formed by square barriers and (ii) a sinusoidal electrostatic potential. In both cases, for certain values of V_{osc} the anisotropy ratio can vanish, and in these conditions the superlattice is characterized by an extreme anisotropy such that the propagation along the y-direction is forbidden.

The effective medium formulation (2.54) predicts that the energy stationary states have the dispersion[7,68]:

$$|E - V_{av}| = \hbar v_F \sqrt{k_x^2 + \chi^2 k_y^2}. \tag{2.55}$$

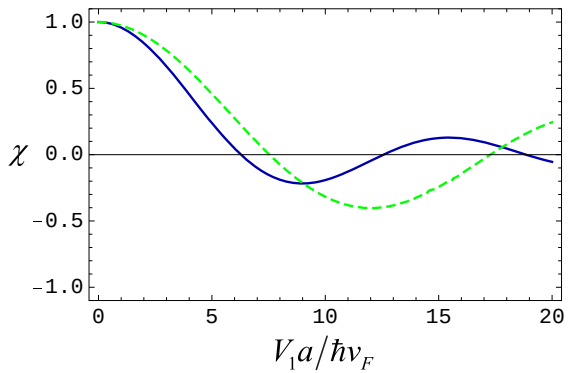


Fig. 2.6. Anisotropy ratio as a function of the amplitude of the oscillating part of the microscopic potential for (i) (solid line) a graphene superlattice with electric potential that alternates between the values $V_2 = -V_1$ in regions with thickness $d_1 = d_2 = a/2$. (ii) (dashed line) a sinusoidal-type electric potential with $V = V_1 \sin(2\pi x/a)$. The period of the superlattice is denoted by a. Adapted from Ref. 7.

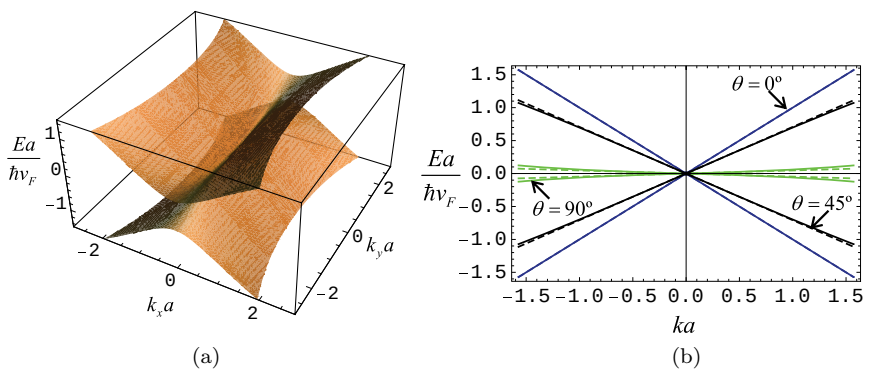


Fig. 2.7. (a) Exact energy dispersion of a graphene superlattice formed by square barriers with $d_1 = d_2 = a/2$, $V_2 = -V_1$ and $V_1 a/\hbar v_F = 6.0$ (b) Dispersion of the energy eigenstates for $\mathbf{k} = k(\cos\theta, \sin\theta)$ calculated with (i) (solid curves) the "exact" microscopic theory. (ii) (dashed curves) the effective medium model based on the parameters V_{av} and χ. Reprinted with permission from Ref. 7.

Figure 2.7(a) depicts the exact energy dispersion of a graphene superlattice with a Krönig–Penney type microscopic potential with $V_1 a/\hbar v_F = 6.0$. As seen, consistent with the fact that for $V_1 a/\hbar v_F = 6.0$ the anisotropy ratio is near zero, the graphene superlattice is

strongly anisotropic and the usual Dirac cone of pristine graphene is stretched along the y-direction. The exact energy dispersion is compared with the effective medium result (2.55) in Fig. 2.7(b), revealing an excellent agreement between the two formalisms.

As discussed in Section 2.3, the effective Hamiltonian describes the time evolution of macroscopic electronic states. To illustrate this property, the time evolution of an initial state was numerically determined using a finite-difference time-domain method using both the microscopic Hamiltonian (2.6) and the effective medium approach (Eq. (2.54)). The superlattice has a sinusoidal-type profile with period $a = 10\,\text{nm}$. The mean electric potential is set equal to zero ($V_{\text{av}} = 0$). The initial electronic state (the bright spot in Fig. 2.8(a)) is taken as a Gaussian wave packet with radial width $R_G = 2.82a$. The quasi-momentum of the wave packet is set so that it propagates along the $x-$direction with the quasi-energy $E_0 a/\hbar v_F \approx 1.9$.[69]

Figures 2.8(b), 2.8(c) and 2.8(d) show the calculated transverse profiles of the probability density function sampled at different instants of time, for pristine graphene ($\chi = 1$) and for superlattices with anisotropy ratio $\chi = 0.7$ and $\chi = 0$, respectively. As expected, for pristine graphene (Fig. 2.8(b)), the time evolution of the initial electronic state causes the wave packet to diffract and increase its characteristic size. In pristine graphene, the group velocity is independent of the direction of propagation, and hence there is no preferred direction of motion.

Quite differently, in a graphene superlattice with $\chi = 0$ (Fig. 2.8(d)) the electronic state is unaffected by diffraction and the shape of the wave front does not change with time. Crucially, the time evolution predicted by the exact microscopic Hamiltonian (dashed lines) is practically coincident with that predicted by the effective medium formulation (solid lines). Indeed, as further demonstrated in the example of Fig. 2.9(a), the effective medium Hamiltonian can determine very accurately the time evolution of an initial electronic state that is less localized than the period a. When the characteristic size of the initial state is comparable or less than a, the effective

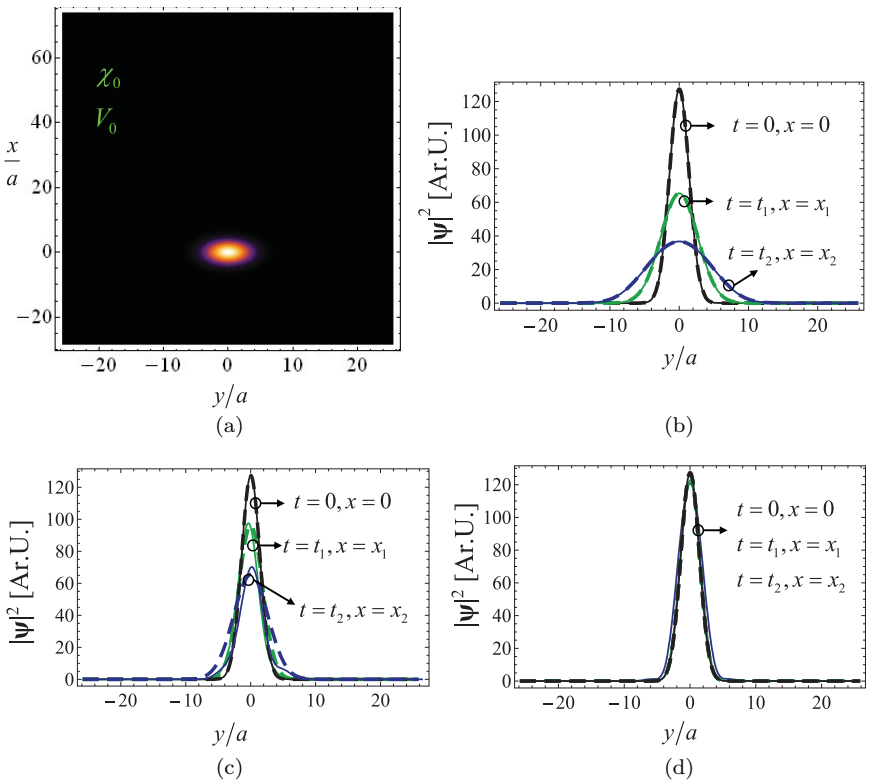


Fig. 2.8. (a) Sketch of the initial macroscopic electronic state (the bright spot). (b)–(d) Transverse profile of the probability density function (normalized to arbitrary units) at $x = 0$ and sampled at the time instant $t = 0$, at $x_1 = 12.35a$ and $t = t_1 = 2000\Delta_t$, and at $x_3 = 24.73a$ and $t = t_2 = 4000\Delta_t$, being $\Delta_t = 0.62$ as. Panel (b) is for pristine graphene ($\chi = 1$,$V = 0$), panel (c) is for a superlattice with $\chi = 0.7$ (in the microscopic model $V_{\rm osc}a/\hbar v_F \approx 3.58$), and panel (d) is for a superlattice with $\chi = 0$ (in the microscopic model $V_{\rm osc}a/\hbar v_F \approx 7.55$). In all the plots, the dashed lines represent the microscopic theory results, and the solid thick lines represent the effective medium results. The microscopic potential has a sinusoidal profile with period $a = 10\,$nm. Reprinted with permission from Ref. 69.

medium theory and the microscopic theory give diverging results (Figs. 2.9(b) and 2.9(c)).

Importantly, the effective medium formulation can be extended to graphene heterostructures formed by non-uniform (with

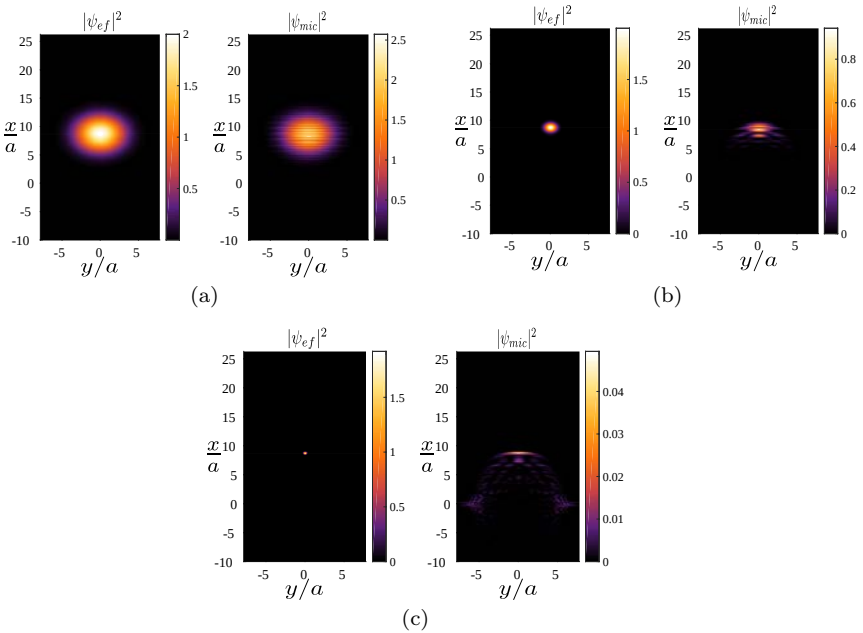


Fig. 2.9. Profiles of the probability density function after the initial state (initially centered at the origin) propagates during $1000\Delta_t$ seconds. Left: effective medium result. Right: microscopic theory result; The superlattice is characterized by an extreme anisotropy $\chi = 0$ (in the microscopic model $V_{\rm osc}a/\hbar v_F \approx 7.55$) and has the same parameters as in Fig. 2.8. The initial state has the radial width (a) $R_G = 4a$ (b) $R_G = a$ and (c) $R_G = 0.25a$. Adapted from Ref. 69.

parameters varying in space) superlattices, as well as to characterize the scattering of electron waves by graphene nanostructures. Detailed examples are reported in Refs. 68 and 69.

2.5. Homogenization Near the Corners of the Brillouin Zone

In Section 2.4, the "low energy" electronic states are always associated with a wavevector near the origin (Γ point, $\mathbf{k} = 0$). This property also applies to the graphene superlattice example. Indeed, within the validity of the microscopic Hamiltonian (2.6), the wavevector $\mathbf{k}$ is measured with respect to the K point of the Brillouin zone, and

hence it is near zero in the vicinity of the Dirac point. Importantly, at a more fundamental level one should take into account that graphene is itself a 2D periodic material described by the Schrödinger equation, with an electric potential V that has the honeycomb symmetry with two carbon atoms per unit cell.[35] Within this more fundamental description, the physics of graphene is determined by the K and K' points at the edges of the Brillouin zone, which are the points where the conduction and valence bands of graphene meet.[35]

The objective of this section is to discuss how to homogenize systems wherein the relevant wave phenomena are determined by points of the $\mathbf{k}$-space different from the origin. To better illustrate the ideas, the discussion is focused on the problem of homogenization of "artificial graphene": a 2D electron gas (2DEG) nanopatterned with scattering centers organized in a lattice with the honeycomb symmetry (Fig. 2.1(c)).[33] Such a superlattice has an electronic band structure analogous to graphene and exhibits linearly dispersing bands near the K and K' points at some Dirac energy, E_D.[33] The electron wave propagation in artificial graphene may be described by the Hamiltonian (2.5), m being the effective electron mass in the electron gas (which is assumed independent of the position) and V is a periodic function taken to be V_0 inside the circular scattering centers, and zero outside.[33,66] The circular scattering centers have radius R, and the nearest neighbor distance is a.

The effective Hamiltonian $\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k}, E)$ of the modulated 2DEG can be determined in a straightforward manner following the ideas of Section 2.3.3. The calculation of $\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k}, E)$ requires the use of numerical methods.[66] It should be noted that $\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k}, E)$ is a scalar function because the wave function in Eq. (2.5) is a scalar. Moreover, since the relevant physics is determined by the corners of the first Brillouin zone, the set B.Z. in Eq. (2.11) should be taken equal to a translated version of the first Brillouin such that both the K and K' points are interior to the set.

In order to describe the wave dynamics of the low-energy states near the K point, it is tempting to mimic the ideas of Section 2.4 and expand $\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k}, E)$ in a Taylor series around $\mathbf{k} = K$. Unfortunately,

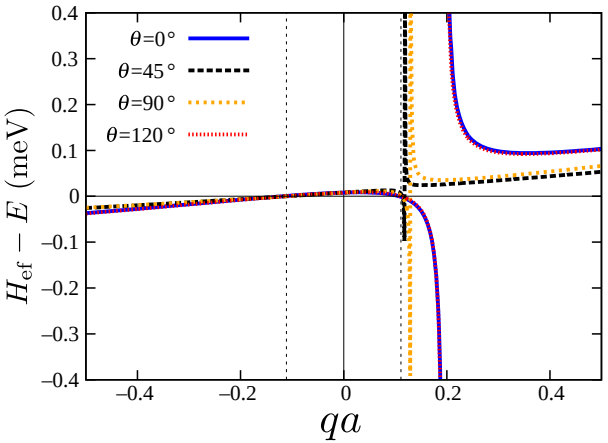


Fig. 2.10. Effective Hamiltonian near the K point as a function of the normalized wave vector $\mathbf{q} = q(\cos\theta, \sin\theta)$ for different directions of propagation and $E = -0.33\,\text{meV}$, $V_0 = -0.8\,\text{meV}$, and $R/a = 0.35$. The vertical dashed lines indicate the zeros of $\boldsymbol{\mathcal{H}}_{\rm ef} - E$. Reprinted with permission from Ref. 66.

such an approach does not work. Indeed, it turns out that the effective Hamiltonian $\boldsymbol{\mathcal{H}}_{\rm ef}(\mathbf{k}, E)$ has a direction-dependent resonant behavior near the K point and therefore the Taylor series is useless. This property is illustrated in Fig. 2.10, which depicts $\boldsymbol{\mathcal{H}}_{\rm ef}(\mathbf{k}, E) - E$ as a function of $\mathbf{q}$, $\mathbf{q} = \mathbf{k} - K$ being the wavevector measured with respect to the K point. Thus, it follows that the system is strongly spatially dispersive at the corners of the Brillouin zone.

Remarkably, it is possible to avoid the strong spatial dispersion effects by considering a pseudo-spinor description of the wave propagation, such that the macroscopic wave function becomes a two-component vector.[66] This can be done with a generalization of the theory of Section 2.3, as described next.

The strong spatial dispersion of $\boldsymbol{\mathcal{H}}_{\rm ef}(\mathbf{k}, E)$ can be attributed to the fact that the wave function can have significant fluctuations within each unit cell because there are two inequivalent scattering sites per cell. This property suggests that the definition of macroscopic state may be too restrictive for the system under study. Indeed,

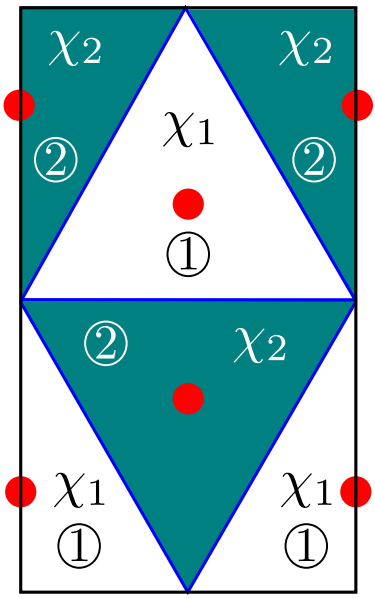


Fig. 2.11. Partition of space determined by the characteristic functions χ_1 and χ_2. The circular dots represent the scattering centers of the honeycomb lattice. Reprinted with permission from Ref. 66.

a macroscopic electronic state cannot be more localized than the lattice period, and hence the two sub-lattices of the modulated 2DEG are not distinguished in the effective medium theory. Inspired by the pseudo-spinor formalism of graphene,[35] next we introduce generalized macroscopic states that allow for the discrimination of the two sub-lattices.

As illustrated in Fig. 2.11, the idea is to consider a partition of the unit cell into two regions ($i = 1, 2$). Each region is described by a characteristic periodic function, $\chi_i(\mathbf{r})$, such that $\chi_1 + \chi_2 = 1$. The function χ_i assumes the value 1 in the ith region and is zero otherwise. By definition, a *generalized macroscopic state* is of the form $\psi = \psi_1\chi_1 + \psi_2\chi_2$ with $\psi_i = \{\psi_i\}_{\text{av}}$ ($i = 1, 2$). Evidently, this definition extends that of Section 2.3 and allows the macroscopic states to be as localized as each individual scattering center.

To obtain an effective medium description of the generalized macroscopic states, the Schrödinger equation (2.1) is rewritten as[66]

$$\underbrace{\begin{pmatrix} \chi_1\hat{H}\chi_1 & \chi_1\hat{H}\chi_2 \\ \chi_2\hat{H}\chi_1 & \chi_2\hat{H}\chi_2 \end{pmatrix}}_{\hat{H}_g} \cdot \psi_g = i\hbar\frac{\partial}{\partial t}\psi_g, \quad \psi_g = \begin{pmatrix} \chi_1\psi \\ \chi_2\psi \end{pmatrix}, \tag{2.56}$$

where $\hat{H}_g$ is a generalized two-component microscopic Hamiltonian. The two-component wave function ψ_g may be regarded as a microscopic pseudo-spinor. Each of the components of ψ_g is associated with a specific sub-lattice of the modulated 2DEG.

Similar to Section 2.3, the goal is to introduce a generalized effective Hamiltonian $\hat{H}_{g,\mathrm{ef}}$ that describes the time evolution of generalized macroscopic initial states $\psi_{t=0} = \psi_{1,t=0}\chi_1 + \psi_{2,t=0}\chi_2$. The effective Hamiltonian must ensure that $(\hat{H}_{g,\mathrm{ef}} \cdot \Psi_g)(\mathbf{r}, t) = \{(\hat{H}_g \cdot \psi_g)(\mathbf{r}, t)\}_{\mathrm{av}}$for any initial macroscopic state (Fig. 2.2). Here, $\Psi_g = \{\psi_g\}_{\mathrm{av}}$ is the macroscopic pseudo-spinor.

To determine the effective Hamiltonian in the Fourier domain, $\boldsymbol{\mathcal{H}}_{g,\mathrm{ef}}(\mathbf{k}, E)$, it is supposed that $-i\hbar\psi_{i,t=0} = f_i e^{i\mathbf{k}\cdot\mathbf{r}}$ where the weighting factors, f_i, can be chosen arbitrarily. Then, calculating the unilateral Fourier (Laplace) transform in time of Eq. (2.56) it is found that

$$(\hat{H}_g - E) \cdot \psi_g = \begin{pmatrix} \chi_1 f_1 \\ \chi_2 f_2 \end{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}}. \tag{2.57}$$

Because the effective Hamiltonian must guarantee that $(\hat{H}_g \cdot \psi_g)_{\mathrm{av}} = \boldsymbol{\mathcal{H}}_{g,\mathrm{ef}}(\mathbf{k}, E) \cdot \psi_{g,\mathrm{av}}$, the above equation implies that[66]

$$[\boldsymbol{\mathcal{H}}_{g,\mathrm{ef}}(\mathbf{k}, E) - E] \cdot \psi_{g,\mathrm{av}} = \frac{1}{2}\begin{pmatrix} f_1 \\ f_2 \end{pmatrix}, \tag{2.58}$$

where $\psi_{g,\mathrm{av}}$ and $(\hat{H}_g \cdot \psi_g)_{\mathrm{av}}$ are given by formulas analogous to Eq. (2.24). Thus, by solving the microscopic problem (2.57) for two independent sets of weighting factors f_i and by finding the corresponding $\psi_{g,\mathrm{av}}$ using Eq. (2.24), it is possible to compute the

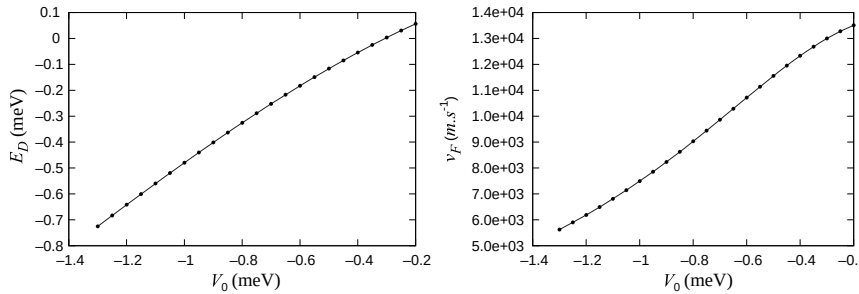


Fig. 2.12. Dirac energy (left) and Fermi velocity (right) as a function of the potential V_0 calculated using the effective medium theory for the Dirac cone near the K point. The radius of the scattering centers is $R/a = 0.35$. Adapted from Ref. 66.

two-component effective Hamiltonian with Eq. (2.58). These ideas are developed in detail in Ref. 66.

Importantly, in contrast with the single-component Hamiltonian (Fig. 2.10), the two-component Hamiltonian $\boldsymbol{\mathcal{H}}_{g,\mathrm{ef}}(\mathbf{k}, E)$ is a smooth function of the wave vector near the Dirac points K and K', and thus it can be approximated by a first-order Taylor series around the Dirac points. It was demonstrated in Ref. 66 that such an expansion leads — after a suitable renormalization of the pseudo-spinor — to the same massless 2D Dirac equation (2.6) that describes the physics of graphene. Hence, the generalized effective medium theory described here gives a "first principles" demonstration that the low-energy electron wave propagation in "artificial graphene" is described by the Dirac massless equation. Moreover, from the coefficients of the Taylor expansion of $\boldsymbol{\mathcal{H}}_{g,\mathrm{ef}}(\mathbf{k}, E)$ it is possible to compute the dependence of the equivalent Fermi velocity v_F and of the Dirac energy E_D as a function of the height of the potential wells $-V_0$ associated with the scattering centers.[66] An illustrative example is reported in Fig. 2.12.

2.6. Quadratic Forms

In the following, we discuss how to determine some quadratic forms of the electromagnetic fields (in case of light waves) or of the wave function (in case of quantum systems) relying on the effective

medium framework. Examples of such quadratic forms are the electromagnetic energy density or the probability density function.

2.6.1. *Hermitian symmetry*

A key property of the Hamiltonian operator $\hat{H}$ in quantum physics is that it is Hermitian. The Hermitian symmetry guarantees that the wave function $\psi(t)$ differs from the initial state $\psi(0)$ by a unitary transformation, and hence the norm $\langle\psi|\psi\rangle$ is independent of time. In particular, $|\psi(\mathbf{r},t)|^2$ is the probability density function for the particle position, and its integral over the space coordinates is independent of time.

The effective Hamiltonian $\hat{H}_{\text{ef}}$ defined in Section 2.3 describes the time dynamics of the envelope-function $\Psi(\mathbf{r},t)$. As previously mentioned, in contrast with the microscopic Hamiltonian, the action of the effective Hamiltonian on the macroscopic state vector is not "local" in time and depends on the past history of the state vector. Thus, the operator $\hat{H}_{\text{ef}}$ has a completely different nature than the microscopic Hamiltonian, and the notion of Hermitian symmetry does not directly apply. Yet, in the Fourier domain the Hermitian symmetry is recovered, and specifically for quantum systems, $\boldsymbol{\mathcal{H}}_{\text{ef}}(\mathbf{k},E)$ is Hermitian symmetric.

To prove such a property, we rewrite Eq. (2.25) as

$$(\hat{H}-E)\cdot\boldsymbol{\psi}=\mathbf{1}_{S\times S}\,e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{2.59}$$

where $\mathbf{1}_{S\times S}$ is the identity matrix of rank S and $\boldsymbol{\psi}=[\psi^{(1)}\ \cdots\ \psi^{(S)}]$ should be understood as a $S\times S$ matrix whose columns are the solutions $\psi^{(l)}$ of Eq. (2.25) for $-i\hbar\psi^{(l)}_{t=0}=e^{i\mathbf{k}\cdot\mathbf{r}}\mathbf{u}_l$. Multiplying both sides of Eq. (2.59) by $\boldsymbol{\psi}^\dagger$ and integrating the resulting expression over the unit cell Ω, it is found that

$$\boldsymbol{\psi}^\dagger_{\text{av}}=[\langle\psi^{(i)}|\hat{H}-E|\psi^{(j)}\rangle]_{i,j}. \tag{2.60}$$

The right-hand side represents a $S\times S$ matrix with generic i,j element given by $\langle\psi^{(i)}|\hat{H}-E|\psi^{(j)}\rangle$, where

$$\langle f|g\rangle=\frac{1}{V_{\text{cell}}}\int_\Omega d^N\mathbf{r}\,f^*(\mathbf{r})g(\mathbf{r}) \tag{2.61}$$

is the canonical inner product in the unit cell. The matrix $\boldsymbol{\psi}_{\mathrm{av}}$ is given by $\boldsymbol{\psi}_{\mathrm{av}} = [\psi_{\mathrm{av}}^{(1)} \quad \cdots \quad \psi_{\mathrm{av}}^{(S)}]$ where the column vectors are defined as in Section 2.3.4.

For quantum systems, the operator $\hat{H}$ is Hermitian with respect to the canonical inner product, and hence the matrix on the right-hand side of Eq. (2.60) is Hermitian. This property shows that $\boldsymbol{\psi}_{\mathrm{av}}$ is also a Hermitian matrix, and from Eq. (2.26) it finally follows that $\boldsymbol{\mathcal{H}}_{\mathrm{ef}}(\mathbf{k}, E)$ is Hermitian symmetric, as we wanted to prove. It is important to highlight that the Hermitian symmetry does not apply to the electromagnetic case, because the operator $\hat{H}$ for light waves (Eq. (2.4)) is not Hermitian with respect to the canonical inner product.

2.6.2. *The probability density function and the probability current for quantum systems*

A crucial observation is that the operation of spatial averaging does not commute with the multiplication operation. Thus, in general the averaged probability density function calculated with the microscopic theory cannot be identified with the squared amplitude of the envelope-function $\Psi(\mathbf{r}, t) = \{\psi(\mathbf{r}, t)\}_{\mathrm{av}}$:

$$\Psi^* \cdot \Psi(\mathbf{r}, t) \neq \{\psi^* \cdot \psi(\mathbf{r}, t)\}_{\mathrm{av}}. \tag{2.62}$$

In other words, typically $|\Psi(\mathbf{r}, t)|^2$ cannot be regarded as the probability density function for the particle position, and thus its integral over the spatial coordinates may be time dependent.[7]

This result may seem at first surprising, but actually is a consequence of the fact that $\Psi(\mathbf{r}, t)$ only describes the envelope of the microscopic wave function. Even though when the initial state is macroscopic one has $\psi = \Psi$ at the time origin $t = 0$, as time passes $\psi(\mathbf{r}, t)$ may depart significantly from a pure macroscopic state and hence it can have strong fluctuations on the scale of the unit cell. In contrast, $\Psi(\mathbf{r}, t)$ always varies slowly on the scale of the unit cell and this explains the property (2.62).

Ultimately, Eq. (2.62) is a consequence of the dispersive nature of the effective Hamiltonian, which is usually energy dependent. If

$\boldsymbol{\mathcal{H}}_{\rm ef}(\mathbf{k}, E)$ is weakly dependent on E, then from the analysis of Section 2.6.1 one may conclude that $\hat{H}_{\rm ef}$ is also Hermitian, and in that case $|\Psi(\mathbf{r}, t)|^2$ can be identified with the probability density function.

Importantly, for the energy stationary states the averaged probability density function is precisely determined by the effective Hamiltonian. Specifically, it can be proven that for Bloch stationary states (ψ) of the microscopic Hamiltonian, one has the following exact result (see Ref. 66 and the supplementary materials of Ref. 68):

$$\{\psi^* \cdot \psi\}_{\rm av} = \Psi^* \cdot \left(1 - \frac{\partial \boldsymbol{\mathcal{H}}_{\rm ef}}{\partial E}\right) \cdot \Psi. \tag{2.63}$$

Thus, for Bloch states the averaged probability density function can be written in terms of the macroscopic wave function and the energy derivative of the effective Hamiltonian. In the absence of energy dispersion, one obtains $\{\psi^* \cdot \psi\}_{\rm av} = \Psi^* \cdot \Psi$, consistent with the previous discussion.

Moreover, supposing that the microscopic Hamiltonian is of the form $\hat{H} = \hat{H}(-i\nabla, \mathbf{r})$ it is possible to demonstrate that for Bloch stationary states the spatially averaged probability current $\mathbf{j} = \psi^* \cdot \hbar^{-1} \partial \hat{H}/\partial \mathbf{k} \cdot \psi$ with $\mathbf{k} = -i\nabla$ exactly satisfies:

$$\left\{\psi^* \cdot \frac{1}{\hbar} \frac{\partial \hat{H}}{\partial \mathbf{k}}(-i\nabla, \mathbf{r}) \cdot \psi\right\}_{\rm av} = \Psi^* \cdot \frac{1}{\hbar} \frac{\partial \boldsymbol{\mathcal{H}}_{\rm ef}}{\partial \mathbf{k}} \cdot \Psi. \tag{2.64}$$

Hence, in the macroscopic framework, the spatially averaged probability current can be identified with the right-hand side of Eq. (2.64). Applications of Eqs. (2.63)–(2.64) are discussed in Refs. 66, 68, 70.

In summary, for Bloch stationary states both the probability density function and the probability current are rigorously determined by the effective medium theory. Note that for Bloch waves one has

$$\{\psi^* \cdot \psi\}_{\rm av} = \frac{1}{V_{\rm cell}} \int_\Omega \psi^* \cdot \psi \, d^N \mathbf{r}. \tag{2.65}$$

A similar expression can be written for $\left\{\psi^* \cdot \frac{1}{\hbar} \frac{\partial \hat{H}}{\partial \mathbf{k}}(-i\nabla, \mathbf{r}) \cdot \psi\right\}_{\rm av}$.

2.6.3. *The energy density and the Poynting vector for electromagnetic systems*

The light counterparts of the probability density function and of the probability current are the electromagnetic energy density and the Poynting vector. Interestingly, similar to the case of electron waves, such quadratic forms of the fields can be rigorously determined in a time-harmonic stationary regime. Specifically, let us consider a generic periodic electromagnetic metamaterial that is described at the microscopic level by a lossless permittivity function $\varepsilon = \varepsilon(\mathbf{r}, \omega)$ and by the permeability $\mu = \mu_0$. As discussed in Section 2.3.5, in the effective medium approach such systems are fully characterized by a non-local dielectric function $\overline{\varepsilon}_{\text{ef}}(\mathbf{k}, \omega)$. Let us consider a generic Bloch electromagnetic mode of the metamaterial with a time-harmonic variation $e^{-i\omega t}$ and associated with the complex microscopic electromagnetic fields $\mathbf{e}$ and $\mathbf{b}$. The spatially averaged electromagnetic energy density and Poynting vector are defined as follows:

$$W_{\text{av}} = \frac{1}{4V_{\text{cell}}} \int_{\Omega} \frac{|\mathbf{b}|^2}{\mu_0} d^3\mathbf{r} + \frac{1}{4V_{\text{cell}}} \int_{\Omega} \frac{\partial}{\partial \omega}(\omega\varepsilon)|\mathbf{e}|^2 d^3\mathbf{r}, \tag{2.66}$$

$$\mathbf{S}_{\text{av}} = \frac{1}{V_{\text{cell}}} \int_{\Omega} \frac{1}{2}\text{Re}\left\{\mathbf{e} \times \frac{\mathbf{b}^*}{\mu_0}\right\} d^3\mathbf{r}. \tag{2.67}$$

Then, in analogy with Section 2.6.2, it is possible to prove that W_{av} and $\mathbf{S}_{\text{av}}$ can be written in terms of the macroscopic fields and of the effective dielectric function as follows[71–73]:

$$W_{\text{av}} = \frac{1}{4}\frac{|\mathbf{B}|^2}{\mu_0} + \frac{1}{4}\mathbf{E}^* \cdot \frac{\partial}{\partial \omega}(\omega\overline{\varepsilon}_{\text{ef}}) \cdot \mathbf{E}, \tag{2.68}$$

$$\mathbf{S}_{\text{av}} \cdot \hat{l} = \frac{1}{2}\text{Re}\left\{\mathbf{E} \times \frac{\mathbf{B}^*}{\mu_0}\right\} \cdot \hat{l} - \frac{1}{4}\omega\mathbf{E}^* \cdot \frac{\partial \overline{\varepsilon}_{\text{ef}}}{\partial k_l}(\mathbf{k}, \omega) \cdot \mathbf{E}, \quad l = x, y, z. \tag{2.69}$$

In the above, $\mathbf{E} = \{\mathbf{e}\}_{\text{av}}$ and $\mathbf{B} = \{\mathbf{b}\}_{\text{av}}$ represent the macroscopic fields associated with the Bloch mode with wavevector $\mathbf{k}$. It is emphasized that Eqs. (2.68)–(2.69) are exact for lossless electromagnetic systems.[71,73]

Moreover, if the metamaterial response can be assumed to a good approximation local so that Eq. (2.36) holds, then Eqs. (2.68)–(2.69) reduce to

$$W_{\text{av}} = \frac{1}{4}\frac{\partial}{\partial\omega}(\omega\mu_0\mu_L)|\mathbf{H}_L|^2 + \frac{1}{4}\frac{\partial}{\partial\omega}(\omega\varepsilon_0\varepsilon_L)|\mathbf{E}|^2, \tag{2.70}$$

$$\mathbf{S}_{\text{av}} = \frac{1}{2}\text{Re}\{\mathbf{E}\times\mathbf{H}_L^*\}, \tag{2.71}$$

where $\mathbf{H}_L \equiv \mu_0^{-1}\mu_L^{-1}\mathbf{B}$ is by definition the macroscopic local magnetic field. It is relevant to highlight that the formulas (2.68)–(2.69) for the stored energy density and for the Poynting vector are coincident with well-known formulas for an electromagnetic spatially dispersive continuum.[48] Similarly, Eqs. (2.70)–(2.71) are coincident with the classical textbook formulas for the energy density and for the Poynting vector in dispersive isotropic dielectrics.[48] Thus, the theory of Refs. 71–73 proves that these classical textbook formulas for a continuum describe precisely the spatially averaged microscopic energy density and Poynting vector within the effective medium theory discussed here.

To illustrate the described ideas, next we consider a 2D metamaterial formed by a square array of high-index dielectric cylindrical inclusions with radius $R/a = 0.435$, permittivity $\varepsilon_d = 50.47$ and permeability $\mu = 1$ embedded in a host material characterized by a lossless Drude dispersion model with $\varepsilon_h = 1 - \omega_p^2/\omega^2$. The lattice period is a and the normalized plasma frequency is taken equal to $\omega_p a/c = 1.0$. The geometry of the unit cell is depicted in the inset of Fig. 2.13(d). The magnetic field is polarized along the z-direction (parallel to the cylinders axes) and the electric field is in the xoy plane. Similar to the example of Section 2.4.1, to a good approximation this metamaterial has a local response and is characterized by local effective parameters that satisfy $\varepsilon_L \approx \mu_L$ in a broad frequency range (see Fig. 2.13(a)). In particular, below the plasma frequency $\omega < \omega_p$ the material behaves as a double-negative material with simultaneously negative permittivity and permeability.[73]

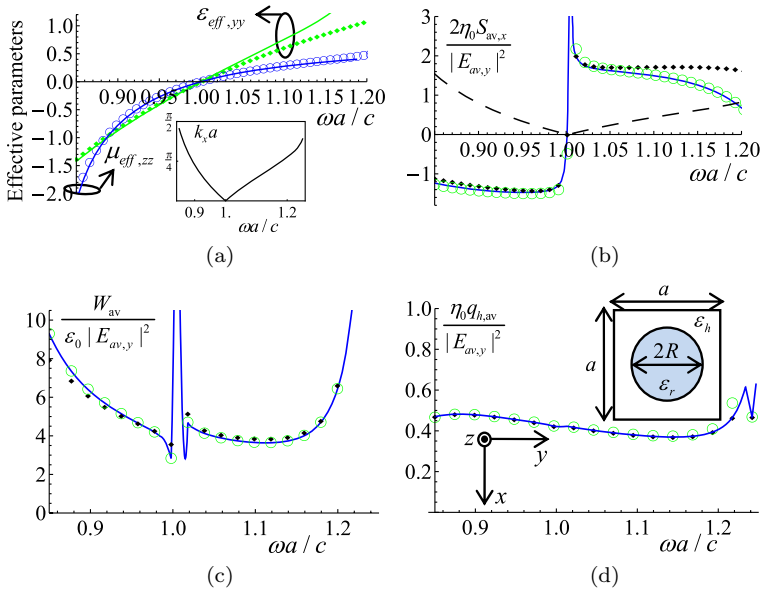


Fig. 2.13. (a) Local permittivity ε_L and local permeability μ_L versus the normalized frequency $\omega a/c$. The discrete symbols are calculated with the effective medium theory and the solid lines are calculated using the Clausius–Mossotti formulas. The inset shows the wavevector k_x as a function of frequency. (b) x-component of the Poynting vector calculated using: (i) Averaged microscopic Poynting vector (solid line) [Eq. (2.67)]; (ii) Non-local homogenization (circles) [Eq. (2.69)]; (iii) Local effective parameters (diamonds) [Eq. (2.71)]; (iv) Result obtained with the erroneous definition of the Poynting vector of Ref. 74 (dashed line). (c) Electromagnetic energy density calculated using: (i) Averaged microscopic energy density (solid line) [Eq. (2.66)]; (ii) Non-local homogenization (circles) [Eq. (2.68)]; (iii) Local effective parameters (diamonds) [Eq. (2.70)]. (d) Heating rate calculated using: (i) Averaged microscopic heating rate (solid line) [Eq. (2.72)]; (ii) Non-local homogenization (circles) [Eq. (2.73)]; (iii) Local effective parameters (diamonds) [Eq. (2.74)]. The unit cell geometry is shown in the inset. Reprinted with permission from Ref. 73.

Figures 2.13(b) and 2.13(c) show a comparison between the numerically calculated spatially averaged Poynting vector and energy density (Eqs. (2.66)–(2.67)) and the same quantities determined using the non-local electromagnetic continuum formulas (Eqs. (2.68)–(2.69)). The non-local dielectric function is obtained as explained in Section 2.3.5 and the macroscopic fields are found by

spatially averaging the microscopic Bloch modes. As seen, the numerical results confirm that Eqs. (2.66)–(2.67) and Eqs. (2.68)–(2.69) give coincident results. Figures 2.13(b) and 2.13(c) also reveal that the results predicted by the local approximation (Eqs. (2.70)–(2.71)) are quite accurate, especially below the plasma frequency. In contrast, the Poynting vector computed with the erroneous formula $\mathbf{S}_{\text{av}} = 1/2\text{Re}\{\mathbf{E} \times \mathbf{B}^*\}$ proposed in Ref. 74 to describe the energy density flux in metamaterials gives a completely disparate result. Indeed, such a formula neglects the artificial magnetism induced by the high-permittivity cylinders, and incorrectly implies that negative refraction and backward propagation are impossible in the metamaterial.[71,72,74]

It is emphasized that the equivalence between Eqs. (2.66)–(2.67) and Eqs. (2.68)–(2.69) is only observed for lossless materials. In case of material loss, the equivalence is only approximate.[73] Interestingly, in the lossy case the effective medium theory can predict exactly the heating rate due to the material absorption in time-harmonic regime. Specifically, defining the spatially averaged microscopic heating rate as

$$q_{\text{av}} = \frac{1}{V_{\text{cell}}} \int_{\Omega} \frac{\omega}{2} \varepsilon''(\mathbf{r}) |\mathbf{e}(\mathbf{r})|^2 d^3\mathbf{r}, \tag{2.72}$$

it can be shown that it is exactly coincident with the result predicted by the corresponding continuum formula[71–73]:

$$q_{\text{av}} = \frac{1}{2}\text{Re}\{-i\omega \mathbf{E}^* \cdot \overline{\varepsilon}_{\text{ef}}(\mathbf{k}, \omega) \cdot \mathbf{E}\}. \tag{2.73}$$

When the material response is approximately local, the continuum formula becomes simply

$$q_{\text{av}} = \frac{1}{2}\omega\varepsilon_0\varepsilon_L''(\omega)|\mathbf{E}|^2 + \frac{1}{2}\omega\mu_0\mu_L''(\omega)|\mathbf{H}_L|^2. \tag{2.74}$$

Figure 2.13(d) shows the comparison among these three definitions (2.72)–(2.74) for a host medium with normalized damping frequency $\Gamma/\omega_p = 0.1$ and a cylinder permittivity $\varepsilon_d = 50.47 + 0.1i$.

The numerical results confirm that Eqs. (2.72)–(2.73) give coincident results. Figure 2.13(d) also reveals that Eq. (2.74) is nearly exact for this example.

2.7. Summary

We described a completely general self-consistent approach to characterize the wave propagation in periodic systems from an effective medium perspective. The theory relies on the introduction of an effective Hamiltonian operator that regards the system as a continuum. The effective Hamiltonian describes exactly the time evolution of the wave packet envelope when the initial state is less localized than the lattice period. In addition, the effective Hamiltonian determines completely the band diagram of the time-stationary states of the periodic system. The described theory can be applied to a wide range of physical systems. Here, we illustrated its application to the homogenization of electromagnetic metamaterials and to the homogenization of semiconductor and graphene superlattices. In particular, it was highlighted that the effective medium description can often be simplified based on the extraction of local effective parameters from the effective Hamiltonian, even for systems wherein the relevant physics is determined by points at the corners of the Brillouin zone. Finally, it was shown that the effective Hamiltonian determines exactly some quadratic forms related to the energy density and the energy transport.

References

1. van Kranendonk, J. and Sipe, J. E. (1977). Foundations of the macroscopic electromagnetic theory of dielectric media, chapter 5 in *Progress in Optics XV* (North-Holland, New York) edited by E. Wolf.
2. Mahan, G. D. and Obermair, G. (1969). Polaritons at surfaces, *Phys. Rev.*, **183**, p. 834.
3. Sipe, J. and van Kranendonk, J. (1974). Macroscopic electromagnetic theory of resonant dielectrics, *Phys. Rev. A* **9**, p. 1806.
4. Lamb, W., Wood, D. M. and Ashcroft, N. W. (1980). Long-wavelength electromagnetic propagation in heterogeneous media, *Phys. Rev. B* **21**, p. 2248.
5. Ashcroft, N. W. and Mermin, N. D. (1976). *Solid State Physics* (Brooks Cole, California).

6. Silveirinha, M. G. (2007). A metamaterial homogenization approach with application to the characterization of microstructured composites with negative parameters, *Phys. Rev. B* **75**, p. 115104.
7. Silveirinha, M. G. and Engheta, N. (2012). Effective medium approach to electron waves: Graphene superlattices, *Phys. Rev. B* **85**, p. 195413.
8. Engheta, N. and Ziolkowski, R. W. (2006). *Electromagnetic Metamaterials: Physics and Engineering Explorations* (Wiley-IEEE Press, Berlin).
9. Smith, D. R., Schultz, S., Markos, P. and Soukoulis, C. M. (2002). Determination of effective permittivity and permeability of metamaterials from reflection and transmission coefficients, *Phys. Rev. B* **65**, p. 195104.
10. Smith, D. R. and Pendry, J. B. (2006). Homogenization of metamaterials by field averaging, *J. Opt. Soc. Am. B* **23**, p. 391.
11. Sjoberg, D. (2006). Dispersive effective material parameters, *Microw. Opt. Tech. Lett.* **48**, p. 2629.
12. Simovski, C. R. (2007). Bloch material parameters of magneto-dielectric metamaterials and the concept of Bloch lattices, *Metamaterials* **1**, p. 62.
13. Simovski, C. R. and Tretyakov, S. A. (2007). Local constitutive parameters of metamaterials from an effective-medium perspective, *Phys. Rev. B* **75**, p. 195111.
14. Simovski, C. R. (2010). On electromagnetic characterization and homogenization of nanostructured metamaterials, *J. Opt.* **13**, p. 013001.
15. Smith, D. R. (2010). Analytic expressions for the constitutive parameters of magnetoelectric metamaterials, *Phys. Rev. E* **81**, p. 036605.
16. Ortiz, G. P., Martínez-Zérega, B. E., Mendoza, B. S. and Mochán, W. L. (2009). Effective optical response of metamaterials, *Phys. Rev. B* **79**, p. 245132.
17. Silveirinha, M. G. (2007). Generalized Lorentz-Lorenz formulas for microstructured materials, *Phys. Rev. B* **76**, p. 245117.
18. Silveirinha, M. G., Baena, J. D., Jelinek, L. and Marqués, R. (2009). Nonlocal Homogenization of an Array of Cubic Particles made of Resonant Rings, *Metamaterials*, **3**, p. 115.
19. Costa, J. T., Silveirinha, M. G. and Maslovski, S. I. (2009). Finite-difference frequency-domain method for the extraction of the effective parameters of metamaterials, *Phys. Rev. B* **80**, p. 235124.
20. Silveirinha, M. G. (2011). Time domain homogenization of metamaterials, *Phys. Rev. B* **83**, p. 165104.
21. Silveirinha, M. G. (2009). Nonlocal Homogenization Theory of Structured Materials, chapter in *Theory and Phenomena of Artificial Materials*, Vol. I (CRC press, BocaRaton), edited by F. Capolino.
22. Fietz, C. and Shvets, G. (2010). Current-driven metamaterial homogenization, *Physica B* **405**, p. 2930.
23. Alù, A. (2011). First-principles homogenization theory for periodic metamaterials, *Phys. Rev. B* **84**, p. 075153.
24. Sozio, V., Vallecchi, A., Albani, M. and Capolino, F. (2015). Generalized Lorentz-Lorenz homogenization formulas for binary lattice metamaterials, *Phys. Rev. B* **91**, p. 205127.

25. Yaghjian, A. D., Alù, A. and Silveirinha, M. G. (2013). Homogenization of spatially dispersive metamaterial arrays in terms of generalized electric and magnetic polarizations, *Photonic Nanostruct.* **11**, p. 374.
26. Yaghjian, A. D., Alù, A. and Silveirinha, M. G. (2013). Anisotropic representation for spatially dispersive periodic metamaterial arrays, chapter 13 in *Transformation Electromagnetics and Metamaterials: Fundamental Principles and Applications* (Springer-Verlag, London), edited by D. H. Werner and D.-H. Kwon.
27. Esaki, L. and Tsu, R. (1970). Superlattice and Negative Differential Conductivity in Semiconductors, *IBM Journal of Research and Development*, **14**, p. 61.
28. Döhler, G. H. (1986). Doping superlattices ("n-i-p-i crystals"), *IEEE J. Quant. Electron.* **22**, p. 1682.
29. Kroemer, H. (2001). Nobel lecture: Quasielectric fields and band offsets: Teaching electrons new tricks, *Rev. Mod. Phys.* **73**, p. 783.
30. Capasso, F. (1987). Band-Gap Engineering: From Physics and Materials to New Semiconductor Devices, *Science* **235**, p. 172.
31. Faist, J., Capasso, F., Sivco, D. L., Sirtori, C., Hutchinson, A. L. and Cho, A. Y. (1994). Quantum cascade laser, *Science* **264**, p. 553.
32. Meyer, J. R., Hoffman, C. A., Bartoli, F. J., Han, J. W., Cook, J. W., Schetzina, J. F., Chu, X., Faurie, J. P. and Schulman, J. N. (1988). Ultrahigh electron and hole mobilities in zero-gap Hg-based superlattices, *Phys. Rev. B* **38**, p. 2204(R).
33. Gibertini, M., Singha, A., Pellegrini, V., Polini, M., Vignale, G., Pinczuk, A., Pfeiffer, L. N. and West, K. W. (2009). Engineering artificial graphene in a two-dimensional electron gas, *Phys. Rev. B* **79**, p. 241406(R).
34. Novoselov, K. S., Geim, A. K., Morozov, S. V., Jiang, D., Zhang, Y., Dubonos, S. V., Grigorieva, I. V. and Firsov, A. A. (2004). Electric Field Effect in Atomically Thin Carbon Films. *Science* **306**, p. 666.
35. Castro Neto, A. H., Guinea, F., Peres, N. M. R., Novoselov K. S. and Geim, A. K. (2009). The electronic properties of graphene, *Rev. Mod. Phys.* **81**, p. 109.
36. Novoselov, K. S., Geim, A. K., Morozov, S. V., Jiang, D., Katsnelson, M. I., Grigorieva, I. V., Dubonos, S. V. and Firsov, A. A. (2005). Two-dimensional gas of massless Dirac fermions in graphene, *Nature* **438**, p. 197.
37. Cheianov, V. V., Fal'ko, V. and Altshuler, B. L. (2007). The focusing of electron flow and a Veselago lens in graphene p-n junctions, *Science* **315**, p. 1252.
38. Rozhkov, A. V., Giavaras, G., Bliokh, Y. P., Freilikher, V. and Nori, F. (2011). Electronic properties of mesoscopic graphene structures: Charge confinement and control of spin and charge transport, *Physics Reports* **77**, p. 503.
39. Park, C. H., Yang, L., Son, Y. W., Cohen, M. L. and Louie, S. G. (2008). New Generation of Massless Dirac Fermions in Graphene under External Periodic Potentials, *Phys. Rev. Lett.* **101**, p. 126804.

40. Park, C.-H., Yang, L., Son, Y.-W., Cohen, M. L. and Louie, S. G. (2008). Anisotropic behaviours of massless Dirac fermions in graphene under periodic potentials, *Nat. Phys.* **4**, p. 213.
41. Park, C.-H., Son, Y.-W., Yang, L., Cohen, M. L. and Louie, S. G. (2008). Electron Beam Supercollimation in Graphene Superlattices, *Nano Lett.* **9**, p. 2920.
42. Bliokh, Y. P., Freilikher, V., Savel'ev, S. and Nori, F. (2009). Transport and localization in periodic and disordered graphene superlattices, *Phys. Rev. B* **79**, p. 075123.
43. Wang, L.-G. and Zhu, S.-Y. (2010). Electronic band gaps and transport properties in graphene superlattices with one-dimensional periodic potentials of square barriers, *Phys. Rev. B* **81**, p. 205444.
44. Yankowitz, M., Xue, J., Cormode, D., Sanchez-Yamagishi, J. D., Watanabe, K., Taniguchi, T., Jarillo-Herrero, P., Jacquod, P. and LeRoyet, B. J. (2012). Emergence of superlattice Dirac points in graphene on hexagonal boron nitride, *Nat. Phys.* **8**, p. 382.
45. Ponomarenko, L. A., Gorbachev, R. V., Yu, G. L., Elias, D. C., Jalil, R., Patel, A. A., Mishchenko, A., Mayorov, A. S., Woods, C. R., Wallbank, J. R., Mucha-Kruczynski, M., Piot, B. A., Potemski, M., Grigorieva, I. V., Novoselov, K. S., Guinea, F., Fal'ko, V. I. and Geim, A. K. (2013). Cloning of Dirac fermions in graphene superlattices, *Nature* **497**, p. 594.
46. Russakov, G. (1970). A derivation of the macroscopic Maxwell Equations, *Am. J. Physics* **38**, p. 1188.
47. Silveirinha, M. G. and Engheta, N. (2012). Metamaterial-Inspired Model for Electron Waves in Bulk Semiconductors, *Phys. Rev. B* **86**, p. 245302.
48. Agranovich, V. and Ginzburg, V. (1966). *Spatial Dispersion in Crystal Optics and the Theory of Excitons* (Wiley-Interscience, New York).
49. Pendry, J. B. (2000). Negative refraction makes a perfect lens, *Phys. Rev. Lett.*, **85**, p. 3966.
50. Silveirinha, M. G. and Engheta, N. (2007). Design of matched zero-index metamaterials using nonmagnetic inclusions in epsilon-near-zero media, *Phys. Rev. B* **75**, p. 075119.
51. Costa, J. T. and Silveirinha, M. G. (2011). Mimicking the Veselago-Pendry's lens with broadband matched double negative metamaterials, *Phys. Rev. B* **84**, p. 155131.
52. Lewin, L. (1947). The electrical constants of a material loaded with spherical particles, *Proc. Inst. Elec. Eng.* **94**, p. 65.
53. Voon, L. and Willatzen, M. (2009). *The $K \cdot P$ Method* (Springer, Berlin).
54. Yu, P. Y. and Cardona, M. 2005. *Fundamentals of Semiconductors*, 3rd Ed. (Springer, Berlin).
55. Bastard, G. (1988). *Wave Mechanics Applied to Semiconductor Heterostructures* (John Wiley).
56. Kane, E. O. (1957). Band structure of indium antimonide, *J. Phys. Chem. Solids* **1**, p. 249.

57. Chuang, S. L. (2009). *Physics of Photonic Devices*, 2nd edn. (John Wiley & Sons, Hoboken, NJ).
58. Burt, M. G. (1992). The justification for applying the effective-mass approximation to microstructures, *J. Phys.: Condens. Matter* **4**, p. 6651.
59. Burt, M. G. (1999). Fundamentals of envelope function theory for electronic states and photonic modes in nanostructures, *J. Phys.: Condens. Matter* **11**, R53.
60. Foreman, B. A. (1993). Effective-mass Hamiltonian and boundary conditions for the valence bands of semiconductor microstructures, *Phys. Rev. B* **48**, 4964(R).
61. Bastard, G. (1981). Superlattice band structure in the envelope-function approximation, *Phys. Rev. B* **24**, p. 5693.
62. Bastard, G. (1982). Theoretical investigations of superlattice band structure in the envelope-function approximation, *Phys. Rev. B* **25**, p. 7584.
63. Silveirinha, M. G. and Engheta, N. (2012). Transformation Electronics: Tailoring the Effective Mass of Electrons, *Phys. Rev. B* **86**, 161104(R).
64. Jelinek, L., Baena, J. D., Voves, J. and Marqués, R. (2011). Metamaterial-inspired perfect tunnelling in semiconductor heterostructures, *New J. Phys.* **13** p. 083011.
65. Silveirinha, M. G. and Engheta, N. (2013). Spatial Delocalization and Perfect Tunneling of Matter Waves: Electron Perfect Lens, *Phys. Rev. Lett.*, **110**, p. 213902.
66. Lannebère, S. and Silveirinha, M. G. (2015). Effective Hamiltonian for electron waves in artificial graphene: A first principles derivation, *Phys. Rev. B* **91**, p. 045416.
67. Silveirinha, M. G. and Engheta, N. (2014). Giant nonlinearity in Zero-Gap Semiconductor Superlattices, *Phys. Rev. B* **89**, p. 085205.
68. Fernandes, D. E., Engheta, N. and Silveirinha, M. G. (2014). Wormhole for electron waves in graphene, *Phys. Rev. B* **90**, p. 041406(R).
69. Fernandes, D. E., Rodrigues, M., Falcão, G. and Silveirinha, M. G. (2016). Time evolution of electron waves in graphene superlattices, *AIP Advances* **6**, p. 075109.
70. Hrebikova, I., Jelinek, L. and Silveirinha, M. G. (2015). Embedded energy state in an open semiconductor heterostructure, *Phys. Rev. B* **92**, p, 155303.
71. Silveirinha, M. G. (2009). Poynting vector, heating rate, and stored energy in structured materials: a first principles derivation, *Phys. Rev. B* **80**, p. 235120.
72. Silveirinha, M. G. (2010). Reply to comment on "Poynting vector, heating rate, and stored energy in structured materials: A first principles derivation", *Phys. Rev. B* **82**, p. 037104.
73. Costa, J. T., Silveirinha, M. G. and Alù, A. (2011). Poynting Vector in Negative-Index Metamaterials, *Phys. Rev. B* **83**, p. 165120.
74. Markel, V. A. (2008). Correct definition of the Poynting vector in electrically and magnetically polarizable medium reveals that negative refraction is impossible, *Opt. Express* **16**, p. 19152.

CHAPTER 3

Hyperbolic Metamaterials

IGOR I. SMOLYANINOV

University of Maryland, USA
smoly@umd.edu

Hyperbolic metamaterials were originally introduced to overcome the diffraction limit of optical imaging. Soon thereafter, it was realized that hyperbolic metamaterials demonstrate a number of novel phenomena resulting from the broadband singular behavior of their density of photonic states. These novel phenomena and applications include super-resolution imaging, new stealth technologies, enhanced quantum electrodynamic effects, thermal hyperconductivity, superconductivity and interesting gravitation theory analogs. Here, we briefly review typical material systems which exhibit hyperbolic behavior and outline important applications of hyperbolic metamaterials.

3.1. Hyperbolic Metamaterial Geometries and Basic Properties

Hyperbolic metamaterials are extremely anisotropic uniaxial materials which behave like a metal in one direction and like a dielectric in the orthogonal direction. Originally introduced to overcome the diffraction limit of optical imaging,[1,2] hyperbolic metamaterials demonstrate a number of novel phenomena resulting from the broadband singular behavior of their density of

photonic states,[3] which range from super-resolution imaging[2,4,5] to enhanced quantum electrodynamic effects,[6–8] new stealth technology,[9] thermal hyperconductivity,[10] high critical temperature (T_c) superconductivity,[11,12] and interesting gravitation theory analogs.[3,13–17] In the early days of metamaterial research, it was believed that only artificially structured materials may exhibit hyperbolic properties. However, later on it was realized that quite a few natural materials may exhibit hyperbolic properties in some frequency ranges.[11,18] Moreover, even the physical vacuum may exhibit hyperbolic metamaterial properties if subjected to a very strong magnetic field.[19]

Basic electromagnetic properties of hyperbolic metamaterials may be understood by considering a non-magnetic uniaxial anisotropic material with dielectric permittivities $\varepsilon_x = \varepsilon_y = \varepsilon_1$ and $\varepsilon_z = \varepsilon_2$. Any electromagnetic field propagating in this material may be expressed as a sum of ordinary and extraordinary contributions, each of these being a sum of an arbitrary number of plane waves polarized in the ordinary ($E_z = 0$) and extraordinary ($E_z \neq 0$) directions. Let us define a "scalar" extraordinary wave function as $\varphi = E_z$ so that the ordinary portion of the electromagnetic field does not contribute to φ. Maxwell equations in the frequency domain result in the following wave equation for φ_ω if ε_1 and ε_2 are kept constant inside the metamaterial[3]:

$$\frac{\omega^2}{c^2}\varphi_\omega = -\frac{\partial^2\varphi_\omega}{\varepsilon_1 \partial z^2} - \frac{1}{\varepsilon_2}\left(\frac{\partial^2\varphi_\omega}{\partial x^2} + \frac{\partial^2\varphi_\omega}{\partial y^2}\right). \tag{3.1}$$

While in ordinary elliptic anisotropic media, both ε_1 and ε_2 are positive, in hyperbolic metamaterials ε_1 and ε_2 have opposite signs. These metamaterials are typically composed of multilayer metal–dielectric or metal wire array structures, as shown in Fig. 3.1. The opposite signs of ε_1 and ε_2 lead to two important consequences. For extraordinary waves in a usual uniaxial dielectric metamaterial, the dispersion law

$$\frac{k_{xy}^2}{\varepsilon_2} + \frac{k_z^2}{\varepsilon_1} = \frac{\omega^2}{c^2}, \tag{3.2}$$

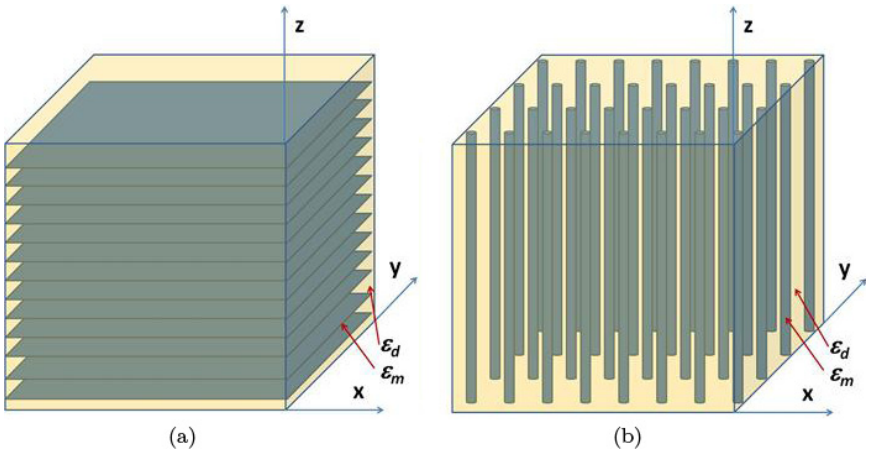


Fig. 3.1. Typical geometries of hyperbolic metamaterials: (a) multilayer metal–dielectric structure, and (b) metal wire array structure.

describes an ellipsoid in the wave momentum (k-) space (which reduces to a sphere if $\varepsilon_1 = \varepsilon_2$, as shown in Fig. 3.2(a)). The absolute value of the k-vector in such a material is finite, which leads to the usual diffraction limit on resolution of regular optics. The phase space volume enclosed between two such equi-frequency surfaces is also finite, corresponding to a finite density of photonic states. However, when one of the components of the dielectric permittivity tensor is negative, Eq. (3.2) describes a hyperboloid in the phase space (Fig. 3.2(b)). As a result, the absolute value of the k-vector is not limited, thus enabling super-resolution imaging with hyperbolic metamaterials. Moreover, the phase space volume between two such hyperboloids (corresponding to different values of frequency) is infinite (see Fig. 3.2(c)). The latter divergence leads to an infinite density of photonic states. While there are many mechanisms leading to a singularity in the density of photonic states, this one is unique as it leads to the infinite value of the density of states for every frequency where different components of the dielectric permittivity have opposite signs. It is this behavior that lies at the heart of the robust performance of hyperbolic metamaterials: while disorder can change the magnitude of the dielectric permittivity components,

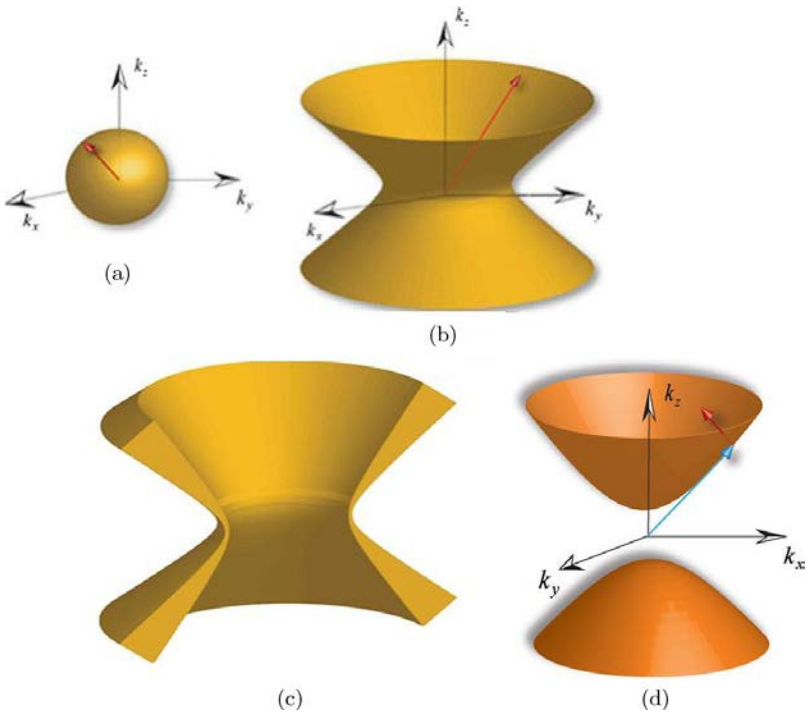


Fig. 3.2. The constant frequency surfaces for (a) isotropic dielectric ($\varepsilon_1 = \varepsilon_2 > 0$) and (b) uniaxial hyperbolic ($\varepsilon_x = \varepsilon_y = \varepsilon_1 > 0, \varepsilon_z = \varepsilon_2 < 0$) metamaterial. (c) The phase space volume between two constant frequency surfaces for the hyperbolic metamaterial.

leading to a deformation of the corresponding hyperboloid in the phase (momentum) space, it will remain a hyperboloid and will therefore still support an infinite density of states. Such effective medium description will eventually fail at the point when the wavelength of the propagating mode becomes comparable to the size of the hyperbolic metamaterial unit cell a, introducing a natural wave number cut-off:

$$k_{\max} = 1/a. \tag{3.3}$$

Depending on the metamaterial design and the fabrication method used, the unit cell size in optical metamaterials runs from $a \sim 10\,\text{nm}$ (in semiconductor[20] and metal–dielectric layered structures[6]) to

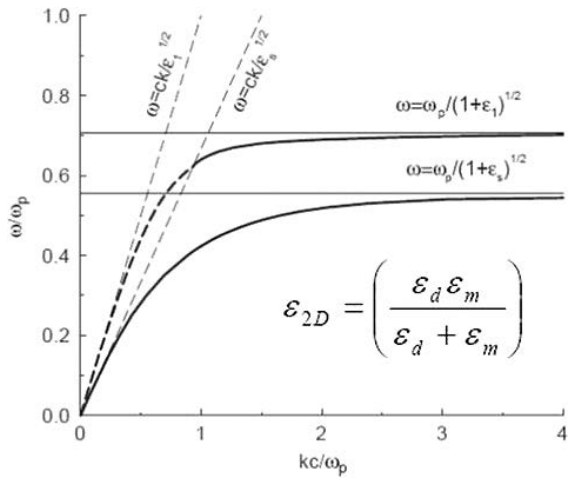


Fig. 3.3. Dispersion law of surface plasmon polaritons in the lossless approximation.

$a \sim 100$ nm (in nanowire composites[21,22]). Since the "hyperbolic" enhancement factor in the density of states[3] scales as

$$\rho(\omega) = \rho_0(\omega) \left(\frac{k_{\max}}{\omega/c} \right)^3, \tag{3.4}$$

where $\rho_0 \sim \omega^2$ is the free-space result, even with the cut-off taken into account, the hyperbolic singularity leads to the optical density of states enhancement by a factor of 10^3–10^5. Physically, the enhanced photonic density of states in the hyperbolic metamaterials originates from the waves with high wave numbers that are supported by the system. Such propagating modes that can achieve X-ray wavelengths at optical frequencies do not have an equivalent in "regular" dielectrics where $k \leq \sqrt{\varepsilon}\omega/c$. Since each of these waves can be thermally excited, a hyperbolic metamaterial shows a dramatic enhancement in the radiative transfer rates.

As has been mentioned above, artificial hyperbolic metamaterials are typically composed of multilayer metal–dielectric or metal wire array structures, as shown in Fig. 3.1. For the multilayer geometry, the diagonal components of the metamaterial permittivity can be

calculated based on the Maxwell–Garnett approximation as follows:

$$\varepsilon_1 = \varepsilon_{xy} = n\varepsilon_m + (1-n)\,\varepsilon_d, \quad \varepsilon_2 = \varepsilon_z = \frac{\varepsilon_m \varepsilon_d}{(1-n)\,\varepsilon_m + n\varepsilon_d}, \tag{3.5}$$

where n is the volume fraction of the metallic phase and $\varepsilon_m < 0$ and $\varepsilon_d > 0$ are the dielectric permittivities of the metal and dielectric, respectively.[23] The validity of Maxwell–Garnett approximation has been clearly demonstrated in Ref. 23. Analytical calculations based on the Maxwell–Garnett approximation performed for periodic array of metal nanolayers were confronted with exact numerical solutions of Maxwell equations. Excellent agreement between numerical simulations and analytical results was demonstrated. The Maxwell–Garnett approximation may also be used for a wire array metamaterial structure.[23] In this case, the diagonal components of the permittivity tensor may be obtained as

$$\varepsilon_1 = \varepsilon_{x,y} = \frac{2n\varepsilon_m\varepsilon_d + (1-n)\varepsilon_d(\varepsilon_d + \varepsilon_m)}{(1-n)\,(\varepsilon_d + \varepsilon_m) + 2n\varepsilon_d},$$

$$\varepsilon_2 = \varepsilon_z = n\varepsilon_m + (1-n)\varepsilon_d. \tag{3.6}$$

Since both ε_m and ε_d depend on frequency, the frequency regions where ε_1 and ε_2 have opposite signs may be typically found for both multilayer and wire array geometries. Depending on the actual signs of ε_1 and ε_2, the phase space shape of the hyperbolic dispersion law may be either a one-sheet ($\varepsilon_2 > 0$ and $\varepsilon_1 > 0$, see Fig. 3.2(b)) or two-sheet ($\varepsilon_2 < 0$ and $\varepsilon_1 > 0$, see Fig. 3.2(d)) hyperboloid. However, in both cases the k-vector is not limited, and the photonic density of states exhibits broadband divergent behavior.

We should also note that it is relatively easy to emulate various 3D hyperbolic metamaterial geometries by planar plasmonic metamaterial arrangements. While rigorous description of such metamaterials in terms of Diakonov surface plasmons (SPs) may be found in Ref. 24, qualitative analogy between 3D and 2D metamaterials may be explained as follows. An SP propagating over a flat metal–dielectric interface may be described by its well-known dispersion

relation shown in Fig. 3.3.

$$k_p = \frac{\omega}{c}\left(\frac{\varepsilon_d \varepsilon_m}{\varepsilon_d + \varepsilon_m}\right)^{1/2}, \tag{3.7}$$

where metal layer is considered to be thick, and $\varepsilon_m(\omega)$ and $\varepsilon_d(\omega)$ are the frequency-dependent dielectric constants of the metal and dielectric, respectively.[25] Thus, similar to the 3D case, we may introduce an effective 2D dielectric constant ε_{2D}, which characterizes the way in which SPs perceive the dielectric material deposited onto the metal surface. By requiring that $k_p = \varepsilon_{2\mathrm{D}}^{1/2}\omega/c$, we obtain

$$\varepsilon_{2\mathrm{D}} = \left(\frac{\varepsilon_d \varepsilon_m}{\varepsilon_d + \varepsilon_m}\right). \tag{3.8}$$

Equation (3.8) makes it obvious that depending on the plasmon frequency, SPs perceive the dielectric material bounding the metal surface (for example a PMMA layer) in drastically different ways. At low frequencies $\varepsilon_{2\mathrm{D}} \approx \varepsilon_d$, so that plasmons perceive a PMMA layer as a dielectric. On the other hand, at high enough frequencies at which $\varepsilon_d(\omega) > -\varepsilon_m(\omega)$ (this happens around $\lambda_0 \sim 500\,\mathrm{nm}$ for a PMMA layer), $\varepsilon_{2\mathrm{D}}$ changes sign and becomes negative. Thus, around $\lambda_0 \sim 500\,\mathrm{nm}$, plasmons perceive a PMMA layer on gold as an "effective metal". As a result, at around $\lambda_0 \sim 500\,\mathrm{nm}$ plasmons perceive a PMMA stripe pattern on gold substrate as a layered hyperbolic metamaterial shown in Fig. 3.1(a). Fabrication of such plasmonic hyperbolic metamaterials in two dimensions requires only very simple and common lithographic techniques.[4]

3.2. Super-Resolution Imaging Using Hyperbolic Metamaterials: The Hyperlens

Optical microscopy is one of the oldest research tools. Its development began in about 1590 with the observation by the Dutch spectacle maker Zaccharias Janssen and his son Hans that a combination of lenses in a tube made small objects appear larger. In 1609, Galileo

Galilei improved on their ideas and developed an *occhiolino* or compound microscope with a convex and a concave lens The acknowledged "father" of microscopy is, however, Anton van Leeuwenhoek (1632–1723) who developed improved grinding and polishing techniques for making short focal length lenses, and he was the first person to consequently see bacteria, protozoa and blood cells.

Although various electron and scanning probe microscopes have long surpassed the compound optical microscope in resolving power, optical microscopy remains invaluable in many fields of science. The practical limit to the resolution of a conventional optical microscope is determined by diffraction: a wave cannot be localized to a region much smaller than half of its vacuum wavelength $\lambda_0/2$. Immersion microscopes introduced by Abbe in the 19th century have slightly improved resolution, on the order of $\lambda_0/2n$ because of the shorter wavelength of light λ_0/n in a medium with refractive index n. However, immersion microscopes are limited by the small range of refractive indices n of available transparent materials. For a while, it was believed that the only way to achieve nanometer-scale spatial resolution in an optical microscope was to detect evanescent optical waves in very close proximity to a studied sample using a near-field scanning optical microscope (NSOM).[26] Although many fascinating results are being obtained with NSOM, such microscopes are not as versatile and convenient to use as regular far-field optical microscopes. For example, an image from a near-field optical microscope is obtained by point-by-point scanning, which is an indirect and a rather slow process, and can be affected by artifacts of the sample.

An important early step to overcome this limitation was made in SP-assisted microscopy experiments,[27] in which 2D image magnification was achieved. In this microscope design, the dispersion behavior of SPPs propagating in the boundary between a thin metal film and a dielectric (3.7) was exploited to use the 2D optics of SPPs with a short wavelength to produce a magnified local image of an object on the surface. If the dispersion curve of SPPs on gold is examined, an excitation wavelength that provides a small group velocity gives rise to a 2D SPP diffraction limit that is on the order of $\lambda_{\mathrm{SPP}} = \lambda/n_g$.

If a 2D "mirror" structure is fabricated in the surface, in the case of Ref. 27 using parabolic droplets on the surface (see Fig. 3.4(b)), then SPPs propagating in the surface reflect or scatter at the boundaries of an object placed on the surface. These reflected or scattered SPPs are then imaged by the surface structure to produce a magnified 2D image. This magnified image can be examined by a far-field microscope by using light scattered from surface roughness or from lithographically generated surface structures that scatter propagating SPPs in to far-field radiation. The increased spatial resolution of microscopy experiments performed with SPPs[27] is based on the "hyperbolic" dispersion law of such waves, which may be written in the form

$$k_{xy}^2 - |k_z|^2 = \frac{\varepsilon_d \omega^2}{c^2}, \tag{3.9}$$

where ε_d is the dielectric constant of the medium bounding the metal surface, which for air is =1, $k_{xy} = k_p$ is the wave vector component in the plane of propagation and k_z is the wave vector component perpendicular to the plane. This form of the dispersion relation originates from the exponential decay of the surface wave field away from the propagation plane.

The "optical hyperlens" design described by Jacob *et al.*[2] extends this idea by using a hyperbolic metamaterial made of a concentric arrangement of metal and dielectric cylinders, which may be characterized by a strongly anisotropic dielectric permittivity tensor in which the tangential ε_θ and the radial ε_r components have opposite signs. The resulting hyperbolic dispersion relation

$$\frac{k_r^2}{\varepsilon_\theta} - \frac{k_\theta^2}{|\varepsilon_r|} = \frac{\omega^2}{c^2}, \tag{3.11}$$

does not exhibit any lower limit on the wavelength of propagating light at a given frequency. Therefore, in a manner similar to the 2D optics of SPPs, there is no usual diffraction limit in this metamaterial medium. Abbe's resolution limit simply does not exist. Optical energy propagates through such a metamaterial in the form of radial

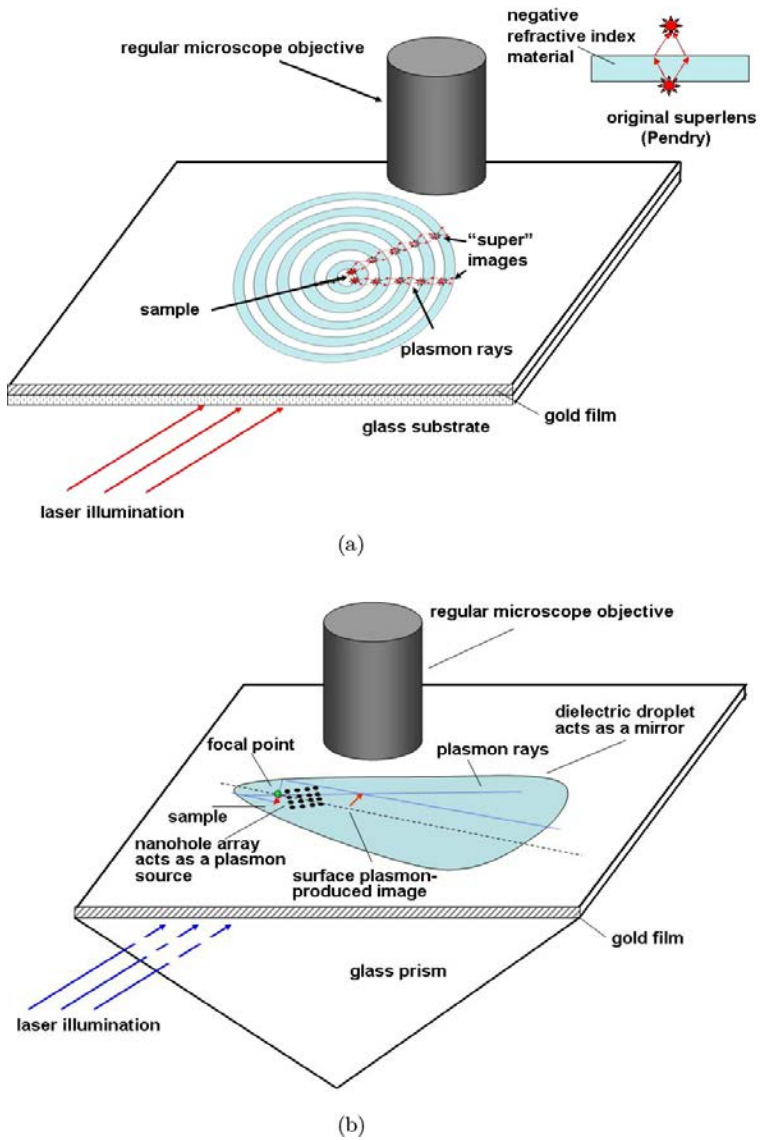
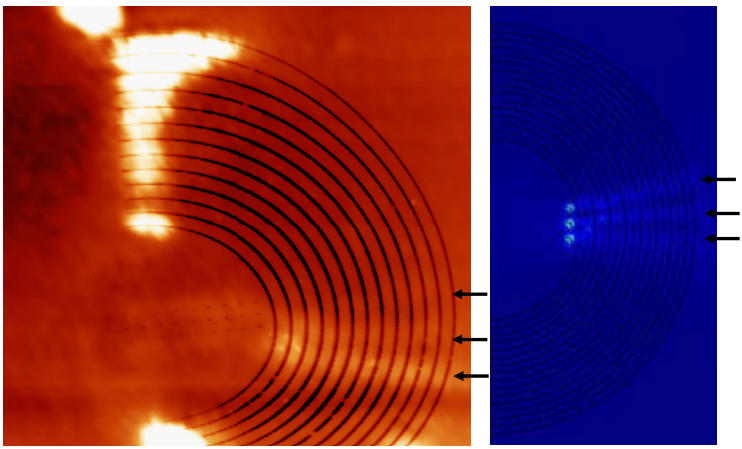


Fig. 3.4. Two modes of operation of a 2D plasmonic microscope. (a) Plasmon microscope operating in the "hyperlens mode": the plasmons generated by the sample located in the center of the hyperlens propagate in the radial direction. The lateral distance between plasmonic rays grows with distance along the radius.

Fig. 3.4. (*Continued*) The images are viewed by a regular microscope. (b) Plasmon microscope operating in the "geometrical optics" mode: nanohole array illuminated by external laser light acts as a source of surface plasmons, which are emitted in all directions. Upon interaction with the sample positioned near the focal point of the parabolically shaped dielectric droplet, and reflection off the droplet edge, the plasmons form a magnified planar image of the sample. The image is viewed by a regular microscope. The droplet edge acts as an efficient plasmon mirror because of total internal reflection.

rays. Moreover, as demonstrated in Section 1, a pattern of polymethyl methacrylate (PMMA) stripes formed on a metal surface (as shown in Fig. 3.4(a)) behaves as a 2D plasmonic equivalent of the 3D hyperbolic metamaterial. Thus, two modes of operation of a 2D plasmonic microscope may be implemented, as shown in Figs. 3.4(a) and 3.4(b). A plasmon microscope may be operated in the "hyperlens mode" (Fig. 3.4(a)) in which the plasmons generated by the sample located in the center of the plasmonic hyperlens propagate in the radial direction. The lateral distance between plasmonic rays grows with distance along the radius. The images are viewed by a regular microscope. Alternatively, a 2D plasmon microscope may be operated in the "geometrical optics" mode as shown in Fig. 3.4(b). A nanohole array illuminated by external laser light may act as a source of SPs, which are emitted in all directions. Upon interaction with the sample positioned near the focal point of the parabolically shaped dielectric droplet, and reflection off the droplet edge, the plasmons form a magnified planar image of the sample. The image is viewed by a regular microscope. The droplet edge acts as an efficient plasmon mirror because of total internal reflection. It appears that both modes of operation exhibit strong evidence of optical super-resolution.

The internal structure of the magnifying hyperlens (Fig. 3.5(a)) consists of concentric rings of PMMA deposited on a gold film surface. The required concentric structures were defined using a Raith E-line electron beam lithography (EBL) system with ∼70 nm spatial resolution. The written structures were subsequently developed using a 3:1 IPA/MIBK solution (Microchem) as developer and imaged using AFM (see Fig. 3.4(a)). According to theoretical proposals in

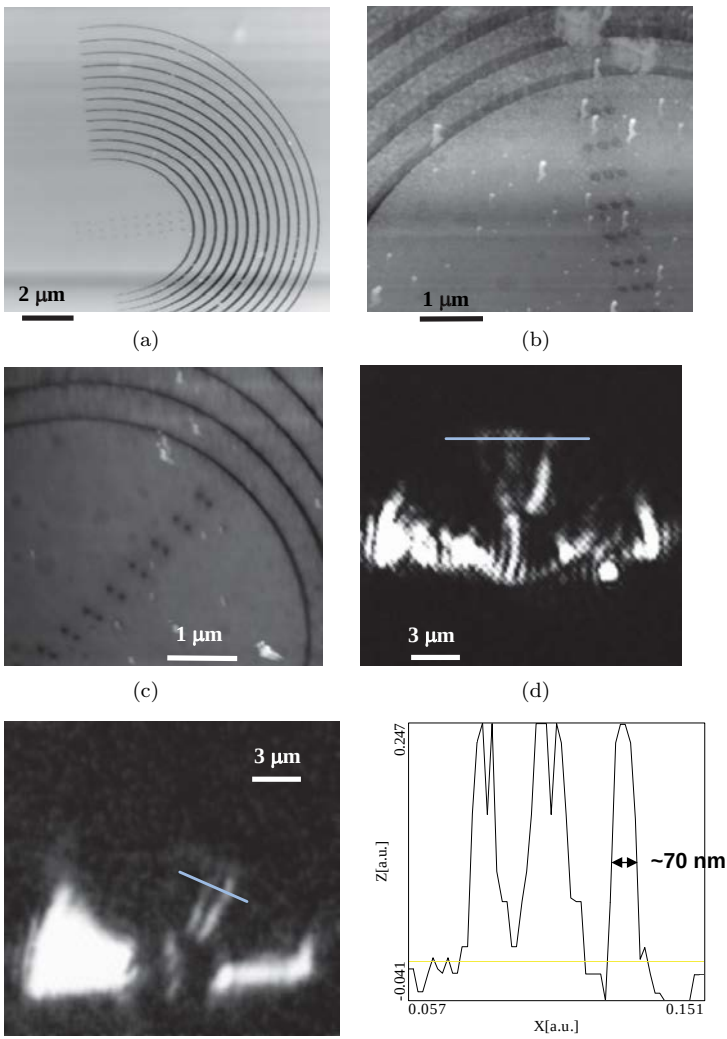


(a) (b)

Fig. 3.5. (a) Superposition image composed of an AFM image of the PMMA on gold plasmonic metamaterial structure superimposed onto the corresponding optical image obtained using a conventional optical microscope illustrating the imaging mechanism of the magnifying hyperlens. Near the edge of the hyperlens, the separation of three rays (marked by arrows) is large enough to be resolved using a conventional optical microscope. (b) Theoretical simulation of ray propagation in the magnifying hyperlens microscope.

Refs. 1 and 2, optical energy propagates through a hyperbolic metamaterial in the form of radial rays. This behavior is clearly demonstrated in Fig. 3.5(b). If point sources are located near the inner rim of the concentric metamaterial structure, the lateral separation of the rays radiated from these sources increases upon propagation towards the outer rim. Therefore, resolution of an "immersion" microscope (a hyperlens) based on such a metamaterial structure is defined by the ratio of inner to outer radii. Resolution appears limited only by losses, which can be compensated by optical gain.

The magnifying superlenses (or hyperlenses) have been independently realized for the first time in two experiments.[4,5] In particular, experimental data obtained using a 2D plasmonic hyperlens (shown in Fig. 3.5(a)) do indeed demonstrate ray-like propagation of subwavelength plasmonic beams emanated by test samples. Far-field optical resolution of at least 70 nm (see Fig. 3.6(f)) has been



(a) (b) (c) (d) (e) (f)

Fig. 3.6. AFM (a–c) and conventional optical microscope (d,e) images of the resolution test samples composed of three (a,b) and two (c) rows of PMMA dots positioned near the center of the magnifying hyperlens. The conventional microscope images presented in (d) and (e) correspond to the samples shown in (b) and (c), respectively. The rows of PMMA dots give rise to either three or two divergent plasmon "rays", which are visible in the conventional optical microscope images. (f) The cross-section of the optical image along the line shown in (d) indicates resolution of at least 70 nm or $\sim\lambda/7$.

demonstrated using such a magnifying hyperlens based on a 2D plasmonic metamaterial design. Rows of either two or three PMMA dots have been produced near the inner ring of the hyperlens (Fig. 3.6(b) and 3.6(c)). These rows of PMMA dots had 0.5 μm periodicity in the radial direction so that phase matching between the incident laser light and SPs can be achieved. Upon illumination with an external laser, the three rows of PMMA dots in Fig. 3.6(b) gave rise to three divergent plasmon "rays", which are clearly visible in the plasmon image in Fig. 3.6(d) obtained using a conventional optical microscope. The cross-section analysis of this image across the plasmon "rays" (Fig. 3.6(f)) indicates resolution of at least 70 nm or $\sim\lambda/7$. The lateral separation between these rays increased by a factor of 10 as the rays reached the outer rim of the hyperlens. This increase allowed visualization of the triplet using a conventional microscope. In a similar fashion, the two rows of PMMA dots shown in Fig. 3.6(c) gave rise to two plasmon rays, which are visualized in Fig. 3.6(e).

The magnifying action and the imaging mechanism of the hyperlens have been further verified by control experiments presented in Fig. 3.7. The image shown in Fig. 3.7(a) presents results of two actual imaging experiments (top portion of Fig. 3.7(a)) performed simultaneously with four control experiments seen at the bottom of the same image. In these experiments, two rows of PMMA dots have been produced near the inner ring of the hyperlens structures seen at the top and at the bottom of Fig. 3.7(a) (the AFM image of the dots is seen in the inset). These rows of PMMA dots had 0.5 μm periodicity in the radial direction so that phase matching between the incident 515-nm laser light and SPs can be achieved. On the other hand, no such PMMA dot structure was fabricated near the control hyperlenses seen in the center of Fig. 3.7(a). Upon illumination with an external laser, the two rows of PMMA dots gave rise to the two divergent "plasmon rays," which are clearly visible in the top portion of the image in Fig. 3.7(a) obtained using a conventional optical microscope. No such rays were observed in the four "control" hyperlenses visible in the bottom portion of the same image. There was no sample to image for the two hyperlenses located in the

center of Fig. 3.7(a). On the other hand, the PMMA dot structure was designed for phase-matched plasmon generation in the "upward" direction as seen in the image. That is why no plasmon rays are visible when the hyperlens structures are inverted, as seen in the bottom of Fig. 3.7(a). When the gold film was replaced with an ITO film in another control experiment performed using the same experimental geometry, no hyperlens imaging occurred since no surface plasmons are generated on ITO surface (see Fig. 3.7(b)). These experiments clearly verify the imaging mechanism and increased spatial resolution of the plasmonic hyperlens.

3.3. Consequences of Singular Photonic Density of States: Radiative Decay Engineering, Thermal Hyperconductivity, and New Stealth Technologies

As we discussed in Section 1, the broadband divergence of photonic density of states in a hyperbolic metamaterial is unique since it leads to the infinite value of the density of states for every frequency where different components of the dielectric permittivity have opposite signs. This very large number of electromagnetic states can couple to quantum emitters leading to such unusual phenomena as the broadband Purcell effect[6] and thermal hyperconductivity.[10] On the other hand, free space photons illuminating a roughened surface of a hyperbolic metamaterial preferentially scatter inside the metamaterial, leading to the surface being "darker than black" at the hyperbolic frequencies.[9] The latter property may find natural applications in stealth technologies.

As a first example of these unusual quantum behaviors, let us consider the broadband Purcell effect, which may become extremely useful for such applications as single-photon sources, fluorescence imaging, biosensing and single molecule detection. In the spirit of Fermi's golden rule, an increased number of radiative decay channels due to the high-k states in hyperbolic media (available for an excited atom) must ensure enhanced spontaneous emission. This enhancement can increase the quantum yield by overcoming

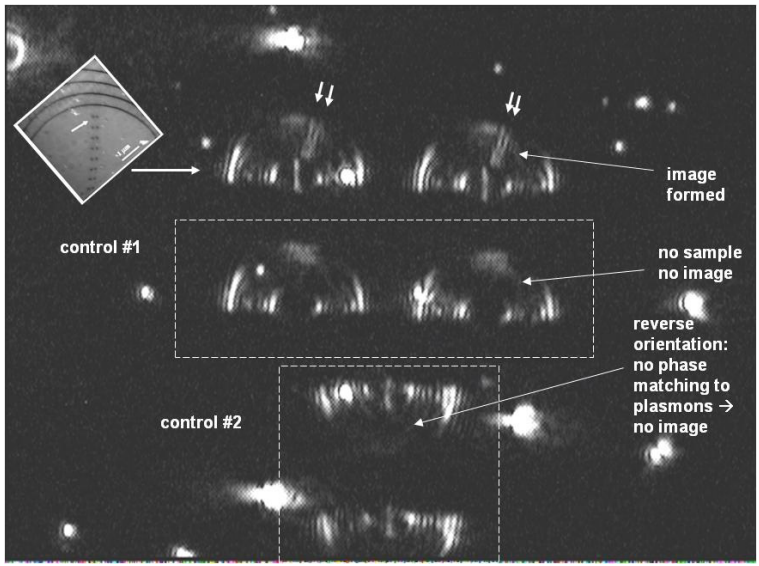
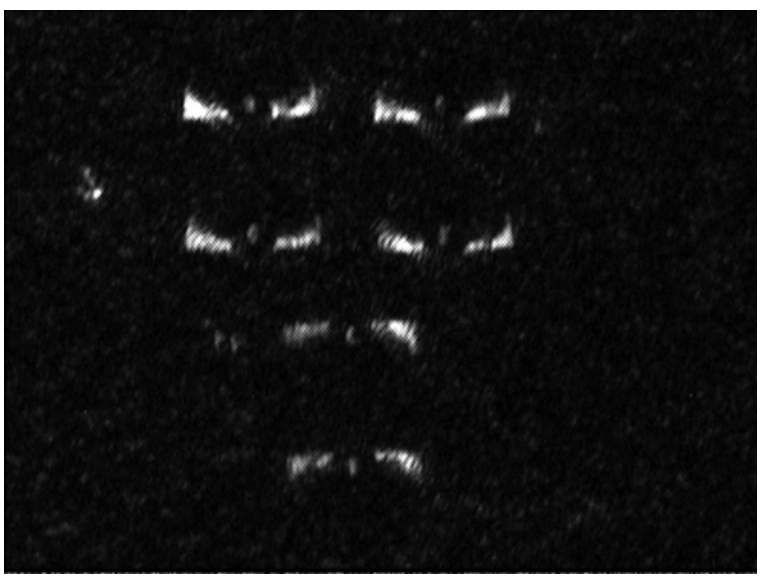


image
formed
control #1
no sample
no image
reverse
orientation:
no phase
matching to
plasmons →
no image
control #2

(a)



(b)

Fig. 3.7.

←

Fig. 3.7. (*figure on facing page*) (a) This image obtained using a conventional optical microscope presents the results of two imaging experiments (top portion of the image) performed simultaneously with four control experiments seen at the bottom of the same image. The rows of PMMA dots shown in the inset AFM image were fabricated near the two top and two bottom hyperlenses. No such pattern was made near the two hyperlenses visible in the center of the image. Upon illumination with an external laser, the two rows of PMMA dots separated by 130 nm gap gave rise to two divergent plasmon rays shown by the arrows, which are clearly visible in the top portion of the image. The four control hyperlenses visible at the bottom do not produce such rays because there is no sample to image for the two hyperlenses in the center, and the two bottom hyperlenses are inverted. (b) Same pattern produced on ITO instead of gold film demonstrates a pattern of ordinary light scattering by the structure without any hyperlens imaging effects.

emission into competing non-radiative decay routes such as phonons. A decrease in lifetime, high quantum yield and good collection efficiency can lead to extraction of single photons reliably at a high repetition rate from isolated emitters.[28] The available radiative channels for the spontaneous photon emission consist of the propagating waves in vacuum, the plasmon on the metamaterial substrate and the continuum of high wave vector waves which are evanescent in vacuum but propagating within the metamaterial. The corresponding decay rate into the metamaterial modes when the emitter is located at a distance $a < d \ll \lambda$ (where a is the metamaterial patterning scale) is[6]

$$\Gamma^{\text{meta}} \approx \frac{\mu^2}{2\hbar d^3} \frac{2\sqrt{\varepsilon_x|\varepsilon_z|}}{(1+\varepsilon_x|\varepsilon_z|)}. \tag{3.12}$$

In the close vicinity of the hyperbolic metamaterial, the power from the dipole is completely concentrated in the large spatial wave vector channels (Fig. 3.8(a) inset). The same evanescent wave spectrum when incident on a lossy metal or dielectric would be completely absorbed, causing a non-radiative decrease in the lifetime of an emitter (quenching). On the contrary, the metamaterial converts the evanescent waves to propagating, and the absorption thus affects

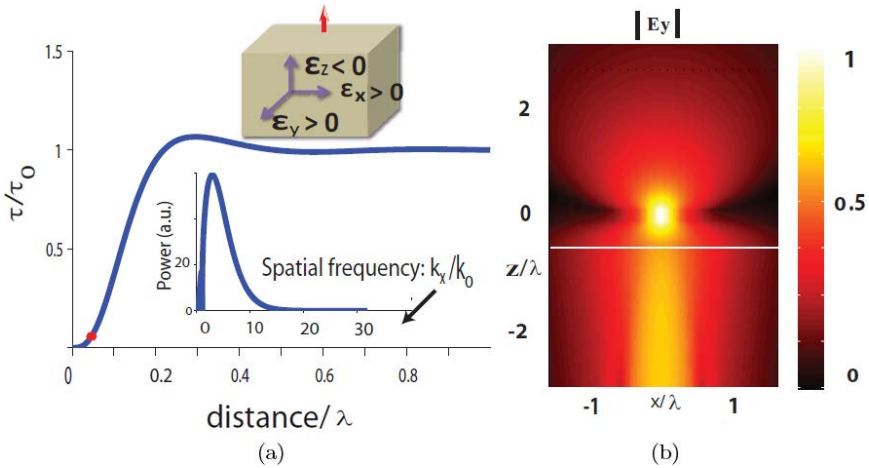


Fig. 3.8. (a) Spontaneous emission lifetime of a perpendicular dipole above a hyperbolic metamaterial substrate (see inset). Note the lifetime goes to zero in the close vicinity of the metamaterial as the photons are emitted nearly instantly. Most of the power emitted by the dipole is concentrated in the large spatial modes (evanescent in vacuum) which are converted to propagating waves within the metamaterial. (inset) (b) False color plot of the field of the point dipole in a plane perpendicular to the metamaterial-vacuum interface (see inset of (a)) depicting the highly directional nature of the spontaneous emission (resonance cone).

the outcoupling efficiency of the emitted photons due to a finite propagation length in the metamaterial.

Along with the reduction in lifetime and high efficiency of emission into the metamaterial, another key feature of the hyperbolic media is the directional nature of light propagation.[6] Figure 3.8(b) shows the field along a plane perpendicular to the metamaterial–vacuum interface exhibiting the beamlike radiation from a point dipole. This is advantageous from the point of view of collection efficiency of light since the spontaneous emitted photons lie within a cone. The group velocity vectors in the medium which point in the direction of the Poynting vector are simply normals to the dispersion curve. For vacuum, these normals point in all directions, and hence the spontaneous emission is isotropic in nature. In contrast to this behavior, the hyperbolic dispersion medium allows wave vectors only

within a narrow region defined by the asymptotes of the hyperbola. Hence, the group velocity vectors lie within the resonance cone, giving rise to a directional spontaneously emitted photon propagating within the metamaterial.

Since its theoretical prediction in Ref. 6, the broadband Purcell effect has indeed been observed in multiple experiments, such as Refs. 7 and 8. Virtually the same physics as in Fig. 3.8(b) is also responsible for the "darker than black" behavior of roughened hyperbolic metamaterials.[9] Free space photons illuminating a rough surface of the hyperbolic metamaterial preferentially scatter inside the metamaterial into the bulk high k-vector modes. As a result, the photon probability to scatter back into free space is almost zero and the roughened surface looks black in the hyperbolic frequency bands.

Let us now consider radiative heat transfer inside hyperbolic metamaterials. It appears that the broadband divergence of the photonic density of states described above also leads to giant increase in radiative heat transfer compared to the Stefan–Boltzmann law in vacuum and in usual anisotropic dielectric materials. According to numerical calculations,[10] this radiative thermal "hyperconductivity" may approach or even exceed heat conductivity via electrons and phonons in regular solids, with the additional advantage of radiative heat transfer being much faster. Therefore, this radiative thermal hyperconductivity may potentially be very useful in fast microelectronics heat management applications.[29] In such applications, heat generated by micro- and nanoelectronic circuit components needs to be quickly dissipated at a heat sink, which cannot be located in the immediate vicinity of the electronic component. A hyperbolic metamaterial heat management layer may solve this important technological problem.

Let us start by tracing how the photonic density of states enters the usual Stefan–Boltzmann law. For the sake of simplicity, we will consider vacuum as a typical example of "normal" or "elliptical" material. As usual, we can start by calculating energy density of the blackbody radiation. A well-known textbook derivation can be

summarized as follows:

$$u_{ell} = \frac{U}{V} = \int_0^\infty \frac{\varepsilon}{\exp\left(\frac{\varepsilon}{kT}\right) - 1} g(\varepsilon) d\varepsilon = \frac{4\sigma T^4}{c}, \tag{3.13}$$

where $g(\varepsilon)$ is the photonic density of states. Eq. (3.13) clearly demonstrates that the drastic change in the density of states schematically shown in Fig. 3.9 must lead to the drastic change in the final result. The singular behavior of the photonic density of states in hyperbolic metamaterial takes these media beyond the realm of the Stefan–Boltzmann law, with no ultimate limit on the radiative heat transfer. For the energy flux along the symmetry axis of a uniaxial hyperbolic metamaterial, it was found[10] that

$$S_T = \frac{\hbar c^2 k_{\max}^4}{32\pi^2} \int d\omega \frac{1}{\exp\left(\frac{\hbar\omega}{kT}\right) - 1} \left| \frac{\varepsilon_1 \frac{d\varepsilon_2}{d\omega} - \varepsilon_2 \frac{d\varepsilon_1}{d\omega}}{\det \|\varepsilon\|} \right| \tag{3.14}$$

where the frequency integration is taken over the frequency bandwidth corresponding to the hyperbolic dispersion. Note that the heat flux in Eqn. (3.14) is very sensitive to the dispersion in the hyperbolic metamaterial, $d\varepsilon/d\omega$. Indeed, the derivative of the dielectric permittivity determines the difference in the asymptotic behavior of the k-vector between the two hyperbolic surfaces that determine the phase space volume between the frequencies ω and $\omega + d\omega$ (see Fig. 3.9), and thus defines the actual value of the density of states. While there are many metamaterial designs leading to the hyperbolic dispersion, the most practical and widely used systems rely on either the metal–dielectric layer approach, or incorporate aligned metal nanowire composites (as shown in Fig. 3.1). For the planar layers design, the hyperbolic behavior is observed for the wavelengths above $\sim$10 μm if the system is fabricated using semiconductors,[20] or for the wavelengths above $\sim$1 μm if the metamaterial is composed of metal–dielectric layers.[7] For the nanowire-based approach, the hyperbolic dispersion is present at $\lambda > 1\mu$m.[21] As a result, with either of these conventional metamaterial designs, the desired hyperbolic behavior

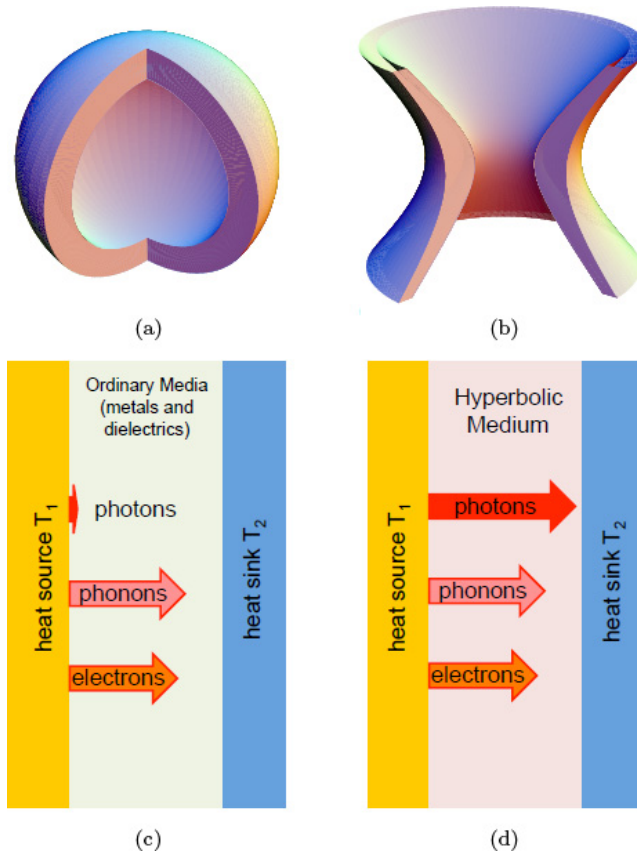


Fig. 3.9. The phase space volume between two constant frequency surfaces for (a) dielectric (elliptical) and (b) hyperbolic material with $\varepsilon_2 > 0$ and $\varepsilon_1 < 0$ (cut-out view). Panels (c) and (d) schematically illustrate different thermal conductivity mechanisms in (c) regular media (metals and dielectric) and (d) hyperbolic media. Giant radiative contribution to thermal conductivity in hyperbolic media can dominate the thermal transport.

covers the full range of wavelength relevant for the radiative heat transfer. As a result, the following estimates on the thermal energy flux in hyperbolic metamaterials have been obtained[10]:

$$S_T \approx \frac{\varepsilon_d}{4(1-n)} S_T^{(0)} \left(\frac{k_{\max}}{k_p} \right)^4 \tag{3.15}$$

for the layered material design, and

$$S_T \approx \frac{5}{16\pi^2} S_T^{(0)} \left(\frac{k_{\max}^2}{k_T k_p} \right)^2 \tag{3.16}$$

for the wire array design, where $S_T^{(0)}$ is the blackbody thermal energy flux for emission into the free space, k_p is the plasma momentum and k_T is the thermal momentum. In both cases, the numerical values of S_T exceed $S_T^{(0)}$ by 4–5 orders of magnitude, thus firmly placing hyperbolic metamaterials in the realm of practical applications for radiative heat transfer and thermal management. A similar enhancement may be also expected in thermal conductivity.

3.4. Photonic Hypercrystals

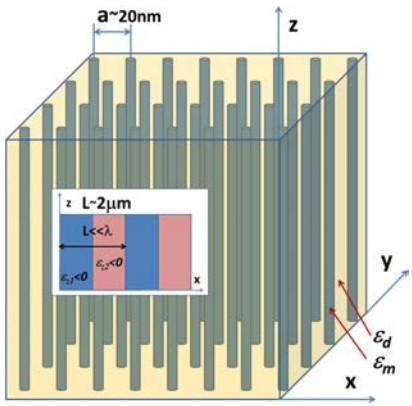
Explosive development of research on hyperbolic metamaterials also resulted in the recent demonstration of a novel artificial optical material, the "photonic hypercrystal",[30] which combines the most interesting features of hyperbolic metamaterials and photonic crystals. Similar to hyperbolic metamaterials, photonic hypercrystals exhibit broadband divergence in their photonic density of states due to the lack of usual diffraction limit on the photon wave vector. On the other hand, similar to photonic crystals, hyperbolic dispersion law of extraordinary photons is modulated by forbidden gaps near the boundaries of photonic Brillouin zones. 3D self-assembly of photonic hypercrystals has been achieved by application of external magnetic field to a cobalt nanoparticle-based ferrofluid. Unique spectral properties of photonic hypercrystals lead to extreme sensitivity of the material to monolayer coatings of cobalt nanoparticles, which should find numerous applications in biological and chemical sensing.

Over the last few decades, considerable progress has been made in developing artificial optical materials with novel and often counterintuitive properties. Revolutionary research by Yablonovitch and John on photonic crystals[31,32] was followed by the development of the electromagnetic metamaterial paradigm by Pendry.[33] Even though

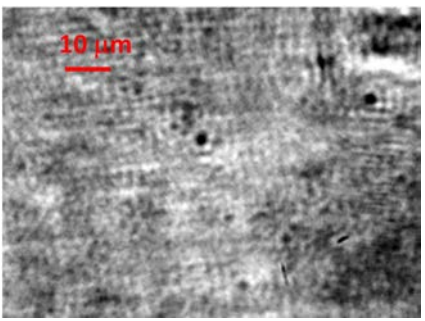
considerable difficulties still exist in fabrication of 3D photonic crystals and metamaterials, both fields exhibit considerable experimental progress.[34,35] On the other hand, on the theoretical side, these fields are believed to be complementary but mutually exclusive. Photonic crystal effects typically occur in artificial optical media which are periodically structured on the scale of free space light wavelength λ, while electromagnetic metamaterials are required to be structured (not necessarily in a periodic fashion) on the scale, which is much smaller than the free space wavelength of light. For example, in metal nanowire-based hyperbolic metamaterials schematically shown in Fig. 3.1(b) the interwire distance must be much smaller than λ. Experimental realization of 3D "photonic hypercrystals"[30] bridges this divide by combining the most interesting properties of hyperbolic metamaterials and photonic crystals.

The concept of the photonic hypercrystal is based on the fact that dispersion law of extraordinary photons in hyperbolic metamaterials (3.2) does not exhibit the usual diffraction limit. Existence of large k-vector modes in a broad range of frequencies means that periodic modulation of hyperbolic metamaterial properties on a scale $L \ll \lambda$ (see inset in Fig. 3.10(a)) would lead to Bragg scattering of extraordinary photons and formation of photonic band structure no matter how small L is.[30] Thus, so-formed "photonic hypercrystals" would combine the most interesting properties of hyperbolic metamaterials and photonic crystals. For example, similar to classic photonic crystal effect predicted by John,[32] strong localization of photons may occur in photonic hypercrystals. However, unlike usual photonic crystals where light localization occurs on a scale $\sim \lambda$, photonic hypercrystals may exhibit light localization on deep subwavelength scale. Similar to surface plasmon resonance (SPR)[25] and (surface-enhanced Raman (SERS))[36]-based sensing, engineered localization of light on deep subwavelength scale in photonic hypercrystals should find numerous applications in biological and chemical sensing.

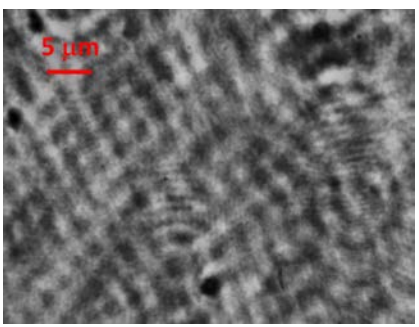
Band structure and field distribution inside a photonic hypercrystal may be obtained in a straightforward manner. While both



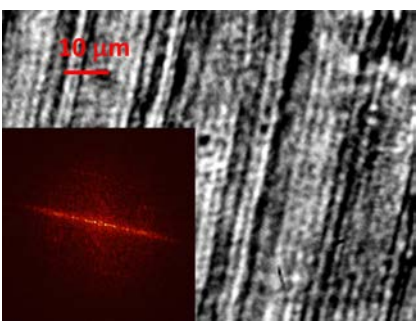
(a)



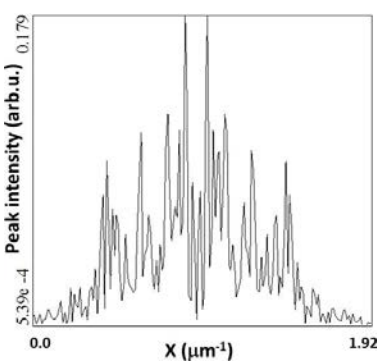
(b)



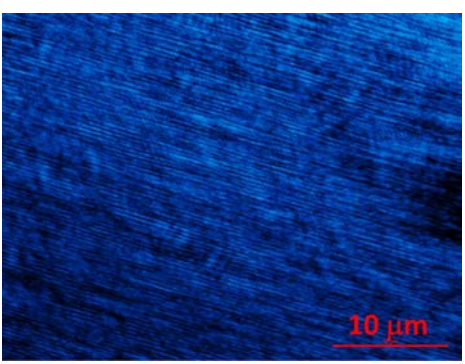
(c)



(d)



(e)



(f)

Fig. 3.10.

←

Fig. 3.10. (*figure on facing page*) (a) Experimental geometry of the ferrofluid-based hyperbolic metamaterial. The array of self-assembled cobalt nanocolumns has typical separation $a \sim 20\,\text{nm}$ between the nanocolumns. The inset shows a photonic hypercrystal structure formed by periodic arrangement of cobalt-rich and cobalt-sparse regions with typical periodicity $L \sim 2\mu\text{m}$, so that periodic modulation of hyperbolic metamaterial properties on a scale $L \ll \lambda$ is achieved in the LWIR spectral range where $\lambda \sim 10\mu\text{m}$. Since photon wave vector in hyperbolic metamaterials is not diffraction-limited, periodic modulation of hyperbolic metamaterial properties on a scale $L \ll \lambda$ would lead to Bragg scattering and formation of band structure. (b–f) Microscopic images of cobalt nanoparticle-based ferrofluid reveal subwavelength modulation of its spatial properties: frames (b) and (d) show microscopic images of the diluted cobalt nanoparticle-based ferrofluid before and after application of external magnetic field. The pattern of self-assembled stripes visible in image (d) is due to phase separation of the ferrofluid into cobalt-rich and cobalt-poor phases. The stripes are oriented along the direction of magnetic field. The inset shows Fourier transform image of frame (d). Its cross-section presented in panel (e) shows a histogram of different periods present in the image. A microscopic image of the sample taken along the axes of the nanowires is shown in frame (c). Panel (f) demonstrates that the original undiluted ferrofluid exhibits similar phase separation in external magnetic field, although on amuch smaller scale.

ε_{xy} and ε_z may exhibit periodic spatial dependencies, let us consider the relatively simple case of coordinate-independent $\varepsilon_{xy} > 0$ and periodic $\varepsilon_z(z) < 0$ with a period $L \ll \lambda$. Aside from the relative mathematical simplicity of this model, it also corresponds to the most readily available low-loss realizations of hyperbolic metamaterials such as the composites formed by metallic nanowires in a dielectric membrane[21] (where $\varepsilon_{xy} > 0$ and $\varepsilon_z(z) < 0$), and planar layered metal–dielectric and semiconductor metamaterials.[7,20] Taking into account the translational symmetry of the system in x and y directions, we can introduce the in-plane wave vector (k_x, k_y) so that the propagating waves can be expressed as

$$\begin{aligned} E_\omega(\vec{r}) &= E(z)\exp(ik_x x + ik_y y), \\ D_\omega(\vec{r}) &= D(z)\exp(ik_x x + ik_y y), \\ B_\omega(\vec{r}) &= B(z)\exp(ik_x x + ik_y y). \end{aligned} \tag{3.17}$$

The uniaxial symmetry of this medium reduces the ordinary and extraordinary waves to, respectively, the TE ($\vec{E}\perp\hat{z}$) and TM ($\vec{B}\perp\hat{z}$)-polarized modes. Introducing the wavefunction $\psi(\vec{r})$ as the z-component of the electric displacement field of the TM wave

$$\psi(\vec{r}) = D_z(\vec{r}) = \varepsilon_z(z)E_z(\vec{r}) = -\frac{c}{\omega}k_x B \tag{3.18}$$

for the wave equation we obtain

$$-\frac{\partial^2\psi}{\partial z^2} + \frac{\varepsilon_{xy}}{\varepsilon_z(z)}\psi = \varepsilon_{xy}\frac{\omega^2}{c^2}\psi. \tag{3.19}$$

In this wave equation, the periodic $\varepsilon_{xy}/\varepsilon_z$ ratio acts as a periodic effective potential. As usual, solutions of Eq. (3.4) may be found as Bloch waves

$$\psi(z) = \sum_{m=0}^{\infty} \psi_m \exp\left(i\left(k_z + \frac{2\pi}{L}m\right)z\right), \tag{3.20}$$

where k_z is defined within the first Brillouin zone $-\pi/L < k_z < \pi/L$. Strong Bragg scattering is observed near the Brillouin zone boundaries at $k_z \sim \pi/L \gg \pi/\lambda$, leading to the formation of photonic band gaps in both the wavenumber and the frequency domains. This behavior is illustrated in Fig. 3.11, where we compare the dispersion diagram for an example of such a nanowire-based photonic hypercrystal to its effective medium counterpart. The material parameters of photonic hypercrystals are based on the parameters of stratified ferrofluid described below. Similar to the usual photonic crystals,[32] adiabatic chirping of L leads to strong field enhancement which, unlike that in the conventional photonic crystals, occurs on a deep subwavelength scale. We should also mention that the most interesting case appears to be the epsilon-near-zero (ENZ) situation where ε_z approaches zero near a periodic set of planes. As has been demonstrated in Ref. 16, electric field of the extraordinary wave diverges in these regions. These periodic field divergences appear to be most beneficial for sensing applications.

Validation of the photonic hypercrystal concept has been achieved using an experimental technique based on 3D self-assembly of cobalt nanoparticles in the presence of external magnetic field.[30] Magnetic nanoparticles in a ferrofluid are known to form nanocolumns aligned along the magnetic field.[37] Moreover, depending on the magnitude of magnetic field, nanoparticle concentration and solvent used, phase separation into nanoparticle-rich and nanoparticle-poor phases may occur in many ferrofluids.[38] This phase separation occurs on a 0.1–1 μm scale. Therefore, it can be used to fabricate a self-assembled photonic hypercrystal.

These experiments used cobalt magnetic fluid 27-0001 from Strem Chemicals composed of 10-nm cobalt nanoparticles in kerosene coated with sodium dioctylsulfosuccinate and a monolayer of LP4 fatty acid condensation polymer. The average volume fraction of cobalt nanoparticles in this ferrofluid is $p = 8.2\%$. Cobalt behaves as an excellent metal in the long-wavelength infrared (LWIR) range, as evident by Fig. 3.12(a): the real part of its refractive index, n, is much smaller than its imaginary part, k.[39] Thus, real part of ε, $\mathrm{Re}\varepsilon = n^2 - k^2$, is negative, and its absolute value is much larger than its imaginary part, $\mathrm{Im}\varepsilon = 2nk$. Therefore, it is highly suitable for fabrication of hyperbolic metamaterials. The structural parameter of such a metamaterial falls in the range of a few nanometers: the cobalt nanoparticle size is 10 nm, while average interparticle distance at 8.2% volume fraction is about 19 nm. Therefore, the metamaterial properties may be described by effective medium parameters on spatial scales $\sim$100 nm. On the other hand, as demonstrated below, ferrofluid begins to exhibit hyperbolic behavior in the range of free space wavelengths $\sim$10,000 nm and above — in the so-called LWIR frequency range. Thus, in between 100 nm and 10,000 nm there exists an ample range of spatial scales which enable photonic hypercrystal behavior described above. For example, if the effective medium parameters of ferrofluid are modulated on the scale of $\sim$2000 nm (as evident from Fig. 3.10(d)), the large k-vector modes which exist in the metamaterial in the LWIR range will experience Bragg scattering due to this modulation.

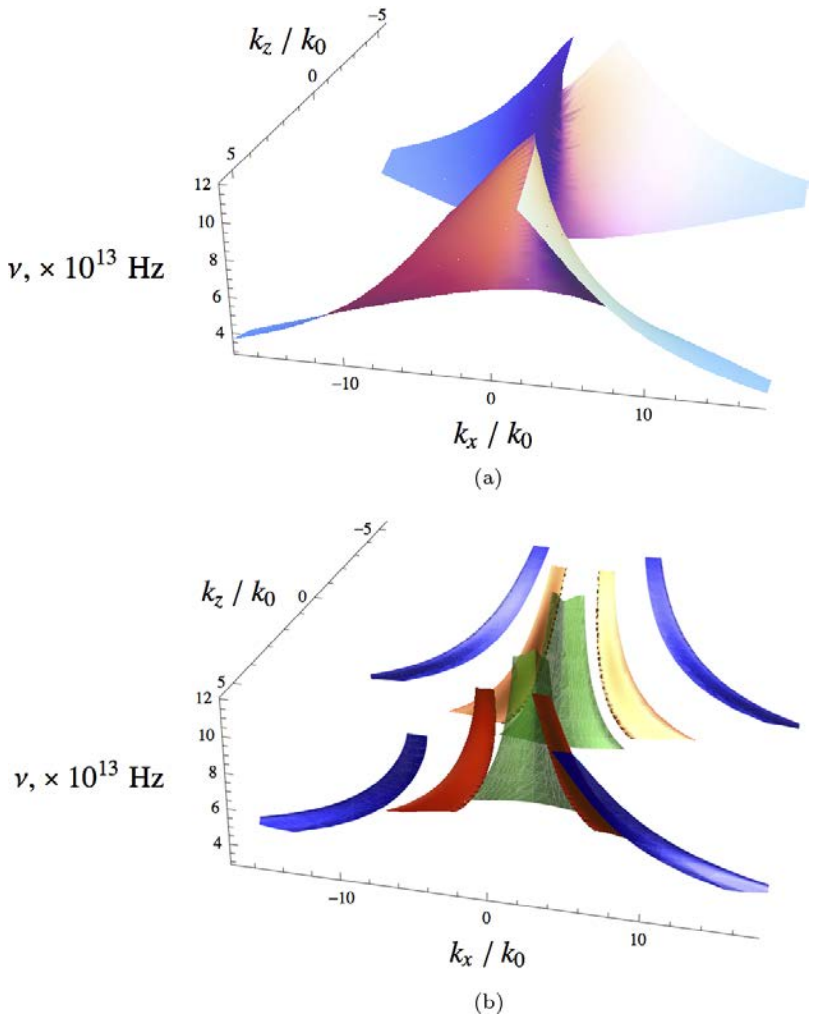
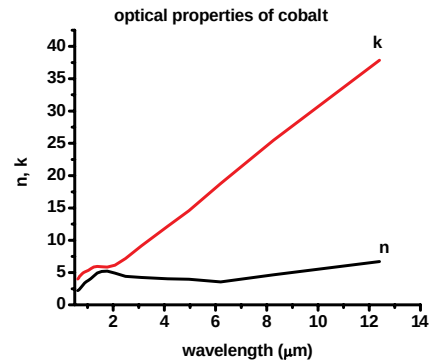


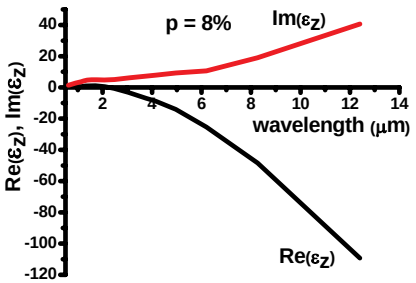
Fig. 3.11. Comparison of the effective medium dispersion of a nanowire-based hyperbolic metamaterial (a) with the exact solution for photonic crystal (b). The hyper-crystal unit cell in (b) is assumed to be 1000 nm, with one half of the cell filled with the same ferrofluid while another half is pure kerosene. The ferrofluid aligned by external magnetic field applied in the z-direction, is characterized by the material parameters described in Fig. 3.12. The stratified ferrofluid is assumed to form layers with the normal along the x-direction. The effective medium dispersion in panel (a) uses the dielectric permittivity tensor obtained

Fig. 3.11. (*Continued*) by the homogenization of the electromagnetic response of the hypercrystal unit cell. This is an artificial result that would be expected if we could experimentally isolate magnetic field-induced hyperbolic behavior of the ferrofluid from its photonic hypercrystal behavior. In both panels, the wave vector components are given in units of the free-space wavenumber k_0.

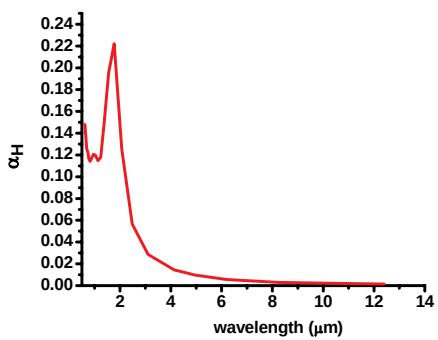
Electromagnetic properties of these metamaterials may be understood based on the Maxwell–Garnett approximation via the dielectric permittivities ε_m and ε_d of cobalt and kerosene, respectively, as illustrated in Fig. 3.12. Volume fraction of cobalt nanoparticles aligned into nanocolumns by external magnetic field, $\alpha(B)$, depends on the field magnitude. At very large magnetic fields, all nanoparticles are aligned into nanocolumns, so that $\alpha(\infty) = p = 8.2\%$. At smaller fields, the difference $\alpha(\infty) - \alpha(B)$ describes cobalt nanoparticles, which are not aligned and distributed homogeneously inside the ferrofluid. Using this model, the diagonal components of the ferrofluid permittivity may be calculated and measured as a function of magnetic field. The value of α_H as a function of wavelength is plotted in Fig. 3.12(b). This plot indicates that the original ferrofluid diluted with kerosene at a 1:10 ratio remains a hyperbolic medium above $\lambda = 5\mu$m. More interestingly, such a diluted ferrofluid develops very pronounced phase separation into periodically aligned cobalt-rich and cobalt-poor phases (with periodicity $L \sim 2\,\mu$m) if subjected to external magnetic field. Optical microscope images of the diluted ferrofluid before and after application of external magnetic field are shown in Figs. 3.10(b) and 3.10(d). The periodic pattern of self-assembled stripes visible in image 3.10(d) appears due to phase separation. The stripes are oriented along the direction of magnetic field. The stripe periodicity $L \sim 2\,\mu$m appears to be much smaller than the free space wavelength in the hyperbolic frequency range. Therefore, created self-assembled optical medium appears to be a photonic hypercrystal. We should also note that the original undiluted ferrofluid exhibits similar phase separation in external magnetic field, although on a much smaller $\sim$0.3 μm spatial scale (see Fig. 3.10(f)).



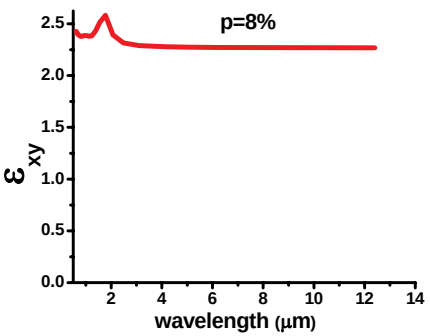
(a)



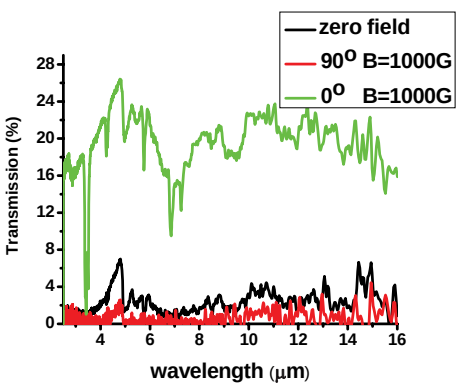
(c)



(b)



(d)



(e)

Fig. 3.12. (a) Optical properties of cobalt as tabulated in Ref. 22: real (n) and imaginary (k) parts of the cobalt refractive index are plotted in the LWIR range. (b) Critical volume fraction of cobalt nanoparticles corresponding to ferrofluid

Fig. 3.12. (*Continued*) transition to hyperbolic metamaterial phase shown as a function of free space light wavelength. (c,d) Wavelength dependencies of ε_z and ε_{xy} at $\alpha(\infty) = 8.2\%$. While ε_{xy} stays positive and almost constant, ε_z changes sign around $\lambda = 3\,\mu$m. (e) Polarization-dependent transmission spectra of 200μm-thick ferrofluid sample measured using FTIR spectrometer are consistent with hyperbolic character of ε tensor.

Polarization dependencies of ferrofluid transmission as a function of magnetic field and nanoparticle concentration measured in a broad 0.5–16 μm wavelength range conclusively prove hyperbolic crystal character of ferrofluid anisotropy in the LWIR range at large enough magnetic field. Fig. 3.13(e) shows polarization-dependent transmission spectra of 200 μm thick undiluted ferrofluid sample obtained using (FTIR) spectrometer. These data are consistent with hyperbolic character of ε tensor of the ferrofluid in $B = 1000$ G. Ferrofluid transmission is large for polarization direction perpendicular to magnetic field (perpendicular to cobalt nanoparticle chains), suggesting dielectric character of ε in this direction. On the other hand, ferrofluid transmission falls to near zero for polarization direction along the chains, suggesting metallic character of ε in this direction. However, these measurements are clearly affected by numerous ferrofluid absorption lines.

Fabricated photonic hypercrystals exhibit all the typical features associated with the hyperbolic metamaterials. For example, absorption spectra measured using FTIR spectrometer with and without external magnetic field are consistent with the decrease of the radiation lifetime of kerosene molecules in the hyperbolic state. In addition, Fig. 3.13 clearly illustrates the photonic hypercrystal potential in chemical and biological sensing, which is made possible by spatially selective field enhancement effects described above. This potential is revealed by detailed measurements of magnetic field-induced transmission of photonic hypercrystals in the broad IR spectral range presented in Figs. 3.13(a) and 3.13(b). FTIR spectral measurements are broadly accepted as a powerful "chemical fingerprinting" tool in chemical and biosensing. Therefore, broadly available magnetic

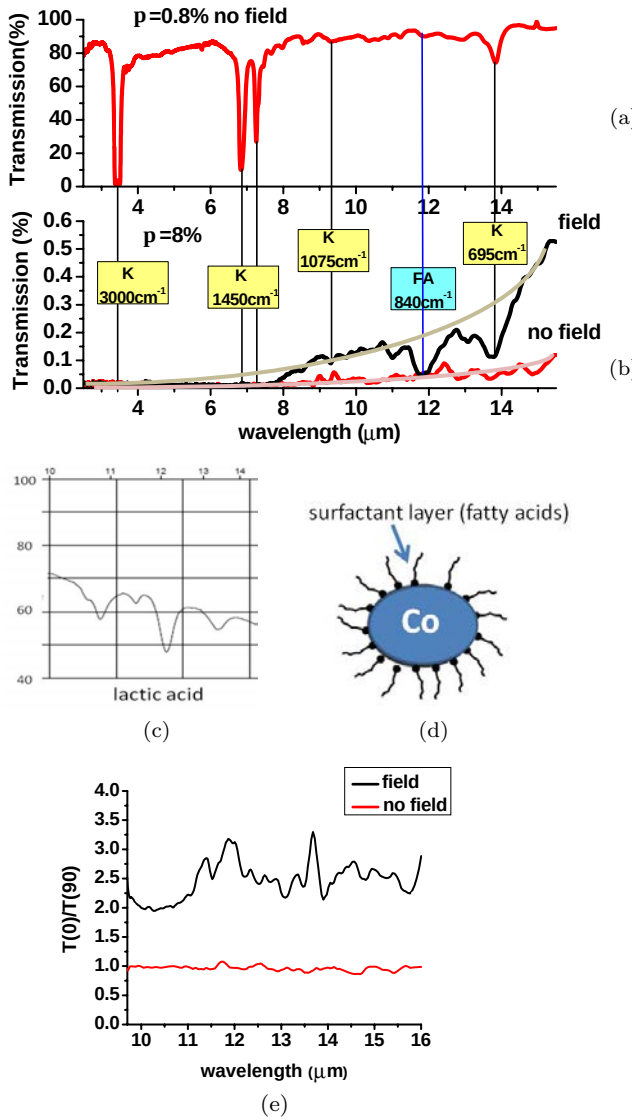


Fig. 3.13. (a) FTIR transmission spectrum of diluted ($\alpha(\infty) = 0.8\%$) ferrofluid exhibits a clear set of kerosene absorption lines. (b) Transmission spectra of the $\alpha(\infty) = 8.2\%$ ferrofluid measured with and without application of external magnetic field. Magnetic field-induced transmission spectrum contains a very pronounced absorption line at $\lambda \sim 12\,\mu$m ($\sim 840\,\text{cm}^{-1}$), which can be attributed

Fig. 3.13. (*Continued*) to lactic acid. Kerosene absorption lines are marked with yellow boxes, while the fatty acid line at 840 cm^{-1} is marked with a green box. (c) The 10–14 μm portion of lactic acid FTIR absorption spectrum. (d) Schematic view of cobalt nanoparticle coated with a monolayer of fatty acids, such as lactic and oleic acid. (e) Extinction coefficient (the ratio T(0)/T(90) of ferrofluid transmissions at 0° and 90°) of ferrofluid subjected to external magnetic field exhibits pronounced resonances around the fatty acid absorption line at $\lambda \sim 12\,\mu$m and the kerosene absorption line at $\lambda \sim 14\,\mu$m. These resonances provide clear evidence of field enhancement by cobalt nanoparticle chains.

field-tunable photonic hypercrystals operating in the IR range open up new, valuable opportunities in chemical analysis. The experimental data presented in Fig. 3.13 clearly illustrate this point. The FTIR transmission spectrum of the diluted ($p = 0.8\%$) ferrofluid in Fig. 3.13(a) exhibits a clear set of kerosene absorption lines, which is consistent with other published data (see for example Ref. 40). On the other hand, magnetic field-induced transmission spectrum of the $p = 8.2\%$ ferrofluid shown in Fig. 3.13(b) contains a very pronounced absorption line at $\lambda \sim 12\,\mu$m ($\sim$840 cm^{-1}), which cannot be attributed to kerosene. Quite naturally, this absorption line may be attributed to fatty acids, since cobalt nanoparticles are coated with a monolayer of surfactant composed of various fatty acids, such as lactiacid, oleicacid, etc. as shown in Fig. 3.13(d). A detailed comparison of Fig. 3.13(b) with the 10–14 μm portion of lactic acid FTIR absorption spectrum shown in Fig. 3.13(c) indeed indicates a close match. The fatty acid line appears to be about as strong as $\lambda \sim 14\,\mu$m ($\sim$695 cm^{-1}) line of kerosene, even though the oscillator strength of these molecular lines is about the same, while the amount of kerosene in the sample is ~ 2 orders of magnitude larger (a monolayer coating of fatty acids on a 10-nm cobalt nanoparticle occupies no more than 1% of ferrofluid volume). This paradoxical situation clearly indicates local field enhancement effects. Another strong evidence of field enhancement is provided by measurements of extinction coefficient of the ferrofluid presented in Fig. 3.6(e). Ferrofluid subjected to magnetic field exhibits pronounced resonances around the fatty acid absorption line at $\lambda \sim 12\,\mu$m and the kerosene

absorption line at $\lambda \sim 14$ μm. These resonances provide clear evidence of field enhancement by cobalt nanoparticle chains. We expect that further optimization of photonic hypercrystals geometry will lead to much stronger sensitivity of their optical properties to chemical and biological inclusions, indicating a very strong potential of photonic hypercrystals in biological and chemical sensing.

3.5. Superconducting Hyperbolic Metamaterials

Superconducting properties of a material, such as electron–electron interactions and the critical temperature of superconducting transition, can be expressed via the effective dielectric response function $\varepsilon_{\text{eff}}(q, \omega)$ of the material. Such a description is valid on the spatial scales below the superconducting coherence length (the size of the Cooper pair), which equals $\sim$100 nm in a typical BCS superconductor. Searching for natural materials exhibiting larger electron–electron interactions constitutes a traditional approach to high-temperature superconductivity research. However, not long ago it was pointed out that the recently developed field of electromagnetic metamaterials deals with the somewhat related task of dielectric response engineering on sub-100 nm scale, and that the metamaterial approach to dielectric response engineering may considerably increase the critical temperature of a composite superconductor–dielectric metamaterial.[41] Moreover, it appears that many high T_c superconductors exhibit hyperbolic metamaterial properties in substantial portions of the electromagnetic spectrum of relevance to electron–electron interaction,[42] so that their hyperbolicity may be partially responsible for their high T_c behavior.

Electromagnetic properties are known to play a very important role in the pairing mechanism and charge dynamics of high T_c superconductors.[43] Moreover, shortly after the original work by Bardeen, Cooper and Schrieffer (BCS),[44] Kirzhnits *et al.* formulated a complementary description of superconductivity in terms of the dielectric response function of the superconductor.[45] The latter work was motivated by a simple argument that phonon-mitigated

electron–electron interaction in superconductors may be expressed in the form of effective Coulomb potential

$$V(\vec{q},\omega) = \frac{4\pi e^2}{q^2 \varepsilon_{\mathit{eff}}(\vec{q},\omega)}, \tag{3.21}$$

where $V = 4\pi e^2/q^2$ is the usual Fourier-transformed Coulomb potential in vacuum and $\varepsilon_{\text{eff}}(q,\omega)$ is the linear dielectric response function of the superconductor treated as an effective medium. Based on this approach, Kirzhnits *et al.* derived simple expressions for the superconducting gap Δ, critical temperature T_c and other important parameters of the superconductor. While thermodynamic stability condition implies[46] that $\varepsilon_{\text{eff}}(q,0) > 0$, the dielectric response function at higher frequencies and spatial momenta is large and negative, which accounts for the weak net attraction and pairing of electrons in the superconducting condensate. In their paper, Kirzhnits *et al.* noted that this effective medium consideration assumes "homogeneous system" so that "the influence of the lattice periodicity is taken into account only to the extent that it may be included into $\varepsilon_{\text{eff}}(q,\omega)$".

In the 40 years which have passed since this very important remark, we have learned that the "homogeneous system" approximation may remain valid even if the basic structural elements of the material are not simple atoms or molecules. Now we know that artificial "metamaterials" may be created from much bigger building blocks, and the electromagnetic properties of these fundamental building blocks ("meta-atoms") may be engineered at will.[47] Since the superconducting coherence length (the size of the Cooper pair) is $\xi \sim 100\,\text{nm}$ in a typical BCS superconductor, we have an opportunity to engineer the fundamental metamaterial building blocks in such a way that the effective electron–electron interaction (3.21) will be maximized, while homogeneous treatment of $\varepsilon_{\text{eff}}(q,\omega)$ will remain valid. In order to do this, the metamaterial unit size must fall within a rather large window between $\sim$0.3 nm (given by the atomic scale) and $\xi > 100\,\text{nm}$ scale of a typical Cooper pair. However, this task is much more challenging than typical applications of superconducting

metamaterials suggested so far,[48] which only deal with metamaterial engineering on scales which are much smaller than the microwave or radiofrequency wavelength. Nevertheless, these experimental difficulties have been overcome in two recent successful demonstrations of metamaterial superconductors.[49,50] In the case of aluminum, its superconducting temperature was more than tripled by the metamaterial engineering.[50]

Let us demonstrate that hyperbolic metamaterial geometry offers a natural way to increase attractive electron–electron interaction in a layered dielectric–superconductor metamaterial. Since hyperbolic metamaterials exhibit considerable dispersion, let us work in the frequency domain and write macroscopic Maxwell equations in the presence of "external" electron density ρ_ω and current J_ω as

$$\begin{aligned} \frac{\omega^2}{c^2}\vec{D}_\omega &= \vec{\nabla}\times\vec{\nabla}\times\vec{E}_\omega - \frac{4\pi i\omega}{c^2}\vec{J}_\omega, \\ \vec{\nabla}\cdot\vec{D}_\omega &= \rho_\omega, \quad \text{and} \quad \vec{D}_\omega = \overleftrightarrow{\varepsilon}_\omega\,\vec{E}_\omega, \end{aligned} \tag{3.22}$$

where the frequency ω is assumed to fall within the hyperbolic frequency band of the metamaterial. Let us solve Eq. (10.22) for the z-component of electric field. After straightforward transformations, we obtain

$$\frac{\omega^2}{c^2}E_z = \frac{4\pi}{\varepsilon_1\varepsilon_2}\frac{\partial\rho}{\partial z} - \frac{4\pi i\omega}{c^2\varepsilon_2}J_z - \frac{\partial^2 E_z}{\varepsilon_1\partial z^2} - \frac{1}{\varepsilon_2}\left(\frac{\partial^2 E_z}{\partial x^2} + \frac{\partial^2 E_z}{\partial y^2}\right). \tag{3.23}$$

Since $E_z = \partial\phi/\partial z$, and the second term on the right side of Eq. (3.23) may be neglected compared to the first one (since $v/c \ll 1$), we obtain

$$\frac{\omega^2}{c^2}\phi + \frac{\partial^2\phi}{\varepsilon_1\partial z^2} + \frac{1}{\varepsilon_2}\left(\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2}\right) = \frac{4\pi}{\varepsilon_1\varepsilon_2}\rho. \tag{3.24}$$

Taking into account that $V = -e\phi$, and neglecting the first term in Eq. (3.24) in the low frequency limit, we find that the effective

Coulomb potential from Eq. (3.21) assumes the form

$$V(\vec{q},\omega) = \frac{4\pi e^2}{q_z^2\varepsilon_2(\vec{q},\omega) + (q_x^2+q_y^2)\varepsilon_1(\vec{q},\omega)} \tag{3.25}$$

in a hyperbolic metamaterial. Since $\varepsilon_{xx} = \varepsilon_{yy} = \varepsilon_1$ and $\varepsilon_{zz} = \varepsilon_2$ have opposite signs, the effective Coulomb interaction of two electrons may become attractive and very strong in the hyperbolic frequency bands. The obvious condition for such a strong interaction to occur is

$$q_z^2\varepsilon_2(\vec{q},\omega) + (q_x^2+q_y^2)\varepsilon_1(\vec{q},\omega) \approx 0 \tag{3.26}$$

which indicates that the superconducting order parameter must be strongly anisotropic. This indeed appears to be the case in such hyperbolic high T_c superconductors as BSCCO.[42,43] In order to be valid, the metamaterial "effective medium" description requires that the structural parameter of the metamaterial (in this particular case, the interlayer distance) must be much smaller than the superconducting coherence length. If the structural parameter approaches 1 nm scale, Josephson tunneling across the dielectric layers will become very prominent in such an anisotropic layered superconducting hyperbolic metamaterial.

The diagonal dielectric permittivity components of the layered superconductor–dielectric metamaterial may be calculated using Maxwell–Garnett approximation using Eq. (10.5). In order to obtain hyperbolic properties, ε_d of the dielectric needs to be very large, since ε_m of the superconducting component given by the Drude model is negative and very large in the far infrared and THz ranges

$$\varepsilon_m = \varepsilon_{m\infty} - \frac{\omega_p^2}{\omega^2} \approx -\frac{\omega_p^2}{\omega^2}. \tag{3.27}$$

This is consistent with the measured dielectric behavior of the parent BSCCO perovskite compound.[43] Moreover, if the high frequency behavior of ε_d may be assumed to follow the Debye model[51]:

$$\mathrm{Re}\varepsilon_d = \frac{\varepsilon_d(0)}{1+\omega^2\tau^2} \approx \frac{\varepsilon_d(0)}{\omega^2\tau^2}, \tag{3.28}$$

broadband hyperbolic behavior arise due to similar $\sim \omega^{-2}$ functional behavior of ε_d and ε_m in the THz range.

As follows from Eq. (10.5), if the volume fraction of metallic phase n is kept constant, the hyperbolic behavior may occur only within the following range of plasma frequency ω_p^2 of the metallic phase:

$$\frac{\omega_p^2 \tau^2}{\varepsilon_d(0)} \in \left[\frac{n}{1-n}; \frac{1-n}{n}\right]. \tag{3.29}$$

Otherwise, either ε_1 and ε_2 will be both positive if ω_p^2 is too small, or both negative if ω_p^2 is too large. Interestingly enough, the boundaries of superconducting and hyperbolic states in high T_c cuprates seem to overlap. Moreover, the crystallographic lattice of BSCCO shown in Fig. 3.14 looks very similar to the geometry of a multilayer hyperbolic metamaterial. Indeed, in BSCCO the anisotropy of DC conductivity may reach 10^4 for the ratio of

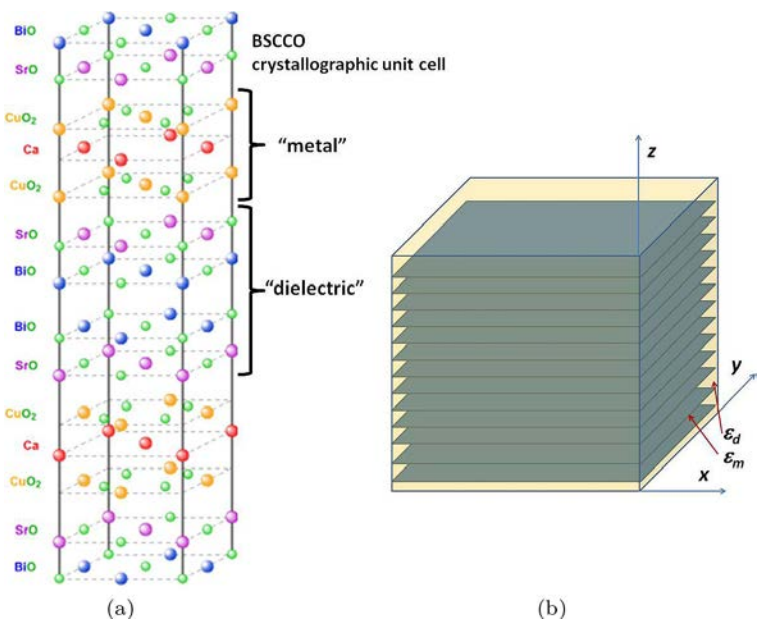


Fig. 3.14. Comparison of the crystallographic unit cell of a BSCCO high T_c superconductor (a) and geometry of a layered hyperbolic metamaterial (b).

in-plane to out-of-plane conductivity in high-quality single crystal samples. Polarization-dependent AC reflectance spectra measured in the THz and far-infrared frequency ranges[43] also indicate extreme anisotropy. In the normal state of high T_c superconductors, the in-plane AC conductivity exhibits Drude-like behavior with a plasma edge close to 10,000 cm^{-1}, while AC conductivity perpendicular to the copper oxide planes is nearly insulating. Extreme anisotropy is also observed in the superconducting state. The typical values of measured in-plane and out of plane condensate plasma frequencies in high T_c superconductors are $\omega_{p,ab} = 4000 - 10{,}000\,\text{cm}^{-1}$, and $\omega_{p,c} = 1 - 1000\,\text{cm}^{-1}$, respectively.[43] The measured anisotropy is the strongest in the BSCCO superconductors. These experimental measurements strongly support the qualitative picture of BSCCO structure as a layered hyperbolic metamaterial (Fig. 3.14(b)) in which the copper oxide layers may be represented as metallic layers, while the SrO and BiO layers may be represented as the layers of dielectric. Based on these measured material parameters, the diagonal components of BSCCO dielectric tensor may be calculated[41,42] as shown in Fig. 3.15. The appearance of hyperbolic bands is quite generic in

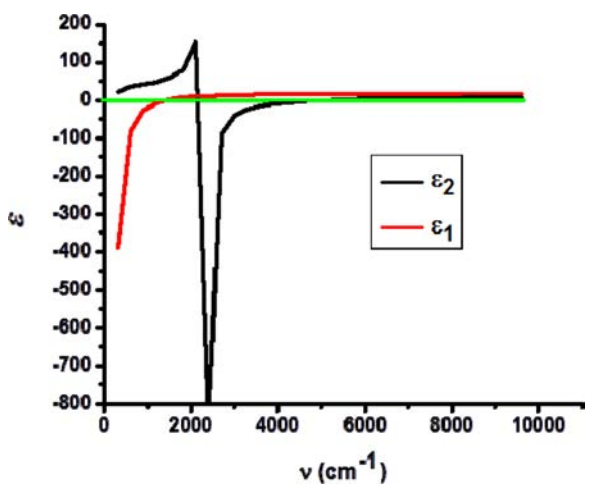


Fig. 3.15. Diagonal components of the permittivity tensor of a high Tc BSCCO superconductor calculated as a function of frequency. The hyperbolic bands appear at $200\,\text{cm}^{-1} < \omega < 1200\,\text{cm}^{-1}$ and $2400\,\text{cm}^{-1} < \omega < 4800\,\text{cm}^{-1}$.

high T_c superconductors. While examples of such natural hyperbolic high T_c superconductors appear to fit well into the metamaterial scheme described above, it would be interesting to try and follow the metamaterial recipe in making novel "designer" superconductors.

3.6. Gravitation Theory Analogs Based on Hyperbolic Metamaterials

Modern developments in gravitation research strongly indicate that classic general relativity is an effective macroscopic field theory, which needs to be replaced with a more fundamental theory based on yet unknown microscopic degrees of freedom. On the other hand, our ability to obtain experimental insights into the future fundamental theory is strongly limited by low energy scales available to terrestrial particle physics and astronomical observations. The emergent analog space-time program offers a promising way around this difficulty. Looking at such systems as superfluid helium and cold atomic Bose–Einstein condensates, physicists learn from nature and discover how macroscopic field theories arise from known well-studied atomic degrees of freedom. Another exciting development along this direction is recent introduction of metamaterials and transformation optics. The latter field is not limited by the properties of atoms and molecules given to us by nature. "Artificial atoms" used as building blocks in metamaterial design offer much more freedom in constructing analogs of various exotic space-time metrics, such as black holes,[52–56] wormholes,[57,58] spinning cosmic strings,[59] and even the metric of Big Bang itself.[15] Explosive development of this field promises new insights into the fabric of space-time, which cannot be gleaned from any other terrestrial experiments.

On the other hand, compared to standard general relativity, metamaterial optics gives more freedom to design an effective space-time with very unusual properties. Light propagation in all static general relativity situations can be mimicked with positive $\varepsilon_{ik} = \mu_{ik}$,[60] while the allowed parameter space of the metamaterial optics is broader. Thus, flat Minkowski space-time with the usual

$(-,+,+,+)$ signature does not need to be a starting point. Other effective signatures, such as the "two times" (2T) physics $(-,-,+,+)$ signature may be realized.[3] Theoretical investigation of the 2T higher-dimensional space-time models had been pioneered by Paul Dirac.[61] More recent examples can be found in Refs. 62 and 63 Metric signature change events (in which a phase transition occurs between say $(-,+,+,+)$ and $(-,-,+,+)$ space-time signature) are being studied in Bose–Einstein condensates and in some modified gravitation theories (see Ref. 64, and the references therein). It is predicted that a quantum field theory residing on a space-time undergoing a signature change reacts violently to the imposition of the signature change. Both the total number and the total energy of the particles generated in a signature change event are formally infinite.[64] While optics of bulk hyperbolic metamaterials provides us with ample opportunities to observe metric signature transitions,[3] even more interesting physics arise at the metamaterial interfaces. Very recently, it was demonstrated that mapping of monochromatic extraordinary light distribution in a hyperbolic metamaterial along some spatial direction may model the "flow of time" in a 3D $(2+1)$ effective Minkowski space-time.[15] If an interface between two metamaterials is engineered so that the effective metric changes signature across the interface, two possibilities may arise. If the interface is perpendicular to the time-like direction z, this coordinate does not behave as a "time-like" variable any more, and the continuous "flow of time" is interrupted. This situation (which cannot be realized in classic general relativity) may be called the "end of time". It appears that optics of metamaterials near the "end of time" event is quite interesting and deserves a detailed study. For example, in the lossless approximation all the possible "end of time" scenarios lead to field divergences, which indicate quite interesting linear and nonlinear optics behavior near the "end of time". On the other hand, if the metamaterial interface is perpendicular to the space-like direction of the effective $(2+1)$ Minkowski space-time, a Rindler horizon may be observed (Rindler metric approximates space-time behavior near the black hole event horizon[60]).

Let us briefly summarize Refs. 3 and 15, which demonstrated that a spatial coordinate may become "time-like" in a hyperbolic metamaterial. To better understand this effect, let us start with a non-magnetic uniaxial anisotropic material with dielectric permittivities $\varepsilon_x = \varepsilon_y = \varepsilon_1$ and $\varepsilon_z = \varepsilon_2$, and assume that this behavior holds in some frequency range around $\omega = \omega_0$. Let us consider the case of constant $\varepsilon_1 > 0$ and $\varepsilon_2 < 0$, and assume that the metamaterial is illuminated by coherent CW laser field at frequency ω_0. We will study spatial distribution of the extraordinary field φ_ω at this frequency. Under these assumptions, Eq. (3.1) may be rewritten in the form of 3D Klein–Gordon equation describing a massive scalar φ_ω field:

$$-\frac{\partial^2\varphi_\omega}{\varepsilon_1\partial z^2} + \frac{1}{|\varepsilon_2|}\left(\frac{\partial^2\varphi_\omega}{\partial x^2} + \frac{\partial^2\varphi_\omega}{\partial y^2}\right) = \frac{\omega_0^2}{c^2}\varphi_\omega = \frac{m^{*2}c^2}{\hbar^2}\varphi_\omega \qquad (3.30)$$

in which the spatial coordinate $z = \tau$ behaves as a "time-like" variable. Therefore, Eq. (3.30) describes world lines of massive particles which propagate in a flat $(2+1)$ Minkowski space-time. When a metamaterial is built and illuminated with a coherent extraordinary CW laser beam, the stationary pattern of light propagation inside the metamaterial represents a complete "history" of a toy $(2+1)$-dimensional space-time populated with particles of mass m^*. This "history" is written as a collection of particle world lines along the "time-like" z coordinate. Note that in the opposite situation in which $\varepsilon_1 < 0$ and $\varepsilon_2 > 0$, Eq. (3.30) would describe world lines of tachyons[65] having "imaginary" mass $m^* = i\mu$. Eq. (3.30) exhibits effective Lorentz invariance under the coordinate transformation

$$z' = \frac{1}{\sqrt{1 - \dfrac{\varepsilon_{xy}}{(-\varepsilon_z)}\beta}}(z - \beta x), \qquad (3.31)$$

$$x' = \frac{1}{\sqrt{1 - \dfrac{\varepsilon_{xy}}{(-\varepsilon_z)}\beta}}\left(x - \beta\frac{\varepsilon_{xy}}{(-\varepsilon_z)}z\right),$$

where β is the effective boost. Similar to our own Minkowski space-time, the effective Lorentz transformations in the xz and yz planes form the Poincare group together with translations along x, y and z axis, and rotations in the xy plane.

The world lines of particles described by Eq. (3.30) are straight lines, which is easy to observe in the experiment.[15] If adiabatic variations of ε_1 and ε_2 are allowed inside the metamaterial, world lines of massive particles in some well-known curvilinear space-times can be emulated, including the world line behavior near the "beginning of time" at the moment of the Big Bang, as illustrated in Fig. 3.16.[15] Thus, mapping of monochromatic extraordinary light distribution in a hyperbolic metamaterial along some spatial direction may model the "flow of time" in an effective 3D $(2 + 1)$ space-time. Since the parameter space of metamaterial optics is broader than the parameter space of general relativity, we can also engineer the "end of time" event if an interface between two metamaterials is prepared so that the effective metric changes signature at the interface. In such a case, the spatial coordinate does not behave as a "time-like" variable any more, and the continuous "flow of time" is suddenly interrupted, as shown in Fig. 3.17(a). This situation (which cannot be realized in classic general relativity) may be called the "end of time". It appears that optics of metamaterials near the "end of time" event is quite interesting and deserves a detailed study.[16] It appears that the optical images of field distribution over the sample surface indicate considerable field enhancement near the presumed plasmonic "end of time" events, as indicated by an arrow in Fig. 3.18(f). On the other hand, a hyperbolic metamaterial interface, which is oriented perpendicular to the "space-like" direction (Fig. 3.17(b)) behaves as a Rindler event horizon.

It is also interesting to note that nonlinear light propagation through a hyperbolic metamaterial may be formulated in a similar fashion as general relativity.[13] Sub-wavelength confinement of light in nonlinear hyperbolic metamaterials due to formation of spatial solitons has attracted much recent attention because of its seemingly counter-intuitive behavior.[66,67] In order to achieve self-focusing in a

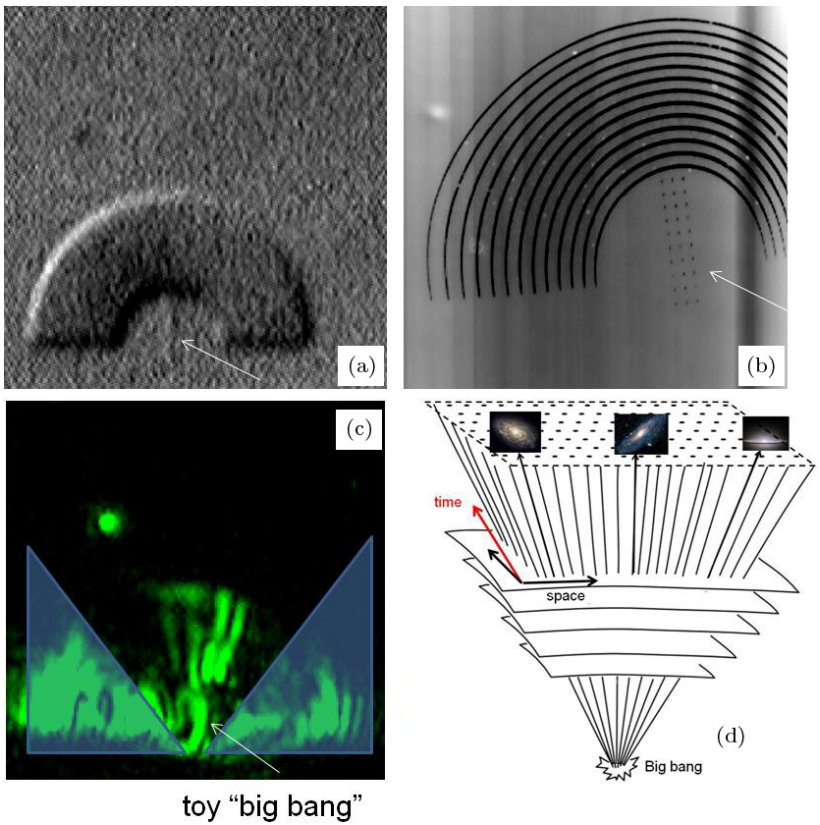


Fig. 3.16. Experimental demonstration of world line behavior in an "expanding universe" using a plasmonic hyperbolic metamaterial: Optical (a) and AFM (b) images of the plasmonic hyperbolic metamaterial based on PMMA stripes on gold. The defect used as a plasmon source is shown by an arrow. (c) Plasmonic rays or "world lines" increase their spatial separation as a function of "time-like" radial coordinate. The point (or moment) $r = \tau = 0$ corresponds to a toy "big bang". For the sake of clarity, light scattering by the edges of the PMMA pattern is partially blocked by semi-transparent triangles. (d) Schematic view of world lines behavior near the Big Bang.

hyperbolic wire medium, a nonlinear self-defocusing Kerr medium must be used as a dielectric host. Reference 13 demonstrated that this behavior finds natural explanation in terms of analog gravity. Since the wave equation describing propagation of extraordinary light

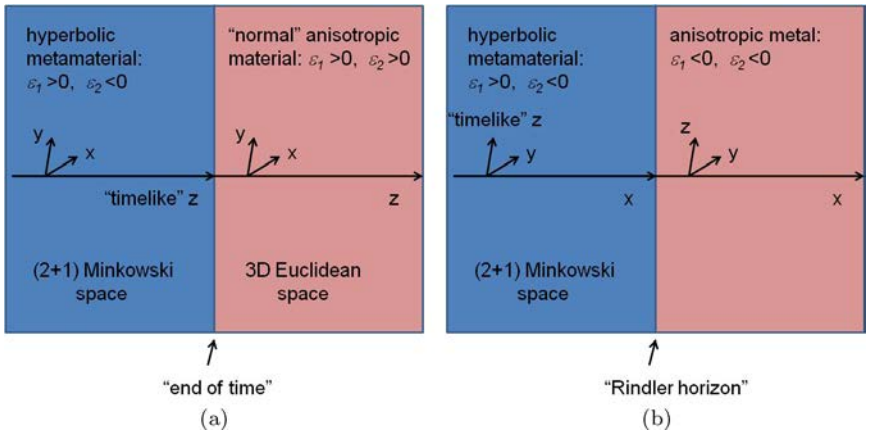


Fig. 3.17. (a) Schematic representation of the "end of time" model in metamaterials: the spatial coordinate z does not behave as a "time-like" variable any more, and the continuous "flow of time" is suddenly interrupted at the interface of two metamaterials. (b) Metric signature change across a "space-like" direction leads to appearance of a Rindler horizon.

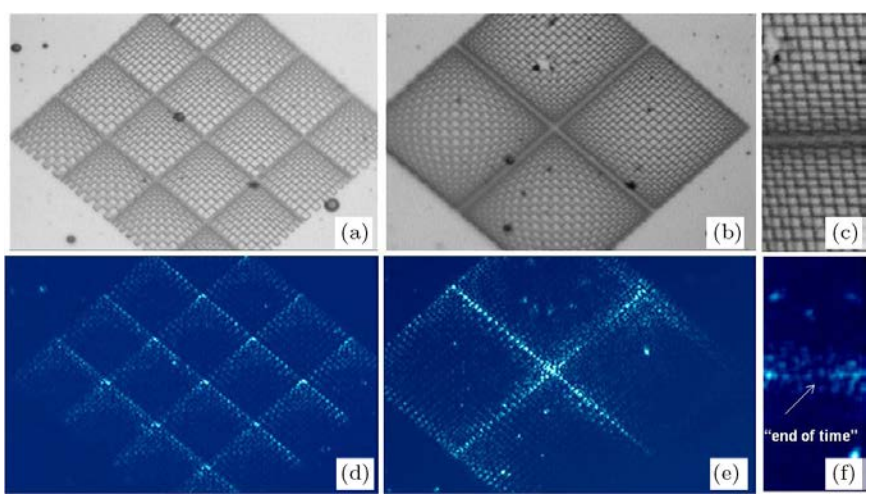


Fig. 3.18. Experimental observation of the "end of time event" in a plasmonic hyperbolic metamaterial illuminated with 488 nm light.

inside hyperbolic metamaterials exhibits $2+1$-dimensional Lorentz symmetry, we may assume that nonlinear optical Kerr effect "bends" this space-time resulting in effective gravitational force between extraordinary photons. In order for the effective gravitational constant to be positive, negative self-defocusing Kerr medium must be used as a host. If gravitational self-interaction is strong enough, spatial soliton may collapse into a black hole analog.

When the nonlinear optical effects become important, they are described in terms of various order nonlinear susceptibilities $\chi^{(n)}$ of the metamaterial:

$$D_i = \chi^{(1)}_{ij} E_j + \chi^{(2)}_{ijl} E_j E_l + \chi^{(3)}_{ijlm} E_j E_l E_m + \cdots . \tag{3.32}$$

Taking into account these nonlinear terms, the dielectric tensor of the metamaterial (which defines its effective metric) may be written as

$$\varepsilon_{ij} = \chi^{(1)}_{ij} + \chi^{(2)}_{ijl} E_l + \chi^{(3)}_{ijlm} E_l E_m + \cdots . \tag{3.33}$$

It is clear that Eq. (3.33) provides coupling between the matter content (photons) and the effective metric of the metamaterial "space-time". However, in order to emulate gravity, the nonlinear susceptibilities $\chi^{(n)}$ of the metamaterial need to be engineered in some particular way. In the weak gravitational field limit, the Einstein equation

$$R^k_i = \frac{8\pi\gamma}{c^4}\left(T^k_i - \frac{1}{2}\delta^k_i T\right) \tag{3.34}$$

is reduced to

$$R_{00} = \frac{1}{c^2}\Delta\phi = \frac{1}{2}\Delta g_{00} = \frac{8\pi\gamma}{c^4} T_{00}, \tag{3.35}$$

where ϕ is the gravitational potential.[60] Since in our effective Minkowski space-time g_{00} is identified with $-\varepsilon_1$, comparison of Eqs. (3.33) and (3.34) indicates that all the second-order nonlinear susceptibilities $\chi^{(2)}_{ijl}$ of the metamaterial must be equal to zero, while the third-order terms may provide correct coupling between the effective metric and the energy–momentum tensor. These terms are associated with the optical Kerr effect. All the higher order $\chi^{(n)}$

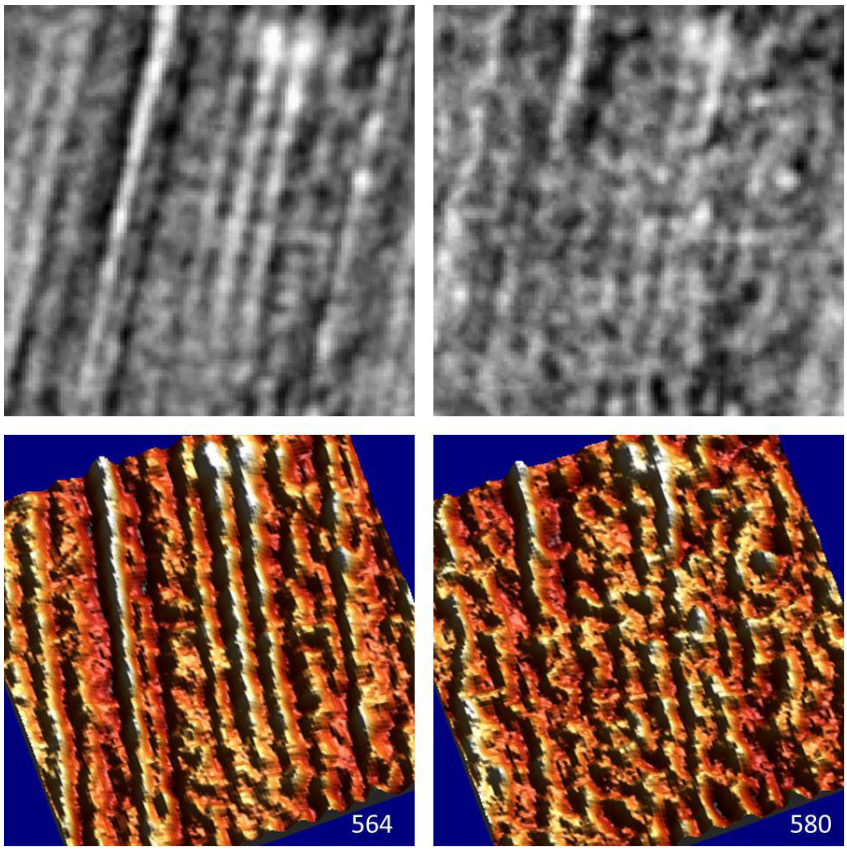


Fig. 3.19. Magnified images of the ferrofluid taken as a function of external magnetic field reveal fine details of the effective Minkowski space-time melting. Top row shows these images in the usual greyscale format, while the quasi-3D representation of the same images in the bottom row provides better visualization of the actual nanoparticle filaments.

terms must be zero at $n > 3$. Indeed, detailed analysis indicates[13] that Kerr effect in a hyperbolic metamaterial leads to effective gravity if the dielectric medium used as a metamaterial host exhibits self-defocusing nonlinearity, so that the magnetic ferrofluids described above appear to be ideal candidates to exhibit such phenomena.

Unlike other typical metamaterial systems, such ferrofluid-based macroscopic self-assembled 3D metamaterials may also exhibit

reach physics associated with topological defects[14,68] and phase transitions. Therefore, as was pointed out recently by Mielczarek and Bojowald,[62,63] the properties of self-assembled magnetic nanoparticle-based hyperbolic metamaterials exhibit strong similarities with the properties of some microscopic quantum gravity models, such as loop quantum cosmology. As described in Section 4, in the presence of an external magnetic field the ferrofluid forms a self-assembled hyperbolic metamaterial, which may be described as an effective "3D Minkowski space-time" for extraordinary photons. If the magnetic field is not strong enough, this effective Minkowski space-time gradually melts under the influence of thermal fluctuations, as illustrated in Fig. 3.19. Thus, unlike other systems exhibiting analog gravity behaviour, the unique feature of the ferrofluid consists in our ability to directly visualize the effective Minkowski space-time formation at the microscopic level.

3.7. Summary

The diverse physical properties and applications of hyperbolic metamaterials outlined above clearly demonstrate that this subfield of electromagnetic metamaterials already extended far beyond its original goal to enable subdiffraction super-resolution imaging. Hyperbolic metamaterials demonstrate a large number of novel phenomena resulting from the broadband singular behavior of their density of photonic states, which encompass enhanced quantum electrodynamic effects, new stealth technology, thermal hyperconductivity, high T_c superconductivity, and very interesting gravitation theory analogs. Moreover, hyperbolic metamaterial behavior appears to be compatible with photonic crystal effects, such as deep sub-wavelength light localization and giant field enhancement, resulting in a fascinating new class of artificial optical media — the photonic hypercrystals. Since all this progress has been demonstrated within a few years from their inception, the future of hyperbolic metamaterials and their applications looks really bright.

References

1. Podolskiy, V. A. and Narimanov, E. E. (2005). Strongly anisotropic waveguide as a nonmagnetic left-handed system, *Phys. Rev. B* **71**, p. 201101.
2. Jacob, Z., Alekseyev, L. V. and Narimanov, E. (2006). Optical hyperlens: Far-field imaging beyond the diffraction limit, *Optics Express* **14**, pp. 8247–8256.
3. Smolyaninov, I. I. and Narimanov, E. E. (2010). Metric signature transitions in optical metamaterials, *Phys. Rev. Lett.* **105**, p. 067402.
4. Smolyaninov, I. I., Hung, Y. J. and Davis, C. C. (2007). Magnifying superlens in the visible frequency range, *Science* **315**, pp. 1699–1701.
5. Liu, Z., Lee, H., Xiong, Y., Sun, C. and Zhang, X. (2007). Far-field optical hyperlens magnifying sub-diffraction-limited objects, *Science* **315**, p. 1686.
6. Jacob, Z., Smolyaninov, I. I. and Narimanov, E. E. (2012). Broadband Purcell effect: Radiative decay engineering with metamaterials, *Appl. Phys. Lett.* **100**, p. 181105.
7. Jacob, Z., Kim, J.-Y., Naik, G. V., Boltasseva, A., Narimanov, E. E. and Shalaev, V. M. (2010). Engineering photonic density of states using metamaterials, *App. Phys. B* **100**, p. 215.
8. Noginov, M. A., Li, H., Yu. A., Barnakov, D., Dryden, G., Nataraj, G., Zhu, C., Bonner, E., Mayy, M., Jacob, Z. and Narimanov, E. E. (2010). Controlling spontaneous emission with metamaterials, *Opt. Lett.* **35**, p. 1863.
9. Narimanov, E., Noginov, M. A., Li, H. and Barnakov, Y. (2010). Darker than black: Radiation-absorbing metamaterial, in *Quantum Electronics and Laser Science Conference, OSA Technical Digest (CD)* (Optical Society of America), paper QPDA6.
10. Narimanov, E. E. and Smolyaninov, I. I. Beyond Stefan-Boltzmann law: Thermal hyper-conductivity, arXiv:1109.5444.
11. Smolyaninov, I. I. (2014). Quantum topological transition in hyperbolic metamaterials based on high Tc superconductors, *J. Phys.: Condensed Matt.* **26**, p. 305701.
12. Smolyaninov, I. I. and Smolyaninova, V. N. (2014). Is there a metamaterial route to high temperature superconductivity?, *Adv. Condensed Matt. Phys.* **2014**, p. 479635.
13. Smolyaninov, I. I. (2013). Analogue gravity in hyperbolic metamaterials, *Phys. Rev. A* **88**, p. 033843.
14. Smolyaninov, I. I., Yost, B., Bates, E. and Smolyaninova, V. N. (2013). Experimental demonstration of metamaterial "multiverse" in a ferrofluid, *Opti. Expr.* **21**, pp. 14918–14925.
15. Smolyaninov, I. I. and Hung, Y. J. (2011). Modeling of time with metamaterials, *JOSA B* **28**, pp. 1591–1595.
16. Smolyaninov, I. I., Hwang, E. and Narimanov, E. E. (2012). Hyperbolic metamaterial interfaces: Hawking radiation from Rindler horizons and spacetime signature transtions, *Phys. Rev. B* **85**, p. 235122.

17. Smolyaninov, I. I. (2014). Holographic duality in nonlinear hyperbolic metamaterials, *J. Opti.* **16**, p. 075101.
18. Narimanov, E. E. and Kildishev, A. V. (2015). Metamaterials: Naturally hyperbolic, *Nat. Photon.* **9**, pp. 214–216.
19. Smolyaninov, I. I. (2011). Vacuum in strong magnetic field as a hyperbolic metamaterial, *Phys. Rev. Lett.* **107**, p. 253903.
20. Hoffman, A. J., Alekseyev, L., Howard, S. S., Franz, K. J., Wasserman, D., Podolskiy, V. A., Narimanov, E. E., Sivco, D. L. and Gmachl, C. (2007). Negative refraction in semiconductor metamaterials, *Nat. Mater.* **6**, p. 946.
21. Noginov, M. A., Barnakov, Yu. A., Zhu, G., Tumkur, T., Li, H. and Narimanov, E. E. (2009). Bulk photonic metamaterial with hyperbolic dispersion, *Appl. Phys. Lett.* **94**, p. 151105.
22. Yao, J., Liu, Z., Liu, Y., Wang, Y., Sun, C., Bartal, G., Stacy A. and Zhang, X. (2008). Optical negative refraction in bulk metamaterials, *Science* **321**, p. 930.
23. Wangberg, R., Elser, J., Narimanov, E. E. and Podolskiy, V. A. (2006). Nonmagnetic nanocomposites for optical and infrared negative-refractive-index media, *J. Opt. Soc. Am. B* **23**, pp. 498–505.
24. Jacob, Z. and Narimanov, E. E. (2008). Optical hyperspace for plasmons: Dyakonov states in metamaterials, *Appl. Phys. Lett.* **93**, p. 221109.
25. Zayats, A. V., Smolyaninov, I. I. and Maradudin, A. (2005). Nano-optics of surface plasmon-polaritons, *Phys. Reports* **408**, pp. 131–314.
26. Pohl, D. W. and Courjon, D. (eds.) (1993). *Near Field Optics*, NATO ASI-E Series, (Kluwer Academic Publishers, The Netherlands).
27. Smolyaninov, I. I., Elliott, J., Zayats, A. V. and Davis, C. C. (2005). Far-field optical microscopy with nanometer-scale resolution based on the in-plane image magnification by surface plasmon polaritons, *Phys. Rev. Lett.* **94**, p. 057401.
28. Lounis, B. and Orrit, M. (2005). Single-photon sources, *Rep. Prog. Phys.* **68**, p. 1129.
29. Morel, Y. and Smolyaninov, I. I. Hyperbolic metamaterials as distributed Bragg mirrors for high power VCSEL devices, US Patent # 8,831,058.
30. Smolyaninova, V. N., Yost, B., Lahneman, D., Narimanov, E. and Smolyaninov, I. I. (2014). Self-assembled tunable photonic hyper-crystals, *Scientific Reports* **4**, p. 5706.
31. Yablonovitch, E. (1987). Inhibited spontaneous emission in solid-state physics and electronics, *Phys. Rev. Lett.* **58**, pp. 2059–2062.
32. John, S. (1987). Strong localization of photons in certain disordered dielectric superlattices, *Phys. Rev. Lett.* **58**, pp. 2486–2489.
33. Pendry, J. B. (2000). Negative refraction makes a perfect lens, *Phys. Rev. Lett.* **85**, pp. 3966–3969.
34. Krauss, T. F., DeLaRue, R. M. and Brand, S. (1996). Two-dimensional photonic-bandgap structures operating at near-infrared wavelengths, *Nature* **383**, pp. 699–702.

35. Schurig, D. *et al.* (2006). Metamaterial electromagnetic cloak at microwave frequencies, *Science* **314**, pp. 977–980.
36. Fleischmann, M., Hendra, P. J. and McQuillan, A. J. (1974). Raman spectra of pyridine adsorbed at a silver electrode, *Chem. Phys. Lett.* **26**, pp. 163–166.
37. Gao, Y. *et al.* (2010). Optical negative refraction in ferrofluids with magnetocontrollability, *Phys. Rev. Lett.* **104**, p. 034501.
38. Zhang, H. and Widom, M. (1993). Spontaneous magnetic order in strongly coupled ferrofluids, *J. Magnet. Magnet. Mater.* **122**, pp. 119–122.
39. Lide, D. R. (Ed.) (2005). *CRC Handbook of Chemistry and Physics*, (CRC Press, Boca Raton).
40. Al-Ghouti, M. A., Al-Degs, Y. S. and Amer, M. (2008). Determination of motor gasoline adulteration using FTIR spectroscopy and multivariate calibration, *Talanta* **76**, pp. 1105–1112.
41. Smolyaninov, I. I. and Smolyaninova, V. N. (2014). Is there a metamaterial route to high temperature superconductivity?, *Adv. Condensed Matt. Phys.* **2014**, p. 479635.
42. Smolyaninov, I. I. (2014). Quantum topological transition in hyperbolic metamaterials based on high Tc superconductors, *J. Phys.: Condensed Matt.* **26**, p. 305701.
43. Basov, D. N. and Timusk, T. (2005). Electrodynamics of high-T_c superconductors, *Rev. Mod. Phys.* **77**, p. 721.
44. Bardeen, J., Cooper, L. N. and Schrieffer, J. R. (1957). Theory of superconductivity, *Phys. Rev.* **108**, p. 1175.
45. Kirzhnits, D. A., Maksimov, E. G. and Khomskii, D. I. (1973). The description of superconductivity in terms of dielectric response function, *J. Low Temp. Phys.* **10**, p. 79.
46. Landaum, L. D. and Lifshitz, E. M. (2000). *Macroscopic Electrodynamics* (Elsevier, Oxford).
47. Pendry, J. B., Schurig, D. and Smith, D. R. (2006). Controlling electromagnetic fields, *Science* **312**, p. 1780.
48. Anlage, S. M. (2011). The physics and applications of superconducting metamaterials, *J. Opt.* **13**, p. 024001.
49. Smolyaninova, V. N., Yost, B., Zander, K., Osofsky, M. S., Kim, H., Saha, S., Greene, R. L. and Smolyaninov, I. I. (2014). Experimental demonstration of superconducting critical temperature increase in electromagnetic metamaterials, *Sci. Reports* **4**, p. 7321.
50. Smolyaninova, V. N., Zander, K., Gresock, T., Jensen, C., Prestigiacomo, J. C., Osofsky, M. S. and Smolyaninov, I. I. Using metamaterial nanoengineering to triple the superconducting critical temperature of bulk aluminum, arXiv:1506.00258 [physics.optics].
51. Kittel, C. (2004). *Introduction to Solid State Physics* (Wiley, New York).
52. Smolyaninov, I. I. (2003). Surface plasmon toy-model of a rotating black hole, *New J. Phys.* **5**, p. 147.

53. Smolyaninov, I. I. (2011). Critical opalescence in hyperbolic metamaterials, *J. Opt.* **13**, p. 125101.
54. Genov, D. A., Zhang, S. and Zhang, X. (2009). Mimicking celestial mechanics in metamaterials, *Nat. Phys.*, **5**, pp. 687–692.
55. Narimanov, E. E. and Kildishev, A. V. (2009). Optical black hole: Broadband omnidirectional light absorber, *Appl. Phys. Lett.* **95**, p. 041106.
56. Cheng, Q. and Cui, T. J. An electromagnetic black hole made of metamaterials, arXiv:0910.2159v3.
57. Greenleaf, A., Kurylev, Y., Lassas, M. and Uhlmann, G. (2007). Electromagnetic wormholes and virtual magnetic monopoles from metamaterials, *Phys. Rev. Lett.* **99**, p. 183901.
58. Smolyaninov, I. I. (2010). Metamaterial Multiverse, *J. Opt.* **13**, p. 024004.
59. Mackay, T. G. and Lakhtakia, A. (2010). Towards a metamaterial simulation of a spinning cosmic string, *Phys. Lett. A* **374**, pp. 2305–2308.
60. Landau, L. and Lifshitz, E. (2000). *The Classical Theory of Fields* (Elsevier, Oxford).
61. Dirac, P. A. M (1936). Wave equations in conformal space, *Ann. Math.* **37**, p. 429.
62. Mielczarek, J. Signature change in loop quantum cosmology, arXiv:1207.4657 [gr-qc].
63. Bojowald, M. and Mielczarek, J. Some implications of signature-change in cosmological models of loop quantum gravity, arXiv:1503.09154 [gr-qc].
64. White, A., Weinfurtner, S. and Visser, M. Signature change events: A challenge for quantum gravity? arXiv:0812.3744v3.
65. Feinberg, G. (1967). Possibility of faster-than-light particles, *Phys. Rev.* **159**, pp. 1089–1105.
66. Ye, F., Mihalache, D., Hu, B. and Panoiu, N. C. (2011). Subwavelength vortical plasmonic lattice solitons *Opt. Lett.* **36**, p. 1179.
67. Silveirinha, M. G. (2013). Theory of spatial optical solitons in metallic nanowire materials, *Phys. Rev. B* **87**, p. 235115.
68. Smolyaninov, I. I., Smolyaninova, V. N. and Smolyaninov, A. I. Experimental model of topological defects in Minkowski spacetime based on disordered ferrofluid: Magnetic monopoles, cosmic strings and the spacetime cloak, *Phil. Trans. of Royal Soc. A*.

CHAPTER 4

Circuit and Analytical Modelling of Extraordinary Transmission Metamaterials

FRANCISCO MEDINA*, **FRANCISCO MESA, RAÚL RODRÍGUEZ-BERRAL and CARLOS MOLERO**

University of Sevilla, Spain

Extraordinary transmission of electromagnetic waves through periodic arrays of electrically small apertures has attracted the attention of many researchers since the discovery of such phenomenon at the end of the 1990s. The explanation of the existence of such frequency-selective enhanced transmission behavior has been linked to the interaction of the impinging uniform plane wave with the so-called (spoof) surface plasmon-polaritons supported by the periodically structured surface. However, an alternative model has also been proposed in the literature which is based on the consideration of the unit cell of the periodic system as a waveguide discontinuity problem. This problem is very common in the classical microwave-engineering literature and is usually treated under the standpoint of equivalent-circuit modelling. Circuit models have a number of obvious advantages related to their analytical or quasi-analytical nature. Thus, the computational effort is negligible

*Corresponding author: medina@us.es

when compared with standard numerical approaches. But, more importantly, these models provide a simplified conceptual frame that makes the understanding of the underlying physical phenomena easier. This chapter provides a review of circuit–like models available in the literature to deal with the transmission of electromagnetic waves through periodically structured screens.

4.1. Introduction

Extraordinary optical transmission (EOT) through a (2D) periodic distribution of subwavelength cylindrical holes made in an optically thick metal slab was first reported by T. W. Ebbesen and co-workers in 1998.[1] Unexpected significant transmission peaks at wavelengths close to the period of the structure were observed, in spite of the electrically small size of the apertures. The amount of transmitted power at those specific frequencies was much larger (at least two orders of magnitude larger) than the one expected from Bethe's small apertures theory.[2] In the first stage, the plasma–like behavior of metals in the optical range was considered to be of primordial importance to the physics of the phenomenon. Although nothing similar had been reported for similar structures in the microwave/millimeter wave regimes (the so-called bandpass frequency selective surfaces[3]), some important experiments[4,5] revealed that a similar extraordinary-transmission behavior was reproduced in the microwave and millimeter-wave frequency ranges when working with thick perforated aluminum slabs. This fact made it apparent that the existence of the EOT peaks is mostly related to the periodic nature of the structure. The plasma–like behavior of metals at optical frequencies introduces some differences with respect to the behavior of metals at lower frequencies, but it is not essential to the explanation of the phenomenon. Actually, metals behave as quasi-perfect conductors at microwave and millimeter-wave frequencies, in such a way that smooth metal surfaces cannot support surface plasmon-polaritons (SPPs), in contrast with the physical situation arising at metal–dielectric interfaces excited by electromagnetic fields in the optical frequency range. Thus, it is more appropriate to use

the acronym ET (Extraordinary Transmission), avoiding the specific term "optical", in what follows.

The ET through perforated or slotted metal screens has been studied in depth by many researchers during the last 17 years, and a well-established theory based on the interaction of the impinging uniform plane wave with the so-called "spoof" (or "designer") surface SPPs[6,7] supported by periodically structured metal surfaces is nowadays the most widely accepted paradigm. The "spoof" (or "designer") plasmons are surface leaky waves supported by a periodically structured metallic surface exhibiting a dispersion curve similar to the one characterizing genuine SPPs propagating along flat or smoothly curved metal–dielectric interfaces in the optical range. The equivalent plasma frequency in the case of periodic structures is controlled by the period rather than by the electrical parameters of the metal (volume density of electrons). Due to this reason, ET can be induced at almost any frequency region of the electromagnetic spectrum as long as the metal can be described as a continuous medium. A number of comprehensive reviews providing a detailed description of this theory can be found in the scientific literature,[8–11] and the role of this kind of SPPs in ET through periodically structured opaque surfaces is nowadays well established. However, it is important to clarify that obtaining accurate dispersion curves for the SPPs is not a trivial task since the computational effort required to generate those dispersion curves is greater than the one involved in the straightforward computation of the transmission and reflection coefficients. Moreover, the SPP model itself does not provide quantitative values of the scattering coefficients at arbitrary frequency points, but only the location of the critical high-transmission frequency points. Fortunately, an alternative point of view can be adopted to interpret the results of the ET experiments. The main concept behind that alternative approach is to consider that the unit cell of a periodic structure illuminated by a uniform plane wave can be seen as a virtual waveguide system with metallic diaphragms inside. Indeed, it is a problem of this kind that is usually posed and solved when a numerical approach is employed to obtain

the scattering parameters of the periodic structure.[12–14] Interestingly, since early times,[15] very efficient methods have been employed by microwave engineers for the characterization of such kind of physical systems (waveguide discontinuity problems). The application of such methods typically leads to equivalent circuits whose electrical parameters can be analytically, numerically or experimentally determined. Once such parameters are known, the original electromagnetic 3D and vectorial problem is reduced to a much simpler electric circuit involving lumped and distributed (transmission lines) elements. The application of this kind of electrical models to the analysis of periodic structures, far from being a new topic, is a rather classical topic in the microwaves and antennas literature,[16–18] and it is also a standard technique in the modelling of frequency-selective surfaces (FSSs) and related structures.[3,19–25] The approximate calculation of the circuit parameters is a lot quicker than the numerical characterization of the structure over a desired frequency band. This calculation can be done without sacrificing accuracy if the topology of the circuit model is conveniently derived. Apart from the obvious computational advantages, the circuit-model approach sheds light on the physics of the problem and makes it easy to understand the periodic structure behavior. However, although circuit models were well-known tools for the description of the behavior of periodic structures, they were not used to deal with ET phenomena until several years after its discovery. Following traditional circuit-modelling guidelines, after introducing some appropriate modifications to account for dynamic effects necessary to include ET in the formulation, three circuit models for 2D rectangular arrays of small holes were reported in Ref. 26 for the first time (the circuit models correspond to thin, moderately thick and arbitrarily thick metal slabs, respectively). In Ref. 26, ET is conceived as an impedance matching problem, without explicit mentioning of SPPs. The derivation of that circuit model is described in detail in this chapter, and the implications of the model are discussed.

It should be mentioned that ET has not only been observed/studied in 2D arrays of holes but also, for instance, in simpler structures based on one-dimensional (1D) arrays of slits in

metal slabs (bar gratings). Thus, one of the first detailed analytical descriptions of the ET phenomenon was based on the in-depth study of the transmissivity of electromagnetic waves through 1D arrays of infinitely long slits made in a thick metal slab, which was laid on an infinitely thick supporting dielectric substrate.[27] To give proper credit to pioneer works on this topic, it is worth mentioning that Ukrainian researchers published some papers at the end of the 1960s about the scattering properties of bar gratings[28,29] and, in those papers, ET peaks can clearly be visualized. In the context of this chapter, it can again be said that a simple explanation of the origin of this ET peak in 1D metal gratings is given in Ref. 30, where an extremely simple circuit model is proposed. The model also accounts for more conventional Fabry–Pérot (FP) resonances and explains why free-standing zero-thickness metal strips cannot support ET peaks.[27] An interesting situation arises when the unit cell of the grating consists of several slits per period (the so-called *compound gratings*). These structures were studied in depth about 10 years ago,[31] although relevant results on the behavior of such kinds of gratings can be found in early former Soviet Union scientific literature.[32] The compound grating structures exhibit interesting complex transmission and reflection spectra due to the interactions between the closely spaced slits forming a group. Following Ref. 31, phase resonances accompanied by narrowband transmission dips located in the middle of wider bandwidth (FP) resonances can be observed. The previously cited paper[30] includes a quasi-static circuit model for that kind of compound gratings. This model will be discussed in this chapter, and its extension to account for dynamical effects will be sketched following the guidelines reported in Ref. 33. The inclusion of metal losses and its influence will also be explained in the frame of the circuit model, as it was done in a previous paper by some of the authors of this chapter.[34]

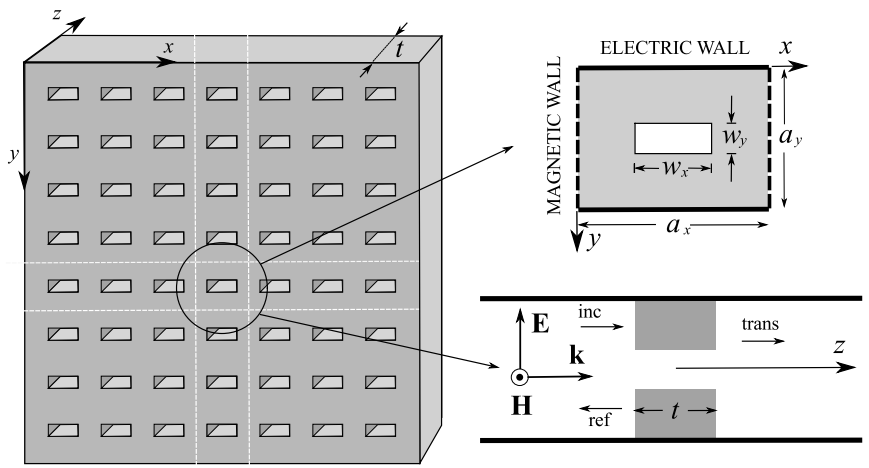
The theory describing ET phenomena in this chapter actually deals with a relatively exotic behavior of diaphragm–like discontinuities in closed virtual waveguides. The virtual waveguides are defined by the periodicity of the problem under study. However,

there is no reason for not obtaining a similar ET behavior through electrically small diaphragms located inside real closed waveguides. This situation will also be discussed in this chapter, and some relevant conclusions will be extracted from this study.

Note that the structures mentioned in the previous paragraphs only involve metallic structures and a single dielectric medium around the metal regions. The presence of dielectric layers, however, is expected to give rise to interesting additional phenomena, since the dielectric slabs provide new paths to the excitation of surface waves, which seem to be closely related to the appearance of ET peaks. The inclusion of these dielectric slabs in the frame of the equivalent–circuit modelling methodology will also be discussed at the end of this chapter.

4.2. Modelling of 2D Arrays of Holes in Metal Plates

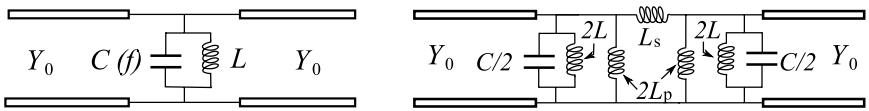
The most paradigmatic structure exhibiting ET is, possibly, the 2D array of electrically small apertures made in an opaque metal film or slab shown in Fig. 4.1(a). The geometry of the holes has been considered to be rectangular in order to facilitate the calculation of the transmission and reflection coefficients using a standard implementation of the mode-matching approach. The structure is excited by an impinging uniform transverse electromagnetic (TEM) wave that propagates along the z direction (normal incidence is considered to simplify the discussion). The solution of the periodic problem reduces to the analysis of a single unit cell, as shown in Fig. 4.1(b) (front and side views). Note that the unit cell problem can be seen as a discontinuity problem caused by the presence of a thick diaphragm placed inside an ideal parallel-plate waveguide (PPW, a closed waveguide whose boundaries are two horizontal electric walls and two vertical magnetic walls; the fundamental mode is of TEM type[35]). The role of the impinging TEM plane wave in the original periodic problem is now played by the incident TEM mode in the PPW. The presence of the diaphragm gives place to the excitation of an infinite number of high-order transverse electric (TE)



(a) (b)

Fig. 4.1. (a) Metal slab (considered as a perfect electric conductor, PEC, in this chapter) of thickness t with a periodic distribution (period a_x and a_y along the x and y directions) of rectangular holes (width w_x and height w_y). (b) Unit cell problem for normal incidence of TEM waves (the top panel is a front view and the bottom panel is a longitudinal view).

and transverse magnetic (TM) modes, which are necessary to satisfy the boundary conditions for the electric and magnetic fields at the discontinuity boundaries (plane surfaces separating the PPW region and the rectangular hole region). If the operation frequency is below the first Rayleigh–Wood (RW) frequency ($f_{\mathrm{RW}} = c/a_y$), which defines the onset of the grating lobes (i.e. the onset of the diffraction regime), all the higher-order modes will operate below cutoff. As is well known,[35,36] the evanescent TE modes store magnetic energy whereas the evanescent TM modes store electrical energy. If one is interested in the calculation of the reflection (S_{11}) and transmission (S_{21}) coefficients for the fundamental TEM mode, the effect of all the high-order modes can be combined in properly defined lumped circuit elements:[15] inductances for TE modes and capacitances for TM modes. Two different situations, depending on the thickness of the metal plate where the holes have been perforated, will be studied next.



(a) (b)

Fig. 4.2. (a) Equivalent circuit for the zero-thickness ($t = 0\,\text{mm}$) screen (PEC). (b) Equivalent circuit including the inductive effects of the TE field inside the hole (electrically thin PEC screens). Figures reprinted from Ref. 26 with permission of IEEE.

4.2.1. *2D arrays of holes in zero-thickness screens*

First, the most basic situation exhibiting ET will be analyzed: a 2D array of electrically small apertures made in a zero-thickness PEC screen ($t = 0$ in Fig. 4.1). In this case, we have the standard problem of a rectangular diaphragm of negligible thickness located inside a PPW. This is a classical problem in waveguide theory,[15,35,36] for which a simple equivalent circuit is available and easily understandable, as shown in Fig. 4.2(a). The uniform (along z) PPW regions are modelled with transmission lines whose characteristic admittances are given by $Y_0 = (a_x/a_y)/\eta_0$ ($\eta_0 \approx 120\pi\,\Omega$ is the free-space impedance). The diaphragm can be represented by a LC-tank circuit connected with the transmission lines (see Fig. 4.2(a)). As previously mentioned, L accounts for the effect of all the scattered evanescent TE modes, whereas C catches the effect of the evanescent TM modes.

This simple circuit will be shown to qualitatively and quantitatively account for the most relevant features of the electrical response of the discontinuity. Thus, at resonance, the LC-tank is an open circuit, in such a way that total transmission is expected from the circuit model at such frequency. More precisely, the circuit predicts a passband behavior around that resonance frequency. For electrically large enough apertures ($w_x > a_x/2$), the system behaves as a standard passband FSS and nothing new is actually contributed with respect to the well-established circuit modelling of this class of systems.[15] In that case, the values of L and C are independent (or almost independent) of the operation frequency in the range of

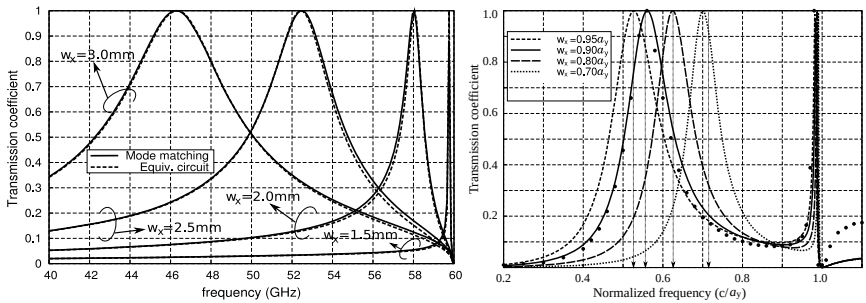
frequencies of interest, which used to be significantly below the onset of the diffraction regime. If this rationale is now applied to electrically small apertures ($w_x < a_x/2$), keeping the values of L and C frequency independent (these values could have been obtained either from analytical expression or extracted from full-wave simulations carried out at relatively low frequencies), the expected resonance frequency would be above f_{RW}. Thus, for small enough apertures, no resonant transmission should appear in the non-diffraction regime. However, it is known that an array of this kind exhibits a total transmission peak at some frequency slightly below f_{RW}, followed by a transmission zero at exactly f_{RW}[9] (this is called a Fano-like resonance). This reported behavior can easily be understood using the equivalent circuit in Fig. 4.2(a) provided the frequency dependence of C associated with the excitation of the first TM mode of the structure is explicitly taken into account (TM_{02} in this case, since odd-order modes are not compatible with the chosen excitation and symmetries). Considering the dependence with the frequency of the modal admittances of the TM modes[35] and assuming that the relative level of excitation of the scattered modes is frequency-independent (this assumption is equivalent to consider that the field profile in the aperture is mostly frequency-independent, which is found to be a quite good approximation), the capacitance in Fig. 4.2(a) can be written as follows[26]:

$$C(\omega) = \frac{A_{\mathrm{TM}_{02}}}{\eta_0 \sqrt{\omega_{\mathrm{TM}_{02}}^2 - \omega^2}} + C_{\mathrm{ho}}, \tag{4.1}$$

where $\omega_{\mathrm{TM}_{02}}$ is the cutoff angular frequency of the TM_{02} mode and C_{ho} is the capacitance associated with the remaining higher-order (ho) TM modes. The capacitance C_{ho} is considered to be frequency-independent because the operation frequency in the range of interest (non-diffraction regime) is much smaller than the cutoff frequencies of the involved TM modes. The dimensionless constant $A_{\mathrm{TM}_{02}}$ depends on the relative level of excitation of the TM_{02} mode which, in turn, depends on the profile of the electric field in the aperture (i.e. depends on its shape and size).

The interesting feature of (4.1) lies on the singular behavior of the TM_{02} contribution to the capacitance at $\omega_{TM_{02}}$, which is usually ignored in the standard circuit modelling of FSSs. Since this capacitance reaches very high values in the close proximity of $\omega_{TM_{02}}$, the LC-tank circuit in Fig. 4.2(a) always behaves as an open circuit at some frequency slightly below $\omega_{TM_{02}}$, independently of how small the inductance L is (i.e. independently of how small the aperture is). Moreover, at exactly $\omega_{TM_{02}}$, $C(\omega)$ is singular and the LC-tank circuit now behaves as a short-circuit, thus reflecting all the impinging power and easily explaining the transmission zero observed at f_{RW}. In brief, the model predicts the Fano-like resonance reported by many authors[9,10] for this kind of structures in an extremely simple way.

Although the values of $A_{TM_{02}}$ and C_{ho} can be approximately computed by assuming a physically acceptable field profile in the aperture (the field profile of the TE_{10} mode of the rectangular waveguide having the dimensions of the aperture would be a good choice, as it has recently been discussed[37]), those values were extracted in Ref. 26 from a mode-matching full-wave simulation of the structure at a few frequency points. In Fig. 4.3(a) a few examples of the circuit model and mode-matching predictions for several rectangular apertures having different sizes are shown. Circuit model and full-wave data are almost indistinguishable. When the considered apertures are electrically large ($w_x = 3.0\,\text{mm}$ or $w_x = 2.5\,\text{mm}$), a typical FSS relatively broadband transmission peak is observed. The location of the transmission peak is mainly controlled by the size of the aperture along the x direction. If the width of the slot is reduced to what can be denominated "subwavelength" size, the transmission peak is not moved to frequencies above the RW anomaly point. A relatively narrow band transmission peak is always obtained below f_{RW} and close to this frequency. The bandwidth of this transmission peak is drastically reduced as the slot width is decreased and, in theory (if losses are totally neglected), there is always a total transmission peak below f_{RW} no matter how small the aperture is. This is the ET peak, which appears even if the size of the aperture is small enough to preclude the existence of the regular FSS resonance within the non-diffraction



(a) (b)

Fig. 4.3. (a) Magnitude of the transmission coefficient ($|S_{21}|$) through a zero-thickness PEC screen with a periodic 2D distribution of rectangular apertures of different widths ($w_x = 3.0, 2.5, 2.0, 1.5\,\text{mm}$, $w_y = 0.5\,\text{mm}$). A square lattice of period $a_x = a_y = 5.0\,\text{mm}$ is considered (this corresponds to $f_{\text{RW}} = 59.9585\,\text{GHz}$). The cases $w_x = 3.0$ and $2.5\,\text{mm}$ correspond to regular FSS operation. The cases $w_x = 2.0$ and $1.5\,\text{mm}$ can be considered examples of extraordinary transmission. (b) Transmissivity ($|S_{12}|^2$) through a perforated PEC finite-thickness plate for several widths of the rectangular holes ($w_x = 0.70, 0.80, 0.90, 0.95\,a_y$, $w_y = 0.2\,a_y$). FDTD data[38] (black circles) have been included for comparison purposes. Dimensions: $a_x = a_y$ and $t = 0.2\,a_y$. Figures reproduced from Ref. 26 with IEEE permission.

frequency span. The behavior in the proximity of the transmission zero associated with the RW anomaly is correctly predicted thanks to the singular behavior of the TM capacitance used in the proposed modelling technique. Another interesting feature of these Fano-like resonances (reported by many authors[9,10]) is the existence of a strong enhancement of the electromagnetic field at resonance. This observation is consistent with the fact of the existence of high current levels flowing through the LC-tank circuit in Fig. 4.2(a) when L is relatively small (due to the small size of the represented aperture). Field enhancement is then also qualitatively accounted for by the proposed model.

4.2.2. *The effect of metal thickness in 2D arrays of holes*

Most of the structures considered in the available literature about ET makes use of relatively thick metal layers, with this thickness having

relevant consequences on the transmission spectrum.[4,9,38–40] One of the most relevant facts derived from the non-zero metal thickness case is the existence of two ET peaks (instead of only one). Material losses might mask the existence of one of those two peaks in an experimental verification, but the numerical analysis of the problem clearly demonstrates the existence of a pair of peaks when losses are neglected or are very low. This situation cannot be accounted for by the circuit in Fig. 4.2(a). The reason is that the existence of a second transmission peak is closely related to the presence of magnetic energy stored inside the hole when the metal slab thickness is relevant. Note that the interior of the hole can be seen as a section of length t of a rectangular waveguide. This waveguide section would operate below cutoff if ET conditions are assumed (small apertures), and it would work above cutoff over a certain region of the range of frequencies below the diffraction regime if the aperture is large enough. In the case of ET problems, due to the small electrical size of the aperture, a single evanescent mode can be considered enough to account for the total field inside the hole (the TE_{10} mode in our case, since the aperture has a rectangular shape). From the circuit theory point of view, as long as the value of t is small enough, a Π-circuit made of inductors can be used to account for the effect of the non-vanishing thickness of the metal slab.[26] The equivalent circuit shown in Fig. 4.2(b) is then proposed to include this effect.[26] The values of the parallel (L_p) and series (L_s) inductances forming the Π inductive circuit can be extracted from the analysis of the circuit under even and odd excitation conditions, as it is explained in Ref. 26. The values of L and C corresponding to the external problem (scattered modes in the input and output virtual PPWs) are very similar to the ones computed for the zero thickness case. The analysis of the response of the circuit in Fig. 4.2(b) predicts the existence of two total transmission (ET) frequencies at

$$\omega_e = \frac{1}{\sqrt{C(\omega_e)L_{\text{eq}}^e}}; \quad \omega_o = \frac{1}{\sqrt{C(\omega_o)L_{\text{eq}}^o}}, \tag{4.2}$$

where

$$L_{\text{eq}}^{e} = \frac{LL_p}{L + L_p}; \quad L_{\text{eq}}^{o} = \frac{2LL_pL_s}{4LL_p + L_s(L + L_p)}. \tag{4.3}$$

The superscripts "e" and "o" in the previous expressions stand for even and odd excitations; i.e. L_{eq}^{e} and L_{eq}^{o} are the equivalent inductances connected in parallel with the capacitors, C, in the even- and odd- circuits derived from the circuit in Fig. 4.2(b) under even and odd excitation conditions. The values of those inductances can easily be extracted from the phase of the reflection coefficients under even and odd excitations and can be considered to be frequency-independent. Note that Eqs. (4.2) are a set of implicit equations providing the ET peaks from the values of the capacitance in (4.1) and the inductances in (4.3).

It should be noted that the previous analysis has to be slightly modified if the TE_{10} mode in the hole is working above its cutoff frequency (i.e. under regular FSS operation instead of ET operation). In that case, the contribution of a short section of waveguide of length $t/2$ to the odd excitation parameters is still of inductive nature. However, the contribution of the propagating TE_{10} mode to the even excitation equivalent circuit becomes of capacitive nature. Fortunately, this circumstance can easily be incorporated to the proposed model. For the even excitation case, the inductive contribution to the circuit model comes only from the external TE modes (the TE evanescent modes excited in the input and output PPWs). Actually, the resonance associated with the even excitation is the one commonly found in conventional FSS operation. Differently, the resonance associated with the odd excitation only appears provided the thickness of the screen is not negligible, since the involved inductance is the one corresponding to a short section of waveguide terminated with a short circuit (this is the virtual short circuit, at the level of the vertical symmetry plane of the unit cell, induced by the anti-symmetrical nature of the excitation). This inductance is typically much smaller than the external inductance, L, and shunt-connected

to it, thus prevailing in the determination of the resonance properties. In spite of the relatively low value of that inductance, a resonance peak can always be observed slightly below f_{RW} because of the singular behavior of $C(\omega)$ in (4.1).

An example of the influence of the metal plate thickness on the transmission spectrum is shown in Fig. 4.3(b). The transmission spectra corresponding to several widths of a rectangular slot are obtained using an in-house mode-matching code and, for one of the cases, the FDTD results reported in the literature[38] are included for comparison purposes. In this figure, the frequency is normalized to f_{RW}. The circuit model results do not explicitly appear since they overlap the mode matching data. It can be observed that the considered rectangular slots, which are electrically large, exhibit a first resonance of relatively large bandwidth whose location is mainly controlled by the length of the slot (i.e. the value of w_x). An estimation of that resonance frequency based on the length of the slots is included in the figure in the form of vertical arrows. This resonance also appears if the thickness of the metal slab is considered to be vanishing, and it is the well-known conventional FSS resonance.[3] In Ref. 38 and elsewhere, this transmission peak is considered to be a "localized resonance". Interestingly, a second narrow resonance peak occurring at a frequency slightly below f_{RW} appears for all the considered aperture sizes. The existence of this resonance (which does not appear in Fig. 4.3(a), for instance) is associated with the nonvanishing thickness of the screen. This resonance is not included in conventional circuit models for FSS structures available in the literature (probably because, in the microwave range, the thickness of the metalizations used in printed circuit technology is negligible). This resonance peak is a true ET phenomenon, and it has been predicted by studying the electrical response of the circuit in Fig. 4.2(a), whose electrical parameters can easily be extracted from a few values of the scattering parameters obtained with any numerical simulator or from an approximation of the electric field over the hole aperture.

4.2.3. *Distributed circuit model for arbitrary thickness screens*

In spite of the good results provided by the circuit model in Fig. 4.2(b) when used to account for the effects of the metal screen thickness, the model has some flaws that deserve attention. First, the model cannot easily account for the physical fact that thick enough screens should completely block transmission if the waveguide formed by the hole is operating below cutoff. Moreover, the model cannot explain the dependence of L_p and L_s with the thickness t of the screen (these inductances are obtained in Ref. 26 for several values of t for a given aperture shape and size, and their dependence on t is discussed). These drawbacks come from the use of a lumped-element equivalent circuit to model the effects of the fields inside the hole. The problems can be solved if a distributed circuit model is considered instead.[26] A distributed model of this type is shown in Fig. 4.4(a). This model takes into account that the thickness of the screen can be of the same order of magnitude as the characteristic distances associated with the variation of the fields along the longitudinal direction (z) inside the hole. For small rectangular holes operating below cutoff (this is the case of interest for ET phenomena), the characteristic distance is governed by the attenuation factor of the fundamental TE_{10} mode inside the hole ($\alpha_{\mathrm{TE}_{10}}$). This evanescent mode is also characterized by a purely imaginary characteristic impedance ($jZ_{\mathrm{TE}_{10}}$). The transform ratio n of the transformers in the model accounts for the coupling between the external TEM and the internal TE modes. This transform ratio can be extracted from full-wave simulations (as done in Ref. 26) or computed from an estimated field profile over the aperture (the one corresponding to the TE_{10} mode in the hole would be a reasonable choice [37]). The values of C and L are, in practice, the same ones that were derived for the zero-thickness screens since, once the field in the aperture is fixed, the internal problem is decoupled from the external one. The circuit in Fig. 4.4(a) can be conveniently solved in terms of the even- and odd- excitation circuits in Fig. 4.4(b) by using the superposition principle. Note that the dependence on

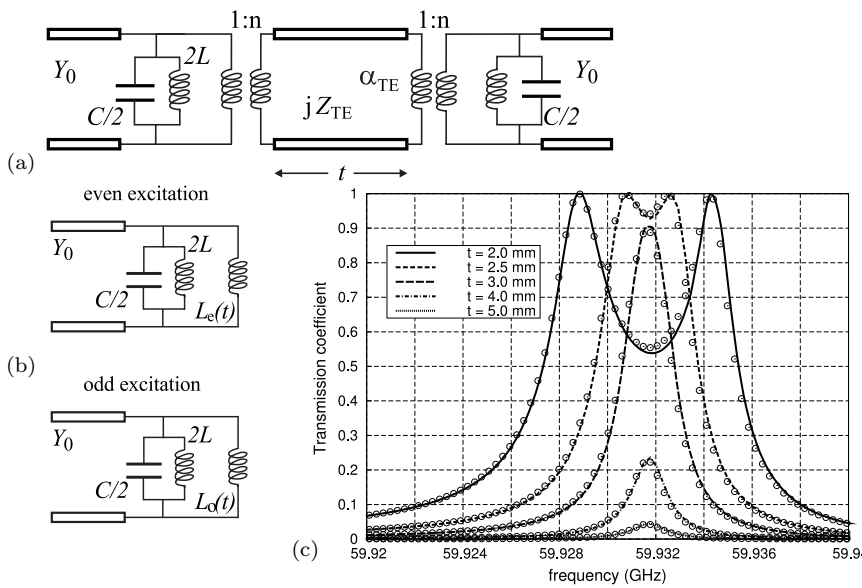


Fig. 4.4. (a) Equivalent circuit for arbitrarily thick screens with subwavelength holes. The circuit accounts for distributed effects inside the hole. (b) Even- and odd-excitation equivalent circuits derived from (a). (c) Influence of the progressive increasing of screen thickness on the evolution of the two-peaks transmission coefficient ($|S_{21}|$). Solid lines: mode-matching numerical results. Circles: circuit model predictions. Dimensions: $a_x = a_y = 5.0\,\text{mm}$; $w_x = 1.5\,\text{mm}$, $w_y = 0.5\,\text{mm}$. Figures reproduced from Ref. 26 with permission of IEEE.

t of the even and odd inductances of the circuits in Fig. 4.4(b) is known in closed form:[26]

$$L_e = \frac{Z_{\text{TE}_{10}}}{2\pi n f_{\text{RW}}} \coth\left(\alpha_{\text{TE}_{10}} \frac{t}{2}\right); \quad L_o = \frac{Z_{\text{TE}_{10}}}{2\pi n f_{\text{RW}}} \tanh\left(\alpha_{\text{TE}_{10}} \frac{t}{2}\right), \tag{4.4}$$

in such a way that the analysis is not to be repeated when the thickness is changed. The formulas (4.4) are consistent with the values of L_p and L_s obtained for different values of the parameter t using a full-wave formulation and the circuit model in Fig. 4.2(b), as it is demonstrated in Ref. 26.

An important advantage of the distributed circuit model (Fig. 4.4(a)) is that such a model can explain what happens when the two peaks of ET collapse due to increasing metal slab thickness. This process cannot be incorporated to the lumped-element circuit model in Fig. 4.2(b). This feature of the distributed model is clearly illustrated by the results reported in Fig. 4.4(c). In that figure, the thickness t of the screen is progressively increased until the two ET peaks collapse into a single peak. The data have been obtained in the absence of losses, in such a way that the double peaks reach a transmission coefficient whose magnitude is the unity. However, the transmission for the single peak is partial; i.e. part of the power is reflected (since no losses have been included in the model) and decreases to zero as expected when the thickness of the screen becomes significantly larger than $1/\alpha_{\mathrm{TE}_{10}}$. It is worth emphasizing that the circuit model data perfectly match the numerical results obtained with the mode-matching approach. Moreover, the values of the circuit parameters have been determined only once for all the values of the metal plate thickness.

So far, it has been established that standard waveguide discontinuity theory can be used to explain the phenomenon of ET in periodically perforated PEC screens (2D case). The employed methodology automatically leads to an equivalent circuit model which is characterized by using very few parameters. The correct incorporation of the frequency dependence of one of those parameters (the capacitance associated with the first excited high-order TM mode) is essential to account for the ET phenomenon using the circuit model. In particular, it has been shown that the main role in the phenomenon is played by the first TM mode when the operation frequency is slightly below and close to its cutoff. Note that this model puts emphasis, when providing an explanation for the ET phenomenon, on the possibility of having a *perfect matching condition* rather than on the excitation of spoof surface plasmons. In the authors' opinion, this point of view is advantageous because the whole transmission (or reflection) spectrum is predicted, not only the location of the ET and RW frequencies.

4.3. Extraordinary Transmission in Hollow-Pipe Metallic Waveguides

In the preceding section, the existence of ET peaks through 2D periodic arrays of subwavelength apertures made in thin or thick metal plates has been related to operation frequencies in close proximity to the cutoff frequency of an evanescent TM mode; namely, the first high-order TM mode supported by the virtual PPW defined by the unit cell of the periodic problem (for normal incidence, it was the TM_{02} mode). Note that, in the frame of the proposed approach, the role of the periodic nature of the structure is to generate such virtual waveguide, whose high-order TM modes would be responsible for extraordinary resonance peaks (even within the diffraction regime, small peaks are observed in the immediate vicinity of the onset of new TM modes[26]). Certainly, this point of view immediately suggests that ET should also be expected through electrically small diaphragms located inside a regular hollow-pipe metallic waveguide. In such situation, it is obvious that periodicity could not be invoked as the reason behind ET. Of course, SPPs are not defined for this physical system so that its presence cannot be used to explain ET peaks (as commonly done in the literature for the case of periodic systems). In the frame of the waveguide problem with diaphragms made of electrically small apertures, eigenmodes of the class discussed in some papers[41–43] would play the role reserved to SPPs by the standard ET theories (periodic case). Apart from those papers, the possibility of observing the ET phenomenon through small apertures located inside closed waveguides has also been reported in other works,[26,44] which show, through numerical simulations or analytical developments, that a phenomenon equivalent to ET in periodic structures can be observed in metallic waveguide systems. In this section, the ET phenomenon will be theoretically and experimentally studied for the case of off-centered small diaphragms placed inside a circular waveguide system.[45,46]

There are two advantages of using circular waveguides (instead of rectangular waveguides[44]) and off-centered diaphragms: (i) the possibility of ruling out the periodicity as a necessary condition for ET,

and (ii) the chance of creating an electromagnetic modal situation which is analogous to the one found when dealing with 2D periodic distributions of holes. It should be noticed that the first TM mode that can be excited in a *rectangular* waveguide has its cutoff frequency above the cutoff frequencies of two TE modes, which would be launched before reaching the onset of the first TM mode. This means that single-mode operation would not be possible when trying to observe ET peaks, in contrast with what happens in the periodic system problem, where a single TEM mode is propagating inside the virtual PPW system at the ET frequency point. Multimode operation does not only invalidate the analogy between the closed waveguide and the previous periodic structure, but also would be a problem from an experimental point of view. Fortunately, the first TM mode supported by a circular waveguide (the TM_{01} mode) is preceded by a single TE mode (the fundamental TE_{11} mode). Thus, before the onset of the TM_{01} mode, single mode TE_{11} operation is possible. In that case, we can use exactly the same reasoning line previously employed for the analysis of 2D periodic distributions of holes. Indeed, the same circuit models can also be employed after minor modifications. It is only necessary to replace the nondispersive TEM mode supported by the virtual PPWs in Fig. 4.1(b) with the frequency-dispersive fundamental TE_{11} mode supported by the circular waveguide; the characteristic admittance, $Z_{TE_{11}}(\omega)$, and propagation constant, $\beta_{TE_{11}}(\omega)$, should be used in the circuit models instead of $Z_0 = Y_0^{-1}$ and k_0. Note that, due to the symmetry properties of the TM_{01} mode, a diaphragm with a centered symmetrical slot would not excite this mode when illuminated with the impinging TE_{11} mode. This problem is easily overcome by just offcentering the location of the aperture. The geometry of the small apertures (diaphragms) used in the experimental setups in Refs. 45 and 46) are depicted in Fig. 4.5(a). Two different geometries were used in Ref. 46: rectangular and elliptical. In this chapter, only the rectangular geometry will be considered.

The scattering parameters of the discontinuity represented in Fig. 4.5(b) can be computed by using the mode matching method

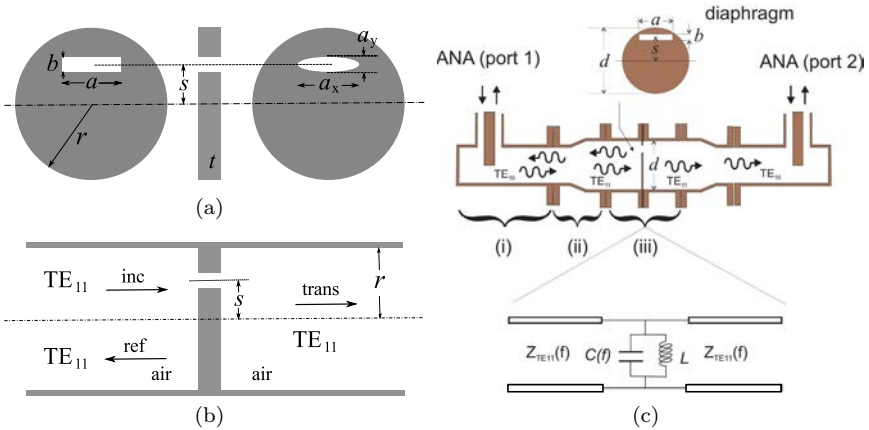
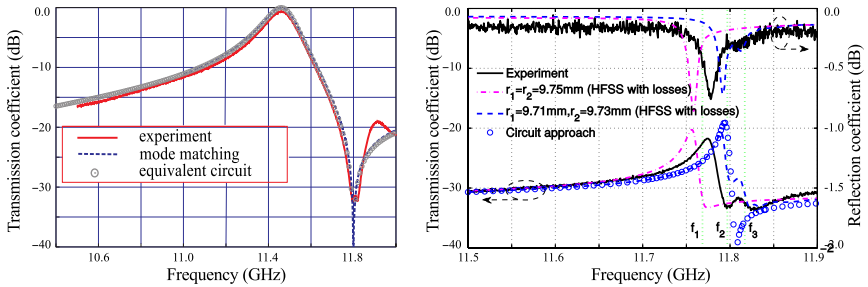


Fig. 4.5. Schematic view of the subwavelength diaphragms used in the experimental study of ET in circular waveguides reported in Refs. 45 and 46: (a) front and lateral views of rectangular and elliptical diaphragms (the diaphragms are off-centered). Rectangular diaphragm dimensions: $a = 10.0\,\text{mm}$, $b = 2.0\,\text{mm}$, $s = 7.0\,\text{mm}$ diaphragm dimensions: $a_x = 13.0\,\text{mm}$, $a_y = 3.0\,\text{mm}$, $s = 6.0\,\text{mm}$. (the elliptical case will not be discussed here). The nominal waveguide radius is $r = 9.75\,\text{mm}$. (b) Schematics of the involved discontinuity problem. (c) Schematics of the experimental setup, including coax-to-waveguide adapters (i), rectangular-to-circular waveguide transitions (ii) and the diaphragm section (iii). A basic equivalent circuit for zero-thickness diaphragms is also shown. Reprinted from Refs. 45 and 46 with permission of AIP and IEEE.

following the guidelines reported in Refs. 47 and 48. To simplify the discussion, it will be assumed that the thickness of the metal plate supporting the electrically small aperture is negligible. The equivalent circuit that models the effect of the diaphragm is then the one sketched in Fig. 4.2(a) or at the bottom of Fig. 4.5(c); i.e. higher-order mode contributions are accounted for by a frequency-independent inductance, L, and a frequency-dependent capacitance, C. The values of L and the two frequency-independent parameters defining $C(\omega)$ (see Eq. (4.1)) can be obtained once again from a few values of the numerically determined scattering parameters.[46] If the thickness of the metal plate where the diaphragm has been realized is not negligible, the circuit model in Fig. 4.2(b) should be used instead. The internal inductances, L_s and L_p, can be extracted

from full-wave simulations carried out at few frequency points (a detailed study of the consequences of using thick diaphragms can be found in Ref. 46). Since the capacitance accounting for the TM_{01} contribution to the scattered field has a singular behavior at the cut-off frequency of this mode (11.77 GHz for the nominal value of the circular waveguide radius considered in this example, $r = 9.75\,\text{mm}$), an extraordinary transmission peak at a frequency slightly below the TM_{01} cutoff frequency is expected. In Fig. 4.6(a), the results measured for the rectangular diaphragm in Fig. 4.5(a) are plotted and the full-wave mode-matching data are included for comparison purposes, as well as data obtained with the proposed circuit model (the thickness of the diaphragm is neglected). A typical Fano–like resonance is observed, with a maximum transmission peak at about 11.42 GHz followed by a transmission zero at the frequency of the onset of the first TM mode of the circular waveguide (around 11.8 GHz). According to these results, it is possible to conclude that the theory presented in this chapter provides a unified conceptual frame to explain ET occurring in periodic structures and in hollow-pipe metallic waveguides. This is a direct consequence of using the impedance matching point of view rather than relying on the excitation of SPPs.

An interesting experimental observation with relevant implications is shown in Fig. 4.6(b). In that figure, the same experiment is carried out but using a smaller aperture size ($a = 5.7\,\text{mm}$ instead of $10.0\,\text{mm}$). The experimental data show that a transmission peak still appears in the transmission coefficient, which is associated with a small dip in the reflection coefficient. The transmission peak (and reflection dip) are very close to the TM_{01} cutoff frequency. However, although the peak exists, most of the impinging power is reflected or absorbed in the diaphragm plate and transmission enhancement is residual. It seems that a high level of transmission at frequencies very close to the onset of the TM mode (this is the situation when the size of the aperture is really small) is not compatible with the presence of metal losses. This experimental fact can be explained in the frame of the equivalent circuit model introduced in this section after noting that the presence of metal losses can be phenomenologically



(a) (c)

Fig. 4.6. (a) Extraordinary transmission for the rectangular diaphragm considered in Fig. 4.5(a). Analytical (circuit model), numerical (mode-matching) and experimental results agree very well (results reproduced from Ref. 45). (b) Measurements, full-wave simulations with HFSS (this is a finite-elements-based commercial solver) and circuit model results for a very small aperture, $a = 5.7$ mm (results reprinted from Ref. 46). Losses have been added to the circuit model by means of a small resistor, R, connected in series with the inductance, L, of the aperture. Reprinted from Refs. 45 and 46 with permission of AIP and IEEE.

included in the model by adding a resistor, R, connected in series with the inductance, L, of the aperture. If this resistor is considered, the impedance of the LC-tank circuit at resonance is not infinity, but approximately $Z_{\text{res}} = L/(RC)$. A good transmission level at resonance is only achieved if $Z_{\text{res}} \gg Z_{\text{TE}_{11}}$. For typical values of R associated with metallic diaphragms and the values of L and C associated with apertures not too small, the above-mentioned condition is reasonably fulfilled. However, if the size of the aperture is reduced, L is reduced and C at resonance increases, in such a way that Z_{res} might be even much smaller than $Z_{\text{TE}_{11}}$. When this happens, even at resonance, most of the impinging power is reflected and only a small amount is transmitted and/or absorbed. After including the resistor R in the equivalent circuit, the analytical results basically capture the main details of the measured spectrum in Fig. 4.6(b). This experimentally observed fact can be extrapolated to the case of 2D periodic distributions of holes. Thus, although ideal materials (PEC) would allow for the existence of a peak of ET at any distance of the RW anomaly frequency, real material losses would destroy the ET peak because of strong reflection at resonance. This phenomenon always

happens for small enough apertures, which are associated with very low L and high C, thus having small values of Z_{res}. In the specific example given in Fig. 4.6(b), 84.1 % of the power was reflected, 0.6 % was transmitted and 15.3 % was dissipated. A more detailed study on this topic can be found in Ref. 46.

4.4. Modelling of 1D Arrays of Slits: Simple and Compound Gratings

4.4.1. *Circuit models for simple metal slit gratings*

In the introduction to this chapter, it was mentioned that ET has been reported in periodic systems simpler than the ones treated in the previous sections. In particular, 1D metallic gratings with very small apertures[27] and non-vanishing metal thickness under TM illumination exhibit a transmission peak at a wavelength close to the period of the structure. ET peaks of this kind can also be observed in some plots included in some works[28,29] published by Ukrainian researchers by the end of 1960s about thick metal gratings. The basic geometry of a transmission grating of this type is sketched in Fig. 4.7(a), where the cross-section of an infinite periodic array of metal slabs separated by dielectric filled gaps is shown.

In contrast with the 2D case, high transmissivity peaks associated with FP resonances of the cavities formed by the slits are possible no matter how small the slit aperture (a) is. These resonances appear when the thickness of the metal plate supporting the slits (h) is close to an integer number of half wavelengths in the filling dielectric at the operation frequency. These kind of resonances are possible for TM illumination because the mode excited in the PPW formed by the conducting boundaries of each slit region is of TEM nature (with no cutoff frequency). In contrast, small 2D holes operating below the cutoff frequency of the lowest-order TE mode supported by the corresponding waveguide (typical situation found in ET operation) cannot give place to FP resonances. It is worth mentioning that the FP resonances are not associated with any special behavior of the electromagnetic fields outside the slit regions. Thus, they should not be confused with true ET. However, provided the thickness of the

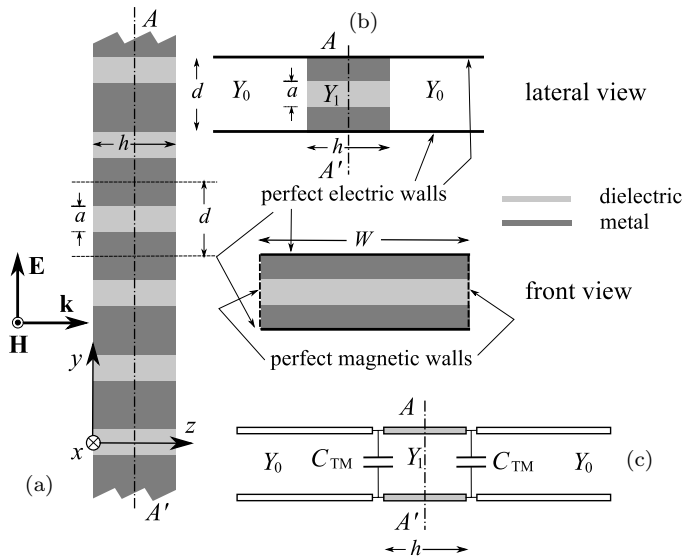


Fig. 4.7. (a) Lateral view of an infinite array of slits of width a made in a metal slab of thickness h. The slits might be filled with a dielectric material. The structure is uniform along the x direction. (b) Lateral and front views of the equivalent discontinuity problem for normal TM illumination. The width W of the equivalent PPW can be arbitrarily chosen. (c) Equivalent circuit for the unit cell problem.

metal plate is not strictly zero, a transmission peak is always observed at the frequency for which the wavelength (in air) is close to the period of the grating (d), independently of the thickness of the metal slab. This resonance is controlled by the periodicity of the structure and is accompanied by a strong perturbation of the electromagnetic field in the region external to the slits. This specific resonance has the same nature as the ET peaks reported for 2D arrays of holes (or for small diaphragms in closed waveguides). In recent years, several authors have reported analytical circuit models for the analysis of this class of structures, and the main results of these works will be summarized in this section.

Let us start our analysis with the case shown in Fig. 4.7(a) of a simple 1D transmission grating made of PEC under normal TM illumination (the effect of metal losses will be added later). Obviously,

the problem is periodic along the y direction, in such a way that the unit cell drawn in Fig. 4.7(b) is the equivalent waveguide problem whose scattering parameters are the same as those of the periodic structure. The propagating mode in the slit region is of TEM nature provided the width of the slit is small enough (for very wide slits, the effect of higher-order modes should be considered). Outside the slits, in the virtual PPWs induced by the periodicity of the problem, a propagating incident, reflected and transmitted TEM mode is always found. If the operation frequency is below the onset of the diffraction regime (i.e. if $\lambda > d$), all the scattered high-order modes (which are of TM nature in this case) are below cutoff. The input admittances of those evanescent modes is of capacitive nature,[35,36] in such a way that a single capacitor, $C_{\mathrm{TM}}(\omega)$, has to be used to account for the influence of the high-order modes in the equivalent circuit shown in Fig. 4.7(c). The capacitance C_{TM} has the same expression as the one given in (4.1); i.e. there is a frequency-dependent contribution with a singularity at the first RW frequency point ($f_{\mathrm{RW}} = c/d$ in this case), which is associated with the first excitable TM mode, in addition to a frequency-independent contribution linked to all the remaining infinite high-order TM modes. The FP-like resonances can easily be explained from the analysis of the circuit in Fig. 4.7(c), even if the dependence on frequency of C_{TM} is ignored. Several transmission peaks of this type could appear before the onset of the diffraction regime if the thickness h of the metal slab is electrically large. A detailed analysis of the location of the FP transmission peaks is reported in Ref. 30. In that paper, it is shown that the effect of the non-negligible value of C_{TM}, provided the slit width (a) is small, can be accounted for by means of an equivalent thickness of the metal slab, h_{eq}, which is slightly larger than the real thickness. The difference between the equivalent and real thickness can be obtained from the ratio between C_{TM} and the per unit length (p.u.l.) capacitance of the slit PPW ($C = \varepsilon_r \varepsilon_0 W/d$). This extra thickness explains the red-shift of the FP transmission peaks with respect to the standard formula that implicitly ignores the effect of the diffracted field. The excess thickness approach cannot be applied to relatively wide slits,

although an analytical solution for the transmission spectrum is still provided by the circuit model in Fig. 4.7(c).

Of special interest in the frame of this chapter devoted to ET is the fact that C_{TM} can reach very high values in the close proximity of $f_{\mathrm{RW}} = c/d$. If the thickness of the metal plate is not strictly zero ($h \neq 0$), the inductance provided by the central transmission line section under odd-excitation conditions (in such conditions the AA' symmetry plane is a virtual short-circuit) will always resonate with some large enough value of the capacitance. This value can always be reached (lossless case) by sufficiently approaching the operation frequency to c/d. However, if the metal grating thickness is strictly zero (in practice, if the thickness is very small), no inductance is available and no extraordinary transmission peak is possible. This observation is consistent with the results reported in Refs. 27 and 28.

The values of C_{TM} can be obtained[30] by combining a single full-wave computation at an arbitrary frequency point within the non-diffraction regime with the static values of the edge capacitance of the step discontinuity in Fig. 4.7(b). This latter static value can be obtained from any commercial Laplace's solver or applying the conformal mapping method. Nevertheless, assuming an approximate profile for the y component of the electric field at the slit aperture (which can be as simple as a uniform field), analytical expressions can be obtained for C_{TM}, including the static frequency-independent contribution and the frequency-dependent one associated with the first high-order TM mode. These approximate analytical formulas for C_{TM} can be found, for instance, in a paper by Khavasi and Mehrany:[49]

$$C_{\mathrm{TM}} = C_{\mathrm{stat}} \left[1 - \alpha + \frac{\alpha}{\sqrt{1 - (d/\lambda)^2}} \right], \tag{4.5}$$

where λ is the wavelength in vacuum, C_{stat} is the static limit value of a grid capacitance such as reported in Ref. 21, and α an empirically determined parameter accounting for the relative level of excitation of the first high-order mode. The authors of the aforementioned paper[49] propose the empirical expression $\alpha = \sqrt{1 - F^4}$, with F being

the "filling factor" of the grating, defined as $F = 1 - a/d$. Alternative analytical formulas for $C_{\mathrm{TM}}(\omega)$, which are based on a conformal mapping approach to obtain the quasi-static limit of that capacitance and an approximation for the contribution of the first TM mode, can be found in another paper by some of the authors of this chapter.[34] The results using any of these sets of expressions are almost identical. Actually, using the strategy reported in any of the above-mentioned references,[30,34,49] very accurate results can be obtained. Thus, circuit model and full-wave (mode-matching method) data are compared in Fig. 4.8(a) for two gratings having the same period and thickness but different values of the slit width (normal TM incidence is considered). Three transmission peaks can be observed in this plot. The

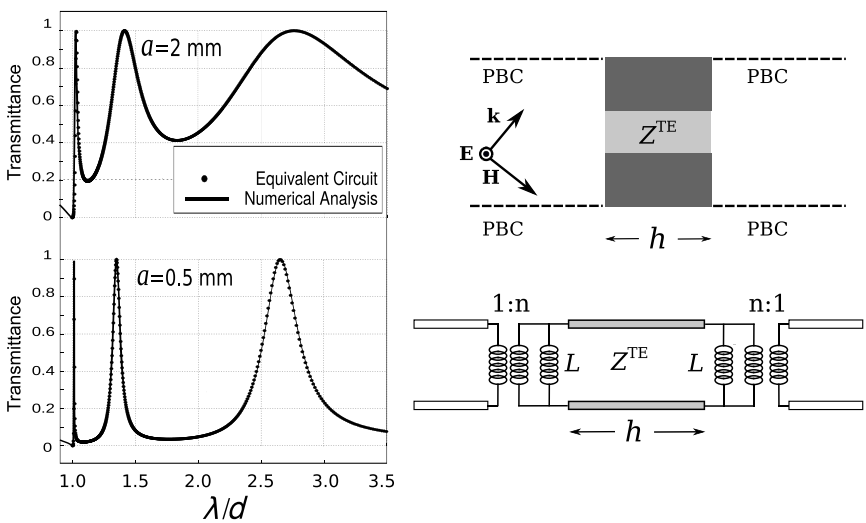


Fig. 4.8. (a) Transmittance calculated for two different metal gratings ($a = 2.0\,\mathrm{mm}$ and $a = 0.5\,\mathrm{mm}$) using a numerical mode-matching approach (solid lines) and the circuit model introduced in this chapter (dots). Dimensions: $d = 5.0\,\mathrm{mm}$, $h = 6.0\,\mathrm{mm}$. Normal TM incidence is considered and air is the dielectric filling all the regions outside metals. Reprinted from Ref. 30 with IEEE permission. (b) Unit cell for the case of TE incidence (top panel) and the corresponding equivalent circuit model.[49] PBC: periodic boundary conditions (magnetic walls in case of normal incidence).

two peaks corresponding to larger wavelengths can be identified as FP resonances of the cavities formed by the slits. The location of those peaks is mainly controlled by the thickness of the metal slab, although a correction due to the external diffracted fields (which depends on the slit width) should be taken into account to generate highly accurate results.[30] Hence, the transmission peaks do not exactly coincide with those frequencies for which the width of the slab is an integer number of $\lambda/2$. The observed red-shift can usually be described in terms of an equivalent slab thickness, as is explained in detail in Ref. 30. In the two analyzed cases, apart from the FP-like peaks there appears a narrow-band transmission peak followed by a transmission zero (Fano resonance) at $\lambda/d \approx 1.0$. This is the true ET peak, which is always present independently of the value of h (except if $h = 0$). The proposed circuit model can reproduce this peak thanks to the introduction of the frequency-dependent singular contribution to C_{TM}.

The authors of Ref. 49 incorporated some interesting additional features to the circuit model for 1D simple metal gratings originally reported in Ref. 30. One of these features was the possibility of dealing with oblique incidence. In this case, the perfect electric walls shown in Fig. 4.7(b) should be replaced by periodic boundary conditions (PBC), in such a way that a *generalized* PPW[50] has to be dealt with. Fortunately, this virtual waveguide also supports non-dispersive TM modes without cutoff frequency, in such a manner that the circuit model in Fig. 4.7(c) is still valid. However, the characteristic admittance $Y_0 = \sqrt{\varepsilon_0/\mu_0}(W/d)$ now has to be divided by the factor $\cos(\theta)$, with θ being the incidence angle, and the propagation constant be corrected including the $\cos(\theta)$ factor. The expression (4.5) should be modified substituting d by $\lambda_C = (1 + \sin\theta)d$ (note that λ_C is the wavelength at which the -1 diffracted order becomes propagative). With these changes, the oblique incidence case is easily incorporated into the circuit model. A more accurate approach can be developed by carrying out an expansion of the aperture electric field as a series of Floquet harmonics of the periodic structure. Using this approach, a better value of C_{TM} can actually be obtained. Note that, in the case of

very narrow slits, the influence of C_{TM} is marginal, in such a way that only the transmission line sections of the circuit model are relevant. Since the characteristic admittance of the slit region is much larger than the characteristic admittance of the external region, relatively narrow-band FP peaks are expected for normal incidence and low angles of incidence. However, due to the $1/\cos(\theta)$ factor weighting the characteristic admittance of the input and output lines in the case of oblique incidence, there always exists a specific value of the incidence angle, θ_B, that yields perfect impedance matching at any frequency below the onset of the first grating lobe. For that angle, almost total transmission can be expected through the grating at any frequency below the onset of the diffraction regime. This angle has been called "plasmonic Brewster angle" in a paper by Alù and co-workers[51] due to the similarity of this phenomenon with the classical total transmission at Brewster angle through a flat interface between two different transparent media.

The case of TE illumination upon a 1D metal grating can also be treated using an appropriate circuit model.[49] In the TE case, the unit cell involves a generalized PPW accounting for the region external to the slits, which has a width equal to the period, d. The height of this PPW can be arbitrarily chosen. The slit region is, under this polarization, a rectangular waveguide (defined by two metal walls and two virtual perfect electric walls, perpendicular to the impinging electric field). It can be assumed that the slit region supports a dispersive TE_{10} mode. A sketch of the unit cell is depicted in the top panel of Fig. 4.8(b). The equivalent circuit for this case would involve a dispersive section of transmission line accounting for the TE mode in the slit region. This waveguide might operate below or above cut-off, depending on the frequency range of interest and the slit width. The discontinuity between the generalized PPW and the rectangular waveguide involves inductances (L) and transformers (n), as shown in the bottom panel of Fig. 4.8(b). As is shown in Figs. 5.8, 5.9 and 5.10 in Ref. 49, the resulting model yields reasonably accurate results even for very wide slits. The authors of this chapter have developed an alternative method to derive the parameters n and L of the circuit

model in Fig. 4.8(b). This model provides more accurate results since the approximations employed in Ref. 49 are avoided. The parameters of the model are obtained by assuming an electric field distribution at the slit aperture plane that corresponds to the TE_{10} mode of the rectangular waveguide that accounts for the slit region.

As a final remark, it should be mentioned that the circuit model for these kinds of gratings can easily be extended in order to deal with the diffraction regime. When the structure is operated at frequencies above the onset of the first grating lobe, one or more Floquet harmonics (or waveguide modes, in the case of normal incidence) would operate above cutoff in the input and output waveguides of the unit cell. In such case, the input admittance of the corresponding Floquet harmonic (oblique incidence) or waveguide mode (normal incidence) becomes a real and dispersive quantity. This means that, apart from the capacitors considered in the analysis reported so far, frequency-dependent resistors that account for the modes (or harmonics) above cutoff should be added to the model. These resistors should be connected in parallel with the capacitors. The power dissipated in the resistors is the power transferred to high-order propagating modes (or grating lobes). The details of this improvement of the reported circuit model are explained in two recent papers.[52,53]

4.4.2. *Circuit model for compound metal slit gratings*

A special class of 1D grating is the so-called *compound* grating. A compound grating is a 1D metallic periodic structure (of the same type considered in the previous subsection) involving several slits per period instead of a single one. When these slits are densely packed, additional interesting resonances appear in the transmission spectrum, as it has been reported in several theoretical[31,32] and experimental[54,55] papers. These resonances have been named as "phase resonances."[31] In all the above papers, the slits were considered air filled. The most basic structure of this type that exhibits phase resonances under normal TM illumination (with all the slits filled with the same dielectric material) is shown in Fig. 4.9(a). In the present analysis, for generality purposes, the slits are assumed

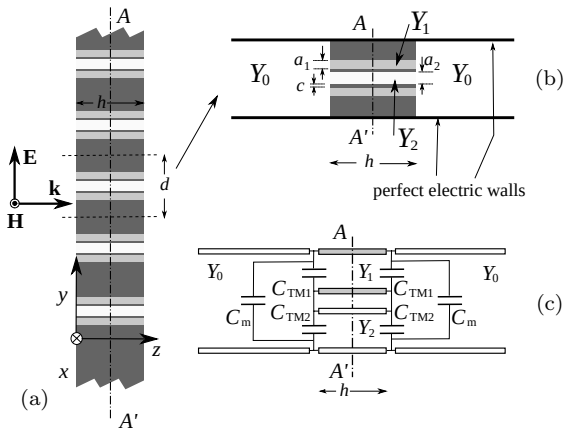
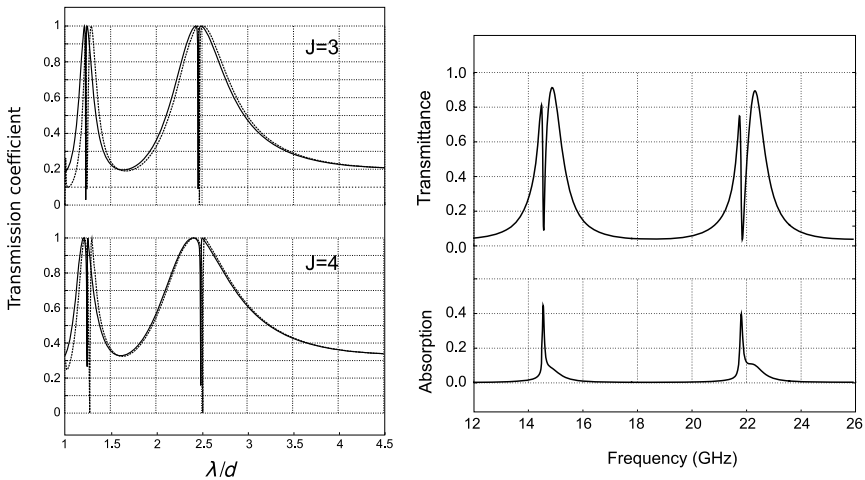


Fig. 4.9. (a) Section of a compound transmission grating whose unit cell is composed by three symmetrically disposed slits per period. (b) Unit cell to be solved (for normal TM incidence). (c) Equivalent circuit for the unit cell. This circuit takes into account the symmetry of the unit cell with respect to the horizontal middle plane. Reprinted from Ref. 30 with IEEE permission.

to be filled with different dielectric media; this feature does not add special difficulties to the analysis. The unit cell to be considered is depicted in Fig. 4.9(b), and the proposed circuit model is the one drawn in Fig. 4.9(c). This structure has been analyzed in Ref. 30 using a quasi-static approach for the treatment of the fields scattered in the regions outside of the slits (input and output waveguides). Due to the quasi-static nature of the employed approach, the capacitors involved in the modelling of the scattered TM fields have frequency-independent capacitance values. These values can be obtained by solving an associated 2D Laplace's problem, as is explained in the cited reference.[30] The presence of three symmetrical slits per period is taken into account by means of the two different transmission lines in Fig. 4.9(c). Obviously, this model incorporates the symmetry of the unit cell with respect to the horizontal symmetry plane of the considered unit cell. Note that the same circuit topology can be used for a structure having four symmetrically disposed slits per period. If each period is composed by five or six slits, an additional transmission line section should be added, as well as new external capacitors

taking into account the coupling between the slits through high-order scattered modes. As the capacitances have been determined using a quasi-static approach, the behavior of the structure at frequencies close to the onset of the first grating lobe are not expected to be captured by this model. Nevertheless, for very thick slabs (relatively large values of h, typically larger than the period d of the structure), interesting phenomena can be observed in the transmission spectrum around each of the FP-like resonances; some of those resonances are far enough from the first RW anomaly, which allows for the use of quasi-static models for the treatment of the scattered high-order modes. In particular, as it is explained in depth in Ref. 30, a strong narrow-band transmission dip (acting as a notch filter) is created close to the central frequency of each FP resonance. The model also predicts more transmission zeros as the number of slits per period is increased (for instance, two transmission dips should be observed if the number of slits per period is five or six). These predictions are totally consistent with the theoretical and numerical results reported by other authors[31] using full-wave computationally intensive numerical approaches.

The circuit model in Fig. 4.9(b) with quasi-static values of the capacitances obtained using a commercial Laplace's solver has been used to generate the curves (solid lines) shown in Fig. 4.10(a). Those curves correspond to the cases of three ($J = 3$) and four ($J = 4$) slits per period, with the slits being identical and uniformly separated. The curves are compared with the data numerically obtained using a full-wave mode-matching approach.[31] The circuit model reproduces with good accuracy the transmission spectrum around $\lambda/d \approx 2.5$. However, although the qualitative response is still well captured, the model is clearly less accurate when dealing with the resonances occurring at about $\lambda/d \approx 1.2$. This is due to the dispersive (frequency-dependent) nature of the capacitances of the model, which is specially significant when the operation frequency is approaching the point $\lambda/d = 1$. It is worth mentioning that this drawback of the model has recently been overcome by using frequency-dependent capacitors whose values are analytically obtained from approximate field



(a) (b)

Fig. 4.10. (a) Transmission spectra for two lossless compound gratings studied in Fig. 4.2 of Ref. 31 (three, $J = 3$, and four, $J = 4$, slits per period). Dimensions: $d = 10\,\text{mm}$, $h = 11.4\,\text{mm}$, $a = 0.8\,\text{mm}$, $c = 0.8\,\text{mm}$. Solid lines: circuit model predictions; dashed lines: numerical data.[31] (b) Transmission spectrum for the structure measured in Fig. 4.2 of Ref. 54 computed by means of the circuit model. Very good agreement with the experimental data[54] has been found. Reprinted from Ref. 30 with IEEE permission.

distributions in the slits apertures.[33] The new model[33] is essentially the dynamic version of the equivalent circuit reported in Ref. 30 for transmission compound gratings. Apart from a much higher accuracy, specially at frequencies close to $\lambda/d = 1$, the dynamic model for compound gratings can also deal with the diffraction regime ($\lambda < d$) and accounts for other interesting high-frequency effects that appear in compound reflection gratings (such as extraordinary specular reflection[56]). Another interesting example of the application of the proposed circuit model to compound gratings is given in Fig. 4.10(b). In that figure, the transmission through a compound grating that was experimentally studied in Ref. 54 is shown. In these plots the absorption is included, since the losses inside the slits have been incorporated into the model by means of a complex characteristic impedance and a complex propagation constant.

These complex quantities have been determined using the strong skin effect approximation, as is explained in another paper by some of the authors.[34] Since the dimensions and the operation frequency range considered in the experimental paper[54] allow for the use of the quasi-static model with confidence, the analytically obtained results almost perfectly agree with the experimental data (some plots where experimental data are superimposed to the analytical data can be found in the previously mentioned paper[33] about a dynamic model for the compound grating problem).

In brief, analytical models based on relatively simple equivalent circuits involving a small number of components (capacitors, inductors and transmission line sections) have shown to be an accurate alternative to numerical methods for studying the electrical response of simple and compound metal gratings.

4.5. Metal Gratings with Finite Thickness Dielectric Slabs

In previous sections, our equivalent-circuit analysis has been restricted to the case of free-standing metallic structures. In the case of 1D gratings, homogeneous dielectrics could be present inside the slits, but this fact does not meaningfully modify the physics of the problem and the kind of equivalent circuit to be handled. However, in many practical cases, the metallic patterned screen is printed on a dielectric slab, or even embedded in a stratified structure with several dielectric layers. The dielectric substrates are usually employed to provide mechanical support, but their electrical influence cannot be neglected. Indeed, if the thickness of the dielectric slabs is of the order of magnitude of the wavelength in the dielectric material, the electrical response of the periodic metallic screen will be strongly affected. The appropriate use of dielectric slabs can also be useful in practical applications, since additional flexibility is provided to the design process.

In microwaves and antenna technologies, structures whose geometry can be described as a periodically patterned metal surface printed on dielectric substrates have been used for long, and they

are known as FSSs.[3,57] In these applications, "ordinary" transmission and reflection processes have been the focus of interest, and no attention has been paid to more "exotic" responses, such as those studied in the previous sections of this chapter. The model that explains the existence of transmission peaks and dips in the transmission spectrum on the basis of the excitation of surface waves would clearly predict that the presence of a layered environment will make such transmission spectrum a lot more involved than the one exhibited by free-standing structures. This is due to the presence of additional surface waves guided or trapped by the dielectric slabs. Due to this reason, significant attention has also been paid to the modelling and study of these kinds of structures in the context of extraordinary transmission phenomena.[58–63]

The purpose of this section is to review some recent analytical equivalent-circuit models useful to deal with a variety of periodic metal structures embedded in stratified dielectric media. In the literature on the topic there are many circuit models dealing with this kind of structures. Classical examples for 1D periodic metal gratings can be found in several papers published in the 1950s.[16–18] More sophisticated circuit-like models for arrays of metal slits placed between two semi-infinite media were later introduced.[64,65] Two-dimensional metallic diffraction grids used as filters in the infrared range of frequencies have been treated with circuit-modelling techniques,[66–68] and similar methods have been successfully applied in other frequency bands. [19,21,24,25,69–72] However, many of the available circuit models have significant restrictions concerning the electrical size of the unit cell. Other methods lack of a systematic procedure to determine the topology of the equivalent circuit and the values of the circuit elements. In particular, a circuit model explaining dynamic effects, such as extraordinary transmission, should incorporate the frequency dependence of the values of some of the components used in the model. Next, a systematic procedure will be described to derive the circuit topology and the values of the circuit components for a class of periodic structures involving one or more dielectric layers. This methodology is comprehensively explained in several papers

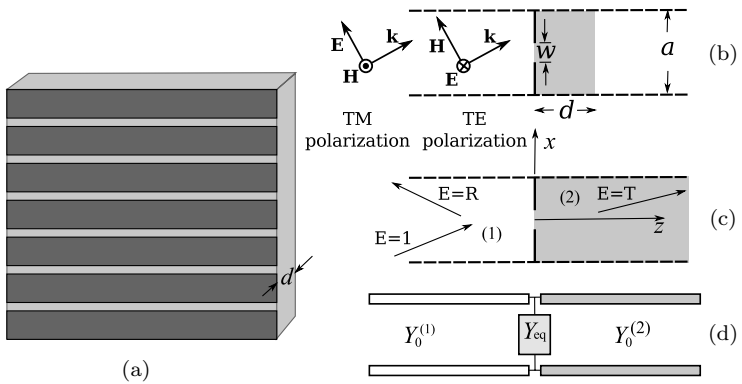


Fig. 4.11. (a) 1D array of metal strips (dark grey) printed on a dielectric substrate of thickness d. The period of the structure is a, the width of the slits is w, and the relative dielectric constant is ε_r. (b) Unit cell of the structure (TM or TE incidence can be considered). (c) Simplified auxiliary unit cell with infinitely thick substrates. (d) Equivalent circuit for the problem in (c).

by some of the authors of this chapter, including both 1D[73–75] and 2D[76,77] periodic structures. The basic guidelines of the approach will be presented for the simplest case (1D metal gratings printed on a dielectric substrate). For more complex geometries, the reader will be referred to the adequate published material.

The structure chosen in this section to exemplify the method employed to derive the equivalent circuits is shown in Fig. 4.11(a). In this structure, an array of metallic strips separated by electrically narrow slots and printed on a dielectric substrate of thickness d is shown. The structure is assumed to be illuminated by a TM (magnetic field parallel to the slots) or TE (electric field parallel to the slots) polarized uniform plane wave. Once again, the periodicity of the structure and the excitation makes it possible to reduce the problem to a waveguide discontinuity problem, which is schematically shown in Fig. 4.11(b). Figure 4.11(c) represents an auxiliary unit cell where the thickness of the dielectric layer is assumed to be infinite. Let us focus our attention on this latter problem, whose equivalent circuit shown in Fig. 4.11(d) is given by means of the known characteristic admittances $Y_0^{(1)}$ and $Y_0^{(2)}$ of the input and

output transmission lines and the unknown equivalent admittance of the discontinuity, $Y_{\rm eq}$. A time harmonic dependence ($\exp(\mathrm{j}\omega t)$) is assumed for the electromagnetic quantities. At the $z = 0$ discontinuity plane, taking an impinging electric field of amplitude $E = 1$, the electric field can be expressed as the following Floquet series of spatial harmonics:

$$E(x) = (1 + R)\mathrm{e}^{-\mathrm{j}k_t x} + \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} E_n \mathrm{e}^{-\mathrm{j}(k_t + k_n)x}, \tag{4.6}$$

where R is the reflection coefficient of the zero-th harmonic (incident wave), $k_t = k_0 \sin\theta$ (tangential-to-the-grating wavevector component of the impinging wave), $k_0 = \omega\sqrt{\varepsilon_0\mu_0}$), and $k_n = 2\pi n/a$. The same expansion coefficients appear at both sides of the grating due to the continuity of the tangential electric field, so that the transmission coefficient is $T = 1 + R$. The TM field at both sides of the grating is then given by

$$H^{(1)}(x) = Y_0^{(1)}(1 - R)\mathrm{e}^{-\mathrm{j}k_t x} - \sum_{n\neq 0} Y_n^{(1)} E_n \mathrm{e}^{-\mathrm{j}(k_t + k_n)x}, \tag{4.7}$$

$$H^{(2)}(x) = Y_0^{(2)}(1 + R)\mathrm{e}^{-\mathrm{j}k_t x} + \sum_{n\neq 0} Y_n^{(2)} E_n \mathrm{e}^{-\mathrm{j}(k_t + k_n)x}, \tag{4.8}$$

where the following harmonic wave admittances have been introduced:

$$Y_n^{(i)} = \begin{cases} \beta_n^{(i)}/\omega\mu_0 & \text{TE harmonics} \\ \omega\varepsilon_i/\beta_n^{(i)} & \text{TM harmonics,} \end{cases} \tag{4.9}$$

with

$$\beta_n^{(i)} = \sqrt{\omega^2\varepsilon_i\mu_0 - (k_t + k_n)^2}, \tag{4.10}$$

being the wavenumber along z of the n-th harmonic in medium (i).

Assuming now that the electric field in the slit is given by $E_{\rm s}(x)$, applying standard Fourier analysis and comparing to (4.6), it is found that

$$1 + R = \frac{1}{a}\widetilde{E}_{\rm s}(k_t); \quad E_n = \frac{1}{a}\widetilde{E}_{\rm s}(k_t + k_n), \tag{4.11}$$

(the symbol $\sim$ denotes Fourier transform) or, equivalently,

$$E_n = (1+R)\frac{\widetilde{E}_{\mathrm{s}}(k_t+k_n)}{\widetilde{E}_{\mathrm{s}}(k_t)}. \tag{4.12}$$

If the continuity of the power through the slit region is imposed,

$$\int_{-w/2}^{w/2} E_{\mathrm{s}}^{*}(x)[H^{(2)}(x)-H^{(1)}(x)]\,\mathrm{d}x = 0 \tag{4.13}$$

and (4.7) and (4.8) are introduced into (4.13), is the following expression for the reflection coefficient finally found:

$$R = \frac{Y_0^{(1)} - Y_0^{(2)} - Y_{\mathrm{eq}}}{Y_0^{(1)} + Y_0^{(2)} + Y_{\mathrm{eq}}}. \tag{4.14}$$

This reflection coefficient relates the incident and reflected plane waves of the equivalent circuit in Fig. 4.11(d), with the equivalent admittance of the discontinuity being given by

$$Y_{\mathrm{eq}} = \sum_{n\neq 0}\left|\frac{\widetilde{E}_{\mathrm{s}}(k_t+k_n)}{\widetilde{E}_{\mathrm{s}}(k_t)}\right|^2 (Y_n^{(1)}+Y_n^{(2)}). \tag{4.15}$$

It can easily be observed in (4.15) that the discontinuity admittance actually corresponds to the parallel connection of the wave admittances of the higher-order spatial harmonics excited at both sides of the metallic grating with a weighting factor related to the spatial spectrum of the field profile at the slit. Note that it is the shape of $E_{\mathrm{s}}(x)$ rather than its absolute values of amplitude and phase that actually affects Y_{eq}. Although the field profile will depend on the operation frequency, for electrically narrow slits (some discussion on the limits of validity are provided in Ref. 33), it is a good assumption to consider a frequency-independent spatial profile, as follows:

$$E_{\mathrm{s}}(x) = A(\omega)f(x). \tag{4.16}$$

The amplitude and phase of the actual electric field excited in the slit [given by the complex factor $A(\omega)$] is a function of the frequency, but this *a posteriori* information is obtained from the solution of the

circuit problem. This is possible given that $A(\omega)$ does not appear in the final expression of the equivalent circuit admittance (4.15) as this admittance only depends on the ratio of $\widetilde{E}_{\mathrm{s}}(k_t + k_n)$ to $\widetilde{E}_{\mathrm{s}}(k_t)$.

If a good estimate of the electric field profile in the slit, $f(x)$, is available, the circuit model in Fig. 4.11(d) with the admittance in (4.15) would provide very accurate values of R and T for the structure in Fig. 4.11(c). A good approximation of the electric field in a narrow slit under TM illumination is given by the function $f(x) = 1/\sqrt{1-(2x/w)^2}$, while for TE illumination $f(x) = \sqrt{1-(2x/w)^2}$ is an appropriate choice. Regardless of the specific field profile chosen to compute Y_{eq}, it is important to take into account that the contributions to (4.15) coming from relatively high-order (ho) harmonics have a very simple dependence on the frequency.[75] Thus, for high-order TM harmonics (only excited under TM incidence), all the admittances are proportional to ω and purely imaginary. It means that the overall contribution of all those harmonics can be represented by means of a regular frequency-independent capacitor, C_{ho}. For TE incidence, only TE modes are excited and, in such case, the contribution to Y_{eq} coming from high-order harmonics can be represented by a regular frequency-independent inductor, L_{ho}. The values of C_{ho} or L_{ho} need to be computed only once for any frequency swapping and angle of incidence. In the low-frequency limit (long-wavelength limit), as the operation frequency is well below the cutoff frequencies of all the harmonics (except the zeroth-order one), a simple capacitor (TM case) or inductor (TE case) is required to account for the effect of the slit discontinuity. This is the approach followed in many works dealing with the circuit modelling of these kind of periodic structures.[21,66] However, when the operation frequency is close to or above the cutoff of some high-order harmonics (either in the dielectric slab or in air), the more complex frequency dependence of Y_{eq} cannot be ignored. Fortunately, the incorporation of the dynamic effects only requires to treat separately very few modes of low order (lo). One or two modes are enough for many practical applications, although some specific situations may require to deal with several low-order modes. For these low-order modes, the exact expression of

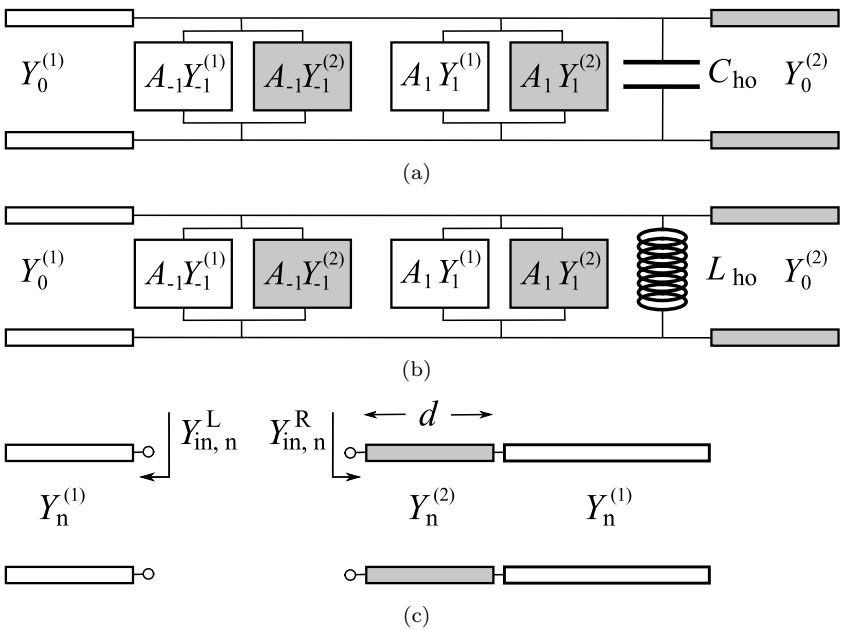


Fig. 4.12. (a) Equivalent circuit for the unit cell in Fig. 4.11(c) under TM incidence with explicit separation of high-order (quasi-static) contributions (C_{ho}) and dynamic low-order harmonics contributions ($n = \pm 1$). (b) The same type of circuit for TE incidence. (c) Equivalent transmission-line circuits defining $Y_{\mathrm{in},n}^{\mathrm{L}}$ and $Y_{\mathrm{in},n}^{\mathrm{R}}$, which are the admittances substituting $Y_n^{(j)}$ when a finite thickness (d) dielectric is present. The extension to a multilayer structure only involved the use of several transmission line sections.

their wave admittance is retained, leaving aside the approximations that led to simple lumped capacitors/inductors. Two examples of equivalent circuits that incorporate dynamic effects are sketched in Figs. 4.12(a) and (b) for TM and TE incidence, respectively. Only the harmonics $n = 1$ and $n = -1$ have been taken as low-order modes, while the remaining infinite contributions give place to the lumped frequency-independent C_{ho} or L_{ho}. If the operation frequency is above the cutoff of any of the considered low-order harmonics, the corresponding admittance would become real and still dispersive. It might happen in the dielectric region, while in the air region the harmonic is still below cutoff. Due to this, the contributions of the

scattered harmonics at both sides of the slit discontinuity have been separated in the drawings.

In the previous discussion, only the structure in Fig. 4.11(c) with two semi-infinite dielectric regions has been considered. The inclusion of a finite thickness (d) dielectric slab can easily be carried out by substituting the wave admittances appearing in the previous formulation ($Y_n^{(j)}$, $n = \pm 1, \pm 2, \ldots$) with the *input* admittances of the system formed by cascaded transmission line sections, each of them representing a dielectric region.[75] These input admittances are easily obtained using conventional transmission line theory,[35] in such a way that the method discussed in this section is actually valid for a metal grating embedded in any layered environment with arbitrary dielectric constants and thicknesses.[75] This idea is graphically shown in Fig. 4.12(c). It should be highlighted that, in contrast with other simplified circuit models, the reported models incorporate all the dynamic features of the transmission/reflection spectrum as long as the field profile in the slit does not appreciably change over the considered frequency range. A simplified version of the models presented in this section has been used by Beruete *et al.*[62,63] to explain the phenomenon known as "anomalous extraordinary transmission," which appears when a TE polarized wave illuminates a strip grating loaded with a finite thickness dielectric slab. The adjective "anomalous" is here used because this kind of structure would not exhibit ET peaks in the absence of the dielectric slab. The aforementioned papers[62,63] are interesting examples of how the circuit-model point of view can shed light on a physical phenomenon whose understanding using other methods would be troublesome.

The derivation of the equivalent circuits shown in Fig. 4.12 has been carried out using an aperture formulation. Therefore, the *a priori* assumption of the shape of the spatial profile of the electric field in the aperture (slit) has been our key premise. Due to the use of this aperture formulation, the topology of the equivalent circuit for the admittance accounting for the effect of the discontinuity is the parallel connection of a number of admittances associated with each of the scattered harmonics. However, if the width of the slit is

large (i.e. w is close to the period, a), in such a way that the width of the metal strips is relatively narrow, the structure is better described as an array of narrow metal strips. In that case, it is easy to have a good estimation of the electric surface current density supported by the metal strips, which makes more convenient the use a formulation where the unknown quantity is the current supported by the strips. This approach is usually known as the "obstacle" formulation of the scattering problem. This option has also been considered in Ref. 75 in order to obtain an appropriate equivalent circuit. When this approach is followed, the equivalent impedance of the discontinuity can be expressed as the series connection of an infinite number of pairs of admittances connected in parallel. In practice, the admittances associated with high-order harmonics contributions collapse into a single capacitor or inductor (depending on the polarization of the impinging wave) and only a few elements of the equivalent circuit (those associated with low-order harmonics) retain a more complex structure, including dynamic and distributed effects. Thus, a complete equivalent-circuit description is available for both slit–like and strip–like structures.[75]

As it has already been mentioned, structures of this type have been treated using circuit models with some degree of simplification with respect to the kind of circuits described in this chapter.[21,25] However, in many relevant practical cases, the use of quasi-static approximate solutions or other simplifying assumptions yield poor results, typically valid within a restricted range of low frequencies. An interesting example to illustrate that ignoring dynamic effects might yield very poor numerical data can be found in Ref. 75. In that work, one of the selected examples shows the phase of the reflection coefficient of a periodic array of narrow metal slits printed on a conductor-backed dielectric substrate. A sketch of the considered problem is depicted in Fig. 4.13(a). For this kind of structure, extremely accurate results (virtually exact) can be obtained by using the method of moments (MoM) with a suitable set of basis functions incorporating the most relevant physical features of the expected solution.[78] If losses are ignored, the magnitude of the reflection coefficient for

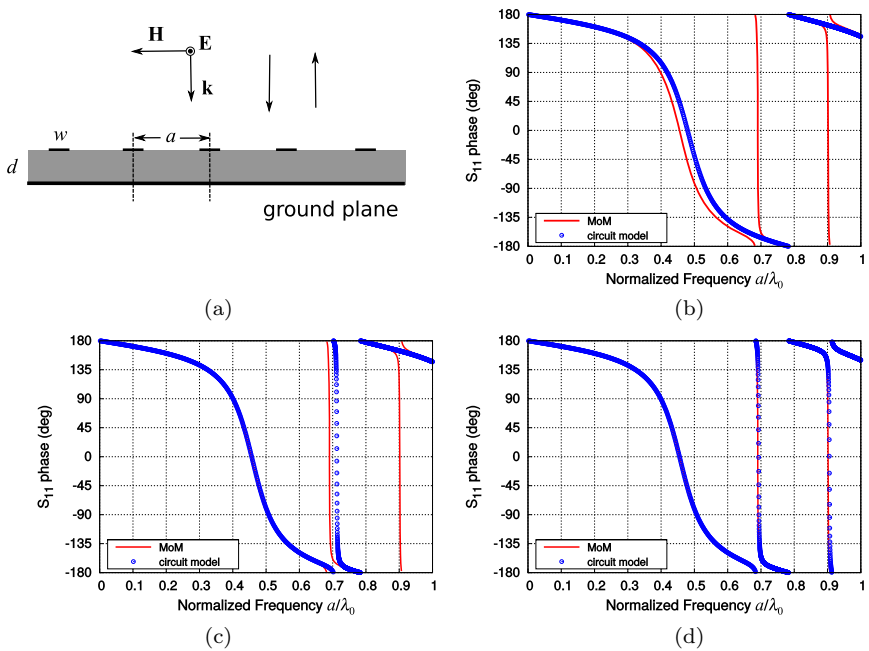


Fig. 4.13. (a) Strip grating with period a made of narrow PEC strips ($w = 0.1a$) printed on a dielectric slab of thickness $d = 0.2a$ and relative permittivity $\varepsilon_r = 10.2$. Normal TE incidence is considered. (b) Phase of the reflection coefficient (S_{11}) using a quasi-static approach to calculate the grating inductance. (c) The same parameter when the exact contribution of the first scattered mode is considered. (d) The same parameter when the exact contribution of the two lowest order modes is incorporated. Reprinted from Ref. 75 with IEEE permission.

frequencies below the onset of the diffraction regime is just 1. Thus, it is the variation of the phase of this coefficient with frequency that has been computed using an exact MoM and the described circuit model. Although it has not been explained before in this chapter, the presence of the ground plane can easily be incorporated into the model by properly short-circuiting the cascade of transmission line sections (see Fig. 4.12(c)) used in the modelling of the structure.[75] The curves shown in Figs. 4.13(b)–(d) correspond to the phase of the reflection coefficient (S_{11}) for the structure under analysis under TE normal illumination in the non-diffractive regime ($a/\lambda_0 \leq 1$); the dimensions and electrical parameters are given in the caption to

Fig. 4.13. The analytical curves (circuit model) in Fig. 4.13(b) have been generated using a quasi-static model for the slit discontinuity (a simple capacitor is used to account for the effect of the slit). It can be clearly observed that the phase is correctly computed only in the low frequency limit ($a/\lambda_0 \lesssim 0.30$). The analytical curves in Fig. 4.13(c) take into consideration the exact dependence with frequency of the first high-order TM mode and those in Fig. 4.13(d) the exact contribution of the first two high-order TE modes. The extraction of two low-order TE modes from the total quasi-static inductance is found to be enough to obtain very accurate results over the whole considered frequency band. The inaccuracy of the quasi-static model is related to the fact of ignoring the interaction of the first two high-order TE modes with the ground plane. If the substrate was significantly thicker, the simplified model considering that only the fundamental mode is reflected by the ground plane would have been sufficient.

The case of a slit grating under TM illumination is considered in Fig. 4.14. A periodic array of narrow slits (see caption for detailed geometrical and electrical parameters) is sandwiched between two different dielectric slabs. The frequency and the angle of incidence

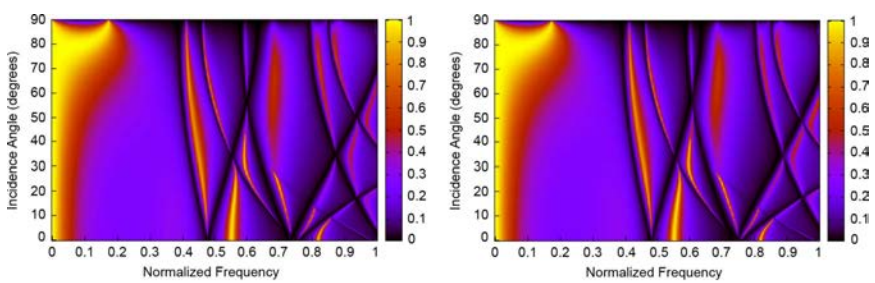


Fig. 4.14. Magnitude of the transmision coefficient for a slit grating (slit width $w = 0.1a$, period a) sandwiched between two dielectric layers of thicknesses $d_1 = 0.4a$ and $d_2 = 0.2a$ and relative permittivities $\varepsilon_r^{(1)} = 2.2$, $\varepsilon_r^{(2)} = 10.2$. The loss tangent of both dielectric layers is $\tan\delta = 0.001$. (a) Circuit model results. (b) MoM results. Normalized frequency: a/λ_0. Reprinted from Ref. 75 with IEEE permission.

are swept along a wide range of values. Note that the differences between the analytical (circuit model) and numerical (MoM) results cannot be appreciated in this color map. Multiple resonances due to the presence of the dielectric layers can be clearly observed, although no grating lobes are allowed.

The analytical methodology presented in this section for a specific class of 1D metal gratings has also been adapted to the analysis of more sophisticated 1D periodic structures, such as the T-shaped and corrugated surfaces studied in Refs. 79 or 80. Thus, a very accurate analytical circuit model for this class of geometries has been developed by some of the authors[81] without the limitations of the analytical approximations given in the aforementioned papers.[79,80] However, more interesting is the fact that the present approach to extract equivalent circuits can equally be used in the case of 2D arrays of holes in metal plates or arrays of metal patches. In the 2D case, the analytical work is more involved, but the underlying rationale behind the extraction of circuit models is the same proposed in this section of the chapter. Once again, it is essential to have a reasonable estimation of the profile of the electric field in the apertures or the electric current on the metal patches. If those functions do not appreciably change in the frequency range of interest, highly accurate circuit models can be obtained. For highly symmetric scatterers or apertures (for instance, rectangular metal patches or holes) under TM or TE illumination with the wavevector of the impinging wave oriented along one of the principal planes of the periodic structure, the reflected and transmitted waves have the same polarization as the incident one. In such case, the topology of the circuit for the computation of the scattering parameters of the fundamental wave (zero-order harmonic) is the same as that proposed in Fig. 4.12. But, in the 2D case, both TE and TM modes (or harmonics) are scattered, in such a way that the asymptotic (quasi-static) contribution associated with the high-order modes involves both a capacitor and an inductor. Indeed, a multimodal network[64,65] can be inferred for 2D arrays of rectangular apertures or patches embedded in or printed on stratified dielectric structures, as it has recently been reported

in Ref. 77. In the multimodal network representation, the input and output modes (harmonics) can be arbitrarily chosen in such a way that illumination using non-uniform plane waves can be treated. In Ref. 77, the case or arbitrary oblique conical incidence is also considered by using a set of properly chosen auxiliary problems. A linear combination of the reflection and transmission coefficients for those auxiliary problems provides the solution for co-pol and cross-pol scattering coefficients. This is an analytical solution to the problem posed by Maci *et al.*,[82] where an elegant semi-numerical approach was reported.

As it has already been pointed out in this chapter, the key point to obtain an appropriate equivalent circuit that incorporates the dynamical features of the model (i.e. frequency-dependent capacitors, inductors or other relevant admittances/impedances) is the availability of a reasonable estimation of the electric field (aperture formulation) or the electric current (obstacle formulation) at the discontinuity plane. When the geometry of the apertures or scatterers is simple, this is not a difficult task. However, for more involved geometries (rings, crosses, Jerusalem crosses, etc.), an approximate analytical description of the aforementioned quantities is not simple any more. Fortunately, a reasonable approximation can be numerically obtained by solving an auxiliary 2D problem consisting on the determination of the modal patterns of the hollow-pipe waveguide whose cross-section matches the geometry of the scatterer. The solution of this translational-symmetry problem is clearly less computationally intensive than the solution of the global scattering problem. Applying this idea, a couple of procedures have recently been reported that are capable of dealing with scatterers having non-trivial shapes.[37,83] In one of those papers,[37] the method reported in this chapter is used in combination with the field pattern extracted from a single run at a relatively low frequency of a numerical code. The other paper[83] makes use of a linear combination of two different modal field patterns (or surface current patterns), thus giving rise to an equivalent circuit that incorporates the first two intrinsic self-resonances of the scatterer. This extended procedure yields an

ultrawideband equivalent circuit for the fundamental modes (harmonics), but cannot lead to a multimodal network such as the one presented in Ref. 77.

4.6. Stacked Structures and Fishnets

A final and interesting situation that has attracted a lot of attention in the metamaterials literature will be analyzed in this section. The structure corresponds to the stacking of two or more periodic arrays of metal patches or apertures of the type studied in previous sections. These so-called fishnet structures have been proposed as a versatile and simple path to achieve negative-index behavior at different wavelength regimes. A detailed description of this kind of system can be found, for instance, in Refs. 84 and 14 and references therein. A fishnet is basically a stack of several closely spaced ET metal layers (i.e. layers or films of metals with 2D periodic arrays of subwavelength apertures). These structures are known to support backward waves along the stacking direction; namely, they present left-handed behavior. Stacked structures for other purposes can also be found in classical and recent literature. Thus, for instance, the stacking of electrically dense metallic grids has been shown to open wide transmission bands in spite of the extremely low transmissivity of each of the individual stacked grids.[72,85] 2D stacked arrays of metal patches are also of interest since they behave as artificial dielectric materials whose effective anisotropic dielectric constant can be tailored by properly adjusting the dimensions of the unit cell, the metal patch, and the dielectric constant and thickness of the slabs separating the arrays of patches.[86,87] Actually, the 1D version of these latter structures (stacked 1D periodic arrays of metal strips) attracted the attention of researchers several decades ago,[17,18,88] since those structures could give place to very high values of the effective permittivity or values lower than unity. In brief, it would be quite useful to extend the equivalent-circuit methodology explained in previous sections of this chapter to the modelling of this class of device. Actually, most of the aforementioned references propose some kind of equivalent circuit, but they are not general enough to deal with

all the situations that can be found in practical implementations. A couple of interesting steps in the appropriate direction are given in Refs. 89 and 90. However, these approaches have limited accuracy if the dimensional parameters are not within a restricted range[89] or require the use of full-wave simulators to extract the circuit parameters.[90] The purpose of this section is to describe the fundamental guidelines to generate equivalent circuits (i.e. analytical solutions) in a systematic manner leading to accurate and fully analytical models for arbitrary values of the geometrical parameters.

For simplicity, the basis of the proposed methodology will be exposed for stacked 1D periodic structures. The cross-section of a 1D structure of this type is sketched in Fig. 4.15(a). In this structure, $N + 1$ identical metal gratings with relatively narrow slits are separated by N dielectric layers of arbitrary permittivity and thickness. The structure can be illuminated along an incidence plane perpendicular to the strips using TM- or TE-polarized plane waves. If the period, p, of the structure is small in comparison with the wavelength and the separation between consecutive metal gratings is large enough, the circuit model for this structure is quite simple and well known. Each metallic grating surface can be modelled by means of a simple capacitor (TM incidence) or inductor (TE incidence), whose values are not affected by the presence of the remaining gratings. The interaction between consecutive gratings is exclusively provided by the single propagating mode (or harmonic), the so-called zero-order harmonic, which is represented in the circuit model by means of a simple transmission line section. These kinds of models have been used, for instance, in the treatment of 2D stacked grids[72,91] or 2D stacked arrays of metal[92] or graphene[93] patches. However, these models are not valid when the interaction between adjacent layers is influenced by high-order modes.

In the first step, our attention will be focused on the simplest case shown in Fig. 4.15(b), where only two periodic metal gratings are interacting through an electrically thin substrate. A general circuit that account for the interactions through modes (harmonics) of any order can always be set in the form of a Π-circuit as the one depicted

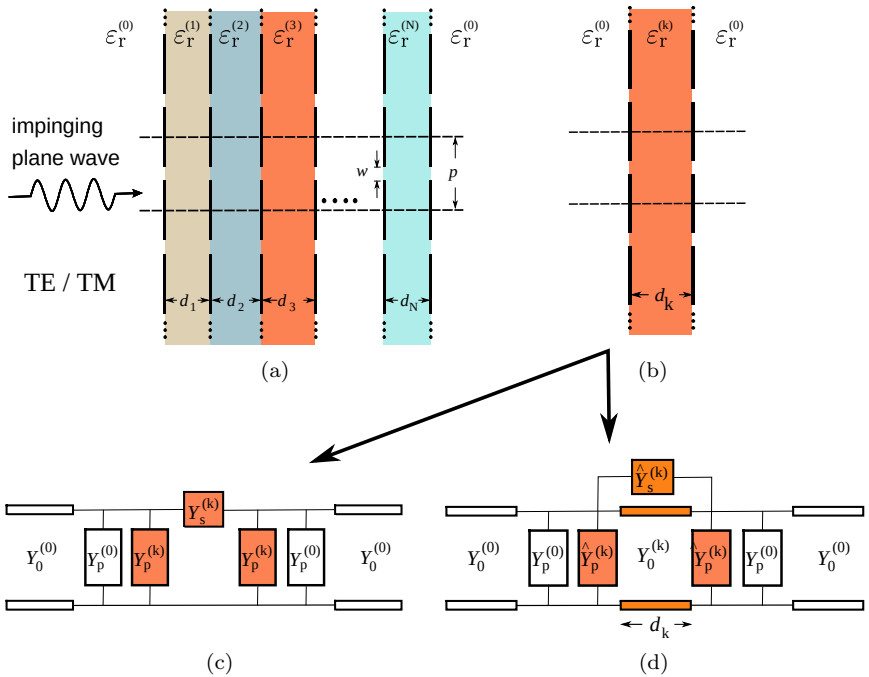


Fig. 4.15. (a) $N+1$ stacked identical 1D metal gratings separated by N dielectric layers of arbitrary thickness (d_k, $k = 1, \ldots, N$) and dielectric permittivities ($\varepsilon_r^{(k)}$). Air is assumed in the external region, but any dielectric material of infinite thickness can be considered. TM/TE oblique incidence can also be accounted for. (b) Elementary structure for the extraction of the basic building blocks to obtain the equivalent circuit of the stacked configuration. (c) General equivalent circuit for the structure with two strips gratings. The colored Π-circuit can be used to build the equivalent circuit of the stacked configuration. (d) Equivalent circuit with the zero-order mode (harmonic) explicitly extracted out.

in Fig. 4.15(c).[94] All the admittances in that circuit can easily be expressed in terms of the aperture (slit) electric field. Of course, this Π-circuit would reduce to a simple transmission line section plus the appropriate slit capacitance (TM incidence) or inductance (TE incidence) associated with the iris discontinuity provided the separation between gratings is large enough. Indeed, the transmission line section corresponding to the zero-order mode (harmonic) contribution, which is always responsible of interaction between gratings, could

be explicitly extracted, as it has been done in Fig. 4.15(d) (other alternative representations can be found in Ref. 94). Nevertheless, the most relevant feature of the circuit in Fig. 4.15(c) concerning the analysis of the stacked structure in Fig. 4.15(a) is the fact that there is an explicit separation of the admittances corresponding to the region between the gratings (colored series and parallel admittances) and the ones corresponding to the region outside the gratings (white parallel admittances). This separation is possible thanks to the use of an "aperture" formulation of the problem, not being straightforward if an "obstacle" formulation was employed. Once the "internal" Π-circuit is known, the characterization of the stacked structure in Fig. 4.15(a) can simply be achieved by stacking the "internal" Π-circuits corresponding to each pair of adjacent gratings. Note that, although all the metal grids have been considered identical in the drawings, no special difficulty arises if the grids are different (the Π-circuit would become asymmetrical). More details about the physical interpretation of the series and parallel admittances of the internal Π-circuit are given in Ref. 94. Roughly speaking, the transmission line section used to represent the dielectric slab between gratings in the simplified circuit models[72,92] has to be supplemented with a new series admittance, connected in parallel with the transmission line section accounting for the zero-order mode (harmonic), which takes into account high-order mode interactions (see Fig. 4.15(d)). Usually these new admittances are frequency-independent capacitors (TM incidence) or inductors (TE incidence), although more complex circuit elements could be required. Thus, for instance, one or more of the high-order modes could operate above cutoff within a frequency range inside some of the dielectric slabs, even if the whole structure is operating in the non-diffraction regime. New transmission line sections accounting for these propagating modes should appear in parallel with the transmission line corresponding to the fundamental zero-order mode. In any case, all those admittances (corresponding to evanescent or propagating contributions) are obtained in analytical form and expressed as the sum of just a few modal terms.

It is worth mentioning that a similar formulation can be used to deal with fishnet-like 2D structures made of stacked 2D periodic distributions of small apertures in metal films. In that case, the interactions through both TM and TE low-order modes have to be simultaneously accounted for. This situation has also been treated in some detail in the recent literature.[95] In that work, it is clearly shown that the introduction in the model of the first high-order TE and TM modes is essential to explain the appearance of transmission bands (including the backward or negative-index transmission band) revealed by the full-wave analysis or by the experiment. The use of standard simplified models that ignore high-order mode (harmonic) interactions does not provide physically correct results.

In order to illustrate the accuracy of the modelling technique reported in this section, a comparison between our analytical circuit model data and numerical HFSS (finite elements method) results is shown in Fig. 4.16(a) for a stacked 1D structure made of four 1D periodic arrays of identical slits separated by three different dielectric layers. The agreement between the computationally intensive calculations based on HFSS and the analytical results given by the equivalent circuit model is excellent over the whole explored frequency band. The unit cell of a 2D version of this structure is schematically shown in Fig. 4.16(b), where four equally spaced extraordinary transmission metal layers are considered. The magnitude of the transmission coefficient computed for normal incidence with the electric field vertically polarized (along the y direction) is reported in Fig. 4.16(c) for two different cases: the perforated metal films are separated by air (top panel) or by a low dielectric constant material ($\varepsilon_r = 1.4$, bottom panel). In these plots, full-wave results computed with HFSS are compared with our circuit model. The curves labeled with EC(0,0) ignore the interaction between consecutive metal layers involving high-order modes. This situation corresponds to the most usual kind of circuit model reported in the literature. It can be seen that, for this specific choice of separation ($d = 0.6P$), only the first transmission band is roughly captured by the model. As discussed in Ref. 95, if the separation between metal layers was meaningfully smaller (say, $d = 0.2P$),

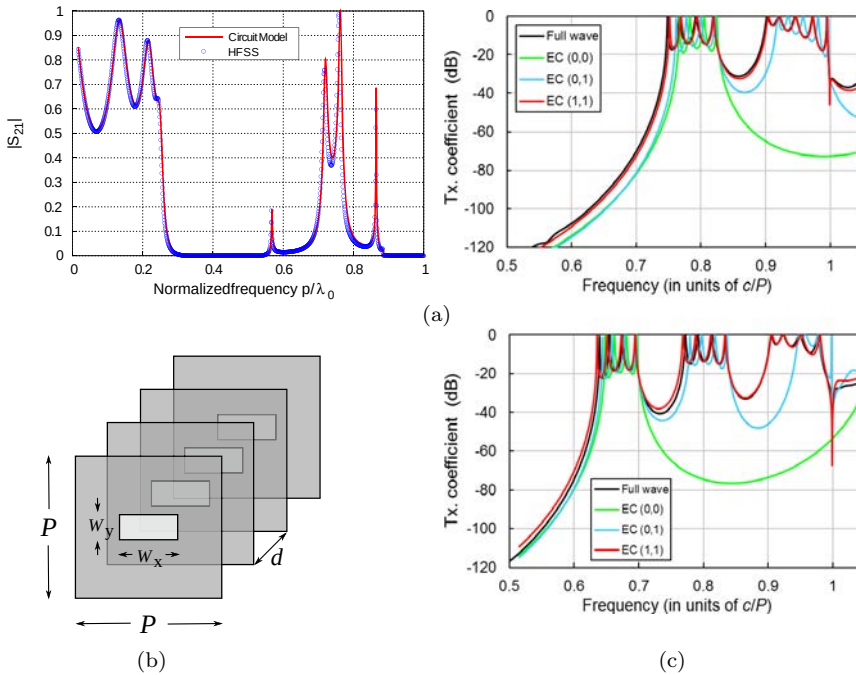


Fig. 4.16. (a) Transmission coefficient magnitude ($|S_{12}|$) for four stacked slit arrays (period p, slit width $w = 0.1p$) separated by three different dielectric slabs and placed in vacuum. Dimensions and electrical parameters: $d_1 = 0.4p$, $d_2 = 0.3p$, $d_3 = 0.2p$, $\varepsilon_r^{(0)} = 1$, $\varepsilon_r^{(1)} = 2.2$, $\varepsilon_r^{(2)} = 4.0$, $\varepsilon_r^{(3)} = 3.0$. The frequency is normalized (p/λ_0) and normal TM incidence is considered. Reprinted from Ref. 94 with APS permission. (b) Unit cell of a stacked 2D structure (fishnet) with a square lattice of period P and apertures of size $W_x = 0.4P$, $W_y = 0.2P$. The separation between perforated metal layers is $d = 0.6P$. (c) Magnitude of the transmission coefficient for the structure in (b) with air (top panel) or dielectric with $\varepsilon_r = 1.4$ (bottom panel) between metal layers. The curve labeled EC(0,0) does not consider high-order modes interactions, EC(0,1) includes the effect of the first high-order TM mode and EC(1,1) adds the first high-order TE mode. Reprinted from Ref. 95 with IEEE permission.

no transmission band would be predicted by this simplified model. However, when the first high-order TM mode is explicitly considered in the interaction, the case labeled as EC(0,1) in Fig. 4.16(c) shows that the results are qualitatively more consistent and one or two new transmission bands are clearly predicted. Finally, the inclusion of

the first high-order TE mode [EC(1,1) curves] provides a much better quantitative agreement between numerical and analytical data. Thus, the circuit model clearly shows the role of the first few high-order modes in the appearance of additional transmission bands, which are completely lost if the interaction through high-order modes is ignored. The computer code written on the basis of the theory reported in this chapter can add as many modes as necessary to properly account high-order modal interactions. More modes would be required when the distance between metal layers is reduced. In any case, the computational effort is always orders of magnitude smaller than that involved in purely numerical calculations.

4.7. Conclusions

The study of the electromagnetic response of periodic metalodielectric planar structures has been a topic of research interest since the first works on diffraction by gratings carried out more than two centuries ago. More recently, the control of electromagnetic waves from RF frequencies to optics using this kind of structure has motivated an explosive regrowth of this research field. The study of ET structures and metamaterials are good examples of the recent interest on the topic. Although powerful commercial full-wave electromagnetic solvers are nowadays available, the development of analytical models to deal with the analysis of this class of devices has obvious advantages from the computational point of view. Moreover, an analytical model will provide more physical insight than purely numerical approaches. Among the analytical models, those based on equivalent circuits have been very popular for decades due to their simplicity and easy implementation, apart from the familiar methodological and conceptual frame that it provides to electrical engineers and physicists. However, most of the equivalent circuit models proposed in the literature have serious limitations, especially when the dynamic frequency-dependent properties of the circuit elements cannot be ignored. In this chapter, a review of recent improvements of the circuit modelling simulation techniques that allow for the accurate analysis (and design) of a variety of periodic structures

that have attracted the interest of researchers working in the field of antennas and microwave engineering and in the field of metamaterials and ET-based structures has been provided.

Acknowledgments

The authors acknowledge the financial support of the Spanish MINECO with European Union FEDER funds (project TEC2013-41913-P) and of the Spanish CEICE of Junta de Andalucía (project P12-TIC-1435). The authors are also indebted to the many colleagues of other research groups involved in joint publications on this topic.

References

1. Ebbesen, T. W., Lezec, H. J., Ghaemi, H. F., Thio, T., and Wolff, P. A. (1998). Extraordinary optical transmission through sub-wavelength hole arrays, *Nature* **391**, pp. 667–669.
2. Bethe, H. A. (1944). Theory of diffraction by small holes, *Phys. Rev.* **66**(7/8), pp. 163–182.
3. Munk, B. A. (2000). *Frequency Selective Surfaces: Theory and Design* (John Wiley and Sons, New York).
4. Beruete, M., Sorolla, M., Campillo, I., Dolado, J. S., Martín-Moreno, L., Bravo-Abad, J. and García-Vidal, F. J. (2004). Enhanced millimeter-wave transmission through subwavelength hole arrays, *Opt. Lett.* **29**(21), pp. 2500–2502.
5. Beruete, M., Sorolla, M., Campillo, I., Dolado, J. S., Martín-Moreno, L., Bravo-Abad J. and García-Vidal F. J. (2005). Enhanced millimeter wave transmission through quasioptical subwavelength perforated plates, *IEEE Trans. Antennas Propag.* **53**(6), pp. 1897–1903.
6. Pendry, J. B., Martín-Moreno, L. and García-Vidal F. J. (2004). Mimicking surface plasmons with structured surfaces, *Science* **305**, pp. 847–848.
7. Hibbins, A. P., Evans, B. R. and Sambles, J. R. (2005). Experimental verification of designer surface plasmons, *Science* **308**, pp. 670–672.
8. Genet, C. and Ebbesen, T. W. (2007). Light in tiny holes, *Nature* **445**, pp. 39–46.
9. García-de-Abajo, F. J. (2007). Colloquium: Light scattering by particle and hole arrays, *Rev. Mod. Phys.* **79**, pp. 1267–1290.
10. García-Vidal, F. J., Martín-Moreno, L., Ebbesen, T. W. and Kuipers, L. (2010). Light passing through subwavelength apertures, *Rev. Mod. Phys.* **82**, pp. 729–787.

11. Gordon, R. Brolo, A. G., Sinton, D. and Kavanagh, K. L. (2010). Resonant optical transmission through hole-arrays in metal films: physics and applications, *Laser Phot. Rev.* **4**(2), pp. 311–335.
12. Schuchinsky, A. G., Zelenchuk, D. E. and Lerer, A. M. (2005). Enhanced transmission in microwave arrays of periodic sub-wavelength apertures, *J. Opt. A: Pure Appl. Opt.* **7**, pp. S102–S109.
13. Schuchinsky, A. G., Zelenchuk, D. E., Lerer, A. M. and Dickie, R. (2006). Full-wave analysis of layered aperture arrays, *IEEE Trans. Antennas Propag.* **54**(2), pp. 490–502.
14. Beruete, M., Campillo, I., Navarro-Cía, M., Falcone, F. and Sorolla, M. (2007). Molding left- or right-handed metamaterials by stacked cutoff metallic hole arrays, *IEEE Trans. Antennas Propag.* **55**(6), pp. 1514–1521.
15. Marcuvitz, N. (1951). *Waveguide Handbook* (McGraw-Hill Book Company, New York, USA).
16. Brown, J. (1950). The design of metallic delay dielectrics, *Proc. of the IEE-Part III: Radio and Com. Eng.* **97**(45), pp. 45–48.
17. Brown, J. (1953). Artificial dielectrics having refractive indices less than unity, *Proc. of the IEE-Part III: Radio and Com. Eng.* **100**(67), pp. 51–62.
18. Brown, J. and Jackson, W. (1955). The properties of artificial dielectrics at centimetre wavelengths, *Proc. of the IEE-Part B: Radio and Elec. Eng.* **102**(1), pp. 11–16.
19. Dubrovka, R., Vázquez J., Parini, C. and Moore, D. (2006). Equivalent circuit method for analysis and synthesis of frequency selective surfaces, *IEE Proc. pt. H, Microw. Antennas Propag.* **153**(3), pp. 213–220.
20. Monni, S., Gerini, G. and Neto, A. (2007). Multimode equivalent networks for the design and analysis of frequency selective surfaces, *IEEE Trans. Antennas Propag.* **55**(10), pp. 2824–2835.
21. Luukkonen, O., Simovski, C., Granet, G., Goussetis, G., Lioubtchenko, D., Raisanen A. V. and Tretyakov, S. A. (2008). Simple and accurate analytical model of planar grids and high-impedance surfaces comprising metal strips or patches, *IEEE Trans. Antennas Propag.* **56**(6), pp. 1624–1632.
22. Luukkonen, O., Costa, F., Simovski, C. R., Monorchio, A. and Tretyakov, S. A. (2009). A thin electromagnetic absorber for wide incidence angles and both polarizations, *IEEE Trans. Antennas Propag.* **57**(10), pp. 3119–3125.
23. Al-Joumayly, M. and Behdad, N. (2009). A new technique for design of low-profile, second-order, bandpass frequency selective surfaces, *IEEE Trans. Antennas Propag.* **57**(2), pp. 452–459.
24. Costa, F., Monorchio, A. and Manara, G. (2012). Efficient analysis of frequency-selective surfaces by a simple equivalent-circuit model, *IEEE Antennas Propag. Mag.* **54**(4), pp. 35–48.
25. Costa, F., Monorchio, A. and Manara, G. (2014). An overview of equivalent circuit modeling techniques of frequency selective surfaces and metasurfaces, *ACES Journal* **29**(12), pp. 960–976.

26. Medina, F., Mesa, F. and Marqués, R. (2008). Extraordinary transmission through arrays of electrically small holes from a circuit theory perspective, *IEEE Trans. Microw. Theory Techn.* **56**(12), pp. 3108–3120.
27. Porto, J. A., García-Vidal F. J. and Pendry, J. B. (2010). Transmission resonances on metallic gratings with very narrow slits. *Phys. Rev. Lett.* **83**, pp. 2845–2848.
28. Sologub, V. G., Schestopalov, V. P. and Polovnikov, G. G. (1967). Diffraction of electromagnetic waves on the grating with narrow slits, *J. Techn. Phys.* **37**(4), pp. 667–679.
29. Sologub, V. G. and Schestopalov, V. P. (1968). The resonance phenomena at the diffraction of the H-polarized waves on the gratings of metal bars, *J. Techn. Phys.* **38**(9), pp. 1505–1520.
30. Medina, F., Mesa, F. and Skigin, D. C. (2010). Extraordinary transmission through arrays of slits: a circuit theory model, *IEEE Trans. Microw. Theory Techn.* **58**(1), pp. 105–115.
31. Skigin, D. C. and Depine, R. A. (2005). Transmission resonances of metallic compound gratings with subwavelength slits, *Phys. Rev. Lett.* **95**, p. 217402.
32. Masalov, S. A., Sirenko, Y. K. and Schestopalov, V. P. (1978). The solution of the problem of the plane wave diffraction on the knife-type grating with compound structure of a period, *Radioteknica and Electronika.* **23**(3), pp. 481–487.
33. Molero, C., Rodríguez-Berral, R., Mesa, F. and Medina, F. (2016). Dynamical equivalent circuit for 1D periodic compound gratings, *IEEE Trans. Microw. Theory Techn.* **64**(4), pp. 1195–1208.
34. Yang, R., Rodríguez-Berral, R., Medina, F. and Hao, Y. (2011). Analytical model for the transmission of electromagnetic waves through arrays of slits in perfect conductors and lossy metal screens, *J. Appl. Phys.* **109**, p. 103107.
35. Pozar, D. M. (2005). *Microwave Engineering* (John Wiley & Sons Inc, New York, USA, 3rd end.)
36. Collin, R. E. (1991). *Field Theory of Guided Waves* (IEEE Press, Piscataway, New Jersey).
37. Mesa, F., García-Vigueras, M., Medina, F., Rodríguez-Berral, R. and Mosig, J. (2015). Circuit model analysis of frequency selective surfaces with scatterers of arbitrary geometry, *IEEE Antennas Wireless Propag. Lett.* **14**, pp. 135–138.
38. Ruan Z. and Qiu, M. (2006). Enhanced transmission through periodic arrays of subwavelength holes: The role of localized waveguide resonances, *Phys. Rev. Lett.* **96**, p. 233901.
39. Martín-Moreno, L., García-Vidal, F. J., Lezec, H. J., Pellerin, K. M., Thio, T., Pendry, J. B. and Ebbesen, T. W. (2001). Theory of extraordinary optical transmission through subwavelength hole arrays, *Phys. Rev. Lett.* **86**, pp. 1114–1117.
40. Martín-Moreno, L. and García-Vidal, F. J. (2004). Optical transmission through circular hole arrays in optically thick metal films, *Opt. Express* **12**(16), pp. 3619–3628.

41. Don, N. G. and Kirilenko, A. A. (2007). Eigenoscillations of plane junctions and total transmission through small holes. In *Proc. 37th Eur. Microw. Conf.*, pp. 24–27, Munich (Germany).
42. Don, N. G., Kirilenko, A. A., and Senkevich, S. L. (2008). New type of eigenoscillations and total-transmission resonance through an iris with below-cutoff hole in a rectangular waveguide, *Radiophys. Quantum Electron.* **51**, p. 101.
43. Kirilenko, A. A. and Perov, A. O. (2008). On the common nature of the enhanced and resonance transmission through the periodical set of holes, *IEEE Trans. Antennas Propag.* **56**(10), pp. 3210–3216.
44. Pang, Y., Hone, A. N., So, P. P. M. and Gordon, R. (2009). Total optical transmission through a small hole in a metal waveguide screen Bethe's aperture, *Opt. Express* **17**(6), pp. 4433–4441.
45. Medina, F., Ruiz-Cruz, J. A., Mesa, F., Rebollar, J. M., Montejo-Garai, J. R. and Marqués R. (2009). Experimental verification of extraordinary transmission without surface plasmons, *Appl. Phys. Lett.* **95**, p. 071102.
46. Medina, F., Mesa, F., Ruiz-Cruz, J. A., Rebollar, J. M. and Montejo-Garai J. R. (2010). Study of extraordinary transmission in a circular waveguide system, *IEEE Trans. Microw. Theory Techn.* **58**(6), pp. 1532–1542.
47. MacPhie, R. H. and Wu, K. L. (1995). Scattering at the junction of a rectangular waveguide and a larger circular waveguide, *IEEE Trans. Microw. Theory Techn.* **43**(9), pp. 2041–2045.
48. Chan, K.-L. and Judah, S. (1997). Two port scattering at an elliptical-waveguide junction, *IEEE Trans. Microw. Theory Techn.* **45**(8), pp. 1255–1262.
49. Khavasi A. and Mehrany, K. (2011). Circuit model for lamellar metallic gratings in the sub-wavelength regime, *IEEE J. Quantum Elect.* **47**, pp. 1330–1335.
50. Varela, J. E. and Esteban, J. (2012). Characterization of waveguides with a combination of conductor and periodic boundary contours: application to the analysis of bi-periodic structures, *IEEE Trans. Microw. Theory Techn.* **60**(3), pp. 419–430.
51. Alu, A., D'Aguanno, G., Mattiucci, N. and Bloemer, M. J. (2011). Plasmonic Brewster angle: broadband extraordinary transmission through optical gratings, *Phys. Rev. Lett.* **106**, p. 123902.
52. Woo, D. W., Muhn, S. J. and Park, W. S. (2012). Simple analytical model of propagation through thick periodic slot, *IEEE Trans. Antennas and Propag.* **60**(11), pp. 5329–5335.
53. Yarmoghaddam, E., Shirmanesh, G. K., Khavasi, A. and Mehrany, K. (2014). Circuit model for a periodic array of slits with multiple propagating diffracted orders, *IEEE Trans. Antennas Propag.* **62**, pp. 4041–4048.
54. Hibbins, A. P., Hooper, I. R., Lockyear, M. I. and Sambles, J. R. (2006). Microwave transmission of a compound gratings, *Phys. Rev. Lett.* **96**, p. 257402.

55. Navarro-Cía, M., Skigin, D. C., Beruete, M. and Sorolla, M. (2009). Experimental demonstration of phase resonances in metallic compound gratings with subwavelength slits in the milimeter wave regime, *Appl. Phys. Lett.* **94**, p. 091107.
56. Fantino, A. N., Grosz, S. I. and Skigin, D. C. (2001). Resonant effects in periodic gratings comprising a finite number of grooves in each period, *Phys. Rev. E* **64**, p. 016605.
57. Vardaxoglou, J. C. (1997). *Frequency selective surfaces: Analysis and Design* (Research Studies Press).
58. Lomakin, V. and Michielssen, E. (2005). Enhanced transmission through metallic plates perforated by arrays of subwavelength holes and sandwiched between dielectric slabs, *Phys. Rev. B* **71**, p. 235117.
59. Moreno, E., Martín-Moreno, L. and García-Vidal, F. J. (2006). Extraordinary optical transmission without plasmons: the s-polarization case, *J. Opt. A: Pure Appl. Opt.* **8**, pp. S94–S97.
60. Lomakin V. and Michielssen, E. (2007). Beam transmission through periodic subwavelength hole structures, *IEEE Trans. Antennas Propag.* **55**(6), pp. 1564–1581.
61. Ortuño, R., García-Meca, C., Rodríguez-Fortu no, F. J., Martí, J. and Martínez, A. (2010). Multiple extraordinary optical transmission peaks from evanescent coupling in perforated metal plates surrounded by dielectrics, *Opt. Express* **18**(8), pp. 7893–7898.
62. Beruete, M., Navarro-Cía, M., Kuznetsov, S. A. and Sorolla, M. (2011). Circuit approach to the minimal configuration of terahertz anomalous extraordinary transmission, *Appl. Phys. Lett.* **98**, p. 014106.
63. Beruete, M., Navarro-Cía and Sorolla, M. (2011). Understanding anomalous extraordinary transmission from equivalent circuit and grounded slab concepts, *IEEE Trans. Microw. Theory Techn.* **59**, pp. 2180–2188.
64. Guglielmi, M. and Oliner, A. A. (1989). Multimode network description of a planar periodic metal-strip grating at a dielectric interface — part I: Rigorous network formulations, *IEEE Trans. Microw. Theory Techn.* **37**(3), pp. 535–541.
65. Guglielmi M. and Oliner, A. A. (1989). Multimode network description of a planar periodic metal-strip grating at a dielectric interface — part II: Small-aperture and small-obstacle solutions, *IEEE Trans. Microw. Theory Techn.* **37**(3), pp. 542–552.
66. Ulrich, R. (1967). Far-infrared properties of metallic mesh and its complementary structure, *Infrared Phys.* **7**(1), pp. 37–55.
67. Ulrich, R. (1967). Effective low-pass filters for far infrared frequencies, *Infrared Phys.* **7**(2), pp. 65–74.
68. Irwin, P. G. J., Ade, P. A. R., Calcutt, S. B., Taylor, F. W., Seeley, J. S., Hunneman, R. and Walton, L. (1993). Investigation of dielectric spaced resonant mesh filter designs for PMIRR, *Infrared Phys.* **34**(6), pp. 549–563.

69. Cunningham, C. T. (1983). Resonant grids and their use in the construction of submillimeter filters, *Infrared Phys.* **23**(4), pp. 207–215.
70. Bayatpur, F. and Sarabandi, K. (2008). Single-layer higher-order miniaturized-element frequency-selective surfaces, *IEEE Trans. Microw. Theory Techn.* **56**(4), pp. 774–781.
71. Bayatpur, F. and Sarabandi, K. (2008). Multipole spatial filters using metamaterial-based miniaturized-element frequency-selective surfaces, *IEEE Trans. Microw. Theory Techn.* **56**(12), pp. 2742–2747.
72. Kaipa, C. S. R., Yakovlev, A. B., Medina, F., Mesa, F., Butler, C. A. M. and Hibbins, A. P. (2010). Circuit modeling of the transmissivity of stacked two-dimensional metallic meshes, *Opt. Express.* **18**(13), pp. 13309–13320.
73. Rodríguez-Berral, R., Medina, F. and Mesa, F. (2010). Circuit model for a periodic array of slits sandwiched between two dielectric slabs, *Appl. Phys. Lett.* **96**, p. 161104.
74. Rodríguez-Berral, R., Medina, F., Mesa, F. and García-Vigueras, M. (2012). Quasi-analytical modeling of transmission/reflection in strip/slit gratings loaded with dielectric slabs, *IEEE Trans. Microw. Theory Techn.* **60**(3), pp. 405–418.
75. Rodríguez-Berral R., Molero, C., Medina, F. and Mesa, F. (2012). Analytical wideband model for strip/slit gratings loaded with dielectric slabs, *IEEE Trans. Microw. Theory Techn.* **60**(12), pp. 3908–3918.
76. García-Vigueras, M., Mesa, F., Medina, F., Rodríguez-Berral, R. and Gómez-Tornero, J. L. (2012). Simplified circuit model for metallic arrays of patches sandwiched between dielectric slabs under arbitrary incidence, *IEEE Trans. Antennas Propag.* **60**(10), pp. 4637–4649.
77. Rodríguez-Berral, R., Mesa, F. and Medina, F. (2015). Analytical multimodal network approach for 2-D arrays of planar patches/apertures embedded in a layered medium, *IEEE Trans. Antennas Propag.* **63**(5), pp. 1969–1984.
78. Florencio, R., Boix, R. R. and Encinar, J. A. (2013). Enhanced mom analysis of the scattering by periodic strip gratings in multilayered substrates, *IEEE Trans. Antennas Propag.* **61**(10), pp. 5088–5099.
79. Kehn, N. N. M. (2012). Moment method analysis of plane-wave scattering from planar corrugated surfaces using parallel-plate cavity Green's function and derivation of analytic reflection-phase formulas for both polarizations and oblique azimut planes, *Radio Sci.* **47**, p. RS3008.
80. Kehn, N. N. M. (2013). Rapid surface-wave dispersion and plane-wave reflection analyses of planar corrugated surfaces by asymptotic corrugations boundary conditions even for oblique azimuth planes, *IEEE Trans. Antennas Propag.* **61**(5), pp. 2695–2707.
81. Molero, C., Rodríguez-Berral, R., Mesa, F. and Medina, F. (2014). Analytical circuit model for 1-D periodic T-shaped corrugated surfaces, *IEEE Trans. Antennas Propag.* **62**(2), pp. 794–803.
82. Maci, S., Caiazzo, M., Cucini, A. and Casaletti, M. (2005). A pole-zero matching method for EBG surfaces composed of a dipole fss printed on a grounded dielectric slab, *IEEE Trans. Antennas Propag.* **53**(1), pp. 70–81.

83. Mesa, F., Rodríguez-Berral, R., García-Vigueras, M., Medina, F. and Mosig, J. R. (2016). A pole-zero matching method for EBG surfaces composed of a dipole fss printed on a grounded dielectric slab, *IEEE Trans. Antennas Propag.* **64**(3), pp. 1106–1111.
84. Kafesaki, M., Tsiapa, I., Katsarakis, N., Koschny, T., Soukoulis, C. M. and Economou, E. N. (2007). Left-handed metamaterials: The fishnet structure and its variations, *Phys. Rev. B* **75**(23), p. 235114.
85. Butler, C. A. M., Parsons, J., Sambles, J. R., Hibbins, A. P. and Hobson, P. A. (2009). Microwave transmissivity of a metamaterial-dielectric stack, *Appl. Phys. Lett.* **95**, p. 174101.
86. Cavallo, D., Syed, W. H. and Neto, A. (2014). Closed-form analysis of artificial dielectric layers-part i: Properties of a single layer under plane-wave incidence, *IEEE Trans. Antennas Propag.* **62**, pp. 6256–6264.
87. Cavallo, D., Syed, W. H. and Neto, A. (2014). Closed-form analysis of artificial dielectric layers-part ii: Extension to multiple layers and arbitrary illumination, *IEEE Trans. Antennas Propag.* **62**, pp. 6265–6273.
88. Kharadly M. M. Z. and Jackson, W. (1953). The properties of artificial dielectrics comprising arrays of conducting elements, *Proc. of the IEE-Part III: Radio and Com. Eng.* **100** (66), pp. 199–212.
89. Marqués, R., Jelinek, L., Mesa, F. and Medina, F. (2009). Analytical theory of wave propagation through stacked fishnet metamaterials, *Opt. Express* **17**, p. 11582.
90. Carbonell, J., Croenne, C., Garet, F., Lheurette, E., Coutaz, J. L. and Lippens, D. (2010). Lumped elements circuit of terahertz fishnet-like arrays with composite dispersion, *J. Appl. Phys.* **108**(1), p. 014907.
91. Padooru, Y. R., Yakovlev, A. B., Kaipa, C. S. R., Medina, F. and Mesa, F. (2011). Circuit modeling of multi-band high-impedance surface absorbers in the microwave regime, *Phys. Rev. B.* **84**(3), p. 035108.
92. Kaipa, C. S. R., Yakovlev, A. B., Medina, F. and Mesa, F. (2012). Transmission through stacked 2-D periodic distributions of square conducting patches, *J. Appl. Phys.* **112**, p. 033101.
93. Padooru, Y. R., Yakovlev, A. B., Kaipa, C. S. R., Hanson, G. W., Medina, F. and Mesa, F. (2013). Dual capacitive-inductive nature of periodic graphene patches: transmission characteristics at low-terahertz frequencies, *Phys. Rev. B* **87**(11), p. 115401.
94. Molero, C., Rodríguez-Berral, R., Mesa, F., Medina, F. and Yakovlev A. B. (2016). Wideband analytical equivalent circuit for 1-D periodic stacked arrays, *Phys. Rev. E* **93**, p. 013306.
95. Torres, V., Mesa, F., Navarro-Cía, M., Rodríguez-Berral, R., Beruete, M. and Medina, F. (2016). Accurate circuit modeling of fishnet structures for negative-index-medium applications, *IEEE Trans. Microw. Theory Techn.* **64**(1), pp. 15–26.

CHAPTER 5

Electromagnetic Metasurfaces: Synthesis, Realizations and Discussions

KARIM ACHOURI* **and CHRISTOPHE CALOZ†**

École Polytechnique de Montréal, Canada
*karim.achouri@polymtl.ca
†christophe.cal2@polymtl.ca

5.1. Introduction

Metasurfaces[1–4] are dimensional reductions of volume metamaterials and functional extensions of frequency selective surfaces.[5] They are composed of 2D arrays of subwavelength scattering particles engineered in such a manner that they transform incident waves into desired reflected and transmitted waves. Compared to volume metamaterials, metasurfaces offer the advantage of being lighter, easier to fabricate and less lossy due to their reduced dimensionality, and while compared to frequency selective surfaces, they provide greater flexibility and functionalities.

In this chapter, we propose a framework that may be used as a rigorous synthesis technique for metasurfaces and which can be divided into two main steps. Firstly, a mathematical description of the metasurface must be obtained by solving an inverse problem

where all the fields around the metasurface (incident, reflected and transmitted fields) are known and the electromagnetic properties, at each point of the metasurface, must be found. The mathematical synthesis technique must be as universal as possible meaning that the specified fields can have arbitrary amplitude, phase, polarization and direction of propagation.

Secondly, the metasurface electromagnetic transfer function (obtained from the first step of the synthesis) is discretized into unit cells corresponding to each lattice site. Each of these unit cells must then be physically implemented in order to build the final metasurface. Therefore, this second part of the synthesis technique consists in finding appropriate and convenient design rules to implement the scattering particles. This second step is usually the most complicated as no universal and systematic techniques exist so far to relate the shape of scattering particles to their effective parameters (or transmission and reflection coefficients), and such analysis is usually performed via parametric simulations of well-chosen structures.[6–14]

To date, only a few metasurface synthesis techniques have been reported in the literature. In this chapter, we will follow the synthesis procedure developed in Ref. 4 which yields the metasurface susceptibility tensors for specified electromagnetic transformations. Alternatively, the methods proposed by Grbic[15–18] and Eleftheriades[19–21] describe the metasurface in terms of impedance tensors, and it can be shown that they represents only a particular case of the general method proposed here. The method proposed by Tretyakov[22,23] relates the waves reflected and transmitted from the metasurface to the polarizabilities of a single scattering element in the case of a normally incident plane wave. In contrast, the method proposed here deals with waves of arbitrary incident angles and arbitrary types. Finally, the method proposed by Salem,[24,25] is a technique called the momentum transformation method, which is a spectral ($\boldsymbol{k}$) method, that is particularly suitable for paraxial wave problems. It can also handle full vectorial problems, but this involves extra complexity compared to the scalar case.

5.2. Metasurface Mathematical Synthesis

5.2.1. *Metasurface synthesis based on susceptibility tensors*

A metasurface is an electromagnetic 2D structure with subwavelength thickness ($\delta \ll \lambda$). The metasurface may be finite, with dimensions $L_x \times L_y$, or infinite. It is typically composed of a non-uniform arrangement of planar scattering particles (full or slotted patches, straight or curved strips, various types of crosses, dielectric resonators, etc.) that transforms incident waves into specified reflected and transmitted waves.

Figure 5.1 shows the synthesis problem to solve. How can one synthesize a metasurface that transforms an arbitrary specified incident wave, $\boldsymbol{\psi}^{\mathrm{i}}(\boldsymbol{r})$, into an arbitrary specified reflected wave, $\boldsymbol{\psi}^{\mathrm{r}}(\boldsymbol{r})$, and an arbitrary specified transmitted wave, $\boldsymbol{\psi}^{\mathrm{t}}(\boldsymbol{r})$, assuming monochromatic waves? Here, the solution will be expressed in terms of the transverse susceptibility tensor functions of $\boldsymbol{\rho} = x\hat{\boldsymbol{x}}+y\hat{\boldsymbol{y}}, \overline{\overline{\chi}}_{\mathrm{ee}}(\boldsymbol{\rho}), \overline{\overline{\chi}}_{\mathrm{mm}}(\boldsymbol{\rho}), \overline{\overline{\chi}}_{\mathrm{em}}(\boldsymbol{\rho})$

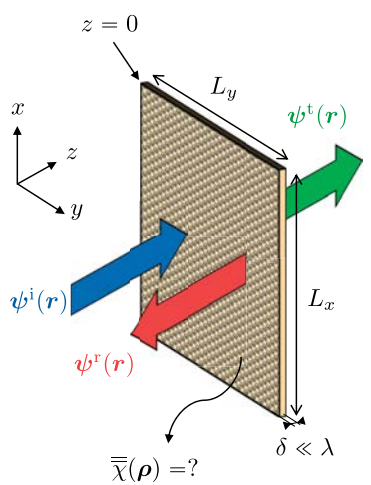


Fig. 5.1. Metasurface synthesis (inverse) problem to solve. A metasurface, generally defined as an electromagnetic 2D non-uniform structure of extent $L_x \times L_y$ with sub-wavelength thickness ($\delta \ll \lambda$), is placed at $z = 0$. Determine the surface susceptibility tensors $\overline{\overline{\chi}}(\boldsymbol{\rho})$ of the metasurface transforming an arbitrary specified incident wave $\boldsymbol{\psi}^{\mathrm{i}}(\boldsymbol{r})$ into an arbitrary specified reflected wave $\boldsymbol{\psi}^{\mathrm{r}}(\boldsymbol{r})$ and an arbitrary specified transmitted wave $\boldsymbol{\psi}^{\mathrm{t}}(\boldsymbol{r})$.

and $\overline{\overline{\chi}}_{\mathrm{me}}(\boldsymbol{\rho})$, which represent the electric/magnetic (e/m) transverse polarization responses (first subscript) to transverse electric/magnetic (e/m) field excitations (second subscript).

The synthesis procedure will always yield $\overline{\overline{\chi}}_{\mathrm{ee}}(\boldsymbol{\rho})$, $\overline{\overline{\chi}}_{\mathrm{mm}}(\boldsymbol{\rho})$, $\overline{\overline{\chi}}_{\mathrm{em}}(\boldsymbol{\rho})$ and $\overline{\overline{\chi}}_{\mathrm{me}}(\boldsymbol{\rho})$ results, but will not guarantee that these results can be practically implemented using planar scattering particles. For instance, if the susceptibilities exhibit multiple spatial variations per wavelength, they may be difficult or impossible to realize. In such cases, one has to determine whether some features may be neglected or one may have to relax the design constraints (e.g. allow higher reflection or increase the metasurface dimensions).

The complete synthesis of a metasurface typically consists of two steps: (1) determination of the mathematical transfer function of the metasurface producing the specified fields, which is generally a continuous function of the transverse dimensions of the metasurface; (2) discretization of the transfer function obtained in (1) according to a 2D lattice and determination of the scattering particles realizing the corresponding transfer function at each lattice site. Step (2) involves a full-wave parametric analysis of judiciously selected scattering particles, from which magnitude and phase maps are established to find the appropriate particle geometries for building the metasurface using the periodic boundary condition approximation.[26]

5.2.2. *Metasurface boundary conditions in terms of surface susceptibility tensors*

A metasurface may be considered as an electromagnetic *discontinuity* in space. Conventional textbook boundary conditions do not apply to such a discontinuity. As was pointed out by Schelkunoff,[27] the mathematical formulation of such conventional boundary conditions is not rigorous in the case of field discontinuities caused by sources such as surface charges and currents, although it yields satisfactory results away from the discontinuities. Assuming an interface at $z = 0$, the conventional boundary conditions relate the fields at $z = 0^{\pm}$, but fail to describe the field behavior at the discontinuity itself ($z = 0$). This discrepancy is due to the fact that Stokes and Gauss theorems

used to derive them assume field continuity in all the regions they apply to, including the interface, whereas the fields may be discontinuous due to the presence of sources. For instance, consider the conventional boundary condition for the normal component of the displacement vector $\boldsymbol{D}$ in the presence of surface charges ρ_s,

$$\hat{\boldsymbol{z}} \cdot \boldsymbol{D}|_{z=0^-}^{0^+} = \rho_s. \tag{5.1}$$

This relation is derived by applying Gauss theorem, $\iiint_V \nabla \cdot \boldsymbol{D} dV = \oiint_S \boldsymbol{D} \cdot \hat{\mathbf{n}} dS$, to a volume V enclosed by the surface S including the interface discontinuity with $\hat{\mathbf{n}}$ the normal unit vector to S. This theorem rigorously applies only if $\boldsymbol{D}$ is continuous inside the entire volume V, whereas in the case of a discontinuous $\boldsymbol{D}$, as in (5.1), its projection onto S is not defined at the interface and application of this theorem is not rigorously correct. Thus, since a metasurface is equivalent to an array of Huygens sources,[15] the correct field behavior on the metasurface cannot be determined using the conventional boundary conditions, and rigorous boundary conditions, namely the generalized sheet transition conditions (GSTCs), must be applied, as will be done next. It should be noted that, from a physical perspective, a metasurface structure is not a single interface but rather a thin inhomogeneous slab, and may be naturally treated as such. However, it is much simpler to treat the metasurface as a single interface using rigorous GSTCs, which is allowed by the fact that it is electromagnetically thin.

Rigorous GSTCs, treating discontinuities in the sense of distributions, were derived by Idemen.[28] The corresponding relations pertaining to this work, first applied by Kuester *et al.* to metasurfaces,[1] are derived in section A for the sake of clarity and completeness. They may be written as[a]

$$\hat{z} \times \Delta \boldsymbol{H} = j\omega \boldsymbol{P}_{\parallel} - \hat{z} \times \nabla_{\parallel} M_z, \tag{5.2a}$$

[a]Throughout this chapter, the medium surrounding the metasurface is assumed to be vacuum, with permittivity and permeability ϵ_0 and μ_0, respectively. Therefore, for notational compactness, and also to avoid confusion between the subscript '0' meaning 'vacuum' or meaning 'first-order discontinuity' (section A), the subscript '0' for 'vacuum' is suppressed everywhere.

$$\Delta \boldsymbol{E} \times \hat{z} = j\omega\mu \boldsymbol{M}_{\parallel} - \nabla_{\parallel}\left(\frac{P_z}{\epsilon}\right) \times \hat{z}, \tag{5.2b}$$

$$\hat{z} \cdot \Delta \boldsymbol{D} = -\nabla \cdot \boldsymbol{P}_{\parallel}, \tag{5.2c}$$

$$\hat{z} \cdot \Delta \boldsymbol{B} = -\mu \nabla \cdot \boldsymbol{M}_{\parallel}. \tag{5.2d}$$

In these relations, the terms in the left-hand sides represent the differences between the fields on the two sides of the metasurface, whose cartesian components are defined as

$$\Delta \Psi_u = \hat{\boldsymbol{u}} \cdot \Delta \boldsymbol{\Psi}(\boldsymbol{\rho})\Big|_{z=0^-}^{0^+} = \Psi_u^{\mathrm{t}} - (\Psi_u^{\mathrm{i}} + \Psi_u^{\mathrm{r}}),\ u = x, y, z, \tag{5.3}$$

where $\boldsymbol{\Psi}(\boldsymbol{\rho})$ represents any of the fields $\boldsymbol{H}$, $\boldsymbol{E}$, $\boldsymbol{D}$ or $\boldsymbol{B}$, and where the superscripts i, r, and t denote incident, reflected and transmitted fields, and $\boldsymbol{P}$ and $\boldsymbol{M}$ are the electric and magnetic surface polarization densities, respectively. They are expressed in terms of the average fields on both sides of the metasurface and read

$$\boldsymbol{P} = \epsilon \overline{\overline{\chi}}_{\mathrm{ee}} \cdot \boldsymbol{E}_{\mathrm{av}} + \overline{\overline{\chi}}_{\mathrm{em}} \sqrt{\mu\epsilon} \cdot \boldsymbol{H}_{\mathrm{av}}, \tag{5.4a}$$

$$\boldsymbol{M} = \overline{\overline{\chi}}_{\mathrm{mm}} \cdot \boldsymbol{H}_{\mathrm{av}} + \overline{\overline{\chi}}_{\mathrm{me}} \sqrt{\frac{\epsilon}{\mu}} \cdot \boldsymbol{E}_{\mathrm{av}}, \tag{5.4b}$$

where the average fields are defined as

$$\Psi_{u,\mathrm{av}} = \hat{\boldsymbol{u}} \cdot \boldsymbol{\Psi}_{\mathrm{av}}(\boldsymbol{\rho}) = \frac{\Psi_u^{\mathrm{t}} + (\Psi_u^{\mathrm{i}} + \Psi_u^{\mathrm{r}})}{2}, \quad u = x, y, z, \tag{5.5}$$

where $\boldsymbol{\Psi}(\boldsymbol{\rho})$ represents either $\boldsymbol{H}$ or $\boldsymbol{E}$. Note that the utilization of susceptibilities, which represent the actual macroscopic quantities of interest, allows for an easier description of the metasurface than, for instance, particle polarizabilities.

The surface may be infinite or finite with dimensions $L_x \times L_y$. The two problems are automatically solved by specifying the fields Ψ_u^{i}, Ψ_u^{r} and Ψ_u^{t} in (5.3) and (5.5) to be of infinite or finite $L_x \times L_y$ extent in the former and latter cases, respectively. In the finite case, truncation practically corresponds to placing a sheet of absorbing material around the metasurface. This operation neglects diffraction at the edges of the metasurface, as is safely allowed by the fact that a

metasurface is generally electrically very large, but properly accounts for the finiteness of the aperture via the GSTCs (5.2a) and (5.2b).

5.2.3. *Synthesis method*

5.2.3.1. *Assumptions*

The proposed synthesis method solves the inverse problem depicted in Fig. 5.1, where the electromagnetic fields are specified everywhere (for all $\boldsymbol{\rho}$) in the $z = 0$ plane on both sides of the metasurface and the properties of the metasurface are the unknowns to be determined. We specifically aim at finding the susceptibilities that transform specified incident waves into specified transmitted and reflected waves. The method essentially consists of solving Eqs. (5.2) for the components of the susceptibility tensors in (5.4).

The last terms in (5.2a) and (5.2b) involve the transverse derivatives of the normal components of the polarization densities, namely $\nabla_{\parallel} M_z$ and $\nabla_{\parallel} P_z$. Solving the inverse problem for non-zero M_z and/or P_z would be quite involved since this would require solving the set of coupled non-homogenous partial differential equations formed by (5.2a) and (5.2b) with non-zero $\nabla_{\parallel} M_z$ and $\nabla_{\parallel} P_z$. Although such a problem could be generally addressed by means of numerical analysis, we enforce here $P_z = M_z = 0$, which will lead to convenient closed-form solutions for the susceptibilities.[b] As shall be seen next, this restriction still allows the metasurface to realize a large number of operations, given the large number of degrees of freedom provided by combinations of its bianisotropic susceptibility tensor components. Note that the problem of non-zero normal polarization densities has been addressed in, Ref. 4 allowing the proposed method to also handle the most general case including normal components, which may be of practical interest (e.g. conducting rings in the metasurface plane, producing M_z contributions).

[b]This restriction may limit the physical realizability of the metasurface in some cases, in the sense that the corresponding synthesized susceptibilities might be excessively difficult to realize with practical scattering particles. In such cases, the restriction might be removed without changing the main spirit of the method but at the cost of losing the closed-form nature of the solution.

The method needs considering only (5.2a) and (5.2b) as these two equations involve all the transverse field components, which is sufficient to completely describe the fields at each side of the metasurface according to the uniqueness theorem. These two equations, with $P_z = M_z = 0$, represent a set a four linear equations relating the transverse electric and magnetic fields to the effective surface susceptibilities.

5.2.3.2. *General solution for surface susceptibilities*

As mentioned above, the four susceptibility tensors in (5.4) are restricted to their four transverse components, and these components will be determined for the specified fields using (5.2a) and (5.2b) with $P_z = M_z = 0$ and using the notation in (5.3) and (5.5). The metasurface continuity conditions read

$$\hat{z} \times \Delta \boldsymbol{H} = j\omega\epsilon \overline{\overline{\chi}}_{\mathrm{ee}} \cdot \boldsymbol{E}_{\mathrm{av}} + jk \overline{\overline{\chi}}_{\mathrm{em}} \cdot \boldsymbol{H}_{\mathrm{av}}, \tag{5.6a}$$

$$\Delta \boldsymbol{E} \times \hat{z} = j\omega\mu \overline{\overline{\chi}}_{\mathrm{mm}} \cdot \boldsymbol{H}_{\mathrm{av}} + jk \overline{\overline{\chi}}_{\mathrm{me}} \cdot \boldsymbol{E}_{\mathrm{av}}. \tag{5.6b}$$

This system can also be written in matrix form to simplify the synthesis procedure. The matrix equivalent of (5.6) is given by

$$\begin{pmatrix} \Delta H_y \\ \Delta H_x \\ \Delta E_y \\ \Delta E_x \end{pmatrix} = \begin{pmatrix} \widetilde{\chi}_{\mathrm{ee}}^{xx} & \widetilde{\chi}_{\mathrm{ee}}^{xy} & \widetilde{\chi}_{\mathrm{em}}^{xx} & \widetilde{\chi}_{\mathrm{em}}^{xy} \\ \widetilde{\chi}_{\mathrm{ee}}^{yx} & \widetilde{\chi}_{\mathrm{ee}}^{yy} & \widetilde{\chi}_{\mathrm{em}}^{yx} & \widetilde{\chi}_{\mathrm{em}}^{yy} \\ \widetilde{\chi}_{\mathrm{me}}^{xx} & \widetilde{\chi}_{\mathrm{me}}^{xy} & \widetilde{\chi}_{\mathrm{mm}}^{xx} & \widetilde{\chi}_{\mathrm{mm}}^{xy} \\ \widetilde{\chi}_{\mathrm{me}}^{yx} & \widetilde{\chi}_{\mathrm{me}}^{yy} & \widetilde{\chi}_{\mathrm{mm}}^{yx} & \widetilde{\chi}_{\mathrm{mm}}^{yy} \end{pmatrix} \begin{pmatrix} E_{x,\mathrm{av}} \\ E_{y,\mathrm{av}} \\ H_{x,\mathrm{av}} \\ H_{y,\mathrm{av}} \end{pmatrix}, \tag{5.7}$$

where the tilde indicates that the susceptibilities in (5.6) have been multiplied by some constant values. The relationship between the susceptibilities in (5.6) and those in (5.7) is

$$\begin{pmatrix} \chi_{\mathrm{ee}}^{xx} & \chi_{\mathrm{ee}}^{xy} & \chi_{\mathrm{em}}^{xx} & \chi_{\mathrm{em}}^{xy} \\ \chi_{\mathrm{ee}}^{yx} & \chi_{\mathrm{ee}}^{yy} & \chi_{\mathrm{em}}^{yx} & \chi_{\mathrm{em}}^{yy} \\ \chi_{\mathrm{me}}^{xx} & \chi_{\mathrm{me}}^{xy} & \chi_{\mathrm{mm}}^{xx} & \chi_{\mathrm{mm}}^{xy} \\ \chi_{\mathrm{me}}^{yx} & \chi_{\mathrm{me}}^{yy} & \chi_{\mathrm{mm}}^{yx} & \chi_{\mathrm{mm}}^{yy} \end{pmatrix} = \begin{pmatrix} \frac{j}{\omega\epsilon}\widetilde{\chi}_{\mathrm{ee}}^{xx} & \frac{j}{\omega\epsilon}\widetilde{\chi}_{\mathrm{ee}}^{xy} & \frac{j}{k}\widetilde{\chi}_{\mathrm{em}}^{xx} & \frac{j}{k}\widetilde{\chi}_{\mathrm{em}}^{xy} \\ -\frac{j}{\omega\epsilon}\widetilde{\chi}_{\mathrm{ee}}^{yx} & -\frac{j}{\omega\epsilon}\widetilde{\chi}_{\mathrm{ee}}^{yy} & -\frac{j}{k}\widetilde{\chi}_{\mathrm{em}}^{yx} & -\frac{j}{k}\widetilde{\chi}_{\mathrm{em}}^{yy} \\ -\frac{j}{k}\widetilde{\chi}_{\mathrm{me}}^{xx} & -\frac{j}{k}\widetilde{\chi}_{\mathrm{me}}^{xy} & -\frac{j}{\omega\mu}\widetilde{\chi}_{\mathrm{mm}}^{xx} & -\frac{j}{\omega\mu}\widetilde{\chi}_{\mathrm{mm}}^{xy} \\ \frac{j}{k}\widetilde{\chi}_{\mathrm{me}}^{yx} & \frac{j}{k}\widetilde{\chi}_{\mathrm{me}}^{yy} & \frac{j}{\omega\mu}\widetilde{\chi}_{\mathrm{mm}}^{yx} & \frac{j}{\omega\mu}\widetilde{\chi}_{\mathrm{mm}}^{yy} \end{pmatrix}. \tag{5.8}$$

Assuming single incident, reflected and transmitted wave set (only one wave of each of the three types), the system (5.7) contains 4 equations for 16 unknown susceptibility components. It is thus *underdetermined* as such, and it can be solved only by restricting the number of *independent* susceptibilities to 4. This single-transformation underdetermination reveals two important facts: (i) *Many different combinations of susceptibilities produce the same fields*; (ii) A metasurface has the *fundamental capability to simultaneously manipulate several linearly independent incident, reflected and transmitted waves*. Specifically, a metasurface, as defined by (5.7), can in principle manipulate *up to 4 sets of incident, reflected and transmitted waves*. If T ($T = 1, 2, 3, 4$) waves are to be manipulated, corresponding to $4T$ independent equations obtained by writing the 4 equations in (5.7) for each of the field sets $\boldsymbol{\Psi}_n(\boldsymbol{\rho})$ ($n = 1, \ldots, T$, $\boldsymbol{\Psi}$ representing either $\boldsymbol{E}$ or $\boldsymbol{H}$), $4T$ $(4, 8, 12, 16)$ susceptibilities have to be specified.

Two approaches may be considered to reduce the number of independent unknown susceptibilities when $T < 4$. A first approach could consist in using more than $4T$ $(4, 8, 12)$ susceptibilities but enforcing relationships between some of them to ensure a maximum of $4T$ independent unknowns. For example, the conditions of reciprocity and losslessness would be a possible way to link some susceptibilities together, if this is compatible with design specifications. According to Kong[29] and Lindell,[30] the conditions for reciprocity are

$$\overline{\overline{\chi}}_{\mathrm{ee}}^{\mathrm{T}} = \overline{\overline{\chi}}_{\mathrm{ee}}, \quad \overline{\overline{\chi}}_{\mathrm{mm}}^{\mathrm{T}} = \overline{\overline{\chi}}_{\mathrm{mm}}, \quad \overline{\overline{\chi}}_{\mathrm{me}}^{\mathrm{T}} = -\overline{\overline{\chi}}_{\mathrm{em}} \tag{5.9}$$

and the conditions for reciprocity, passivity and losslessness are

$$\overline{\overline{\chi}}_{\mathrm{ee}}^{\mathrm{T}} = \overline{\overline{\chi}}_{\mathrm{ee}}^{*}, \quad \overline{\overline{\chi}}_{\mathrm{mm}}^{\mathrm{T}} = \overline{\overline{\chi}}_{\mathrm{mm}}^{*}, \quad \overline{\overline{\chi}}_{\mathrm{me}}^{\mathrm{T}} = \overline{\overline{\chi}}_{\mathrm{em}}^{*}, \tag{5.10}$$

where the superscripts T and $*$ denote the matrix transpose and complex conjugate operations, respectively. Enforcing conditions between susceptibilities also enforces conditions on the fields on both sides of the metasurface. Therefore, this approach restricts the diversity of electromagnetic transformations achievable with the metasurface.

A second approach, representing a more general synthesis method for quasi-arbitrary electromagnetic transformations, is then generally

preferred. This approach consists in selecting only T susceptibility tensor components in each of the $4T$ equations included in (5.7). The number of different combinations of susceptibilities able to perform a given transformation can be very large. However, most combinations may correspond to unphysical solutions or physical solutions requiring unnecessary active or lossy elements or even non-reciprocal features. For example, rotating by 90° the polarization of a normally incident plane wave can be achieved using: chiral bianisotropy, combination of active and lossy scattering particles, birefringence or even non-reciprocal gyrotropy as in Faraday devices. All of these solutions could be obtained from the synthesis technique by considering different combinations of susceptibilities. Consequently, the choice of susceptibility combinations will depend on the required application.

Note that these considerations hold in the most general case of transformations where the amplitude, phase and polarization of the incident field are all modified by the metasurface. Under such conditions, the system in (5.6) can indeed handle up to $T = 4$ independent wave sets. In the particular case of single-set transformation, only 4 susceptibilities (2 electric and 2 magnetic) are generally required, as will be shown shortly. However, as will be seen thereafter, depending on the transformation and choice of susceptibilities, only 2 susceptibilities (1 electric and 1 magnetic) may be sufficient.

Since many different transformations and susceptibility combinations are possible, we will now, without loss of generality, restrict our attention to the case of single transformation (section 5.2.3.3) and briefly address the case of multiple transformations by considering the case of double transformation with monoanisotropic metasurfaces. The solutions for bianisotropic, triple-wave and quadruple-wave metasurfaces can be obtained by following the same procedure.

5.2.3.3. *Single transformation*

We consider here the problem of single ($T = 1$) transformation [only one specified wave set: ($\mathbf{\Psi}^{\mathrm{i}}$, $\mathbf{\Psi}^{\mathrm{r}}$, $\mathbf{\Psi}^{\mathrm{t}}$)] for the most simple case of a monoanisotropic ($\overline{\overline{\chi}}_{\mathrm{em}} \equiv \overline{\overline{\chi}}_{\mathrm{me}} = 0$) and diagonal ($\chi_{\mathrm{ee}}^{xy} \equiv \chi_{\mathrm{ee}}^{yx} \equiv \chi_{\mathrm{mm}}^{xy} \equiv \chi_{\mathrm{mm}}^{yx} = 0$), and hence non-gyrotropic and reciprocal, metasurface.

Under such simplifications, the system (5.7) reduces to

$$\begin{pmatrix} \Delta H_y \\ \Delta H_x \\ \Delta E_y \\ \Delta E_x \end{pmatrix} = \begin{pmatrix} \widetilde{\chi}_{\mathrm{ee}}^{xx} & 0 & 0 & 0 \\ 0 & \widetilde{\chi}_{\mathrm{ee}}^{yy} & 0 & 0 \\ 0 & 0 & \widetilde{\chi}_{\mathrm{mm}}^{xx} & 0 \\ 0 & 0 & 0 & \widetilde{\chi}_{\mathrm{mm}}^{yy} \end{pmatrix} \begin{pmatrix} E_{x,\mathrm{av}} \\ E_{y,\mathrm{av}} \\ H_{x,\mathrm{av}} \\ H_{y,\mathrm{av}} \end{pmatrix}, \tag{5.11}$$

which straightforwardly yields, using (5.8), the following simple relations for the 4 susceptibilities:

$$\chi_{\mathrm{ee}}^{xx} = \frac{-\Delta H_y}{j\omega\epsilon E_{x,\mathrm{av}}}, \tag{5.12a}$$

$$\chi_{\mathrm{ee}}^{yy} = \frac{\Delta H_x}{j\omega\epsilon E_{y,\mathrm{av}}}, \tag{5.12b}$$

$$\chi_{\mathrm{mm}}^{xx} = \frac{\Delta E_y}{j\omega\mu H_{x,\mathrm{av}}}, \tag{5.12c}$$

$$\chi_{\mathrm{mm}}^{yy} = \frac{-\Delta E_x}{j\omega\mu H_{y,\mathrm{av}}}, \tag{5.12d}$$

where, according to (5.3) and (5.5), $\Delta H_y = H_y^{\mathrm{t}} - (H_y^{\mathrm{i}} + H_y^{\mathrm{r}})$, $E_{x,\mathrm{av}} = (E_x^{\mathrm{t}} + E_x^{\mathrm{i}} + E_x^{\mathrm{r}})/2$, and so on.

Upon synthesis, a metasurface with the susceptibilities given by (5.12) will produce exactly the specified reflected and transmitted transverse components of the fields when the metasurface is illuminated by the specified incident field. Since the longitudinal fields are completely determined from the transverse components, according to the uniqueness theorem, the complete specified electromagnetic fields are exactly generated by the metasurface. Note that, due to the orthogonality between x- and y-polarized waves, the susceptibilities in (5.12) can be separated into two subsets corresponding to Eqs. (5.12a) and (5.12d), and Eqs. (5.12b) and (5.12c), respectively. These two sets of susceptibilities are able to independently and simultaneously transform x- and y-polarized waves, which corresponds to the simplest case of double transformation. This contrasts with the more general case of double transformation, that will be discussed in

section 5.2.3.4, which allows transformation of quasi-arbitrary waves and not just orthogonal x- and y-polarized waves as is the case here.

What has been described so far in this section represents the first step of the synthesis procedure. As mentioned in section 5.2.1, the second step consists in determining the scattering particles realizing the transfer function corresponding to the synthesized susceptibilities. In this second step, one computes the full-wave scattering parameters for an isolated unit cell within 2D periodic boundary conditions, where periodicity is an approximation of typically slowly varying scattering elements in the plane of the metasurface.[13,16,30,31]

The periodic boundary conditions in full-wave analysis are generally restricted to rectilinearly propagating waves. Now, the prescribed waves may change directions at the metasurface (e.g. case of generalized refraction). In such cases, "rectilinear" periodic boundary conditions cannot directly describe the physics of the problem. However, the results they provide correspond to a rigorous mapping with the physical problem, and they may thus be rigorously used in the synthesis.

In order to enable the second step of the synthesis, we now need to establish the relationships existing between the susceptibilities and the scattering parameters. The forthcoming methodology for single transformation is analogous to that proposed in Refs. 16, 18, 32, while the corresponding methodologies for multiple transformation, to be presented in the next subsection, are more general.

In the plane wave approximation, which is naturally valid when the source of the incident wave is far enough from the metasurface and when the metasurface is made of an array of identical scattering particles (uniform), the response of each scattering particle may be expressed in terms of its reflection and transmission coefficients.[13,15,31] The problem splits into an x-polarized incident plane wave problem and a y-polarized incident plane wave problem, whose fields at normal incidence are respectively given by

$$\boldsymbol{E}^{\mathrm{i}} = \hat{x}, \qquad \boldsymbol{E}^{\mathrm{r}} = R_x \hat{x}, \qquad \boldsymbol{E}^{\mathrm{t}} = T_x \hat{x}, \tag{5.13a}$$

$$\boldsymbol{H}^{\mathrm{i}} = \frac{1}{\eta}\hat{y}, \qquad \boldsymbol{H}^{\mathrm{r}} = -\frac{R_x}{\eta}\hat{y}, \qquad \boldsymbol{H}^{\mathrm{t}} = \frac{T_x}{\eta}\hat{y} \tag{5.13b}$$

and

$$\boldsymbol{E}^{\mathrm{i}} = \hat{y}, \qquad \boldsymbol{E}^{\mathrm{r}} = R_y \hat{y}, \qquad \boldsymbol{E}^{\mathrm{t}} = T_y \hat{y}, \tag{5.14a}$$

$$\boldsymbol{H}^{\mathrm{i}} = -\frac{1}{\eta}\hat{x}, \qquad \boldsymbol{H}^{\mathrm{r}} = \frac{R_y}{\eta}\hat{x}, \qquad \boldsymbol{H}^{\mathrm{t}} = -\frac{T_y}{\eta}\hat{x}, \tag{5.14b}$$

where R_u and T_u $(u = x, y)$ represent reflection and transmission coefficients, respectively.[c] Inserting (5.13) and (5.14) into (5.6) with the four non-zero susceptibilities given in (5.12) leads to the transmission and reflection coefficients

$$T_x = \frac{4 + \chi_{\mathrm{ee}}^{xx}\chi_{\mathrm{mm}}^{yy}k^2}{(2 + jk\chi_{\mathrm{ee}}^{xx})(2 + jk\chi_{\mathrm{mm}}^{yy})}, \tag{5.15a}$$

$$R_x = \frac{2jk\,(\chi_{\mathrm{mm}}^{yy} - \chi_{\mathrm{ee}}^{xx})}{(2 + jk\chi_{\mathrm{ee}}^{xx})\,(2 + jk\chi_{\mathrm{mm}}^{yy})} \tag{5.15b}$$

and

$$T_y = \frac{4 + \chi_{\mathrm{ee}}^{yy}\chi_{\mathrm{mm}}^{xx}k^2}{(2 + jk\chi_{\mathrm{ee}}^{yy})(2 + jk\chi_{\mathrm{mm}}^{xx})}, \tag{5.16a}$$

$$R_y = \frac{2jk\,(\chi_{\mathrm{mm}}^{xx} - \chi_{\mathrm{ee}}^{yy})}{(2 + jk\chi_{\mathrm{ee}}^{yy})\,(2 + jk\chi_{\mathrm{mm}}^{xx})}, \tag{5.16b}$$

where $k = \omega\sqrt{\mu\epsilon} = \frac{2\pi}{\lambda}$. These relations may be used in the second step of the synthesis to determine the scattering parameters corresponding to the synthesized susceptibilities. Solving (5.15) and (5.16) for the susceptibilities yields

$$\chi_{\mathrm{ee}}^{xx} = \frac{2j\,(T_x + R_x - 1)}{k\,(T_x + R_x + 1)}, \tag{5.17a}$$

$$\chi_{\mathrm{ee}}^{yy} = \frac{2j\,(T_y + R_y - 1)}{k\,(T_y + R_y + 1)}, \tag{5.17b}$$

[c]The waves in (5.13) and (5.14) are defined as *rectilinear* (i.e. they do not change direction at the metasurface) for consistency with periodic boundary conditions to be used in full-wave simulations for the second step of the synthesis (see comment at the end of the last paragraph of the present section).

$$\chi_{\mathrm{mm}}^{xx} = \frac{2j\,(T_y - R_y - 1)}{k\,(T_y - R_y + 1)}, \tag{5.17c}$$

$$\chi_{\mathrm{mm}}^{yy} = \frac{2j\,(T_x - R_x - 1)}{k\,(T_x - R_x + 1)}. \tag{5.17d}$$

In (5.16) and (5.17), the reflection and transmission coefficients are associated with scattering parameters S_{ij} with $i, j = 1, \ldots, 4$ accounting for the two ports (incident and transmitted waves) and two polarizations (x and y). Specifically, assigning ports 1, 2, 3 and 4 to x-polarized input, y-polarized input, x-polarized output and y-polarized output, respectively, one has $R_x = S_{11}$, $T_x = S_{31}$, $R_y = S_{22}$ and $T_y = S_{42}$, while the other 12 scattering parameters are not required since the chosen tensors are uniaxial so that the metasurface is not gyrotropic (i.e. does not involve transformations between x-polarized and y-polarized waves).

As will be shown in section 5.4, the synthesized susceptibilities generally have both real and imaginary parts, which may represent gain or loss. Consequently, the metasurface is, in many cases, lossy or active or even a combination of both depending on which susceptibility components are considered. It is therefore generally necessary to explicitly compute the bianisotropic Poynting theorem to determine whether there is loss or gain. Assuming the convention $e^{j\omega t}$, the time-average bianisotropic Poynting theorem is given by[29]

$$\nabla \cdot \langle \boldsymbol{S} \rangle = -\langle \boldsymbol{I}_{\mathrm{Je}} \rangle - \langle \boldsymbol{I}_{\mathrm{Jm}} \rangle - \langle \boldsymbol{I}_{\mathrm{P}} \rangle - \langle \boldsymbol{I}_{\mathrm{M}} \rangle, \tag{5.18}$$

where $\langle \cdot \rangle$ denote the time-average, $\boldsymbol{S}$ is the Poynting vector and $\boldsymbol{I}_{\mathrm{Je}}$, $\boldsymbol{I}_{\mathrm{Jm}}$, $\boldsymbol{I}_{\mathrm{P}}$ and $\boldsymbol{I}_{\mathrm{M}}$ are loss (or gain) contributions emerging from the electric currents, magnetic currents, electric polarization and magnetic polarization, respectively. The terms in (5.18) read

$$\langle \boldsymbol{S} \rangle = \frac{1}{2} \mathrm{Re}(\boldsymbol{E} \times \boldsymbol{H}^*), \tag{5.19a}$$

$$\langle \boldsymbol{I}_{\mathrm{Je}} \rangle = \frac{1}{4} \mathrm{Re}(j\omega\epsilon \boldsymbol{E}^* \cdot (\overline{\overline{\chi}}_{\mathrm{ee}} - \overline{\overline{\chi}}_{\mathrm{ee}}^*) \cdot \boldsymbol{E}), \tag{5.19b}$$

$$\langle \boldsymbol{I}_{\mathrm{Jm}} \rangle = \frac{1}{4} \mathrm{Re}(j\omega\mu \boldsymbol{H}^* \cdot (\overline{\overline{\chi}}_{\mathrm{mm}} - \overline{\overline{\chi}}_{\mathrm{mm}}^*) \cdot \boldsymbol{H}), \tag{5.19c}$$

$$\langle \boldsymbol{I}_{\mathrm{P}} \rangle = \frac{1}{4} \mathrm{Re}[j\omega\epsilon(\boldsymbol{E}^* \cdot (\overline{\overline{\chi}}_{\mathrm{ee}} - \overline{\overline{\chi}}_{\mathrm{ee}}^{\dagger}) \cdot \boldsymbol{E} + 2\eta \boldsymbol{E}^* \cdot \overline{\overline{\chi}}_{\mathrm{em}} \cdot \boldsymbol{H})], \tag{5.19d}$$

$$\langle \boldsymbol{I}_{\mathrm{M}} \rangle = \frac{1}{4} \mathrm{Re}[j\omega\mu(\boldsymbol{H}^* \cdot (\overline{\overline{\chi}}_{\mathrm{mm}} - \overline{\overline{\chi}}_{\mathrm{mm}}^{\dagger}) \cdot \boldsymbol{H} - 2\boldsymbol{E}^* \cdot \overline{\overline{\chi}}_{\mathrm{me}}^{\dagger} \cdot \boldsymbol{H}/\eta)], \tag{5.19e}$$

where † is the conjugate transpose operator. From these relations, it is straightforward to obtain the conditions for passivity and losslessness that were given in (5.10). Note that the terms $\boldsymbol{E}$ and $\boldsymbol{H}$ in (5.19) are the fields acting on the metasurface and can be replaced by the corresponding average fields. While the relations (5.10) can be easily used to determine if the metasurface is lossless and passive, it is only by computing (5.18) that one can determine whether the metasurface is more active or more lossy and which of the contributions in (5.19) are dominant: If the terms (5.19b) to (5.19e) are positive, the metasurface is lossy, if they are negative, the metasurface is active. Moreover, as will be seen in section 5.4, it is possible to have an apparently passive and lossless metasurface while, in fact, having a combination of electric gain and magnetic loss (or vice versa) that would perfectly cancel each other.

In order to simplify the metasurface unit cell design procedure, which is usually performed via full-wave simulation, one may consider "ideal" unit cells, i.e. unit cells made of lossless dielectric substrates and PEC metallic patterns. Applying relations (5.17) to compute the susceptibilities corresponding to such "ideal" unit cells will necessarily yield purely real susceptibilities. As a consequence, the exact complex susceptibilities can not be realized, which will lead to results diverging from the specified transformation. A solution to minimize discrepancies between the specified response and the approximated response considering only real susceptibilities is to set the imaginary parts of the susceptibilities to zero while optimizing their real parts so that the response of the metasurface follows the specified response as closely as possible. In that case, the amplitude of the transmission coefficients (5.15) and (5.16) (or reflection coefficients if the metasurface is used in reflection) is approximatively 1, while the phase of

these coefficients remains unchanged. This means that, in the case of refractive metasurfaces, the physical realization of these types of metasurfaces often consists in implementing a full transmission structure inducing a given phase shift according to the specification.

5.2.3.4. *Multiple transformations*

In the previous section, we have seen how to synthesize a metasurface in the case of a single transformation. However, as mentioned in 5.2.3.2, the general system of equations (5.7) has the capability to perform more than a single transformation given its large number of degrees of freedom, i.e. its 16 susceptibility components. Here, we will see how the system (5.7) can be solved for several transformations including incident waves coming from one side only or both sides of the metasurface. To accommodate for the additional degrees of freedom, three additional wave transformations are added, so that (5.7) transforms to

$$\begin{pmatrix} \Delta H_{y1} & \Delta H_{y2} & \Delta H_{y3} & \Delta H_{y4} \\ \Delta H_{x1} & \Delta H_{x2} & \Delta H_{x3} & \Delta H_{x4} \\ \Delta E_{y1} & \Delta E_{y2} & \Delta E_{y3} & \Delta E_{y4} \\ \Delta E_{x1} & \Delta E_{x2} & \Delta E_{x3} & \Delta E_{x4} \end{pmatrix} = \begin{pmatrix} \widetilde{\chi}_{\mathrm{ee}}^{xx} & \widetilde{\chi}_{\mathrm{ee}}^{xy} & \widetilde{\chi}_{\mathrm{em}}^{xx} & \widetilde{\chi}_{\mathrm{em}}^{xy} \\ \widetilde{\chi}_{\mathrm{ee}}^{yx} & \widetilde{\chi}_{\mathrm{ee}}^{yy} & \widetilde{\chi}_{\mathrm{em}}^{yx} & \widetilde{\chi}_{\mathrm{em}}^{yy} \\ \widetilde{\chi}_{\mathrm{me}}^{xx} & \widetilde{\chi}_{\mathrm{me}}^{xy} & \widetilde{\chi}_{\mathrm{mm}}^{xx} & \widetilde{\chi}_{\mathrm{mm}}^{xy} \\ \widetilde{\chi}_{\mathrm{me}}^{yx} & \widetilde{\chi}_{\mathrm{me}}^{yy} & \widetilde{\chi}_{\mathrm{mm}}^{yx} & \widetilde{\chi}_{\mathrm{mm}}^{yy} \end{pmatrix} \times \begin{pmatrix} E_{x1,\mathrm{av}} & E_{x2,\mathrm{av}} & E_{x3,\mathrm{av}} & E_{x4,\mathrm{av}} \\ E_{y1,\mathrm{av}} & E_{y2,\mathrm{av}} & E_{y3,\mathrm{av}} & E_{y4,\mathrm{av}} \\ H_{x1,\mathrm{av}} & H_{x2,\mathrm{av}} & H_{x3,\mathrm{av}} & H_{x4,\mathrm{av}} \\ H_{y1,\mathrm{av}} & H_{y2,\mathrm{av}} & H_{y3,\mathrm{av}} & H_{y4,\mathrm{av}} \end{pmatrix}, \tag{5.20}$$

where the subscripts 1, 2, 3 and 4 indicate the electromagnetic fields corresponding to four distinct and independent wave

transformations. As previously done, the susceptibilities can be obtained by matrix inversion conjointly with (5.8).

As an example, we now consider the problem of double ($T = 2$) transformation [two specified wave sets: ($\boldsymbol{\Psi}_1^{\mathrm{i}}$, $\boldsymbol{\Psi}_1^{\mathrm{r}}$, $\boldsymbol{\Psi}_1^{\mathrm{t}}$) and ($\boldsymbol{\Psi}_2^{\mathrm{i}}$, $\boldsymbol{\Psi}_2^{\mathrm{r}}$, $\boldsymbol{\Psi}_2^{\mathrm{t}}$)] for a monoanisotropic ($\overline{\overline{\chi}}_{\mathrm{em}} \equiv \overline{\overline{\chi}}_{\mathrm{me}} = 0$) but not uniaxial and hence gyrotropic metasurface. In that case, the system (5.20) reduces to

$$\begin{pmatrix} \Delta H_{y1} & \Delta H_{y2} \\ \Delta H_{x1} & \Delta H_{x2} \\ \Delta E_{y1} & \Delta E_{y2} \\ \Delta E_{x1} & \Delta E_{x2} \end{pmatrix} = \begin{pmatrix} \widetilde{\chi}_{\mathrm{ee}}^{xx} & \widetilde{\chi}_{\mathrm{ee}}^{xy} & 0 & 0 \\ \widetilde{\chi}_{\mathrm{ee}}^{yx} & \widetilde{\chi}_{\mathrm{ee}}^{yy} & 0 & 0 \\ 0 & 0 & \widetilde{\chi}_{\mathrm{mm}}^{xx} & \widetilde{\chi}_{\mathrm{mm}}^{xy} \\ 0 & 0 & \widetilde{\chi}_{\mathrm{mm}}^{yx} & \widetilde{\chi}_{\mathrm{mm}}^{yy} \end{pmatrix} \begin{pmatrix} E_{x1,\mathrm{av}} & E_{x2,\mathrm{av}} \\ E_{y1,\mathrm{av}} & E_{y2,\mathrm{av}} \\ H_{x1,\mathrm{av}} & H_{x2,\mathrm{av}} \\ H_{y1,\mathrm{av}} & H_{y2,\mathrm{av}} \end{pmatrix}, \tag{5.21}$$

which, after matrix inversion, yields the following eight susceptibilities:

$$\chi_{\mathrm{ee}}^{xx} = \frac{j}{\epsilon\omega} \frac{(E_{y1,\mathrm{av}}\Delta H_{y2} - E_{y2,\mathrm{av}}\Delta H_{y1})}{(E_{x2,\mathrm{av}}E_{y1,\mathrm{av}} - E_{x1,\mathrm{av}}E_{y2,\mathrm{av}})}, \tag{5.22a}$$

$$\chi_{\mathrm{ee}}^{xy} = \frac{j}{\epsilon\omega} \frac{(E_{x2,\mathrm{av}}\Delta H_{y1} - E_{x1,\mathrm{av}}\Delta H_{y2})}{(E_{x2,\mathrm{av}}E_{y1,\mathrm{av}} - E_{x1,\mathrm{av}}E_{y2,\mathrm{av}})}, \tag{5.22b}$$

$$\chi_{\mathrm{ee}}^{yx} = \frac{j}{\epsilon\omega} \frac{(E_{y2,\mathrm{av}}\Delta H_{x1} - E_{y1,\mathrm{av}}\Delta H_{x2})}{(E_{x2,\mathrm{av}}E_{y1,\mathrm{av}} - E_{x1,\mathrm{av}}E_{y2,\mathrm{av}})}, \tag{5.22c}$$

$$\chi_{\mathrm{ee}}^{yy} = \frac{j}{\epsilon\omega} \frac{(E_{x1,\mathrm{av}}\Delta H_{x2} - E_{x2,\mathrm{av}}\Delta H_{x1})}{(E_{x2,\mathrm{av}}E_{y1,\mathrm{av}} - E_{x1,\mathrm{av}}E_{y2,\mathrm{av}})}, \tag{5.22d}$$

$$\chi_{\mathrm{mm}}^{xx} = \frac{j}{\mu\omega} \frac{(H_{y2,\mathrm{av}}\Delta E_{y1} - H_{y1,\mathrm{av}}\Delta E_{y2})}{(H_{x2,\mathrm{av}}H_{y1,\mathrm{av}} - H_{x1,\mathrm{av}}H_{y2,\mathrm{av}})}, \tag{5.22e}$$

$$\chi_{\mathrm{mm}}^{xy} = \frac{j}{\mu\omega} \frac{(H_{x1,\mathrm{av}}\Delta E_{y2} - H_{x2,\mathrm{av}}\Delta E_{y1})}{(H_{x2,\mathrm{av}}H_{y1,\mathrm{av}} - H_{x1,\mathrm{av}}H_{y2,\mathrm{av}})}, \tag{5.22f}$$

$$\chi_{\mathrm{mm}}^{yx} = \frac{j}{\mu\omega} \frac{(H_{y1,\mathrm{av}}\Delta E_{x2} - H_{y2,\mathrm{av}}\Delta E_{x1})}{(H_{x2,\mathrm{av}}H_{y1,\mathrm{av}} - H_{x1,\mathrm{av}}H_{y2,\mathrm{av}})}, \tag{5.22g}$$

$$\chi_{\mathrm{mm}}^{yy} = \frac{j}{\mu\omega} \frac{(H_{x2,\mathrm{av}}\Delta E_{x1} - H_{x1,\mathrm{av}}\Delta E_{x2})}{(H_{x2,\mathrm{av}}H_{y1,\mathrm{av}} - H_{x1,\mathrm{av}}H_{y2,\mathrm{av}})}, \tag{5.22h}$$

where the subscripts 1 and 2 stand for the first and the second wave set transformation, respectively. Applying the conditions (5.9) and (5.10) to (5.22) indicates that in most cases the metasurface will be not only active/lossy but also non-reciprocal. The same argument applies to the more general case of the fully bianisotropic metasurface described by the susceptibilities in (5.20) which may, depending on the choice of transformations, be non-reciprocal and active/lossy. Note that the choice of using the susceptibility tensors $\overline{\overline{\chi}}_{\mathrm{ee}}$ and $\overline{\overline{\chi}}_{\mathrm{mm}}$ in (5.21) was arbitrary and other sets of susceptibilities, for instance including bianisotropic components, may be more suited for some specific transformations.

Now we are interested in establishing a relation between the scattering parameters and the susceptibilities of a metasurface in the most general case of a fully bianisotropic metasurface.[18] By following the procedure used to obtain relations (5.15), (5.16) and (5.17), we assume that the metasurface is illuminated by four different plane waves. The first two plane waves are x- and y-polarized, respectively, and normally incident from the left-hand side of the metasurface. The two remaining plane waves are x- and y-polarized, respectively, and normally incident from the right-hand side of the metasurface. For convenience, the system (5.20) is first written in the following compact form

$$\overline{\overline{\Delta}} = \widetilde{\overline{\overline{\chi}}} \cdot \overline{\overline{A}}_v, \tag{5.23}$$

where $\overline{\overline{\Delta}}$ refers to the matrix of the field differences and $\overline{\overline{A}}_v$ refers to the matrix of the field averages. Defining each wave transformation in a similar fashion as done in (5.13) and (5.14) but this time using the scattering matrices $\overline{\overline{S}}_{11}$, $\overline{\overline{S}}_{22}$, $\overline{\overline{S}}_{12}$ and $\overline{\overline{S}}_{21}$, the matrix $\overline{\overline{\Delta}}$ reads

$$\overline{\overline{\Delta}} = \begin{pmatrix} -\overline{\overline{N}}_2/\eta + \overline{\overline{N}}_2\overline{\overline{S}}_{11}/\eta & -\overline{\overline{N}}_2/\eta + \overline{\overline{N}}_2\overline{\overline{S}}_{12}/\eta \\ + \overline{\overline{N}}_2\overline{\overline{S}}_{21}/\eta & + \overline{\overline{N}}_2\overline{\overline{S}}_{22}/\eta \\ -\overline{\overline{N}}_1\overline{\overline{N}}_2 - \overline{\overline{N}}_1\overline{\overline{N}}_2\overline{\overline{S}}_{11} & \overline{\overline{N}}_1\overline{\overline{N}}_2 - \overline{\overline{N}}_1\overline{\overline{N}}_2\overline{\overline{S}}_{12} \\ + \overline{\overline{N}}_1\overline{\overline{N}}_2\overline{\overline{S}}_{21} & + \overline{\overline{N}}_1\overline{\overline{N}}_2\overline{\overline{S}}_{22} \end{pmatrix} \tag{5.24}$$

and the matrix $\overline{\overline{A}}_v$ reads

$$\overline{\overline{A}}_v = \frac{1}{2}\begin{pmatrix} \overline{\overline{I}} + \overline{\overline{S}}_{11} + \overline{\overline{S}}_{21} & \overline{\overline{I}} + \overline{\overline{S}}_{12} + \overline{\overline{S}}_{22} \\ \overline{\overline{N}}_1/\eta - \overline{\overline{N}}_1\overline{\overline{S}}_{11}/\eta + \overline{\overline{N}}_1\overline{\overline{S}}_{21}/\eta & -\overline{\overline{N}}_1/\eta - \overline{\overline{N}}_1\overline{\overline{S}}_{12}/\eta + \overline{\overline{N}}_1\overline{\overline{S}}_{22}/\eta \end{pmatrix}, \tag{5.25}$$

where the matrices $\overline{\overline{S}}_{nm}$, $\overline{\overline{I}}$, $\overline{\overline{N}}_1$ and $\overline{\overline{N}}_2$ are defined by

$$\overline{\overline{S}}_{nm} = \begin{pmatrix} S_{nm}^{xx} & S_{nm}^{xy} \\ S_{nm}^{yx} & S_{nm}^{yy} \end{pmatrix}, \quad \overline{\overline{I}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$
$$\overline{\overline{N}}_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \overline{\overline{N}}_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.26}$$

The scattering matrices of a given structure can be obtained by full-wave simulations and, subsequently, the susceptibilities of the structure can be retrieved by matrix inversion of (5.23) with (5.24) and (5.25). Alternatively, it is possible to find the scattering parameters of a metasurface with known susceptibilities by solving (5.23) for the scattering parameters.

5.3. Scattering Particle Synthesis and Implementation

As mentioned in the previous sections, the metasurface synthesis technique is composed of two main steps. The first one is the mathematical description of the metasurface, from the specified fields, either in terms of susceptibilities or directly as transmission and reflection coefficients, using relations (5.15) and (5.16) for instance. The second step is the physical implementation of each unit cell of the metasurface. This latter step is, as of today, not a trivial operation. Here, two possible approaches are briefly discussed, while many more exist. In both approaches, the susceptibility functions, obtained by the aforementioned synthesis procedure, are spatially discretized in the $x-y$ plane where each discrete point corresponds to a unit cell to be implemented. The unit cells are simulated (assuming periodic

boundary conditions), one by one or group by group when the structure is fully or partly non-uniform, using commercial software that compute their scattering parameters. The required physical parameters for the scattering particles are obtained by mapping the scattering parameters onto the susceptibility function. Finally, an enhanced design is achieved by tuning the parameters of the scattering particles via parametric analysis or standard optimization techniques.

5.3.1. *Metallic scatterers*

Metasurfaces consisting of metallic scatterers, typically arranged on dielectric substrates, have been the most commonly reported types of metasurfaces. In all cases, using more than one layer is an effective way to increase the available number of degrees of freedom, and hence achieving enhanced properties, including higher bandwidth and larger phase coverage of the structure's unit cell. It was recently shown that three cascaded layers, where the two outer layers are identical, represents the minimal configuration to achieve full transmission and a 2π phase coverage.[9,17] With three layers, the overall thickness of the metasurface generally remains subwavelength (usually in the order of $t \approx \lambda/10$) with negligible loss increase. Further increasing the number of layers may naturally be an approach for even broader bandwidth, at the expense of extra loss and weight. Note that, if the three metallic layers are all different, a more diverse electromagnetic response is achievable at the cost of a more complicated design.

A typical shape for the scattering particles forming the metallic layers is the Jerusalem cross, as shown in Fig. 5.2. The Jerusalem cross has the advantage of featuring fairly well decoupled responses for x and y polarizations, consequently simplifying the implementation. In the structure of Fig. 5.2, relatively strong capacitive coupling in the transverse $(x - y)$ plane offers the benefit of lowering the resonance frequencies[5,36] or, alternatively, of reducing the free-space electrical size of the unit cell ($d \approx \lambda/5$), while introducing more complexity in terms of coupling.

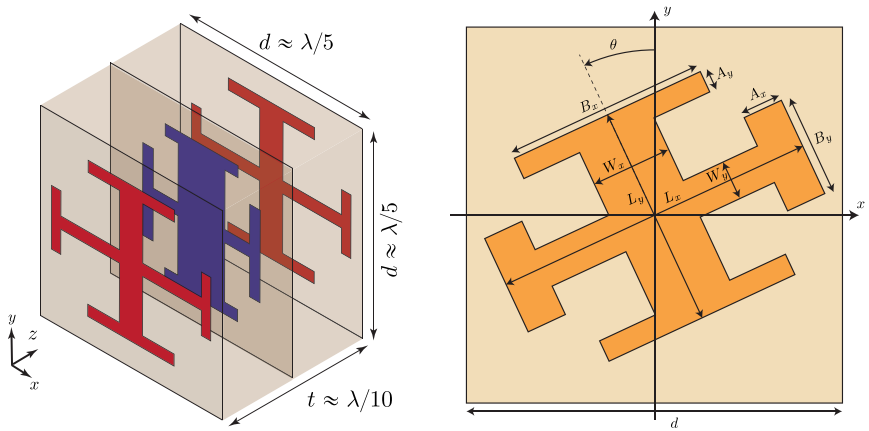


Fig. 5.2. Generic unit cell (a) with three metallic (PEC) Jerusalem crosses separated by dielectric slabs, the outer layers are identical. (b) One metallic layer has usually up to nine different dimensions that can be modified.

The realization of the unit cells is greatly simplified if the longitudinal evanescent mode coupling between the three metallic layers can be minimized. This occurs when the dielectric spacer relative permittivity is decreased or/and its thickness is increased. In that case, each layer can be designed separately and the overall response of the multilayer unit cell can then be found using simple transmission matrix approaches.[17]

It can be easily verified that cascading three metallic layers (with similar outer layers) is sufficient to realize a transmission coefficient of $0 \leq |T| \leq 1$ with a 2π-phase coverage. For a given polarization, each metallic layer can be described by an impedance layer.[17] The entire structure can then be analyzed using the ABCD matrix technique by modelling each dielectric space by a transmission line section as

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ Y_1 & 1 \end{pmatrix} \begin{pmatrix} \cos(\beta d) & j\eta_d \sin(\beta d) \\ \frac{j\sin(\beta d)}{\eta_d} & \cos(\beta d) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ Y_2 & 1 \end{pmatrix} \times \begin{pmatrix} \cos(\beta d) & j\eta_d \sin(\beta d) \\ \frac{j\sin(\beta d)}{\eta_d} & \cos(\beta d) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ Y_1 & 1 \end{pmatrix}, \tag{5.27}$$

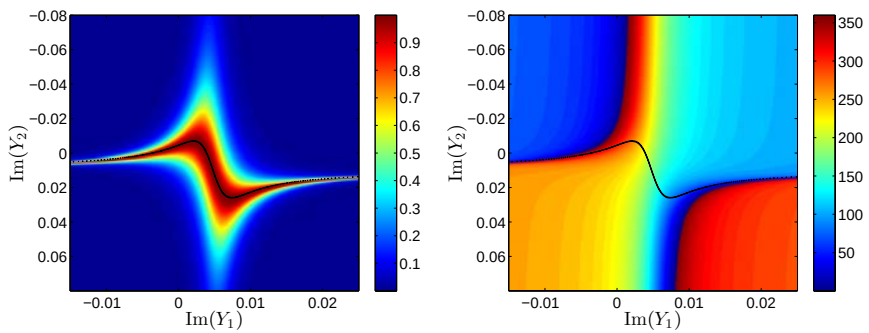


Fig. 5.3. Transmitted (a) power ($|S_{21}|^2$) and (b) phase for the three cascaded metallic layers of Fig. 5.2. The black line indicates that full transmission can be achieved and that the corresponding phase varies between 0 and 2π. The x and y axes respectively correspond to the imaginary parts of the admittance of the outer layers and the middle layer. The real parts of these admittances is zero because these structures are assumed to be lossless.

where β is the propagation constant along z and d and η_d are the thickness and the impedance of the dielectric substrates, respectively. The terms Y_1 and Y_2 correspond to the admittances of the outer layers and the middle layer, respectively. Finally, the ABCD matrix (5.27) can be converted to the following scattering parameter matrix

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \frac{1}{2A + B/\eta_0 + C\eta_0} \begin{pmatrix} B/\eta_0 - C\eta_0 & 2 \\ 2 & B/\eta_0 - C\eta_0 \end{pmatrix}. \tag{5.28}$$

Substituting (5.27) into (5.28) and plotting the magnitude and phase of S_{21}, as presented in Figs. 5.3(a) and 5.3(b), reveals that scanning the transmitted phase from 0 to 2π while maintaining a constant transmission of $|S_{21}| = 1$ is indeed possible.

There exists no direct methods to map the dimensions of the metallic layers onto the impedances Y_1 and Y_2. The general idea is to simulate one structure with a given set of dimensions. From the simulated transmitted phase and magnitude, the equivalent impedances can be found. Then, the impedance Y_2 of the unit cell can be tuned by

modifying the shape of layer 2, whereas changing layers 1 affect both impedances Y_1 and Y_2 due to coupling between the layers. Therefore, to obtain a specified transmitted magnitude and phase, the outer layers 1 are first modified until the impedance Y_1 matches the specifications. Then, the middle layer 2 is varied to reach the required Y_2 impedance. Usually, a few optimization iterations are required to design each unit cell.

5.3.2. *Dielectric scatterers*

The second implementation is based on all-dielectric scattering particles. It has been known for a long time that dielectric resonators exhibit both electric and magnetic resonances.[34,35] More recently, dielectric resonators have been used to realize all-dielectric metasurfaces.[36,37] Such metasurfaces are particularly attractive in the optical regime, where plasma loss associated with metallic materials may be important. A typical unit cell is shown in Fig. 5.4(a) where the particles are dielectric cylinders of circular cross-section with permittivity $\varepsilon_{r,1}$ placed on a substrate with permittivity $\varepsilon_{r,2}$. Other types of particle shapes are naturally also possible. It is possible, by tuning the physical dimensions of the resonator as well as the permittivities ratio $\varepsilon_{r,1}/\varepsilon_{r,2}$, to tune the electric and magnetic resonances to the same frequency. In this scenario, if the two resonances have the same

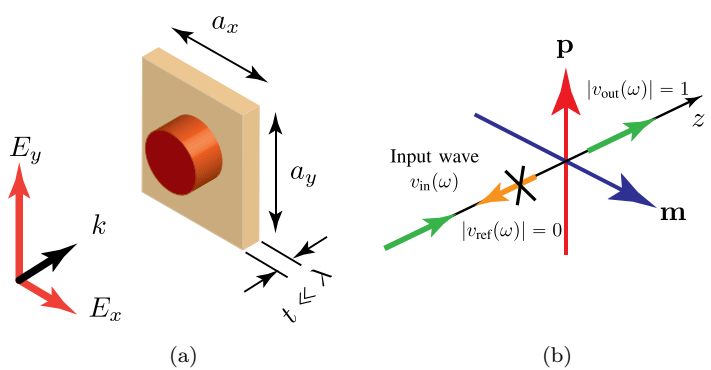


Fig. 5.4. (a) Representation of an all-dielectric all-pass metasurface unit cell consisting of a dielectric resonator ($\varepsilon_{r,1}$) embedded in a host layer of permittivity $\varepsilon_{r,2}$. (b) Operation principle for full transmission (zero reflection).

strength and are associated with orthogonal dipole moments in the transverse plane of the metasurface, as shown in Fig. 5.4(b), reflection may be totally suppressed. This is due to perfect destructive interference of the waves scattered by the electric and magnetic scattering particles in the incident side of the metasurface, and their constructive interference at the transmission side of it.[35] In this case, the transmission is theoretically 100% and flat over a wide bandwidth. Moreover, the transmission phase covers a full 2π range around the resonance frequency ω. This powerful concept may therefore apply to all-pass metasurfaces with controllable phase over a large bandwidth. An example of a dielectric metasurface is shown in Fig. 5.5(a) where dielectric cylinders are held together by thin dielectric connections. It must be noted that, in general, the electric and magnetic resonances have different dispersions,[38] which leads to a transmission amplitude that is not perfectly flat over the bandwidth of interest. However, such small variations in transmission may in general be neglected since the transmission coefficient S_{21} does not drop below -2 dB (assuming no dielectric loss), as can be seen in Fig. 5.5(b).

Structurally, symmetric shapes like cylinders or squares present the same behavior for x- and y-polarized waves. However, using

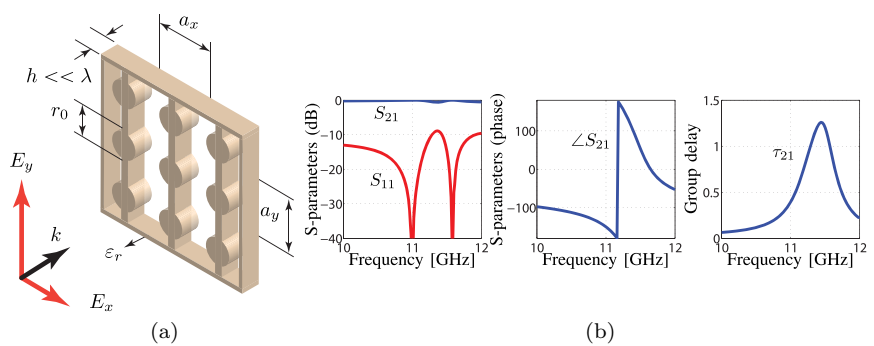


Fig. 5.5. (a) Example of dielectric metasurface. (b) Rigorous coupled wave analysis (RCWA) of the metasurface in (a). The plots show the transmission and reflection coefficients as well as the transmission phase and the transmission group delay, respectively. The field is polarized in the x direction (perpendicular to the dielectric interconnections).

90°-asymmetric shapes, such as ellipses and rectangles, allow for a complete and independent control of the two orthogonal polarizations, as recently demonstrated in Ref. 37. As an additional advantage, dielectric unit cells have a greatly reduced number of physical parameters to adjust compared to the three layer Jerusalem crosses of Fig. 5.2, effectively simplifying the optimization procedure to achieve the specified response.

5.4. Discussions

In this section, we address several points related to the synthesis technique presented previously and discuss the implementation and some limitations of metasurfaces. The synthesis technique can be used to realize many types of different metasurfaces, for instance, polarization rotators are easily synthesized because the metasurface is, in that case, uniform (susceptibilities are not a function of $\boldsymbol{\rho}$), which is a consequence of not changing the direction of propagation of the waves. Polarization rotators can thus be straightforwardly realized with (5.23) in the case of normally incident waves. A more interesting case to discuss is the topic of refractive metasurfaces which induce changes in the direction of propagation of waves and thus require non-uniform susceptibility functions. As we will see, this type of electromagnetic transformation is a perfect candidate to illustrate how the synthesis method works as well as to put into practice several of the concepts developed above.

5.4.1. *Refractive metasurfaces*

Here, we present and discuss a particular case of reflection-less refractive metasurface. This choice of electromagnetic transformation is chosen for the sake of simplicity, but the discussion developed therein is applicable to other types of transformations.

The electromagnetic problem considered here consists in synthesizing a metasurface transforming an obliquely incident plane wave forming a $\pi/9$ angle with respect to z in the $x-z$ plane into a transmitted plane wave with a positive $\pi/4$ "refraction" angle

(p-polarization). Due to the 2D nature of this transformation, the metasurface is assumed to be infinite in the y direction and having a length of 10λ in the x direction. The electric and magnetic fields at the metasurface ($z = 0$) have the general following form:

$$\boldsymbol{E}^{\mathrm{a}} = A^{\mathrm{a}} \left(\hat{\boldsymbol{x}} \frac{k_z^{\mathrm{a}}}{k} - \hat{\boldsymbol{z}} \frac{k_x^{\mathrm{a}}}{k} \right) e^{-jk_x^{\mathrm{a}} x} \quad \text{and} \quad \boldsymbol{H}^{\mathrm{a}} = \frac{A^{\mathrm{a}}}{\eta} \hat{\boldsymbol{y}} e^{-jk_x^{\mathrm{a}} x}, \tag{5.29}$$

where $a = i, t, r$ denotes the incident, transmitted and reflected waves, respectively, A is their amplitude and $\eta = \sqrt{\mu/\epsilon}$ is the intrinsic impedance of the surrounding medium associated with the wavenumber k. In this example, the reflected wave is specified to be zero ($A^{\mathrm{r}} = 0$) and the incident and transmitted waves have unit-amplitude ($A^{\mathrm{i}} = A^{\mathrm{t}} = 1$). The corresponding tangential difference and average fields in (5.3) and (5.5) are, for the components in the plane of the metasurface, given by

$$\Delta E_x = \frac{\sqrt{2}}{2} e^{-jk\frac{\sqrt{2}}{2}x} - \cos(\pi/9) e^{-jk\sin(\pi/9)x}, \tag{5.30a}$$

$$\Delta H_y = \frac{1}{\eta} (e^{-jk\frac{\sqrt{2}}{2}x} - e^{-jk\sin(\pi/9)x}), \tag{5.30b}$$

$$E_{\mathrm{av},x} = \frac{1}{2} \left(\frac{1}{2} e^{-jk\frac{\sqrt{2}}{2}x} + \cos(\pi/9) e^{-jk\sin(\pi/9)x} \right), \tag{5.30c}$$

$$H_{\mathrm{av},y} = \frac{1}{2\eta} (e^{-jk\frac{\sqrt{2}}{2}x} + e^{-jk\sin(\pi/9)x}). \tag{5.30d}$$

The metasurface susceptibilities are then obtained by substituting (5.30) into (5.12a) and (5.12d). They are naturally obtained in closed-form given the closed-forms (5.30), but are not written explicitly here, for conciseness. Instead, the real and imaginary parts of the susceptibilities χ_{ee}^{xx} and χ_{mm}^{yy} are plotted in Figs. 5.6(a) and 5.6(b), respectively. Note that the susceptibilities for this problem only depend on x, since no wave transformation is prescribed in the y direction. Moreover, only two of the four susceptibilities (χ_{ee}^{xx} and χ_{mm}^{yy}) are required (the other two being undefined) since no electric fields exist along y and no magnetic fields exist along x. Also note

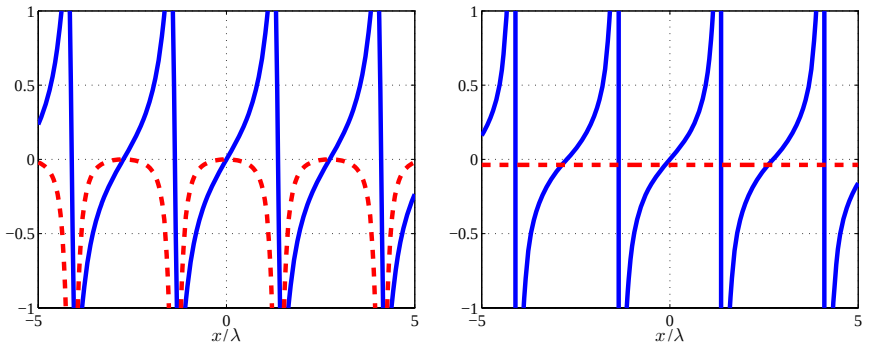
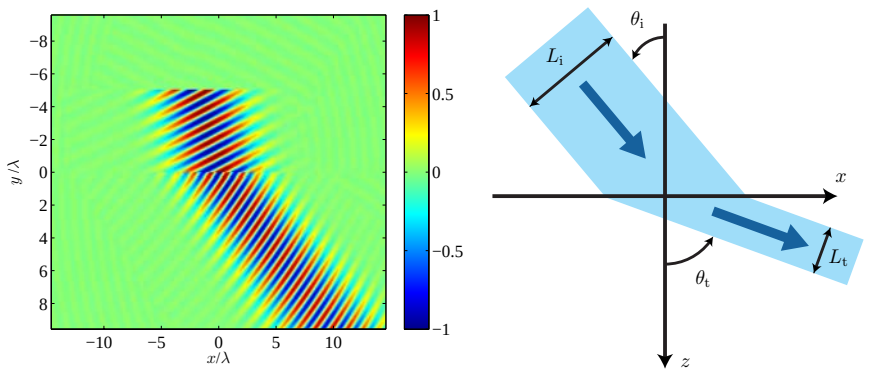


Fig. 5.6. Real (blue line) and imaginary (dashed red line) parts of: (a) χ_{ee}^{xx} and of (b) χ_{mm}^{yy} for a metasurface refracting a plane wave incident with an angle of $\pi/9$ with respect to z in the $x-z$ plane into a transmitted plane wave forming a $\pi/4$ angle.

that the x-periodicity of the susceptibility is larger than λ, suggesting that this metasurface should be easily implementable with simple scattering particles.

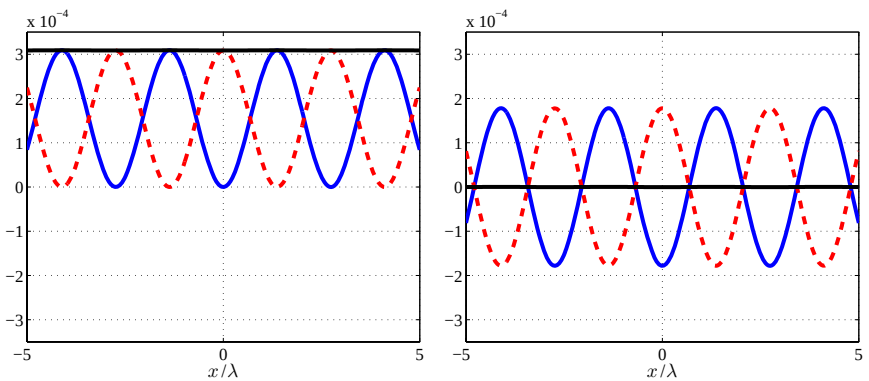
In order to verify that the metasurface performs as expected, a finite-difference frequency domain (FDFD) simulation,[39] illustrating the response of the metasurface when illuminated by the specified Gaussian p-polarized beam, is shown in Fig. 5.7(a). As expected, the metasurface refracts the incident beam at an angle of $\pi/4$ without any reflection.

As can be seen in Fig. 5.6, both electric and magnetic susceptibilities are not only complex functions but their respective imaginary parts are negative, indicating a lossy structure according to the discussion at the end of section 5.2.3.3. It can be easily verified that the required transformation is lossy by considering relations (5.19). While relations (5.19d) and (5.19e) are zero due to the monoanisotropic and diagonal nature of the metasurface, relations (5.19b) and (5.19c) are not zero indicating the presence of power dissipation. Figure 5.8(a) shows the electric and magnetic dissipated (real and positive) power on the metasurface (blue line and dashed red line, respectively) and the total dissipated power (black line).



(a) (b)

Fig. 5.7. Illustrations of the electromagnetic transformation. (a) FDFD simulation of the metasurface with susceptibilities as in Fig. 5.6. (b) Representation of the different beamwidths between the incident and transmitted waves.



(a) (b)

Fig. 5.8. Metasurface power budget including the electric (blue line) and magnetic (dashed red line) contributions (corresponding to Eqs. (5.19b) and (5.19c), respectively) when (a) $A^t = 1$ and (b) $A^t = \sqrt{\cos\theta_\mathrm{i}/\cos\theta_\mathrm{t}}$. In these graphs, positive values correspond to loss and negative values to gain.

The presence of the observed loss might come as a surprise especially when considering that the specified incident and transmitted fields have the same amplitude (transmission coefficient $A^\mathrm{t} = 1$). In this transformation, the presence of loss can be explained[40] by considering

the simple geometrical description shown in Fig. 5.7(b). What might be confusing when considering plane waves is obvious when considering Gaussian beams, the beamwidth of the incident beam (L_i) is not the same as the beamwidth of the transmitted beam (L_t), except in the trivial case where $\theta_\mathrm{i} = \theta_\mathrm{t}$. This means that even if the Poynting vectors of the incident (S_i) and transmitted (S_t) waves are equal, the incident power (defined as $P_\mathrm{i} = S_\mathrm{i} L_\mathrm{i}$) is, in general, not the same as the transmitted power (P_t). This translates into a transmission efficiency (η) given by

$$\eta = \frac{P_\mathrm{t}}{P_\mathrm{i}} = \frac{S_\mathrm{t} L_\mathrm{t}}{S_\mathrm{i} L_\mathrm{i}} = (A^\mathrm{t})^2 \frac{\cos\theta_\mathrm{t}}{\cos\theta_\mathrm{i}}. \tag{5.31}$$

Note that here, the amplitude of the incident wave is assumed to be $A^\mathrm{i} = 1$. The power dissipated or produced by the metasurface (P_m) is easily obtained as the difference of the incident power and the transmitted power

$$P_\mathrm{m} = P_\mathrm{i} - P_\mathrm{t} = \cos\theta_\mathrm{i} - (A^\mathrm{t})^2 \cos\theta_\mathrm{t}, \tag{5.32}$$

which can also be written as

$$P_\mathrm{m} = P_\mathrm{i}(1 - \eta). \tag{5.33}$$

It is clear, from Eq. (5.31), that the transmission coefficient A^t as well as the angles θ_i and θ_t play major roles in determining the transmission efficiency of the metasurface. If $\theta_\mathrm{t} > \theta_\mathrm{i}$ (as in the example considered here), the metasurface remains purely lossy (both electrically and magnetically) as long as the transmission coefficient is limited to the range $0 \leq A^\mathrm{t} \leq 1$. The metasurface is naturally more and more lossy as A^t decreases to 0.

In our example, $A^\mathrm{t} = 1$, which gives a power efficiency of $\eta = 75.25\%$, and the total dissipated power by the metasurface can be obtained by integrating both electric and magnetic contributions (corresponding to the black line in Fig. 5.8(a)) over the size of the metasurface. Increasing the value of A^t beyond 1 will increase the efficiency but at the cost of having a partially active metasurface. In the particular case where $A^\mathrm{t} = \sqrt{\cos\theta_\mathrm{i}/\cos\theta_\mathrm{t}}$, the efficiency (5.31)

is always equal to $\eta = 1$, and consequently $P_{\mathrm{m}} = 0$. But even if the total dissipated or produced power by the metasurface is 0, the metasurface is actually a combination of alternating electric gain and loss and magnetic loss and gain such that electric and magnetic contributions perfectly cancel each other, as illustrated in Fig. 5.8(b). It is therefore wrong to assume that because $P_{\mathrm{m}} = 0$, the metasurface is passive and lossless because, as illustrated here, the metasurface would actually be simultaneously active and lossy. Note that, in the case where $\theta_{\mathrm{t}} < \theta_{\mathrm{i}}$, it was shown Ref. 40 that the metasurface is purely passive (but lossy) when $A^{\mathrm{t}} = \cos\theta_{\mathrm{i}}/\cos\theta_{\mathrm{t}}$.

The transmission and reflection coefficients, given by relations (5.15), which correspond to the transformation described above, are shown in Fig. 5.9. As may be seen in Fig. 5.9(a), the absolute value of the transmission coefficient oscillates around 87% (which, converted in terms of transmitted power, is very close to the previously calculated power efficiency, $\eta = 75.25\%$, while the absolute value of the reflection coefficient oscillates around 13%. The fact that the observed reflection coefficient is non-zero may *a priori* appear contradictory given the prescription of zero reflection. However, remember that the scattering coefficients are computed based on the assumption of rectilinear propagation, which obviously does

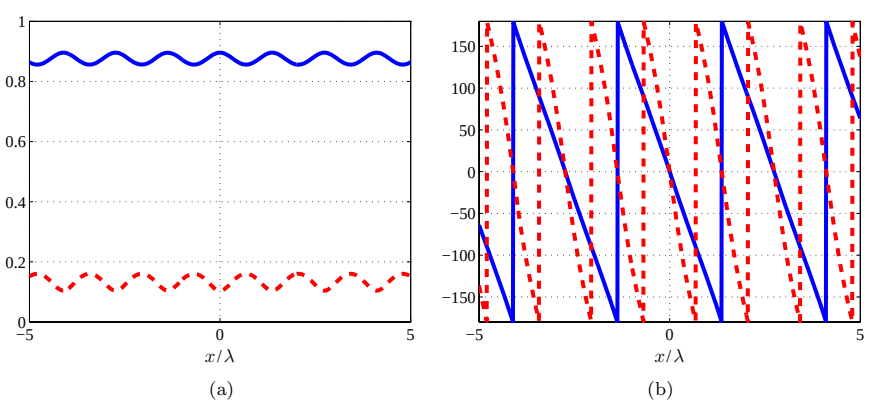


Fig. 5.9. (a) Amplitude and (b) phase of the transmission (blue line) and reflection (dashed red line) coefficients calculated using (5.16) with the susceptibilities in 5.6.

not correspond to the present example. The actual reflection produced by the susceptibilities plotted in Fig. 5.6 is *rigorously zero*, and the non-zero reflection parameter in Fig. 5.9(a) is an artifact of the mapping between the rectilinear scattering parameters and the physical problem.

However, as pointed out in the last paragraph of section 5.2.3.2, these scattering parameters can be directly used for synthesis: full-wave (using periodic boundary conditions) designing the scattering particles so that they produce the same scattering parameters as those obtained using (5.15) will automatically provide the desired physical solution in real (non-rectilinear) conditions.

At this point, it must be noted that the synthesized metasurface can be rigorously described by its susceptibilities, as given in Fig. 5.6, but it can also be alternatively described as an equivalent combination of an amplitude and a phase grating, as given in Fig. 5.9. So far, only the synthesis of the metasurface has been discussed, but considering the metasurface has an amplitude/phase grating can be used as an analysis tool to simply and efficiently predict how the metasurface would scatter a given impinging wave. For instance, using Fourier optics, the reflected and transmitted waves can be respectively approximated by $\psi^{\mathrm{r}} = \psi^{\mathrm{i}} R$ and $\psi^{\mathrm{t}} = \psi^{\mathrm{i}} T$, where ψ^{i} corresponds to the projected phase of the incident wave on the metasurface. Then, the k-vectors of the reflected and transmitted waves can be respectively found by taking the Fourier transform of ψ^{r} and ψ^{t} which gives crucial information on the direction of propagation of these waves. In the particular case where ψ^{i} corresponds to the specified incident wave, ψ^{t} corresponds to a transmitted wave propagating at a $\pi/4$ angle while ψ^{r} corresponds to an induced surface wave. The fact that the reflected wave corresponds to a surface wave can be understood by noting that the period of the phase of R in Fig. 5.9 is twice that of the phase of T. Consequently, the vector component k_x^{r} is not only larger than k_x^{t} but also larger than the wavenumber k, which results in an imaginary k_z^{r} component effectively classifying the reflected wave as a surface wave. This also helps understanding why the metasurface appears to have reflection while the specified

reflected wave was initially set to zero. In the susceptibility approach, there is no reflected wave and part of the energy of the incident wave is absorbed by the metasurface. In the grating approach, what appears to be a reflected wave is in fact a surface wave and the only wave scattered by the metasurface is the specified transmitted wave.

An important deduction is that a monoanisotropic metasurface (i.e. only described by the susceptibility tensors $\overline{\overline{\chi}}_{\mathrm{ee}}$ and $\overline{\overline{\chi}}_{\mathrm{mm}}$), used for refraction or reflection (or both), can never be passive and/or lossless. A question that might arise then is what would happen if only the real values of the susceptibilities were realized instead of the complex ones shown in Fig. 5.6? In that case, the magnitude of the new transmission coefficient (which is not shown here) would be very close to full and quasi-uniform transmission ($T \approx 100\%$ and can be approximated as such) while the phase profile of the transmission coefficient would remain unaltered compared to that computed from the exact susceptibilities presented in Fig. 5.9(b). Since the transmission phase has not been changed from the exact one, the approximate metasurface performs a transformation that essentially follows the specification, with the exception of undesired diffraction orders due to the nullification of the imaginary parts of the susceptibilities. It results that, whether a lossy structure is implemented or an altered version of it where only the real values of the susceptibilities are implemented, the total power being transmitted into the desired direction remains the same because, in the former case, part of the power is absorbed by the metasurface while, in the latter case, part of the power is diffracted into other directions. In practice, it is often the latter case that is preferred as it corresponds to an easier structure to realize because the metasurface is essentially only affecting the phase of the incident wave while keeping its amplitude constant.

There exists, however, an alternative approach to realize passive, lossless and fully efficient refractive metasurfaces. As shown in Ref. 41, the monoanisotropic metasurface discussed above is inherently a symmetric structure and thus, provides the same impedance matching from both sides of the metasurface, which is the reason

why it fails to efficiently refract an incident wave. The alternative approach would be to consider a bianisotropic metasurface which, from its non-zero magnetoelectric coupling coefficients, would be asymmetric. Because rotation of polarization is not required in refractive metasurfaces, the non-zero susceptibilities considered for the synthesis are the diagonal components of $\overline{\overline{\chi}}_{\mathrm{ee}}$ and $\overline{\overline{\chi}}_{\mathrm{mm}}$ and the non-diagonal components of $\overline{\overline{\chi}}_{\mathrm{em}}$ and $\overline{\overline{\chi}}_{\mathrm{me}}$. The introduction of non-zero magnetoelectric coupling coefficients doubles the number of unknowns, which means that the multiple wave transformation technique described in section 5.2.3.4 is used here to obtain the following fully determined system of equations

$$\begin{pmatrix} \Delta H_{y1} & \Delta H_{y2} \\ \Delta H_{x1} & \Delta H_{x2} \\ \Delta E_{y1} & \Delta E_{y2} \\ \Delta E_{x1} & \Delta E_{x2} \end{pmatrix} = \begin{pmatrix} \widetilde{\chi}_{\mathrm{ee}}^{xx} & 0 & 0 & \widetilde{\chi}_{\mathrm{em}}^{xy} \\ 0 & \widetilde{\chi}_{\mathrm{ee}}^{yy} & \widetilde{\chi}_{\mathrm{em}}^{yx} & 0 \\ 0 & \widetilde{\chi}_{\mathrm{me}}^{xy} & \widetilde{\chi}_{\mathrm{mm}}^{xx} & 0 \\ \widetilde{\chi}_{\mathrm{me}}^{yx} & 0 & 0 & \widetilde{\chi}_{\mathrm{mm}}^{yy} \end{pmatrix} \begin{pmatrix} E_{x1,\mathrm{av}} & E_{x2,\mathrm{av}} \\ E_{y1,\mathrm{av}} & E_{y2,\mathrm{av}} \\ H_{x1,\mathrm{av}} & H_{x2,\mathrm{av}} \\ H_{y1,\mathrm{av}} & H_{y2,\mathrm{av}} \end{pmatrix}, \tag{5.34}$$

where the second transformation is the reciprocal of the first one. In other words, if Fig. 5.7(b) represents the first transformation, the second transformation would consist in reversing the direction of propagation of all the waves while maintaining their respective amplitude. For p-polarization, the system (5.34) reduces to

$$\begin{pmatrix} \Delta H_{y1} & \Delta H_{y2} \\ \Delta E_{x1} & \Delta E_{x2} \end{pmatrix} = \begin{pmatrix} \widetilde{\chi}_{\mathrm{ee}}^{xx} & \widetilde{\chi}_{\mathrm{em}}^{xy} \\ \widetilde{\chi}_{\mathrm{me}}^{yx} & \widetilde{\chi}_{\mathrm{mm}}^{yy} \end{pmatrix} \begin{pmatrix} E_{x1,\mathrm{av}} & E_{x2,\mathrm{av}} \\ H_{y1,\mathrm{av}} & H_{y2,\mathrm{av}} \end{pmatrix}, \tag{5.35}$$

which can be solved easily to yield the following susceptibility functions

$$\chi_{\mathrm{ee}}^{xx} = \frac{4\sin(\alpha x)}{\beta\cos(\alpha x) + \sqrt{\beta^2 - \gamma^2}}, \tag{5.36a}$$

$$\chi_{\mathrm{mm}}^{yy} = \frac{\beta^2 - \gamma^2}{4k^2} \frac{4\sin(\alpha x)}{\beta\cos(\alpha x) + \sqrt{\beta^2 - \gamma^2}}, \tag{5.36b}$$

$$\chi_{\mathrm{em}}^{xy} = -\chi_{\mathrm{me}}^{yx} = \frac{2j}{k} \frac{\gamma\cos(\alpha x)}{\beta\cos(\alpha x) + \sqrt{\beta^2 - \gamma^2}}, \tag{5.36c}$$

where $\beta = k_z^{\mathrm{i}} + k_z^{\mathrm{t}}$, $\gamma = k_z^{\mathrm{i}} - k_z^{\mathrm{t}}$ and $\alpha = k_x^{\mathrm{t}} - k_x^{\mathrm{i}}$ and where relations (5.29) have been used with parameters $A^{\mathrm{i}} = 1$ and $A^{\mathrm{t}} = \sqrt{\cos\theta_{\mathrm{i}}/\cos\theta_{\mathrm{t}}}$. Note that this last relation is equivalent to equalizing the longitudinal power flow (Poynting vector) for the waves on both sides of the metasurface and also corresponds to the case discussed above where the monoanisotropic metasurface has an efficiency $\eta = 1$ with a combination of active and lossy elements as was shown in Fig. 5.8(b). Now, in the case of the bianisotropic metasurface, it can be verified that relations (5.36) correspond to not only a reciprocal (see (5.9)) but also a passive/lossless (see (5.10)) metasurface.

In this section, we have seen how a refractive metasurface could be synthesized. We have seen that a symmetric (or monoanisotropic) refractive metasurface is always either lossy or active (or both) and in the case of purely lossy susceptibilities, the transformation is inherently inefficient. It is possible to ignore the imaginary parts of the susceptibilities (which are responsible for the loss) at the expense of additional undesired diffraction orders. We have also addressed the case of asymmetric (or bianisotropic) refractive metasurfaces which, compared to their symmetric counterparts, have perfect transmission efficiency while being reciprocal, lossless and passive. However, the better performances of bianisotropic metasurfaces come at the price of more complex realization requirements due to their non-zero magnetoelectric coupling coefficients. Finally, we have also discussed a simple analysis tool based on Fourier optics to predict the behavior of the metasurface under different illumination angles.

5.4.2. *Fundamental limitations of metasurfaces*

Metasurfaces, and metamaterials in general, are composed of metallic or dielectric subwavelength resonant particles. This fact results in several drawbacks and limitations. Metallic structures usually present high losses especially at optical frequencies. Resonant particles often have limited bandwidth which prevents the implementation of broadband devices. Moreover, it is difficult to design resonant particles that are much smaller than the wavelength because

the smaller the particles are, the less they interact with the incident field and, consequently, the less they can control it. For this reason, metasurfaces having susceptibilities with very fast spatial variations over subwavelength distances (e.g. refractive metasurface with very large angle of refraction) might not be implementable since the discretized metasurface unit cell size would be so small that the scattering particles would not resonate. For instance, the scattering particles, discussed in section 5.3, have a dimension of $\lambda/5$ in the plane of the metasurface. Reaching smaller unit cell size while maintaining a strong control in terms of phase and transmission of the incident field is challenging.

Metasurfaces have been presented as 2D reductions of volume metamaterials and are consequently less bulky, less lossy and easier to fabricate. The synthesis technique presented in this chapter shows that a metasurface has the capability to mathematically transform an arbitrary incident fields into arbitrary reflected and transmitted fields. But, does the reduction of 3D metamaterials to 2D metasurfaces imply reduced processing power or functionalities? Is there anything that a metasurface cannot do, due to its reduced dimensionality, that a 3D metamaterial can achieve? One example could be the electromagnetic cloaks (used in transmission) that require the object to be hidden to be surrounded by a metamaterial structure, and thus metasurfaces would not be appropriate in such situation.

Beam expanders (or telescopes), usually used in optics[42] to increase the width of an incident Gaussian beam, would also be difficult to realize with a metasurface. Beam expanders are easily implemented with the combination of two different lenses separated by the sum of their focal lengths. Such optical component could not be realized with a single zero-thickness passive metasurface because the extent of the transmitted beam in the plane of the metasurface is larger than that of the incident beam and, consequently, the metasurface would have to amplify the incident beam in the region far from the center. In this case, using a 2D metasurface goes at the cost of having an active structure compared to beam expanders that are purely passive.

Another example is the realization of optical analog processing systems[43] based on Fourier optics principles. Often referred to as 4f-systems, they are composed of two lenses, which successively perform the spatial Fourier transforms of the input field. In between them, at the Fourier plane, a mask can be placed to block certain spatial components. Therefore, these systems can be used as spatial filters where the transmitted beam corresponds to the incident beam convolved with the Fourier transform of a given mask function. It is easy to show that such a system can be implemented with a zero-thickness metasurface but with important limitations. Assume that the metasurface is synthesized with the susceptibilities in (5.12). Here, for simplicity, only the component χ_{ee}^{xx} is presented. The spatial filtering metasurface is given by

$$\chi_{\mathrm{ee}}^{xx} = \frac{-\Delta H_y}{j\omega\epsilon E_{x,\mathrm{av}}} = \frac{2j}{\omega\epsilon}\frac{H_y^{\mathrm{t}} - H_y^{\mathrm{i}}}{E_x^{\mathrm{t}} + E_x^{\mathrm{i}}} = \frac{2j}{\omega\epsilon}\frac{\mathcal{F}\{M\} * H_y^{\mathrm{i}} - H_y^{\mathrm{i}}}{\mathcal{F}\{M\} * E_x^{\mathrm{i}} + E_x^{\mathrm{i}}}, \tag{5.37}$$

where M is a mask function. The presence of convolution products in (5.37) shows that the susceptibilities are inherently proportional to the specified incident field. As a consequence, the metasurface acts as a spatial filtering device but only for the specified incident field, so that different incident fields result in different transformations in contrast to what happens in an optical 4f-system that performs the same filtering operation independently of the incident beam. Here again, the reduced dimensionality of metasurfaces is disadvantageous.

These limitations might seem to be detrimental to the future of metasurfaces. However, solutions exist to overcome some of these difficulties. For example, the problems arising due to the reduced dimensionality of metasurfaces can be mitigated by cascading several metasurfaces instead of using just one. In that case, an optical system that has a length of several focal lengths, like the 4f-system for example, corresponding to several thousand of wavelengths could be implemented with a length of only a few tens of wavelengths using metamaterials.[44] The issues caused by the high losses of metallic scattering particles and limited bandwidth of resonant structures

could be addressed by using dielectric resonators and highly coupled particles as done, for instance, in Ref. 45.

5.4.3. *Limitations of the synthesis technique*

The mathematical synthesis technique presented in this document is certainly powerful, but its main drawback is that it describes a zero-thickness metasurface, a fictitious interface that obviously does not exist in reality. In general, we have assumed that if the physical metasurface is very thin compared to the free space wavelength, then it is accurately modeled by the GSTCs. In this section, we will see an example which illustrates the divergence between the ideal model and reality.

When using the synthesis technique to model a subwavelength thick metasurface, the GSTCs provide relatively accurate results only to some extent. The GSTCs are, by definition, intended to deal with discontinuities. Consequently, a material slab (even deeply subwavelength) can not be accurately modelled by such continuity conditions and must rather be analyzed as a two-interface problem using the usual boundary conditions. Usually when ideal susceptibilities, obtained using the synthesis technique, are associated to a subwavelength metasurface, discrepancies appear between the simulated metasurface response and the expected response.

In order to evaluate the discrepancies introduced by the metasurface thickness, a simple numerical experiment is conducted next. A metasurface is synthesized to absorb a normally incident plane wave. The incident plane wave has an electric field defined by $\boldsymbol{E}_{\mathrm{i}} = \hat{\boldsymbol{x}} e^{-jkz}$, and the transmitted plane wave has an electric field defined by $\boldsymbol{E}_{\mathrm{t}} = \hat{\boldsymbol{x}} T e^{-jkz}$, where the transmission coefficient T can vary between 0 and 1. Inserting these fields into (5.12a) and (5.12d) yields

$$\chi = \chi_{\mathrm{ee}}^{xx} = \chi_{\mathrm{mm}}^{yy} = \frac{2j}{k_0} \frac{(T-1)}{(T+1)}. \tag{5.38}$$

The susceptibilities in (5.38) can be easily converted into the electric permittivity, $\epsilon_r = 1 + \chi_{\mathrm{ee}}^{xx}/d$, and the magnetic permeability, $\mu_r = 1 + \chi_{\mathrm{mm}}^{yy}/d$, where d is the thickness of the metasurface. Dividing

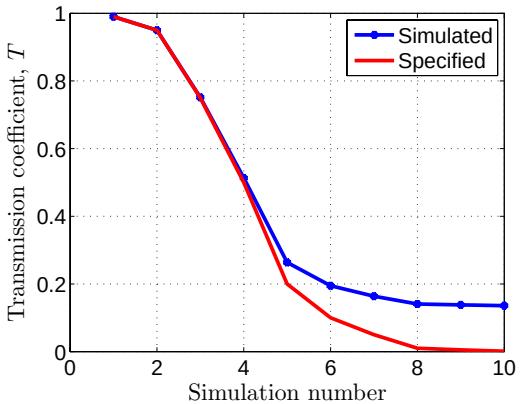


Fig. 5.10. The red curve corresponds to the specified transmission coefficient. The blue curve corresponds to the simulated transmission coefficients for a metasurface of thickness $d = \lambda/100$.

by d dilutes the effect of the susceptibilities over the thickness of the metasurface. It is a valid approximation, as shown in Ref. 28, as long as d remains subwavelength.

Electromagnetic simulations are performed using COMSOL for different values of T and the results are reported in Fig. 5.10. As can be seen, for $T > 0.5$ the simulated transmission is in good agreement with the specification. But for $T < 0.5$, a discrepancy appears, increasing as T is reduced to 0.

In order to understand the results presented in Fig. 5.10, the problem is analyzed as a two-interface problem where the surrounding medium 1 is vacuum with parameters

$$n_1 = 1, \quad Z_1 = \sqrt{\frac{\mu_0}{\epsilon_0}}, \tag{5.39}$$

and where medium 2, corresponding to the metasurface, can be defined easily by noting that the electric and magnetic susceptibilities are equal to each other as shown in (5.38), and consequently we have

$$n_2 = \sqrt{(1+\chi/d)^2} = 1 + \frac{\chi}{d}, \quad Z_2 = \sqrt{\frac{\mu_0}{\epsilon_0}}\sqrt{\frac{1+\chi/d}{1+\chi/d}} = Z_1. \tag{5.40}$$

The transmission coefficient, T_{21}, from medium 1 to medium 2 and the reflection coefficient, R, inside the metasurface are given by

$$T_{21} = \frac{2Z_2}{Z_1 + Z_2} = 1, \quad R = \frac{Z_1 - Z_2}{Z_1 + Z_2} = 0. \tag{5.41}$$

Here, we see that the metasurface is reflectionless as initially specified. Since there is no reflection within the metasurface, the transmitted amplitude simply reads

$$|T_{\text{trs}}| = |e^{-jk_0 n_2 d}| = |e^{-jk_0 d}||e^{-jk_0 \chi}| = \exp\left[2\frac{(T-1)}{(T+1)}\right]. \tag{5.42}$$

The expression (5.42) precisely agrees with the simulated results (blue curve) of Fig. 5.10. It is clear that as T decreases to 0, the total transmission from the subwavelength thick metasurface converges to e^{-2}. This simple example illustrates how the GSTCs fail to accurately model the response of this metasurface. For this very specific case, where only normal incidence and transmission are considered, it is possible to modify the ideal susceptibilities of the zero-thickness metasurface such that the response of the subwavelength thick metasurface exactly matches the response of the zero-thickness one. If the ideal susceptibility, χ, is defined from (5.38), then the "corrected" susceptibility is

$$\chi_{\text{corr}} = \chi\frac{f}{d}, \tag{5.43}$$

where f is a correcting factor. Following the same procedure as used from (5.39) to (5.42) with (5.43) instead of (5.38), the factor f can be found and Eq. (5.38) becomes

$$\chi_{\text{corr}} = \chi\frac{\ln T}{2d}\frac{T+1}{T-1} = \frac{j\ln T}{k_0 d}. \tag{5.44}$$

The corrected susceptibility (5.44) gives the expected transmission coefficient $|T_{\text{trs}}| = |T|$. From (5.44), we see that the corrected susceptibility for the subwavelength thick metasurface is proportional to the natural logarithm of T meaning that, as T decreases to 0, the electric and magnetic susceptibilities both converge to $-j\infty$ whereas the ideal susceptibilities (5.38) converge to $-2j/k_0$. The important

difference between (5.38) and (5.44) illustrates how implementing the ideal susceptibilities can become impracticable when the thickness of the metasurface is taken into account and how much the corrected susceptibilities diverge from the ideal ones.

The particular example presented here yields the simple corrected expression (5.44), but in general no analytical forms would exist for the corrected susceptibilities. One possibility that has been considered, to reduce the aforementioned discrepancies, is to derive boundary conditions of higher orders than the GSTCs. As they stand, the GSTCs only account for a zeroth order discontinuity, meaning that only the discontinuities of the fields are taken into account but not the discontinuities of the derivatives of the fields. Higher-order boundary conditions that take into considerations the first derivative of the fields may therefore yield more accurate results. The developments of such boundary conditions are discussed in Ref. 46.

5.5. Conclusion

A metasurface synthesis method based on transverse susceptibility tensors has been introduced. The technique provides closed-form expressions for selected electric and magnetic susceptibility components to theoretically perform electromagnetic transformations where the incident, reflected and transmitted waves can be specified arbitrarily. The metasurface can be reflection-less or transmission-less and can have an infinite or a finite size. Moreover, it has been shown that, by selecting more transverse susceptibility components, it is possible to perform several (up to four) sets of independent electromagnetic transformations with the same metasurface, thus allowing multi-functionality. The proposed method can handle, among others, reciprocal or non-reciprocal electromagnetic transformations, generalized refraction, polarization rotation and orbital angular momentum multiplexing. In other words, the method can be used to perform any electromagnetic transformation, without needing to resort to case-specific synthesis techniques.

The proposed synthesis method also includes the realization of the physical scattering particles that would correspond to the

synthesized ideal susceptibilities. Note that, in some cases, the implementation of the metasurface might be practically difficult or even impossible to realize. However, even in relatively extreme cases, typically corresponding to fast susceptibility variations in comparison with the wavelength, the proposed synthesis might be used as an initial and insightful step of the complete synthesis. Moreover, if required, one could practically relax some of the used assumptions, including the zero thickness of the metasurface, introducing non-zero longitudinal dipole moments (P_z and M_z), and allowing more non-zero susceptibility tensor components.

Appendix

A. *Distribution-based generalized sheet transition conditions (GSTCs)*[1,28]

A function $f(z)$ that is discontinuous up to the Nth order at $z = 0$ may be expressed in the sense of distributions as

$$f(z) = \{f(z)\} + \sum_{k=0}^{N} f_k \delta^{(k)}(z). \tag{A.1}$$

In this relation, $\{f(z)\}$ and $\sum_{k=0}^{N} f_k \delta^{(k)}(z)$ are the regular and singular parts of $f(z)$, respectively. The regular part is defined for $z \neq 0$ in the sense of usual functions as

$$\{f(z)\} = f_+(z)U(z) + f_-(z)U(-z), \tag{A.2}$$

where $U(x)$ is the unit step function and $f_\pm(z)$ denote the parts of $f(z)$ in the regions $z \gtrless 0$. The singular part, defined at $z = 0$, is a Taylor-type series, where $\delta^{(k)}(z)$ is the k^{th} derivative of the Dirac delta function, and f_k is the corresponding weighting coefficient, which is z-independent.

The function $f(z)$ in (A.1) represents here any of the quantities in Maxwell equations. Since these equations involve spatial derivatives, the question arises as how to compute the z-derivative of $f(z)$. Since f_k does not depend on z, taking the z-derivatives of the singular part of (A.1) only increases the derivative order of the Dirac delta

function, from k to $k+1$. On the other hand, the derivative of the regular part, given by (A.2), involves the derivative of $U(\cdot)$, which may be expressed in the sense of distributions, in connection with a test function ϕ, as

$$\langle U', \phi \rangle = -\langle U, \phi' \rangle = \langle \delta, \phi \rangle, \tag{A.3}$$

where $\langle \cdot, \cdot \rangle$ represents the functional inner product. In (A.3), the first equality was obtained by integrating by part and taking into account the fact that ϕ has a finite support, while the second equality follows from setting the lower bound of the integral to zero for eliminating U, using the fact that the primitive of ϕ' is ϕ, by definition, and again that ϕ has a finite support, and finally applying the sifting property of the Dirac delta function according to which $\phi(0) = \langle \delta, \phi \rangle$. In other words, the derivative of the unit step function is the Dirac delta function. Therefore, using (A.3), the z-derivative of (A.2) is obtained as

$$\begin{aligned} \frac{d}{dz}\{f(z)\} &= \{f'_+(z)U(z) + f'_-(z)U(-z)\} + [f_+(0) - f_-(0)]\delta(z) \\ &= \{f'\} + [[f]]\delta(z), \end{aligned} \tag{A.4}$$

where $\{f'\}$ (curl bracket term in the second equality) represents the regular part of the derivative of $f(z)$, defined at $z \neq 0$, and the term $[[f]]$ (square bracket term in the second equality) represents the singularity, at $z = 0$. Remember that the unit of $\delta(z)$ is $(z)^{-1}$ since $\int_{-\infty}^{+\infty} \delta(z)dz = 1$ is dimensionless.

Rigorous GSTCs can now be derived using (A.1) and (A.4). The derivation is performed here only for Maxwell–Ampère equation, as the derivations for the other Maxwell equations are essentially similar. Maxwell–Ampère equation in the monochromatic regime reads

$$\nabla \times \boldsymbol{H} = \boldsymbol{J} + j\omega \boldsymbol{D}. \tag{A.5}$$

Expressing $\boldsymbol{H}$ in the form of (A.1) and using the transverse-longitudinal decomposition $\nabla = \nabla_{\parallel} + \hat{z}\frac{\partial}{\partial z}$ transforms the left-hand

side of (A.5) into

$$\nabla \times \boldsymbol{H} = \nabla_{\parallel} \times \{\boldsymbol{H}\} + \hat{z} \times \frac{\partial}{\partial z}\{\boldsymbol{H}\} + \sum_{k=0}^{N} \nabla_{\parallel} \times \boldsymbol{H}_k \delta^{(k)}(z) + \sum_{k=0}^{N} \hat{z} \times \frac{\partial}{\partial z} \boldsymbol{H}_k \delta^{(k)}(z). \tag{A.6}$$

In the right-hand side of (A.6), the second term can be evaluated using (A.4), while the derivative in the last term only affects the Dirac delta function since $\boldsymbol{H}_k$ does not depend on z. Therefore, Eq. (A.6) becomes

$$\begin{aligned} \nabla \times \boldsymbol{H} = \nabla_{\parallel} \times \{\boldsymbol{H}\} + \hat{z} \times \left\{ \frac{\partial}{\partial z} \boldsymbol{H} \right\} + \hat{z} \times [[\boldsymbol{H}]] \delta(z) \\ + \sum_{k=0}^{N} \nabla_{\parallel} \times \boldsymbol{H}_k \delta^{(k)}(z) + \sum_{k=0}^{N} \hat{z} \times \boldsymbol{H}_k \delta^{(k+1)}(z), \end{aligned} \tag{A.7}$$

where the first two terms and the last two terms are the regular and singular parts, respectively.

Substituting (A.7) along with the (A.1) expressions of $\boldsymbol{D}$ and $\boldsymbol{J}$ into (A.5) finally transforms Maxwell–Ampère equation into

$$\begin{aligned} \nabla_{\parallel} \times \{\boldsymbol{H}\} + \hat{z} \times \left\{ \frac{\partial}{\partial z} \boldsymbol{H} \right\} + \hat{z} \times [[\boldsymbol{H}]] \delta(z) \\ + \sum_{k=0}^{N} \nabla_{\parallel} \times \boldsymbol{H}_k \delta^{(k)}(z) + \sum_{k=0}^{N} \hat{z} \times \boldsymbol{H}_k \delta^{(k+1)}(z) \\ = \{\mathbf{J}(z)\} + \sum_{k=0}^{N} \mathbf{J}_k \delta^{(k)}(z) + j\omega\{\mathbf{D}(z)\} + j\omega \sum_{k=0}^{N} \mathbf{D}_k \delta^{(k)}(z), \end{aligned} \tag{A.8}$$

where $\{\boldsymbol{J}(z)\}$ is a volume current, measured in $(\mathrm{A/m^2})$, while $\mathbf{J}_k$ represents surface currents, measured in $(\mathrm{A \cdot m^{k-1}})$ since the unit of $\delta^{(k)}(z)$ is $(1/\mathrm{m^{k+1}})$. One may now equate the terms of the same

discontinuity orders, i.e. of same Dirac derivative orders, in this equation and in the other three corresponding Maxwell equations.[d] The result is, for the terms of order $\delta^{(0)}(z) = \delta(z)$,

$$\hat{z} \times [[\boldsymbol{H}]] + \nabla_{\parallel} \times \boldsymbol{H}_0 = \boldsymbol{J}_0 + j\omega \boldsymbol{D}_0, \tag{A.9a}$$

$$\hat{z} \times [[\boldsymbol{E}]] + \nabla_{\parallel} \times \boldsymbol{E}_0 = -\boldsymbol{K}_0 - j\omega \boldsymbol{B}_0, \tag{A.9b}$$

$$\hat{z} \cdot [[\boldsymbol{D}]] + \nabla_{\parallel} \cdot \boldsymbol{D}_0 = \rho_0, \tag{A.9c}$$

$$\hat{z} \cdot [[\boldsymbol{B}]] + \nabla_{\parallel} \cdot \boldsymbol{B}_0 = m_0 \tag{A.9d}$$

and, for the terms of order $\delta^{(k)}(z)$ with $k \geq 1$,

$$\hat{z} \times \boldsymbol{H}_{k-1} + \nabla_{\parallel} \times \boldsymbol{H}_k = \boldsymbol{J}_k + j\omega \boldsymbol{D}_k, \tag{A.10a}$$

$$\hat{z} \times \boldsymbol{E}_{k-1} + \nabla_{\parallel} \times \boldsymbol{E}_k = -\boldsymbol{K}_k - j\omega \boldsymbol{B}_k, \tag{A.10b}$$

$$\hat{z} \cdot \boldsymbol{D}_{k-1} + \nabla_{\parallel} \cdot \boldsymbol{D}_k = \rho_k, \tag{A.10c}$$

$$\hat{z} \cdot \boldsymbol{B}_{k-1} + \nabla_{\parallel} \cdot \boldsymbol{B}_k = m_k. \tag{A.10d}$$

Equation (A.9) are the *universal boundary conditions for monochromatic waves at a planar surface at rest*, while Eqs. (A.10) are *compatibility relations* that must to be recursively applied to determine the unknown terms in (A.9).[28] Note, letting $z \to 0$ in the regular parts of (A.9), the presence of additional terms compared to the case of conventional boundary conditions (e.g. Eq. (A.9a), where $[[\boldsymbol{H}]] = [H(z = 0_+) - H(z = 0_-)]$ and $\mathbf{J}_0$ is the sheet surface current, includes the additional terms $\nabla_{\parallel} \times \mathbf{H}_0$ and $j\omega \mathbf{D}_0$).

Let us now specialize to the case of interest: an *infinitesimal sheet discontinuity in free space*. This means that the quantities $\boldsymbol{J}_k$, $\boldsymbol{K}_k$, ρ_k and m_k exclusively reside at $z = 0$, so that $\boldsymbol{J}_k \equiv \boldsymbol{K}_k \equiv \rho_k \equiv m_k \equiv 0$ for $k \geq 1$, meaning that only the term $k = 0$ survives in the series (A.1) for these quantities. However, the situation is different for the fields $\boldsymbol{E}_k$, $\boldsymbol{E}_k$, $\boldsymbol{D}_k$ and $\boldsymbol{B}_k$, since these fields exist also at $z \neq 0$. Strictly, $N \to \infty$ for these fields. However, since the

[d] Rigorously, the Dirac delta function disappears upon integrating over z terms of equal discontinuity order.

discontinuity is purely concentrated at $z = 0$, the Taylor-type series in (A.1) includes only a small number of significant terms, and the series can be safely truncated at some value of N. Choosing some value for N (e.g. $N = 2$), equations (A.10) may be solved recursively for $k = N$ to $k = 1$, with $\boldsymbol{D}_k = \epsilon \boldsymbol{E}_k$ and $\boldsymbol{B}_k = \mu \boldsymbol{H}_k$. This procedure reduces the compatibility relations to

$$\hat{z} \times \boldsymbol{H}_0 = 0, \tag{A.11a}$$

$$\hat{z} \times \boldsymbol{E}_0 = 0, \tag{A.11b}$$

$$\hat{z} \cdot \boldsymbol{D}_0 = 0, \tag{A.11c}$$

$$\hat{z} \cdot \boldsymbol{B}_0 = 0. \tag{A.11d}$$

One may now introduce the electric and magnetic polarization densities, $\mathbf{P}$ and $\mathbf{M}$, respectively, to account for the action of the scattering particles forming the metasurface. For this purpose, the standard constitutive relations $\boldsymbol{D} = \epsilon \boldsymbol{E} + \boldsymbol{P}$ and $\boldsymbol{B} = \mu(\boldsymbol{H} + \boldsymbol{M})$ are in a form that properly models the first-order surface discontinuity in (A.9), namely

$$\boldsymbol{D}_0 = \epsilon \boldsymbol{E}_0 + \boldsymbol{P}_0, \tag{A.12a}$$

$$\boldsymbol{H}_0 = \frac{1}{\mu} \boldsymbol{B}_0 - \boldsymbol{M}_0, \tag{A.12b}$$

where $\boldsymbol{P}_0$ and $\boldsymbol{M}_0$ represent the (first order) electric and magnetic *surface* polarization densities, respectively. In the absence of sources ($\boldsymbol{J}_0 = \boldsymbol{K}_0 = \rho_0 = m_0 = 0$), substitution of (A.12) and application of (A.11) transforms (A.9) into

$$\hat{z} \times [[\boldsymbol{H}]] = j\omega \boldsymbol{D}_0 - \nabla_{\parallel} \times \boldsymbol{H}_0 = j\omega \boldsymbol{P}_{0,\parallel} + \nabla_{\parallel} \times \boldsymbol{M}_{0,\mathrm{n}}, \tag{A.13a}$$

$$\hat{z} \times [[\boldsymbol{E}]] = -j\omega \boldsymbol{B}_0 - \nabla_{\parallel} \times \boldsymbol{E}_0 = -j\omega\mu \boldsymbol{M}_{0,\parallel} + \frac{1}{\epsilon} \nabla_{\parallel} \times \boldsymbol{P}_{0,\mathrm{n}}, \tag{A.13b}$$

$$\hat{z} \cdot [[\boldsymbol{D}]] = -\nabla_{\parallel} \cdot \boldsymbol{D}_0 = -\nabla_{\parallel} \cdot \boldsymbol{P}_{0,\parallel}, \tag{A.13c}$$

$$\hat{z} \cdot [[\boldsymbol{B}]] = -\nabla_{\parallel} \cdot \boldsymbol{B}_0 = -\mu \nabla_{\parallel} \cdot \boldsymbol{M}_{0,\parallel}, \tag{A.13d}$$

where the subscripts $\parallel$ and n denote transverse and normal components, respectively.

Using the relation $\nabla_\parallel \times (\hat{\boldsymbol{z}}\psi) = -\hat{\boldsymbol{z}} \times \nabla_\parallel \psi$ and the difference notation (5.3), Eqs. (A.13) finally take the form

$$\hat{z} \times \Delta \boldsymbol{H} = j\omega \boldsymbol{P}_\parallel - \hat{z} \times \nabla_\parallel M_z, \tag{A.14a}$$

$$\Delta \boldsymbol{E} \times \hat{z} = j\omega\mu \boldsymbol{M}_\parallel - \nabla_\parallel \left(\frac{P_z}{\epsilon} \right) \times \hat{z}, \tag{A.14b}$$

$$\hat{z} \cdot \Delta \boldsymbol{D} = -\nabla \cdot \boldsymbol{P}_\parallel, \tag{A.14c}$$

$$\hat{z} \cdot \Delta \boldsymbol{B} = -\mu \nabla \cdot \boldsymbol{M}_\parallel. \tag{A.14d}$$

References

1. Kuester, E. F., Mohamed, M., Piket-May, M. and Holloway, C. (2003). Averaged transition conditions for electromagnetic fields at a metafilm, *IEEE Trans. Antennas Propag.* **51**(10), pp. 2641–2651.
2. Holloway, C., Dienstfrey, A., Kuester, E. F., O'Hara, J. F., Azad, A. K. and Taylor, A. J. (2009). A discussion on the interpretation and characterization of metafilms/metasurfaces: The two-dimensional equivalent of metamaterials, *Metamaterials* **3**(2), pp. 100–112.
3. Holloway, C., Kuester, E. F., Gordon, J., O'Hara, J., Booth, J. and Smith, D. (2012). An overview of the theory and applications of metasurfaces: The two-dimensional equivalents of metamaterials, *IEEE Antennas Propag. Mag.* **54**(2), pp. 10–35.
4. Achouri, K., Salem, M. A. and Caloz, C. (2015). General metasurface synthesis based on susceptibility tensors, *IEEE Trans. Antennas Propag.*
5. Munk, B. A. (2000). *Frequency Selective Surfaces: Theory and Design* (John Wiley & Sons, Hoboken).
6. Sabah, C. and Urbani, F. (2013). Experimental analysis of shaped magnetic resonator for mu-negative metamaterials, *Opt. Commun.* **294**, pp. 409–413.
7. Yang, Y., Huang, R., Cong, L., Zhu, Z., Gu, J., Tian, Z., Singh, R., Zhang, S., Han, J. and Zhang, W. (2011). Modulating the fundamental inductive-capacitive resonance in asymmetric double-split ring terahertz metamaterials, *Appl. Phys. Lett.* **98**(12).
8. Vehmas, J., Ra'di, Y., Karilainen, A. and Tretyakov, S. (2013). Eliminating electromagnetic scattering from small particles, *IEEE Trans. Antennas Propag.* **61**(7), pp. 3747–3756.
9. Monticone, F., Estakhri, N. M. and Alù, A. (2013). Full control of nanoscale optical transmission with a composite metascreen, *Phys. Rev. Lett.* **110**, p. 203903.

10. Durn-Sindreu, M., Naqui, J., Bonache, J. and Martn, F. (2012). Split rings for metamaterial and microwave circuit design: A review of recent developments (invited paper), *Int. J. RF Microw. C. E.* **22**(4), pp. 439–458.
11. Arbabi, A. and Faraon, A. (2014). Fundamental limits of ultrathin metasurfaces, *arXiv preprint arXiv:1411.2537*.
12. Albooyeh, M., Kruk, S., Menzel, C., Helgert, C., Kroll, M., Krysinski, A., Decker, M., Neshev, D., Pertsch, T., Etrich, C. *et al.* (2014). Resonant metasurfaces at oblique incidence: Interplay of order and disorder, *Sci. Rep.* **4**.
13. Asadchy, V. S., Faniayeu, I. A., Ra'di, Y. and Tretyakov, S. A. (2014). Determining polarizability tensors for an arbitrary small electromagnetic scatterer, *arXiv preprint arXiv:1401.4930*.
14. Abdelrahman, A., Elsherbeni, A. and Yang, F. (2014). Transmitarray antenna design using cross-slot elements with no dielectric substrate, *IEEE Antennas Wireless Propag. Lett.* **13**, pp. 177–180.
15. Pfeiffer, C., Emani, N. K., Shaltout, A. M., Boltasseva, A., Shalaev, V. M. and Grbic, A. (2014). Efficient light bending with isotropic metamaterial huygens surfaces, *Nano Lett.* **14**(5), pp. 2491–2497.
16. Pfeiffer, C. and Grbic, A. (2013). Metamaterial huygens' surfaces: Tailoring wave fronts with reflectionless sheets, *Phys. Rev. Lett.* **110**, p. 197401.
17. Pfeiffer, C. and Grbic, A. (2013). Millimeter-wave transmitarrays for wavefront and polarization control, *IEEE Trans. Microwave Theory Tech.* **61**(12), pp. 4407–4417.
18. Pfeiffer, C. and Grbic, A. (2014). Bianisotropic metasurfaces for optimal polarization control: Analysis and synthesis, *Phys. Rev. Appl.* **2**, p. 044011.
19. Wong, J. P., Selvanayagam, M. and Eleftheriades, G. V. (2014). Design of unit cells and demonstration of methods for synthesizing huygens metasurfaces, *Photonic. Nanostruct.* **12**(4), pp. 360–375.
20. Selvanayagam, M. and Eleftheriades, G. (2014). Polarization control using tensor huygens surfaces, *IEEE Trans. Antennas Propag.* **62**(12), pp. 6155–6168.
21. Epstein, A. and Eleftheriades, G. (2014). Passive lossless huygens metasurfaces for conversion of arbitrary source field to directive radiation, *IEEE Trans. Antennas Propag.* **62**(11), pp. 5680–5695.
22. Niemi, T., Karilainen, A. and Tretyakov, S. (2013). Synthesis of polarization transformers, *IEEE Trans. Antennas Propag.* **61**(6), pp. 3102–3111.
23. Ra'di, Y., Asadchy, V. and Tretyakov, S. (2013). Total absorption of electromagnetic waves in ultimately thin layers, *IEEE Trans. Antennas Propag.* **61**(9), pp. 4606–4614.
24. Salem, M. A., Achouri, K. and Caloz, C. (2014). Metasurface synthesis for time-harmonic waves: Exact spectral and spatial methods, *Prog. Electromagn. Res.* (*invited*) **149**, pp. 205–216.

25. Salem, M. A. and Caloz, C. (2014). Manipulating light at distance by a metasurface using momentum transformation, *Opt. Express* **22**(12), pp. 14 530–14 543.
26. Yu, N., Genevet, P., Kats, M. A., Aieta, F., Tetienne, J.-P., Capasso, F. and Gaburro, Z. (2011). Light propagation with phase discontinuities: Generalized laws of reflection and refraction, *Science* **334**(6054), pp. 333–337.
27. Schelkunoff, S. (1972). On teaching the undergraduate electromagnetic theory, *IEEE Trans. Educ.* **15**(1), pp. 15–25.
28. Idemen, M. M. (2011). *Discontinuities in the Electromagnetic Field* (John Wiley & Sons, Hoboken).
29. Kong, J. (1986). *Electromagnetic Wave Theory*, ser. A Wiley-Interscience publication (John Wiley & Sons, Hoboken).
30. Lindell, I. (1994). *Electromagnetic Waves in Chiral and Bi-isotropic Media*, ser. The Artech House Antenna Library. (Artech House, London).
31. Asadchy, V. S. and Fanyaev, I. A. (2011). Simulation of the electromagnetic properties of helices with optimal shape, which provides radiation of a circularly polarized wave, *J. Adv. Res. Phys.* **2**(1), p. 011107.
32. Holloway, C., Mohamed, M., Kuester, E. F. and Dienstfrey, A. (2005). Reflection and transmission properties of a metafilm: With an application to a controllable surface composed of resonant particles, *IEEE Trans. Electromagn. Compat.* **47**(4), pp. 853–865.
33. Hong, J.-S. G. and Lancaster, M. J. (2004). *Microstrip Filters for RF/Microwave Applications* (John Wiley & Sons, Hoboken).
34. Mie, G. (1908). Beiträge zur optik trüber medien, speziell kolloidaler metallösungen, *Annalen der Physik* **330**(3), pp. 377–445.
35. Kerker, M. (2013). *The Scattering of Light and Other Electromagnetic Radiation: Physical Chemistry: A Series of Monographs.* (Academic Press, Cambridge).
36. Krasnok, A., Makarov, S., Petrov, M., Savelev, R., Belov, P. and Kivshar, Y. (2015). Towards all-dielectric metamaterials and nanophotonics, *Proc. SPIE* **9502**, pp. 950 203–950 203–17.
37. Arbabi, A., Horie, Y., Bagheri, M. and Faraon, A. (2014). Complete control of polarization and phase of light with high efficiency and sub-wavelength spatial resolution, *arXiv preprint arXiv:1411.1494*.
38. Asadchy, V. S., Faniayeu, I. A., Ra'di, Y., Khakhomov, S. A., Semchenko, I. V. and Tretyakov, S. A. (2015). Broadband reflectionless metasheets: Frequency-selective transmission and perfect absorption, *Phys. Rev. X* **5**, p. 031005.
39. Vahabzadeh, Y., Achouri, K. and Caloz, C. (2016). Simulation of metasurfaces in finite difference techniques, *arXiv preprint arXiv:1602.04086*.
40. Zhu, B. O. and Feng, Y. (2015). Passive metasurface for reflectionless and arbitary control of electromagnetic wave transmission, *IEEE Trans. Antennas Propag.* **63**(12), pp. 5500–5511.
41. Asadchy, V., Albooyeh, M., Tcvetkova, S., Ra'di, Y., and Tretyakov, S. (2016). Metasurfaces for perfect and full control of refraction and reflection, *arXiv preprint arXiv:1603.07186*.

42. Saleh, B. and Teich, M. (2007). *Fundamentals of Photonics*, ser. Wiley Series in Pure and Applied Optics (Wiley, Hoboken).
43. Goodman, J. W. *et al.* (1968). *Introduction to Fourier Optics.* (McGraw-Hill, New York).
44. Silva, A., Monticone, F., Castaldi, G., Galdi, V., Al, A. and Engheta, N. (2014). Performing mathematical operations with metamaterials, *Science* **343**(6167), pp. 160–163.
45. Aieta, F., Kats, M. A., Genevet, P. and Capasso, F. (2015). Multiwavelength achromatic metasurfaces by dispersive phase compensation, *Science* **347**(6228), pp. 1342–1345.
46. Achouri, K., Salem, M. A. and Caloz, C. (2015). Improvement of metasurface continuity conditions, *ISAP*, pp. 123–125.

CHAPTER 6

Metasurfaces for General Control of Reflection and Transmission

SERGEI TRETYAKOV*, VIKTAR ASADCHY and ANA DÍAZ-RUBIO

Aalto University, Finland

This chapter discusses most general linear bianisotropic metasurfaces and reviews various homogenization models and methods to synthesize metasurfaces with desired reflection and transmission properties. In the last part, we give several examples of metasurface synthesis, illustrating a general approach to realizing the desired and optimized response. The focus is on understanding and selecting physical properties of unit cells which are most suitable for realizing the desired response of metasurfaces, and on finding appropriate topologies and dimensions of the unit cells.

6.1. Introduction

Metasurface[1–4] is an electrically thin composite material layer designed and optimized to function as a tool to control and transform electromagnetic waves. The layer thickness is small and can

*Corresponding author: sergei.tretyakov@aalto.fi

be considered as negligible with respect to the wavelength in the surrounding space. The composite structure forming a metasurface is assumed to behave as a *material* in the electromagnetic (optical) sense, meaning that it can be homogenized on the wavelength scale, and the metasurface can be adequately characterized by its effective, surface-averaged properties. Similar to volumetric materials, where the notions of the permittivity and permeability result from volumetric averaging of microscopic currents over volumes which are small compared to the wavelength, the metasurface parameters result from 2D, surface averaging of microscopic currents on the same wavelength scale. This implies that the unit-cell sizes of composite metasurfaces are reasonably small as compared with the wavelength. In terms of the optical response, this means that metasurfaces reflect and transmit plane waves as sheets of homogeneous materials, in contrast to diffraction gratings which produce multiple diffraction lobes.

Due to the presence of small but strongly polarizable inclusions, the fields inside the layer change significantly, so that the effects of the layer on waves incident from surrounding space can be very strong, despite the very small thickness. One of the main challenges is to synthesize and realize metasurfaces which change the incident fields in desired ways. This goal means that we would like to create metasurface structures with fully adjustable and controllable reflection and transmission properties.

The chapter starts from a discussion of Huygens' principle in electromagnetics. Applying this principle, it is possible to replace arbitrary volumetric sources of fields by equivalent surface currents flowing on a theoretically infinitely thin surface surrounding the volume where the volumetric sources are located. We explain how this principle and its generalizations can be used as a general design paradigm for creating metasurfaces with the desired response.

Next, we explain different alternative homogenization models suitable for description and modelling of the electromagnetic properties of metasurfaces. While the properties of volumetric (3D) materials and metamaterials are described by such effective parameters as

permittivity, permeability, chirality parameter, etc., in order to model the properties of effectively homogeneous sheets (2D "materials"), other parameters are needed. The conventional notions of permittivity and permeability lose their meaning simply because there is no volume over which the fields and induced polarizations could be averaged. Surface-averaging approaches lead to introduction of such effective parameters as sheet impedance, collective polarizability and surface susceptibility, which we will explain in this chapter.

In the following section, we discuss methods to find what properties of unit cells are required, if the goal is to create a metasurface with some desired reflection and transmission properties. Knowing the necessary physical properties, it becomes possible to find what unit-cell topologies are suitable for realization of the required response. In the last part, we give several examples of functional single-layer metasurfaces synthesized and created using the theoretical approaches described in this chapter.

6.2. Generalized Huygens' Principle

Electromagnetic fields can be controlled and transformed using engineered materials, often called *metamaterials.* The conventional paradigm of using metamaterials for transformations of electromagnetic fields implies that we engineer artificial materials in such a way that the polarization and conduction currents induced in the material, acting as secondary sources, create the desired fields outside or inside of the metamaterial sample. Within the metasurface paradigm, electromagnetic fields are controlled by polarization and conduction currents in a very thin layer, effectively a sheet of negligible thickness. It appears that an effective approach to understanding and synthesizing metasurfaces can be built upon Huygens' principle and its generalizations. Indeed, this principle tells us that the fields outside of the sample volume can be found as those generated by equivalent surface currents flowing only on the volume surface, as illustrated in Fig. 6.1. Thus, it appears that desired field transformations can be achieved by engineering only surface currents

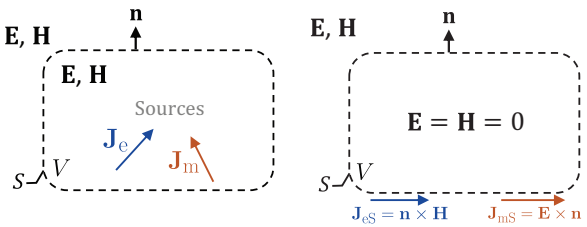


Fig. 6.1. Equivalent surface currents $\mathbf{J}_{eS}$ and $\mathbf{J}_{mS}$ create the same fields outside volume V as the volumetric current sources enclosed by S.

of the volume surface, and there appears to be no reason why the volume enclosed by such an engineered surface could not be made negligibly small.

The Huygens' principle, also called the *equivalence* principle, has a long history. It was introduced by Huygens in 1690 and put into a mathematical form by Fresnel (1818), Helmholtz (1860) and Kirchhoff (1882) for scalar fields. The vector form is usually attributed to Stratton and Chu (1939). Conventional derivations[5] are based on the use of Green's function: Electromagnetic fields created by volumetric sources in volume V are written as an integral over V, which is then transformed into a surface integral over surface S bounding volume V. For our goals, we will use an alternative derivation,[6] which shows that Huygens' principle can be used for understanding and synthesizing general metasurfaces. In this concept, equivalent surface currents are introduced as a means to *modulate* sources and fields in a certain volume according to a prescribed rule.

Let us consider a closed surface S which bounds region V (Fig. 6.1). Following Lindell,[6] we define a *modulation function*

$$P(\mathbf{r}) = 1 \quad \text{if } \mathbf{r} \in V, \qquad P(\mathbf{r}) = 0 \text{ otherwise} \tag{6.1}$$

With this function, we "modulate" the Maxwell equations

$$\nabla \times \mathbf{E} + \frac{\partial}{\partial t}\mathbf{B} = -\mathbf{J}_m, \qquad \nabla \times \mathbf{H} - \frac{\partial}{\partial t}\mathbf{D} = \mathbf{J}_e \tag{6.2}$$

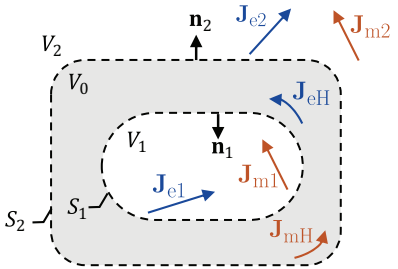


Fig. 6.2. In the generalized Huygens' principle the equivalent currents $\mathbf{J}_{\mathrm{eH}}$ and $\mathbf{J}_{\mathrm{mH}}$ fill volume V_0, which separates V_1 and V_2.

multiplying them by $P(\mathbf{r})$. Obviously, for any differentiable vector function $\mathbf{F}(\mathbf{r})$

$$P(\mathbf{r})\nabla \times \mathbf{F}(\mathbf{r}) = \nabla \times [P(\mathbf{r})\mathbf{F}(\mathbf{r})] - [\nabla P(\mathbf{r})] \times \mathbf{F}(\mathbf{r}) \tag{6.3}$$

Denoting the modulated (in this example, truncated) functions as $\mathbf{F}_{\mathrm{V}} = P(\mathbf{r})\mathbf{F}(\mathbf{r})$, the result reads

$$\nabla \times \mathbf{E}_{\mathrm{V}} + \frac{\partial}{\partial t}\mathbf{B}_{\mathrm{V}} = -\mathbf{J}_{\mathrm{mV}} + [\nabla P(\mathbf{r})] \times \mathbf{E} \tag{6.4}$$

$$\nabla \times \mathbf{H}_{\mathrm{V}} - \frac{\partial}{\partial t}\mathbf{D}_{\mathrm{V}} = \mathbf{J}_{\mathrm{eV}} + [\nabla P(\mathbf{r})] \times \mathbf{H} \tag{6.5}$$

For this choice of $P(\mathbf{r})$, we have $\nabla P(\mathbf{r}) = \mathbf{z}_0(\mathbf{r})\delta_S(\mathbf{r})$, where δ_S is the Dirac delta function: that is, the equivalent electric and magnetic currents are bound to surface S. This result expresses the classical Huygens principle. It is important to note that the derivation is extremely general and is independent from the properties of the medium which fills the space on both sides of surface S.

This general concept of modulating fields and sources using equivalent surface currents can be generalized in many ways, offering important insight into operational principles of metasurfaces as well as giving us effective design tools. For example, let us assume[7] that volume V_1 is separated from the rest of the space (volume V_2), not by a surface, but by a volumetric region V_0, as shown in Fig. 6.2.

One can define two complementary "modulation functions"

$$P_1(\mathbf{r}) = 1 \quad \text{if } \mathbf{r} \in V_1, \qquad P_1(\mathbf{r}) = 0 \quad \text{if } \mathbf{r} \in V_2 \tag{6.6}$$

$$P_2(\mathbf{r}) = 0 \quad \text{if } \mathbf{r} \in V_1, \qquad P_2(\mathbf{r}) = 1 \quad \text{if } \mathbf{r} \in V_2 \tag{6.7}$$

satisfying $P_1(\mathbf{r}) + P_2(\mathbf{r}) = 1$. Equivalent currents

$$\mathbf{J}_{\text{mH}} = -[\nabla P_1(\mathbf{r})] \times \mathbf{E}_1 - [\nabla P_2(\mathbf{r})] \times \mathbf{E}_2 \tag{6.8}$$

$$\mathbf{J}_{\text{eH}} = [\nabla P_1(\mathbf{r})] \times \mathbf{H}_1 + [\nabla P_2(\mathbf{r})] \times \mathbf{H}_2 \tag{6.9}$$

which flow in V_0 "glue" the two sets of fields into one system. They define generalized continuity conditions and the generalized Huygens principle,[7] where the layer carrying equivalent currents can be of any thickness and shape. This derivation of the generalized Huygens principle can be extended to find a way to "modulate" fields in a certain volume of space by filling this volume with a metamaterial with specific material parameters,[8] because the modulating function (6.1) can be any differentiable function.

This discussion brings us to the formulation of the general approach to metasurface synthesis. Suppose that there is a wave coming from volume V_2 into V_1 and we want to change the fields in V_1 by inserting a sheet or a material layer between the two volumes. We start from adding additional volumetric sources in V_2 which make the desired effects in V_1.[a] Next, we replace these additional sources with equivalent Huygens' currents, in a volume or on a surface. These currents are proportional to the fields, so we can use metamaterials or metasurfaces to create them. The design challenge is to find particular topologies of metasurface layers such that the induced surface currents are related to the fields in the desired way. To this end, we will need to understand homogenization models of metasufaces, which is the subject of the next section.

[a] At this stage, it is important to appreciate that our freedom to manipulate fields in volume V_2 using a metasurface between volumes V_2 and V_1 is limited by the fact that we can freely add any source only in region V_1. For example, we cannot create an energy sink or source in V_2 this way. Thus, one needs to be careful when claiming a possibility for "general" or "full" control of fields using metasurfaces.

6.3. Homogenization Models of Metasurfaces

Constructing a homogenization model means determining what *effective parameters* can properly describe the metasurface response to incident fields and finding these parameters from the known properties of the unit cells or from experimental data or simulations of the metasurface response (usually from plane-wave reflection and transmission coefficients). This task is analogous to introduction of the notions of permittivity and permeability of volumetric materials, establishing mixing rules linking the single-molecule polarizabilities to the effective materials parameters and finding methods for material parameter measurements. Obviously, material homogenization is a crucial pre-requisite for understanding and practical usage of electromagnetic materials and metasurfaces. Using effective parameters, we dramatically reduce complexity of the problem, modelling the collective response of extremely many individual molecules (in natural media) or small engineered inclusions (in metamaterials and metasurfaces) by only a few effective parameters, like the permittivity or conductivity. Without homogenization models, understanding, design and optimization of metasurfaces becomes an unrealistically huge problem of brute-force numerical optimization.

It is clear that conventional effective parameters used to model volumetric materials (permittivity, permeability, bianisotropy coefficients) are not applicable to metasurfaces, because the very definitions of these parameters imply averaging over small *volumes* which contain many molecules or inclusions. In the case of a metasurface, this averaging volume is negligibly small, because metasurfaces have negligible thickness. Thus, metasurface homogenization should imply averaging over a small *surface* area, which is small compared to the wavelength but contains many (at least, several) inclusions or unit cells. For plane-wave illumination, averaging over the area of one unit cell is sufficient. The effective models of metasurfaces are written for macroscopic quantities defined as surface-averaged values of electric and magnetic fields and surface-averaged polarizations and currents. The concept of metasurface homogenization is illustrated in Fig. 6.3.

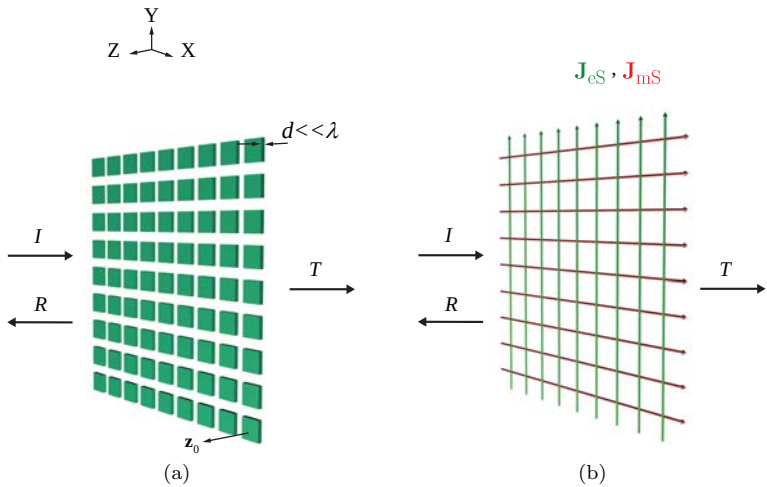


Fig. 6.3. In the homogenization model, an array of electrically small unit cells is replaced by an equivalently responding set of electric and magnetic currrent sheets. While at small distances (comparable to or smaller than the array period), the fields of the two structures are very different, and at distances much larger than the array period, the fields created by the metasurface and by the model pair of two homogenized current sheets are not distinguishable.

In the literature, relations between surface-averaged quantities are usually written in two different but equivalent forms. In the first one,[9,10] the induced surface-averaged electric and magnetic polarization surface densities $\mathbf{P}$ and $\mathbf{M}$ are connected to the so-called "averaged fields" on the metasurface plane $\mathbf{E}^{\mathrm{av}}$ and $\mathbf{H}^{\mathrm{av}}$:

$$\mathbf{P} = \epsilon_0 \overline{\overline{\chi}}_{\mathrm{ee}} \cdot \mathbf{E}^{\mathrm{av}} + \sqrt{\epsilon_0 \mu_0}\, \overline{\overline{\chi}}_{\mathrm{em}} \cdot \mathbf{H}^{\mathrm{av}} \tag{6.10}$$

$$\mathbf{M} = \overline{\overline{\chi}}_{\mathrm{mm}} \cdot \mathbf{H}^{\mathrm{av}} + \sqrt{\frac{\epsilon_0}{\mu_0}}\, \overline{\overline{\chi}}_{\mathrm{me}} \cdot \mathbf{E}^{\mathrm{av}} \tag{6.11}$$

Here, the fields $\mathbf{E}^{\mathrm{av}}$ and $\mathbf{H}^{\mathrm{av}}$ are defined as the arithmetic averages of the (surface-averaged) fields measured at the two opposite sides of the metasurface:

$$\mathbf{E}^{\mathrm{av}} = \frac{1}{2}(\mathbf{E}_+ + \mathbf{E}_-), \qquad \mathbf{H}^{\mathrm{av}} = \frac{1}{2}(\mathbf{H}_+ + \mathbf{H}_-) \tag{6.12}$$

For example, if the metasurface is excited by a plane wave coming from the side marked by "+", $\mathbf{E}_+$ equals the sum of the electric fields

of the incident and reflected plane waves, while $\mathbf{E}_-$ is the electric field of the transmitted plane wave. All these fields are defined at the metasurface plane. It is important not to confuse this averaging of the fields on the two sides of the metasurface with averaging the fields over the unit-cell *area*. All the fields which enter (6.12) are already surface-averaged (macroscopic) fields.

Because the normal (to the metasurface) components of the fields can be expressed in terms of the tangential components using Maxwell's equations in the space surrounding the metasurface, Eqs. (6.10) and (6.11) can be also written with the tangential components of the fields $\mathbf{E}_t^{\text{av}}$ and $\mathbf{H}_t^{\text{av}}$ instead of the total ones. If the *effective surface susceptibilities* $\overline{\overline{\chi}}_{\text{ij}}$ are known, one can solve reflection and transmission problems combining (6.10) and (6.11) with the expressions for plane-wave fields created by the surface polarizations $\mathbf{P}$ and $\mathbf{M}$. These expressions can also be written in the form of *averaged transition conditions*,[9] which relate the jumps of the tangential electric and magnetic fields across the metasurface and the polarization vectors:

$$\mathbf{E}_{t+} - \mathbf{E}_{t-} = j\omega\mu_0 \mathbf{z}_0 \times \mathbf{M}_t - \nabla_t \frac{P_n}{\epsilon} \tag{6.13}$$

$$\mathbf{z}_0 \times (\mathbf{H}_{t+} - \mathbf{H}_{t-}) = j\omega \mathbf{P}_t + \nabla_t \times \mathbf{z}_0 M_n \tag{6.14}$$

Here, $\mathbf{z}_0$ is the unit vector normal to the metasurface and pointing to the direction of the side marked $+$. P_n and M_n are the normal components of the corresponding vectors, and ∇_t is the 2D differentiation operator ($\nabla_t = \nabla - \mathbf{z}_0 \frac{\partial}{\partial z}$, where the axis z is along $\mathbf{z}_0$). Here we assume that the particles forming the metasurface are positioned inside a thin layer of a dielectric material with the permittivity $\epsilon = \epsilon_0 \epsilon_{\text{r}}$.[51]

In the alternative approach, surface-averaged polarization vectors can be related to the incident (external) fields.[11–14,26] Most commonly, these relations are written for dipole moments[b] induced in

[b]In this book, the magnetic dipole moment is defined as in most textbooks, having the dimension $[\text{A} \cdot \text{m}^2]$. Note that in many journal publications, e.g. Refs. 32, 33, 34, 40, 41, 49, the definition is different, including an additional multiplier μ_0.

each unit cell, $\mathbf{p}$ and $\mathbf{m}$, as

$$\mathbf{p} = \overline{\overline{\widehat{\alpha}}}_{\mathrm{ee}} \cdot \mathbf{E}_{\mathrm{inc}} + \overline{\overline{\widehat{\alpha}}}_{\mathrm{em}} \cdot \mathbf{H}_{\mathrm{inc}} \tag{6.15}$$

$$\mathbf{m} = \overline{\overline{\widehat{\alpha}}}_{\mathrm{mm}} \cdot \mathbf{H}_{\mathrm{inc}} + \overline{\overline{\widehat{\alpha}}}_{\mathrm{me}} \cdot \mathbf{E}_{\mathrm{inc}} \tag{6.16}$$

The polarization density vectors equal $\mathbf{P} = \mathbf{p}/S$, $\mathbf{M} = \mathbf{m}/S$, where S is the unit-cell area. The coefficients in (6.15) and (6.16) are called *collective polarizability dyadics*, because they model the unit-cell response when the unit cells are arranged into a periodical lattice, measuring the collective response to the incident fields. Similar to the previous model, the normal (to the metasurface) components of the incident fields can be expressed in terms of the tangential components using Maxwell's equations in the space surrounding the metasurface, and Eqs. (6.15) and (6.16) can be written with the tangential components of the incident fields $\mathbf{E}_{t\,\mathrm{inc}}$ and $\mathbf{H}_{t\,\mathrm{inc}}$ instead of the complete vectors.

For metasurfaces without bianisotropic effects, the homogenization approach based on individual and collective polarizabilities naturally leads to the model in terms of surface (sheet) impedance and admittance, where the surface-averaged electric and magnetic fields are related to the *total* fields at the metasurface.[26,27,50]

The homogenization models expressed via (6.10), (6.11) and (6.15), (6.16) are algebraically equivalent because the parameters of one set can be expressed in terms of the parameters of the other set. To do that, one can start from Eqs. (6.15) and (6.16) and find the amplitudes of plane waves generated by the arrays of dipole moments $\mathbf{p}$ and $\mathbf{m}$, created by some test incident plane wave $\mathbf{E}_{\mathrm{inc}}$, $\mathbf{H}_{\mathrm{inc}}$. Let us write these fields for the normal-incidence illumination. The induced dipole moments (6.15) and (6.16) correspond to the surface-averaged electric and magnetic current sheets with the surface current densities $\mathbf{J}_{\mathrm{e}} = \frac{j\omega\mathbf{p}}{S}$ and $\mathbf{J}_{\mathrm{m}} = \frac{j\omega\mu_0\mathbf{m}}{S}$ that radiate plane waves into the surrounding medium (most commonly, free space). The reflected field is the sum of the fields radiated by these surface currents:

$$\mathbf{E}_{\mathrm{ref}} = -\frac{j\omega}{2S}[\eta_0\mathbf{p} - \mathbf{z}_0 \times \mu_0\mathbf{m}] \tag{6.17}$$

Here, η_0 is the wave impedance of the surrounding isotropic material (most commonly, free space), and $\mathbf{z}_0$ is the unit vector normal to the metasurface plane and pointing towards the source, as in Fig. 6.3. Similarly, the transmitted field is the sum of the incident field and the fields caused by the array:

$$\mathbf{E}_{\mathrm{tr}} = \mathbf{E}_{\mathrm{inc}} - \frac{j\omega}{2S}[\eta_0 \mathbf{p} + \mathbf{z}_0 \times \mu_0 \mathbf{m}] \tag{6.18}$$

Knowing the scattered fields and the incident fields, we in fact know the total fields at both sides of the metasurface ($\mathbf{E}_+ = \mathbf{E}_{\mathrm{inc}} + \mathbf{E}_{\mathrm{ref}}$, $\mathbf{E}_- = \mathbf{E}_{\mathrm{tr}}$, which maintains polarizations $\mathbf{P} = \mathbf{p}/S$, $\mathbf{M} = \mathbf{m}/S$. This way, we can determine the parameters $\overline{\overline{\chi}}_{\mathrm{ij}}$ of the model equations (6.10) and (6.11) in terms of the collective polarizabilities of the model equations (6.15) and (6.16). The result reads

$$\overline{\overline{\chi}}_{\mathrm{ee}} = \frac{1}{\epsilon_0}(\Delta_p^*)^{-1} \cdot \left[\overline{\overline{\widehat{\alpha}}}_{\mathrm{ee}} + \frac{j\omega}{2\eta_0}\overline{\overline{\widehat{\alpha}}}_{\mathrm{em}} \cdot \left(\frac{S}{\mu_0}\overline{\overline{I}}_t - \overline{\overline{\widehat{\alpha}}}_{\mathrm{mm}}\frac{j\omega}{2\eta_0}\right)^{-1} \cdot \overline{\overline{\widehat{\alpha}}}_{\mathrm{me}}\right] \tag{6.19}$$

$$\overline{\overline{\chi}}_{\mathrm{em}} = \frac{1}{\sqrt{\epsilon_0\mu_0}}(\Delta_p^*)^{-1} \cdot \left[\overline{\overline{\widehat{\alpha}}}_{\mathrm{em}} + \frac{j\omega}{2\eta_0}\overline{\overline{\widehat{\alpha}}}_{\mathrm{em}} \cdot \left(\frac{S}{\mu_0}\overline{\overline{I}}_t - \overline{\overline{\widehat{\alpha}}}_{\mathrm{mm}}\frac{j\omega}{2\eta_0}\right)^{-1} \cdot \overline{\overline{\widehat{\alpha}}}_{\mathrm{mm}}\right] \tag{6.20}$$

$$\overline{\overline{\chi}}_{\mathrm{me}} = \sqrt{\frac{\mu_0}{\epsilon_0}}(\Delta_m^*)^{-1} \cdot \left[\overline{\overline{\widehat{\alpha}}}_{\mathrm{me}} + \frac{j\omega\eta_0}{2}\overline{\overline{\widehat{\alpha}}}_{\mathrm{me}} \cdot \left(S\overline{\overline{I}}_t - \overline{\overline{\widehat{\alpha}}}_{\mathrm{ee}}\frac{j\omega\eta_0}{2}\right)^{-1} \cdot \overline{\overline{\widehat{\alpha}}}_{\mathrm{ee}}\right] \tag{6.21}$$

$$\overline{\overline{\chi}}_{\mathrm{mm}} = (\Delta_m^*)^{-1} \cdot \left[\overline{\overline{\widehat{\alpha}}}_{\mathrm{mm}} + \frac{j\omega\eta_0}{2}\overline{\overline{\widehat{\alpha}}}_{\mathrm{me}} \cdot \left(S\overline{\overline{I}}_t - \overline{\overline{\widehat{\alpha}}}_{\mathrm{ee}}\frac{j\omega\eta_0}{2}\right)^{-1} \cdot \overline{\overline{\widehat{\alpha}}}_{\mathrm{em}}\right] \tag{6.22}$$

where

$$\Delta_p^* = S\overline{\overline{I}}_t - \overline{\overline{\widehat{\alpha}}}_{\rm ee}\frac{j\omega\eta_0}{2} + \frac{\omega^2}{4}\overline{\overline{\widehat{\alpha}}}_{\rm em} \cdot \left(\frac{S}{\mu_0}\overline{\overline{I}}_t - \overline{\overline{\widehat{\alpha}}}_{\rm mm}\frac{j\omega}{2\eta_0}\right)^{-1} \cdot \overline{\overline{\widehat{\alpha}}}_{\rm me} \tag{6.23}$$

$$\Delta_m^* = S\overline{\overline{I}}_t - \overline{\overline{\widehat{\alpha}}}_{\rm mm}\frac{j\omega\mu_0}{2\eta_0} + \frac{\mu_0\omega^2}{4}\overline{\overline{\widehat{\alpha}}}_{\rm me} \cdot \left(S\overline{\overline{I}}_t - \overline{\overline{\widehat{\alpha}}}_{\rm ee}\frac{j\omega\eta_0}{2}\right)^{-1} \cdot \overline{\overline{\widehat{\alpha}}}_{\rm em} \tag{6.24}$$

Here, $\overline{\overline{I}}_t$ is the two-dimensional unit dyadic. We note that in the general case, each of the susceptibility dyadics depend on *all* collective polarizabilities. For example, the *electric* susceptibility $\overline{\overline{\chi}}_{\rm ee}$ depends not only on the *electric* collective polarizability $\overline{\overline{\widehat{\alpha}}}_{\rm ee}$ but also on the magnetic polarizability and the magnetoelectric coupling dyadics. Only in some special simple cases is there one-to-one correspondence between the two sets of effective parameters. For example, assuming that $\overline{\overline{\widehat{\alpha}}}_{\rm ee} = \alpha_{\rm ee}\overline{\overline{I}}_t$, $\overline{\overline{\widehat{\alpha}}}_{\rm mm} = \alpha_{\rm mm}\overline{\overline{I}}_t$, and $\overline{\overline{\widehat{\alpha}}}_{\rm em} = \overline{\overline{\widehat{\alpha}}}_{\rm me} = 0$, we get simple relations

$$\chi_{\rm ee} = \frac{\hat{\alpha}_{\rm ee}}{\epsilon_0\left(S - \hat{\alpha}_{\rm ee}\frac{j\omega\eta_0}{2}\right)}, \qquad \chi_{\rm mm} = \frac{\hat{\alpha}_{\rm mm}}{\left(S - \hat{\alpha}_{\rm mm}\frac{j\omega\mu_0}{2\eta_0}\right)} \tag{6.25}$$

In a similar way, we can find the collective polarizabilities if the susceptibilities are known:

$$\overline{\overline{\widehat{\alpha}}}_{\rm ee} = (\Delta_p)^{-1} \cdot \left[\epsilon_0\overline{\overline{\chi}}_{\rm ee} - \epsilon_0\frac{j\omega\mu_0}{2\eta_0}\overline{\overline{\chi}}_{\rm em} \cdot \left(\overline{\overline{I}}_t + \overline{\overline{\chi}}_{\rm mm}\frac{j\omega\mu_0}{2\eta_0}\right)^{-1} \cdot \overline{\overline{\chi}}_{\rm me}\right] \tag{6.26}$$

$$\overline{\overline{\widehat{\alpha}}}_{\rm em} = (\Delta_p)^{-1} \cdot \left[\sqrt{\epsilon_0\mu_0}\,\overline{\overline{\chi}}_{\rm em} - \epsilon_0\mu_0\,\frac{j\omega}{2\eta_0}\overline{\overline{\chi}}_{\rm em} \cdot \left(\overline{\overline{I}}_t + \overline{\overline{\chi}}_{\rm mm}\frac{j\omega\mu_0}{2\eta_0}\right)^{-1} \cdot \overline{\overline{\chi}}_{\rm mm}\right] \tag{6.27}$$

$$\overline{\overline{\widehat{\alpha}}}_{\mathrm{me}} = (\Delta_m \mu_0)^{-1} \cdot \left[\sqrt{\frac{\epsilon_0}{\mu_0}}\,\overline{\overline{\chi}}_{\mathrm{me}} - \epsilon_0 \frac{j\omega}{2}\overline{\overline{\chi}}_{\mathrm{me}} \cdot \left(\overline{\overline{I}}_t + \epsilon_0 \overline{\overline{\chi}}_{\mathrm{ee}} \frac{j\omega\eta_0}{2} \right)^{-1} \cdot \overline{\overline{\chi}}_{\mathrm{ee}} \right] \tag{6.28}$$

$$\overline{\overline{\widehat{\alpha}}}_{\mathrm{mm}} = (\Delta_m)^{-1} \cdot \left[\overline{\overline{\chi}}_{\mathrm{mm}} - \epsilon_0 \frac{j\omega\eta_0}{2}\overline{\overline{\chi}}_{\mathrm{me}} \cdot \left(\overline{\overline{I}}_t + \epsilon_0 \overline{\overline{\chi}}_{\mathrm{ee}} \frac{j\omega\eta_0}{2} \right)^{-1} \cdot \overline{\overline{\chi}}_{\mathrm{em}} \right] \tag{6.29}$$

where

$$\Delta_p = \frac{1}{S}\overline{\overline{I}}_t + \epsilon_0 \overline{\overline{\chi}}_{\mathrm{ee}} \frac{j\omega\eta_0}{2S} + \frac{\epsilon_0 \mu_0 \omega^2}{4S}\overline{\overline{\chi}}_{\mathrm{em}} \cdot \left(\overline{\overline{I}}_t + \overline{\overline{\chi}}_{\mathrm{mm}} \frac{j\omega\mu_0}{2\eta_0} \right)^{-1} \cdot \overline{\overline{\chi}}_{\mathrm{me}} \tag{6.30}$$

$$\Delta_m = \frac{1}{S}\overline{\overline{I}}_t + \overline{\overline{\chi}}_{\mathrm{mm}} \frac{j\omega\mu_0}{2S\eta_0} + \epsilon_0 \frac{\mu_0 \omega^2}{4S}\overline{\overline{\chi}}_{\mathrm{me}} \cdot \left(\overline{\overline{I}}_t + \epsilon_0 \overline{\overline{\chi}}_{\mathrm{ee}} \frac{j\omega\eta_0}{2} \right)^{-1} \cdot \overline{\overline{\chi}}_{\mathrm{em}} \tag{6.31}$$

In the same special case of an isotropic metasurface, assuming $\overline{\overline{\chi}}_{\mathrm{ee}} = \chi_{\mathrm{ee}}\overline{\overline{I}}_t$, $\overline{\overline{\chi}}_{\mathrm{mm}} = \chi_{\mathrm{mm}}\overline{\overline{I}}_t$, and $\overline{\overline{\chi}}_{\mathrm{em}} = \overline{\overline{\chi}}_{\mathrm{me}} = 0$, we get simply

$$\hat{\alpha}_{\mathrm{ee}} = \frac{S\epsilon_0\chi_{\mathrm{ee}}}{1 + \frac{j\omega\eta_0\epsilon_0}{2}\chi_{\mathrm{ee}}}, \qquad \hat{\alpha}_{\mathrm{mm}} = \frac{S\chi_{\mathrm{mm}}}{1 + \frac{j\omega\mu_0}{2\eta_0}\chi_{\mathrm{mm}}} \tag{6.32}$$

Both formalisms described above are based on the relations between the induced polarization vectors and the electromagnetic fields at the metasurface plane. There is another possible and very useful description of general metasurfaces, where the electromagnetic properties are modeled by relations between the tangential components of the electromagnetic fields at the two sides of the metasurface. In this approach, the polarization vectors $\mathbf{P}$ and $\mathbf{M}$ are not used. Instead, the surface-averaged tangential electric fields on the two

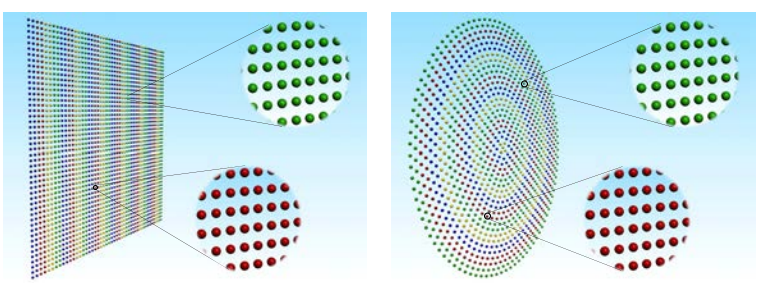
sides of the metasurface are expressed as functions of the tangential components of the magnetic fields (the impedance-matrix model) or the other way around (the admittance matrix model).[15,26] Furthermore, scattering matrix or transmission matrix descriptions are possible. This formalism is basically the same as the common method of modelling microwave components in the theory of microwave circuits. As compared with the models of collective polarizabilities or susceptibilities, the effective parameters of this model are directly measurable (if written in form of a scattering matrix). The transmission matrix form allows very simple calculations for cascaded structures, where a number of different metasurfaces are stacked together (the transmission matrix of the stack is the product of matrices for individual metasurfaces, provided that the separation distance is large compared to the array periods). The impedance matrix model allows interpretation in the form of an equivalent circuit, which helps to gain insight into the physical processes behind the metasurface operation, for example, for absorbing metasurfaces.[41] The effective parameters in the matrix models (dyadic impedances or the elements of the scattering or transmission matrices) can be expressed in terms of the collective polarizabilities of susceptibilities.

6.4. Synthesis of Metasurface Topologies

To synthesize a metasurface for a particular application, we can first determine what collective polarizabilities are required to realize the desired reflected and transmitted fields. Next, we will use the relations between the collective polarizabilities of periodically arranged unit cells and the polarizabilities of the same unit cell but measured in empty space (we call them *individual polarizabilities*). Knowing how a small object should respond to plane-wave illuminations, appropriate topologies can be finally found. The last stage is numerical optimization and fine-tuning of the structure, which is needed because the analytical formulas used in the synthesis process are based on some approximations.

For simplicity, here we will describe the synthesis process for a special case of normally incident plane wave as the metasurface excitation. Actually, in many cases this approach is appropriate even for the synthesis of non-uniform metasurfaces for rather complex field transformations, like wave focusing, thanks to a possibility to use the physical optics approximation, considering the metasurface as a locally periodical structure excited by plane waves. This concept is illustrated in Fig. 6.4.

The starting point is Eqs. (6.17) and (6.18), which define the reflected and transmitted fields in terms of the induced unit-cell electric and magnetic moments. In view of the homogenization-model relations (6.15) and (6.16), they determine the set of collective polarizabilities which we need to realize in order to get the desired reflected and transmitted fields. Indeed, substitution of (6.15) and (6.16) into (6.17) and (6.18) gives the relations between the reflected and transmitted fields in terms of the incident field and the collective polarizabilities. Let us write these expressions explicitly for metasurfaces which are isotropic in the surface plane. In this special case, dyadic coefficients in (6.15) and (6.16) are invariant with respect to rotation



(a) (b)

Fig. 6.4. Within the physical optics approximation, the metasurface can be considered as a locally periodical structure. This way, we can synthesize metasurfaces for various transformations of incident plane waves, such as deflection of incident beams (a) or for focusing (b).

around the unit vector $\mathbf{z}_0$, normal to the layer plane:

$$\overline{\overline{\widehat{\alpha}}}_{ij} = \widehat{\alpha}_{ij}^{\mathrm{co}}\overline{\overline{I}}_{\mathrm{t}} + \widehat{\alpha}_{ij}^{\mathrm{cr}}\overline{\overline{J}}_{\mathrm{t}} \tag{6.33}$$

where $\overline{\overline{I}}_{\mathrm{t}} = \overline{\overline{I}} - \mathbf{z}_0\mathbf{z}_0$ is the transverse unit dyadic and $\overline{\overline{J}}_{\mathrm{t}} = \mathbf{z}_0 \times \overline{\overline{I}}_{\mathrm{t}}$ is the vector-product operator.

Assuming the normal incidence of the incident plane wave (in which case the induced dipole moments are the same in all unit cells), we can find the amplitudes of the reflected and transmitted plane waves as[32]

$$\begin{aligned}
\mathbf{E}_{\mathrm{ref}} = -\frac{j\omega}{2S}\Bigg\{&\left[\eta_0\widehat{\alpha}_{\mathrm{ee}}^{\mathrm{co}} \pm \widehat{\alpha}_{\mathrm{em}}^{\mathrm{cr}} \pm \mu_0\widehat{\alpha}_{\mathrm{me}}^{\mathrm{cr}} - \frac{\mu_0}{\eta_0}\widehat{\alpha}_{\mathrm{mm}}^{\mathrm{co}}\right]\overline{\overline{I}}_{\mathrm{t}} \\
&+ \left[\eta_0\widehat{\alpha}_{\mathrm{ee}}^{\mathrm{cr}} \mp \widehat{\alpha}_{\mathrm{em}}^{\mathrm{co}} \mp \mu_0\widehat{\alpha}_{\mathrm{me}}^{\mathrm{co}} - \frac{\mu_0}{\eta_0}\widehat{\alpha}_{\mathrm{mm}}^{\mathrm{cr}}\right]\overline{\overline{J}}_{\mathrm{t}}\Bigg\} \cdot \mathbf{E}_{\mathrm{inc}}, \\
\mathbf{E}_{\mathrm{tr}} = \Bigg\{&\left[1 - \frac{j\omega}{2S}\left(\eta_0\widehat{\alpha}_{\mathrm{ee}}^{\mathrm{co}} \pm \widehat{\alpha}_{\mathrm{em}}^{\mathrm{cr}} \mp \mu_0\widehat{\alpha}_{\mathrm{me}}^{\mathrm{cr}} + \frac{\mu_0}{\eta_0}\widehat{\alpha}_{\mathrm{mm}}^{\mathrm{co}}\right)\right]\overline{\overline{I}}_{\mathrm{t}} \\
&- \frac{j\omega}{2S}\left[\eta_0\widehat{\alpha}_{\mathrm{ee}}^{\mathrm{cr}} \mp \widehat{\alpha}_{\mathrm{em}}^{\mathrm{co}} \pm \mu_0\widehat{\alpha}_{\mathrm{me}}^{\mathrm{co}} + \frac{\mu_0}{\eta_0}\widehat{\alpha}_{\mathrm{mm}}^{\mathrm{cr}}\right]\overline{\overline{J}}_{\mathrm{t}}\Bigg\} \cdot \mathbf{E}_{\mathrm{inc}}
\end{aligned} \tag{6.34}$$

where η_0 is the free-space wave impedance (for simplicity and without loss of generality, we assume that the surrounding homogeneous and isotropic medium is free space). Here and thereafter, to distinguish between illuminations of the sheet from the two opposite sides, we will use double signs ($\pm$) for these two cases, where the top and bottom signs correspond to the incident plane wave propagating in the $-\mathbf{z}_0$ and $\mathbf{z}_0$ directions, respectively. The components proportional to the transverse unit dyadic $\overline{\overline{I}}_t$ define the co-polarized reflected and transmitted fields $\mathbf{E}_{\mathrm{ref,tr}}^{\mathrm{co}}$, and those proportional to the rotation operator $\overline{\overline{J}}_t$ give the values of the cross-polarized reflected and transmitted fields $\mathbf{E}_{\mathrm{ref,tr}}^{\mathrm{cr}}$, see an illustration in Fig. 6.5.

Equations (6.34) can also be used for finding the effective parameters of the metasurface from calculated or measured response to probe plane waves. It can be done by solving the above equations

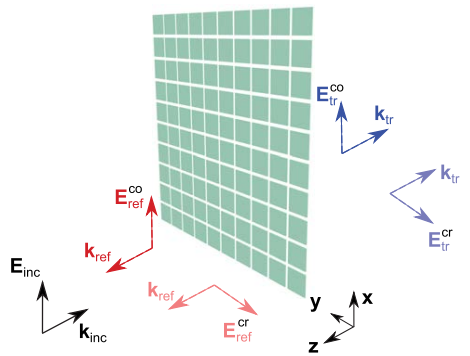


Fig. 6.5. Using uniaxial metasurfaces, we can engineer co- and cross-polarized reflection and transmission coefficients.

for the collective polarizabilities if the reflected and transmitted field created by the known incident field are known.

The next step in the synthesis procedure is to select a set of polarizability values, which would give the desired values of the reflected and transmitted fields. Note that in the general case the metasurface can respond differently for illuminations of its two sides, so in the general case we may need to make sure that reflection and transmission responses for illuminations of both sides of the metasurface are as required for the thought application.

There may be multiple possible choices of polarizabilities which ensure the desired operation. We have eight (complex scalar) polarizability parameters, but the number of reflection/transmission coefficients which we want to tune can be smaller than eight. The choice is usually determined by practical considerations. For example, using non-reciprocal inclusions in metasurface design is usually not desirable. Thus, if no non-reciprocal effects are needed for the target application, it makes sense to limit the choice of polarizabilities to those of reciprocal unit cells. The collective polarizability dyadics of reciprocal metasurfaces satisfy[36,37]

$$\overline{\overline{\widehat{\alpha}}}_{\mathrm{ee}} = \overline{\overline{\widehat{\alpha}}}_{\mathrm{ee}}^{T}, \quad \overline{\overline{\widehat{\alpha}}}_{\mathrm{mm}} = \overline{\overline{\widehat{\alpha}}}_{\mathrm{mm}}^{T}, \quad \overline{\overline{\widehat{\alpha}}}_{\mathrm{em}} = -\mu_0 \overline{\overline{\widehat{\alpha}}}_{\mathrm{me}}^{T} \tag{6.35}$$

Here, T denotes the transpose operation. For our example uniaxial metasurface, these conditions mean that

$$\widehat{\alpha}_{\mathrm{ee}}^{\mathrm{cr}} = \widehat{\alpha}_{\mathrm{mm}}^{\mathrm{cr}} = 0, \qquad \widehat{\alpha}_{\mathrm{em}}^{\mathrm{co}} = -\mu_0 \widehat{\alpha}_{\mathrm{me}}^{\mathrm{co}}, \qquad \widehat{\alpha}_{\mathrm{em}}^{\mathrm{cr}} = \mu_0 \widehat{\alpha}_{\mathrm{me}}^{\mathrm{cr}} \tag{6.36}$$

which reduces the number of free parameters.

The next step is to find out what are the polarizabilities of a single unit cell in free space, which would correspond to collective polarizabilities determined in the previous step. The knowledge of the individual cell response is usually desirable because single-particle polarizabilities are easier to understand and synthesize, as there is no need to model particle interactions in infinite lattices, and for some simple topologies even analytical models are available for polarizabilities of small scatterers in free space. The relations between the collective and individual polarizabilities can be found using the notion of the interaction factor for infinite grids. Namely, we note that each unit cell is actually excited by the local field at the unit-cell position, which is the sum of the incident field and the field created by the currents in all the other unit cells at the position of this particular, reference unit cell. We express this notion in terms of the interaction constant[26,38] β:

$$\mathbf{E}_{\mathrm{loc}} = \mathbf{E}_{\mathrm{inc}} + \overline{\overline{\beta}}_{\mathrm{e}} \cdot \mathbf{p} \tag{6.37}$$

$$\mathbf{H}_{\mathrm{loc}} = \mathbf{H}_{\mathrm{inc}} + \overline{\overline{\beta}}_{\mathrm{m}} \cdot \mu_0 \mathbf{m} \tag{6.38}$$

For the assumed normal-incidence plane-wave excitations, the dipole moments $\mathbf{p}$ and $\mathbf{m}$ of all unit cells are the same. Approximate analytical expressions for the interaction constants of dipolar arrays are available in the literature.[26]

For a single unit cell in free space, the induced moments are proportional to the local fields (since the particle is now alone in empty infinite space, there is no difference between the local fields and the incident fields). We write these relations in terms of the individual polarizabilities as

$$\mathbf{p} = \overline{\overline{\alpha}}_{\mathrm{ee}} \cdot \mathbf{E}_{\mathrm{loc}} + \overline{\overline{\alpha}}_{\mathrm{em}} \cdot \mathbf{H}_{\mathrm{loc}} \tag{6.39}$$

$$\mathbf{m} = \overline{\overline{\alpha}}_{\mathrm{mm}} \cdot \mathbf{H}_{\mathrm{loc}} + \overline{\overline{\alpha}}_{\mathrm{me}} \cdot \mathbf{E}_{\mathrm{loc}} \tag{6.40}$$

Substituting the local fields from (6.37) and (6.38) into (6.15) and (6.16) and solving for the dipole moments, the individual polarizabilities can be expressed in terms of the earlier determined collective polarizabilities.[32]

At this stage, we know how a single inclusion which will be used to form the metasurface with the required response should respond to external fields if tested in free space. Next, we use our knowledge on electromagnetic properties of small scatteres to find suitable realizations. Electric polarizability is easy to control. Basically, any electrically small object made of any material is electrically polarizable, and it is enough to choose the material and dimensions so that the polarizability $\overline{\overline{\alpha}}_{\mathrm{ee}}$ takes the desired value. To access resonant regime, for microwave frequencies metal particles of a proper shape (meander, for instance) can be used. In the infared and visible range, plasmonic nanoparticles offer possible realizations. To realize magnetic response in reciprocal particles, various split-ring topologies can be used in the microwave range, or dielectric particles in the magnetic resonant mode in the visible. Magnetoelectric coupling effects (parameters $\overline{\overline{\alpha}}_{\mathrm{me}}$ and $\overline{\overline{\alpha}}_{\mathrm{em}}$) in reciprocal particles are realized by proper shaping of the inclusions. For example, to control $\alpha_{\mathrm{em}}^{\mathrm{co}}$, various spiral (mirror-asymmetric) shapes are used. Parameter $\alpha_{\mathrm{em}}^{\mathrm{cr}}$ is determined by the symmetry or asymmetry in reversal of the unit vector $\mathbf{z}_0$.

The final step is fine-tuning of the unit-cell parameters using numerical solvers for periodical arrays. This step is usually required because the analytical expressions for the interaction constants β and the analytical models for particle polarizabilities, which are usually used in determining the unit-cell dimensions, are approximate and the result can be improved using accurate simulations. On the other hand, brute-force numerical optimizations without first determining the required topologies from the homogenization-model equations usually does not bring satisfactory results.

It is possible to approach the problem of metasurface synthesis using the homogenization model in terms of surface susceptibilities, expressed in (6.10) and (6.11). In this scenario,[52,53] the solution starts

from finding such sets of susceptibilities which correspond to the desired response in terms of the reflection and transmission coefficients. Usually, there is no unique solution and some restrictions on the susceptibilities are imposed based on considerations of reciprocity and simplicity. Next, using extensive numerical simulations, physical realizations of metasurfaces with the desired susceptibilities are found. Most commonly, a set of simple basic shapes of metal patches (usually, square or rectangular patches, strips, crosses and Jerusalem crosses) is considered. For periodical arrays of simple-shape elements, there is considerable knowledge on what properties one can achieve using different topologies of the unit cell.[29,30] This knowledge helps in intelligent guidance of the optimization process. Finally, S-parameters and transfer matrices of single- and multilayered arrays of such patches are simulated, with variable geometry parameters, in order to find suitable structures and fine-tune their response. A disadvantage of this method as compared to the use of the collective polarizabilities is that there is no reasonably simple possibility to analytically find the required parameters of single unit cells and analytically explore all physically possible approaches to realization. An advantage is that this design approach leads to easily realizable planar structures (but the number of required layers may be large, often five layers and more). On the other hand, the semi-analytical method based on unit-cell polarizabilities assumes that unit cells respond and interact as electric and magnetic dipoles, which is not always an adequate model.

Yet another synthesis method, especially suitable for Huygens' metasurfaces, is based on the idea that each unit cell should be a combination of orthogonal electric and magnetic dipoles.[55,56] Their required amplitudes and phases are found from the solution for plane-wave incidence on an infinite array. Next, a certain topology of a unit cell is selected so that it would be electrically and magnetically polarizable when excited by a plane wave (usually, a kind of symmetric split ring loaded by a capacitance or a set of electric dipoles and loaded rings), and the dimensions are numerically optimized to reach the desired response of the array.

6.5. Examples

In this section, we give several examples of synthesizing single-layer metasurfaces for various applications in order to demonstrate the general approach to the synthesis and design of these structures. Here, we use the method based on the effective polarizabilities model.

6.5.1. *Twist-polarizing metasurfaces*

This example is based on the results published in paper by Niemi *et al.*[32] For a number of applications, it is desirable to realize a thin layer which would create no reflections and rotate the polarization plane of the incident wave by 90°. Such a device is called *twist-polarizer*, see an illustration in Fig. 6.6. The response (absence of reflection and required polarization rotation in transmission) should be the same for any polarization of the incident field, which means that the structure should be isotropic in its plane. Thus, the formulas for uniaxial metasurfaces given above are suitable to synthesize twist-polarizers.

We start from the general expressions for the reflected and transmitted fields (6.34) for a uniaxial metasurface and find out what

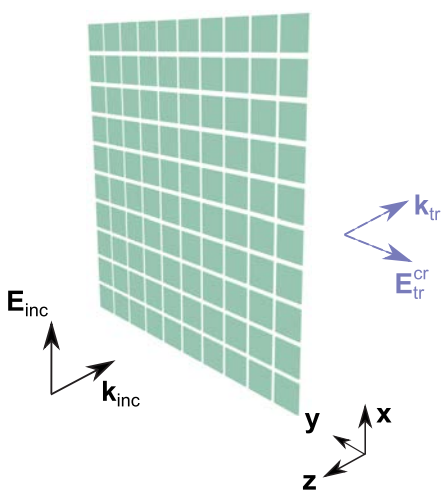


Fig. 6.6. Required functionality of a twist-polarizing metasurface.

should be the values of the collective polarizabilities, so that $\mathbf{E}_{\rm ref} = 0$ and $\mathbf{E}_{\rm tr} = A\,\mathbf{z}_0 \times \mathbf{E}_{\rm inc} = A\,\overline{\overline{J}}_t \cdot \mathbf{E}_{\rm inc}$. The coefficient $A = e^{j\phi}$ can be any complex number having unit absolute value, since we set no restriction on the phase shift in transmission through the twist-polarizing metasurface. Using the general formulas for the reflection and transmission coefficients of single-layer metasurfaces (6.34), we find that the collective polarizabilities of the unit cells must satisfy the following relations:

$$\eta_0 \widehat{\alpha}_{\rm ee}^{\rm co} \pm \widehat{\alpha}_{\rm em}^{\rm cr} \pm \mu_0 \widehat{\alpha}_{\rm me}^{\rm cr} - \frac{\mu_0}{\eta_0} \widehat{\alpha}_{\rm mm}^{\rm co} = 0 \tag{6.41}$$

(co-polarized reflection is absent)

$$\eta_0 \widehat{\alpha}_{\rm ee}^{\rm cr} \mp \widehat{\alpha}_{\rm em}^{\rm co} \mp \mu_0 \widehat{\alpha}_{\rm me}^{\rm co} - \frac{\mu_0}{\eta_0} \widehat{\alpha}_{\rm mm}^{\rm cr} = 0 \tag{6.42}$$

(cross-polarized reflection is absent)

$$1 - \frac{j\omega}{2S} \left(\eta_0 \widehat{\alpha}_{\rm ee}^{\rm co} \pm \widehat{\alpha}_{\rm em}^{\rm cr} \mp \mu_0 \widehat{\alpha}_{\rm me}^{\rm cr} + \frac{\mu_0}{\eta_0} \widehat{\alpha}_{\rm mm}^{\rm co} \right) = 0 \tag{6.43}$$

(co-polarized transmission is absent)

$$-\frac{j\omega}{2S} \left(\eta_0 \widehat{\alpha}_{\rm ee}^{\rm cr} \mp \widehat{\alpha}_{\rm em}^{\rm co} \pm \mu_0 \widehat{\alpha}_{\rm me}^{\rm co} + \frac{\mu_0}{\eta_0} \widehat{\alpha}_{\rm mm}^{\rm cr} \right) = A \tag{6.44}$$

(cross-polarized transmitted field has the desired amplitude and phase).

The first two relations (6.41) and (6.42) express the conditions for realizing a Huygens layer, which does not produce reflections. Equations (6.43) and (6.44) ensure that the transmitted field is twist-polarized. Solving (6.41) together with (6.43), we find the requirements on co-polarized electric and magnetic polarizabilities:

$$\eta_0 \widehat{\alpha}_{\rm ee}^{\rm co} \pm \widehat{\alpha}_{\rm em}^{\rm cr} = \frac{S}{j\omega} \tag{6.45}$$

$$\frac{1}{\eta_0} \widehat{\alpha}_{\rm mm}^{\rm co} \mp \widehat{\alpha}_{\rm me}^{\rm cr} = \frac{S}{j\omega\mu_0} \tag{6.46}$$

This result tells what physical properties of unit cells can be used to realize non-reflecting layers. First, it is obvious that we need to use such unit cells that both electric and magnetic dipole moments can be induced (this result implies that the layer thickness cannot be equal to zero: a planar distribution of surface electric current produces no magnetic moment in the sheet plane). The simplest realization scenario is to use unit cells without antisymmetric bianisotropic response ($\widehat{\alpha}^{\rm cr}_{\rm em} = \widehat{\alpha}^{\rm cr}_{\rm me} = 0$). In this case, we need a unit cell with balanced ("equally strong") electric and magnetic response, and we need to make sure that the normalized collective polarizabilities equal $S/(j\omega)$. Note that the unit-cell area S is a free design parameter; we only need to make sure that the array period is smaller than the wavelength.

Alternatively, it is possible to utilize the magnetoelectric coupling effect, which relaxes the strict requirement on the magnetic polarizability. Theoretically, the magnetic polarizability can even equal zero, if $\widehat{\alpha}^{\rm cr}_{\rm me} = -S/(j\omega\mu_0)$. Especially for optical applications, realizing magnetic response is a challenge, because the effect of inducing magnetic moments by magnetic fields is a weak second-order spatial dispersion effect[37] (in terms of the electrical size of the unit cell). Reciprocal magnetoelectric coupling, measured by $\widehat{\alpha}^{\rm cr}_{\rm me}$, is a stronger, first-order effect, and using that effect it is easier to realize the required balance of induced electric and magnetic moments.[41] The antisymmetric coupling effect can be realized either by making the structure asymmetric with respect to its two transverse dimensions (reciprocal omega coupling, where $\mu_0\widehat{\alpha}^{\rm cr}_{\rm me} = \widehat{\alpha}^{\rm cr}_{\rm em}$), or by introducing non-reciprocal elements but keeping the geometrical symmetry (in that case $\mu_0\widehat{\alpha}^{\rm cr}_{\rm me} = -\widehat{\alpha}^{\rm cr}_{\rm em}$), or combining these two possibilities (in that case, there is no restrictive relation between $\widehat{\alpha}^{\rm cr}_{\rm me}$ and $\widehat{\alpha}^{\rm cr}_{\rm em}$).

Twist-polarizers can be designed to rotate the polarization plane of waves incident on one or both sides of the metasurface. The choice of the $\pm$ and $\mp$ signs in (6.45) and (6.46) determines the side of the metasurface which is matched with free space. If it is required that the twist-polarizer should work the same way for waves incident on

either of its sides, there should be no omega coupling in the unit cells, because we have to demand that $\widehat{\alpha}_{\mathrm{em}}^{\mathrm{cr}} = \widehat{\alpha}_{\mathrm{me}}^{\mathrm{cr}} = 0$.

The remaining two equations (6.42) and (6.44) lead to the conditions

$$\eta_0 \widehat{\alpha}_{\mathrm{ee}}^{\mathrm{cr}} - \widehat{\alpha}_{\mathrm{em}}^{\mathrm{co}} = -\frac{S}{j\omega} A \tag{6.47}$$

$$\eta_0 \widehat{\alpha}_{\mathrm{mm}}^{\mathrm{cr}} + \widehat{\alpha}_{\mathrm{me}}^{\mathrm{co}} = -\frac{S}{j\omega\mu_0} A \tag{6.48}$$

This result indicates that we have the choice of two physical mechanisms (or their combinations) for realizing polarization conversion: non-reciprocal electric and magnetic polarization effect ($\widehat{\alpha}_{\mathrm{ee}}^{\mathrm{cr}} \neq 0$, $\widehat{\alpha}_{\mathrm{mm}}^{\mathrm{cr}} \neq 0$, in this case possibly also $\widehat{\alpha}_{\mathrm{me}}^{\mathrm{cr}} = \widehat{\alpha}_{\mathrm{em}}^{\mathrm{cr}} \neq 0$) or chirality effect ($\widehat{\alpha}_{\mathrm{me}}^{\mathrm{co}} = -\widehat{\alpha}_{\mathrm{em}}^{\mathrm{co}} \neq 0$). The first possibility implies the use of magnetized ferrite or/and magnetized plasma inside the unit cells. Alternatively, we need to use some active devices[42] (e.g. amplifiers) or some external force to modulate the surface parameters in time.[43] The second possibility implies the use of particles with broken mirror-reflection symmetry. In the last case, $\mu_0 \widehat{\alpha}_{\mathrm{me}}^{\mathrm{co}} = -\widehat{\alpha}_{\mathrm{em}}^{\mathrm{co}}$. Again we see that there is considerable design freedom in providing the desired polarization rotation because there are only four relations connecting eight particle polarizabilities.

If we decide to use only reciprocal structures in the design (which is very reasonable from the practical point of view, since the desired response of the metasurface is reciprocal) and require that the twist-polarizer works in the same way for plane waves illuminating both sides of the layer (symmetrical response, $\widehat{\alpha}_{\mathrm{em}}^{\mathrm{cr}} = 0$), the solution for the required collective polarizabilities is unique, and it reads

$$\eta_0 \widehat{\alpha}_{\mathrm{ee}}^{\mathrm{co}} = \frac{\mu_0}{\eta_0} \widehat{\alpha}_{\mathrm{mm}}^{\mathrm{co}} = \frac{S}{j\omega} \tag{6.49}$$

$$\mu_0 \widehat{\alpha}_{\mathrm{me}}^{\mathrm{co}} = -\widehat{\alpha}_{\mathrm{em}}^{\mathrm{co}} = -A \frac{S}{j\omega} \tag{6.50}$$

We see that the unit cells should be chiral particles characterized by these specific values of the polarizabilities. Requirement (6.49) demands that the electric and magnetic responses are balanced, which is known to be realizable for metal-wire spirals of properly chosen dimensions.[44–46] The parameter A is limited only by the requirement of full power transmission, that is, $|A| = 1$. If we set $A = -j$, the relations (6.49) and (6.50) become the same as those for balanced spirals[44–46] and for small Huygens antennas.[47]

Now we are ready to design the unit cells. First, we find the required polarizabilities of a single unit cell in free space using the interaction constant concept, as explained above. For the case of reciprocal helices, the relations between collective and individual polarizabilities can be found in Ref. 32. Next, we select one of the possible chiral shapes, for example, a canonical helix,[32] which is formed by a metal-wire loop connected to two short pieces of straight wires, and make use of the antenna theory to find the relations between the polarizabilities of this helix and its dimensions.[48] Having analytical formulas which connect the helix dimensions and the collective polarizabilities of regular 2D arrays of these helices, it is possible to find dimensions of spirals for which the required values of the collective polarizabilities are realized. As usually, the final design step is a numerical optimization of the unit-cell sizes, varying the dimensions in the vicinity of the analytically found values. Example results, both theoretical and experimental, can be found in Ref. 32.

6.5.2. *Matched transmitarrays*

The material in this section is based on the results of Asadchy *et al.*[40] The functionality of transmitarrays is to shape the transmitted waves by locally controlling the phase of waves passing through the metasurface. Losses due to reflections and absorption in the metasurface should be minimized. Typical applications are extremely thin and (usually) flat lenses. Conventional transmitarrays based on frequency-selective surfaces[15–19] or on metasurfaces[20,21] are multilayer structures with typically three to five layers of patch arrays. Conventional single-layer trasmitarrays[22] produce considerable

reflections. The metasurface concept allows synthesis and realization of single-layer arrays with comparable or better functionalities.

We start from the general expressions for the reflected and transmitted fields (6.34) and demand that the reflected wave is absent,

$$\eta_0\widehat{\alpha}_{\mathrm{ee}}^{\mathrm{co}} \pm \widehat{\alpha}_{\mathrm{em}}^{\mathrm{cr}} \pm \mu_0\widehat{\alpha}_{\mathrm{me}}^{\mathrm{cr}} - \frac{\mu_0}{\eta_0}\widehat{\alpha}_{\mathrm{mm}}^{\mathrm{co}} = 0 \tag{6.51}$$

$$\eta_0\widehat{\alpha}_{\mathrm{ee}}^{\mathrm{cr}} \mp \widehat{\alpha}_{\mathrm{em}}^{\mathrm{co}} \mp \mu_0\widehat{\alpha}_{\mathrm{me}}^{\mathrm{co}} - \frac{\mu_0}{\eta_0}\widehat{\alpha}_{\mathrm{mm}}^{\mathrm{cr}} = 0 \tag{6.52}$$

the transmitted wave has the same polarization as the incident wave

$$\eta_0\widehat{\alpha}_{\mathrm{ee}}^{\mathrm{cr}} \mp \widehat{\alpha}_{\mathrm{em}}^{\mathrm{co}} \pm \mu_0\widehat{\alpha}_{\mathrm{me}}^{\mathrm{co}} + \frac{\mu_0}{\eta_0}\widehat{\alpha}_{\mathrm{mm}}^{\mathrm{cr}} = 0 \tag{6.53}$$

and that the transmitted wave has the same amplitude as the incident wave and its phase is shifted by a desired angle ϕ:

$$1 - \frac{j\omega}{2S}\left(\eta_0\widehat{\alpha}_{\mathrm{ee}}^{\mathrm{co}} \pm \widehat{\alpha}_{\mathrm{em}}^{\mathrm{cr}} \mp \mu_0\widehat{\alpha}_{\mathrm{me}}^{\mathrm{cr}} + \frac{\mu_0}{\eta_0}\widehat{\alpha}_{\mathrm{mm}}^{\mathrm{co}}\right) = e^{j\phi} \tag{6.54}$$

The $\pm$ signs in the above relations refer to the incident waves illuminating the two sides of the metasurface (waves propagating along $\mp\mathbf{z}_0$). If we want to design a transmitarray which works the same way when illuminated from any side, we should make sure that the above relations are satisfied when we take either top or bottom signs. Under this restriction, we should chose unit cells whose parameters obey

$$\widehat{\alpha}_{\mathrm{em}}^{\mathrm{cr}} + \mu_0\widehat{\alpha}_{\mathrm{me}}^{\mathrm{cr}} = 0, \qquad \widehat{\alpha}_{\mathrm{em}}^{\mathrm{cr}} - \mu_0\widehat{\alpha}_{\mathrm{me}}^{\mathrm{cr}} = 0 \tag{6.55}$$

$$\widehat{\alpha}_{\mathrm{em}}^{\mathrm{co}} + \mu_0\widehat{\alpha}_{\mathrm{me}}^{\mathrm{co}} = 0, \qquad \widehat{\alpha}_{\mathrm{em}}^{\mathrm{co}} - \mu_0\widehat{\alpha}_{\mathrm{me}}^{\mathrm{co}} = 0 \tag{6.56}$$

which means that all these coupling coefficients must equal zero and the unit cells should not exhibit any bianisotropic effects.

Thus, for symmetric transmitarrays the above requirements (6.51)–(6.54) simplify to

$$\eta_0 \widehat{\alpha}_{\rm ee}^{\rm co} - \frac{\mu_0}{\eta_0} \widehat{\alpha}_{\rm mm}^{\rm co} = 0 \tag{6.57}$$

$$\eta_0 \widehat{\alpha}_{\rm ee}^{\rm cr} - \frac{\mu_0}{\eta_0} \widehat{\alpha}_{\rm mm}^{\rm cr} = 0, \qquad \eta_0 \widehat{\alpha}_{\rm ee}^{\rm cr} + \frac{\mu_0}{\eta_0} \widehat{\alpha}_{\rm mm}^{\rm cr} = 0 \tag{6.58}$$

$$1 - \frac{j\omega}{2S} \left(\eta_0 \widehat{\alpha}_{\rm ee}^{\rm co} + \frac{\mu_0}{\eta_0} \widehat{\alpha}_{\rm mm}^{\rm co} \right) = e^{j\phi} \tag{6.59}$$

Equations (6.58) mean that $\widehat{\alpha}_{\rm ee}^{\rm cr} = \widehat{\alpha}_{\rm mm}^{\rm cr} = 0$, the electric and magnetic polarizabilities are symmetric dyadics and, thus, the unit cells must be reciprocal (see Eq. (6.36)). The remaining two simple Eqs. (6.57) and (6.59) have a unique solution for the required electric and magnetic collective polarizabilities:

$$\eta_0 \widehat{\alpha}_{\rm ee}^{\rm co} = \frac{\mu_0}{\eta_0} \widehat{\alpha}_{\rm mm}^{\rm co} = \frac{S}{j\omega}(1 - e^{j\phi}) \tag{6.60}$$

Next we use the relations[32] between the collective polarizabilities of the particles in infinite arrays $\widehat{\alpha}_{\rm ee}^{\rm co}$, $\widehat{\alpha}_{\rm mm}^{\rm co}$ and the polarizabilities of the same particles considered as individual, single scatteres in free space. For this simple case of reciprocal non-bianisotropic unit cells, these relations read[32,40]

$$\frac{1}{\eta_0 \alpha_{\rm ee}} = \frac{1}{\eta_0 \widehat{\alpha}_{\rm ee}^{\rm co}} + \frac{\beta_{\rm e}}{\eta_0}, \qquad \frac{1}{\mu_0 \alpha_{\rm mm}/\eta_0} = \frac{1}{\mu_0 \widehat{\alpha}_{\rm mm}^{\rm co}/\eta_0} + \frac{\beta_{\rm e}}{\eta_0} \tag{6.61}$$

From here, we find the required polarizabilities of individual unit cells in free space:

$$\frac{1}{\eta_0 \alpha_{\rm ee}} = \frac{1}{\mu_0 \alpha_{\rm mm}/\eta_0} = \frac{1}{\eta_0} {\rm Re}(\beta_{\rm e}) - \frac{\omega}{2S} \frac{\sin\phi}{1 - \cos\phi} + j \frac{k^3}{6\pi\sqrt{\epsilon_0 \mu_0}} \tag{6.62}$$

This result corresponds to a lossless dipole scatterer, because the imaginary part is only due to radiation damping. The real part of the interaction constant can be estimated analytically for moderate values of the unit-cell sizes and small dipolar particles.[26] For electrically small unit cells, the following quasi-static approximation

is valid:

$$\frac{1}{\eta_0}\mathrm{Re}(\beta_\mathrm{e}) \approx \frac{0.36}{\sqrt{\epsilon_0\mu_0}\,a^3} \tag{6.63}$$

where a is the array period (square unit cells).

To realize such a transmitarray, we need to design electrically small particles with the polarizabilities given by (6.62). The particle should be polarizable by both electric and magnetic fields, have no bianisotropy and be made from a material with negligible losses. A good candidate for realizing low-loss particles with balanced electric and magnetic polarizabilities at microwave frequencies is a metal spiral, see the previous section. However, spirals are chiral objects, but for this application chirality is not allowed (Eqs. (6.56)). A possible solution is to use unit cells containing racemic combinations of two or more balanced spirals. Such arrays behave as non-chiral metasurfaces, because the number of right- and left-handed spirals in each unit cell is the same, and the chirality of individual spirals is compensated.

Note that the same approach can be used to realize single-layer absorbers, which are transparent outside of the absorption band,[40] because for that application also one needs to ensure balanced electric and magnetic response, while chirality is not allowed.

6.5.3. *Metamirrors*

The material in this section is based on the results of Ra'di *et al.* and Asadchy *et al.*[33,34] Let us consider synthesis of metasurfaces which fully reflect incident waves allowing full control over the reflection phase. Often, it is desirable to utilize both sides of the metasurface, so we will seek for possibilities to independently control the reflection phase for illuminations of both sides of the surface. Let us also demand that the polarization state does change upon reflection. Such a metamirror can be used, for example, as a reflectarray antenna, replacing large and heavy parabolic reflectors. Thought realizations as single arrays of small scatterers offer additional application possibilities since metamirrors (in contrast to conventional reflectarrays)

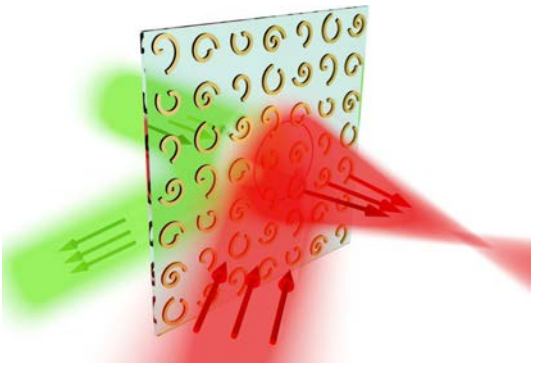


Fig. 6.7. Single-layer metamirrors can independently control reflections from its two sides, while the transmission through the metasurface is negligible.

have no ground plane and are transparent outside of the operational frequency band. Possible functionalities of metamirrors[33] are illustrated in Fig. 6.7.

Similar to the previous examples, we start from the general expressions for the reflection and transmission coefficients (6.34) in terms of the collective polarizabilities of unit cells. Because the desired properties should hold for any polarization of the incident fields, the use of uniaxial structures (with the only preferred direction being the direction normal to the surface) is the only possibility. For the thought application, we demand that

$$1 - \frac{j\omega}{2S}\left(\eta_0\widehat{\alpha}_{\mathrm{ee}}^{\mathrm{co}} \pm \widehat{\alpha}_{\mathrm{em}}^{\mathrm{cr}} \mp \mu_0\widehat{\alpha}_{\mathrm{me}}^{\mathrm{cr}} + \frac{\mu_0}{\eta_0}\widehat{\alpha}_{\mathrm{mm}}^{\mathrm{co}}\right) = 0 \tag{6.64}$$

(co-polarized transmission coefficient is zero)

$$\eta_0\widehat{\alpha}_{\mathrm{ee}}^{\mathrm{cr}} \mp \widehat{\alpha}_{\mathrm{em}}^{\mathrm{co}} \pm \mu_0\widehat{\alpha}_{\mathrm{me}}^{\mathrm{co}} + \frac{\mu_0}{\eta_0}\widehat{\alpha}_{\mathrm{mm}}^{\mathrm{cr}} = 0 \tag{6.65}$$

(cross-polarized transmission coefficient is zero)

$$\eta_0\widehat{\alpha}_{\mathrm{ee}}^{\mathrm{cr}} \mp \widehat{\alpha}_{\mathrm{em}}^{\mathrm{co}} \mp \mu_0\widehat{\alpha}_{\mathrm{me}}^{\mathrm{co}} - \frac{\mu_0}{\eta_0}\widehat{\alpha}_{\mathrm{mm}}^{\mathrm{cr}} = 0 \tag{6.66}$$

(cross-polarized reflection coefficient is zero)

$$-\frac{j\omega}{2S}\left(\eta_0\widehat{\alpha}_{\rm ee}^{\rm co}+\widehat{\alpha}_{\rm em}^{\rm cr}+\mu_0\widehat{\alpha}_{\rm me}^{\rm cr}-\frac{\mu_0}{\eta_0}\widehat{\alpha}_{\rm mm}^{\rm co}\right)=e^{j\phi} \tag{6.67}$$

(co-polarized reflected field has the same amplitude as the incident field and the desired phase shift, ϕ, if the incident wave propagates along $-\mathbf{z}_0$)

$$-\frac{j\omega}{2S}\left(\eta_0\widehat{\alpha}_{\rm ee}^{\rm co}-\widehat{\alpha}_{\rm em}^{\rm cr}-\mu_0\widehat{\alpha}_{\rm me}^{\rm cr}-\frac{\mu_0}{\eta_0}\widehat{\alpha}_{\rm mm}^{\rm co}\right)=e^{j\theta} \tag{6.68}$$

(co-polarized reflected field has the same amplitude as the incident field and the desired phase shift, θ, if the incident wave propagates along $+\mathbf{z}_0$).

Similar to the previous example of twist-polarizing metasurfaces, we see that there are several alternative possibilities to realize metamirrors, using reciprocal or non-reciprocal unit cells. However, in most practical situations the physical properties of inclusions which we can use can be restricted by various considerations. For example, let us consider condition (6.65), which ensures that there is no cross-polarized (with respect to the polarization of the incident wave) transmitted field. Physically, this condition means that the strengths of all physical effects which result in creation of cross-polarized fields behind the metasurface must be balanced so that the total cross-polarized transmission is zero. From (6.65), we see that the cross-polarized transmission can appear if at least one of the following is true:

(1) $\widehat{\alpha}_{\rm ee}^{\rm cr}\neq 0$,

(2) $\widehat{\alpha}_{\rm mm}^{\rm cr}\neq 0$,

(3) $\widehat{\alpha}_{\rm em}^{\rm co}-\mu_0\widehat{\alpha}_{\rm me}^{\rm co}\neq 0$.

In case (1), when $\widehat{\alpha}_{\rm ee}^{\rm cr}\neq 0$, the electric response is non-reciprocal, for example, we have some magnetized plasma filling our unit cells. In case (2), when $\widehat{\alpha}_{\rm mm}^{\rm cr}\neq 0$, the magnetic response is non-reciprocal,

meaning that we have, for example, some magnetized ferrite materials in the unit cells (there are other possibilities to realize non-reciprocal unit cells). In both cases, the polarization transformation in transmission is due to the Faraday effect. In case (3), if the cells are reciprocal, we have chiral (mirror-asymmetric) unit cells, and there is cross-polarizated transmission due to the optical activity of the metasurface.

We see from (6.65) that if we are going to use non-reciprocal unit cells (for example, to have more flexibility in shaping reflections), either *both* electric and magnetic response must be non-reciprocal or, if only one of the parameters $\widehat{\alpha}_{\rm ee}^{\rm cr}$ and $\widehat{\alpha}_{\rm mm}^{\rm cr}$ is non-zero, we must use chiral unit cells, and carefully balance the strength of both effects so that (6.65) is satisfied.

In fact, it is clear that the required functionality (full control over the reflection phase) does not require any non-reciprocal phenomena. Thus, it is most reasonable to use reciprocal structures for realizing metamirrors. Based on these considerations, we can discard possible non-reciprocal realizations, which means that $\widehat{\alpha}_{\rm ee}^{\rm cr} = \widehat{\alpha}_{\rm mm}^{\rm cr} = 0$ (see the reciprocity conditions (6.36)). In this case, to satisfy (6.65), we must ensure that $\widehat{\alpha}_{\rm em}^{\rm co} - \mu_0 \widehat{\alpha}_{\rm me}^{\rm co} = 0$. But the reciprocity condition tells that $\widehat{\alpha}_{\rm em}^{\rm co} = -\mu_0 \widehat{\alpha}_{\rm me}^{\rm co}$; thus, both these coupling coefficients must vanish. Physically, this means that if a metamirror is reciprocal, it must be non-chiral. Similar considerations can be applied to other requirements (6.64)–(6.68). For reciprocal metamirrors, the solution of (6.64)–(6.68) is unique and it reads[33]

$$\begin{aligned}
\eta_0 \widehat{\alpha}_{\rm ee}^{\rm co} &= \frac{S}{j\omega}\left[1 - \frac{e^{j\phi} + e^{j\theta}}{2}\right], \\
\widehat{\alpha}_{\rm em}^{\rm cr} &= \mu_0 \widehat{\alpha}_{\rm me}^{\rm cr} = \frac{-S}{j\omega}\left[\frac{e^{j\phi} - e^{j\theta}}{2}\right], \\
\frac{\mu_0}{\eta_0} \widehat{\alpha}_{\rm mm}^{\rm co} &= \frac{S}{j\omega}\left[1 + \frac{e^{j\phi} + e^{j\theta}}{2}\right]
\end{aligned} \tag{6.69}$$

The next step is to find what are the required individual polarizabilities of a single unit cell in free space (not interacting with the

other particles in the array), see (6.39) and (6.40). This can be done using the concept of the interaction constant $\overline{\overline{\beta}}$, as explained above. The result reads[33]

$$\eta_0\alpha_{\rm ee}^{\rm co}=\frac{1-e^{j(\theta+\phi)}+\frac{j\omega\eta_0}{\beta_{\rm e}S}\left[-\frac{1}{2}\left(e^{j\phi}+e^{j\theta}\right)+1\right]}{-e^{j(\theta+\phi)}+\left(1+j\frac{\omega\eta_0}{\beta_{\rm e}S}\right)^2}\frac{\eta_0}{\beta_{\rm e}},$$

$$\alpha_{\rm em}^{\rm cr}=\mu_0\alpha_{\rm me}^{\rm cr}=\frac{-\frac{j\omega\eta_0}{2\beta_{\rm e}S}\left(e^{j\phi}-e^{j\theta}\right)}{-e^{j(\theta+\phi)}+\left(1+j\frac{\omega\eta_0}{\beta_{\rm e}S}\right)^2}\frac{\eta_0}{\beta_{\rm e}}, \tag{6.70}$$

$$\frac{\mu_0}{\eta_0}\alpha_{\rm mm}^{\rm co}=\frac{1-e^{j(\theta+\phi)}+\frac{j\omega\eta_0}{\beta_{\rm e}S}\left[\frac{1}{2}\left(e^{j\phi}+e^{j\theta}\right)+1\right]}{-e^{j(\theta+\phi)}+\left(1+j\frac{\omega\eta_0}{\beta_{\rm e}S}\right)^2}\frac{\eta_0}{\beta_{\rm e}}$$

Approximate analytical formulas for calculations of the electrical interaction coefficient $\beta_{\rm e}$ can be found in the literature,[26,33] see also Eq. (6.63).

The result shows that the unit cells should be lossless bianisotropic omega particles. The cells should be polarizable electrically ($\alpha_{\rm ee}^{\rm co}\neq 0$) and magnetically ($\alpha_{\rm mm}^{\rm co}\neq 0$), and there should be bianisotropic omega coupling ($\alpha_{\rm em}^{\rm cr}=\mu_0\alpha_{\rm me}^{\rm cr}\neq 0$). The last condition means that electric field applied to the particle in the direction orthogonal to $\mathbf{z}_0$ should induce magnetic moment in the direction orthogonal to the applied electric field and also orthogonal to $\mathbf{z}_0$. The typical topology of such particles is an Ω-shaped piece of a metal wire (appropriate for microwave-frequency realizations). Knowing the required values of the polarizabilities, we can now use the known analytical models of polarizabilities of omega particles[37,39] and find the particle dimensions for which the polarizabilities take the required values. At the final stage, the dimensions can be optimized using full-wave simulations of the designed particles in the infinite array.

Reports on numerical and experimental studies of metamirrors realized as arrays of metal Ω-shaped particles can be found in the

literature.[33–35] Specially shaped dielectric particles can be used for realization of metamirrors for optical applications.[49]

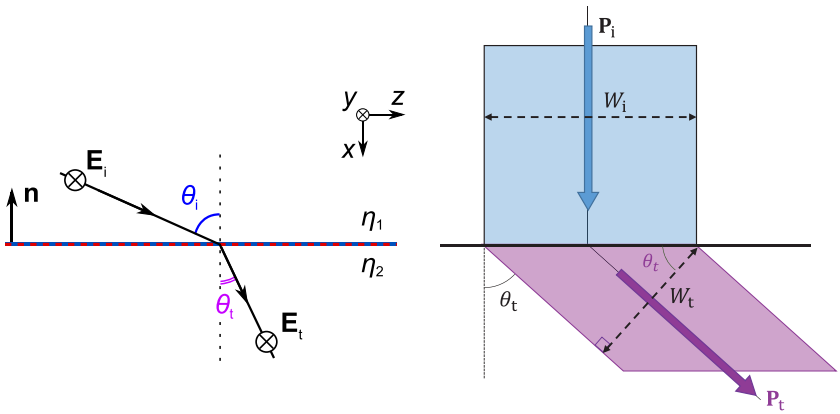
6.6. Perfect Control of Anomalous Transmission and Reflection

The design methodology explained in the previous sections is based on the assumption that each unit cell introduces a phase shift in transmission or reflection without any modification in the wave amplitude, following the same design principle that the conventional phased arrays. Recently, it has been demonstrated that this simplistic method does not allow us to achieve perfect performance due to the energy that is inevitable scattered in undesirable directions.[57–60] These studies show that for realizing perfect performance metasurfaces, which redirect all the energy in the desired direction, the amplitudes of the incident and transmitted/reflected waves have to be different. In this section we present the last advances in the design of perfect metasurfaces for shaping transmitted and reflected waves. We will go into detail about the power considerations in both transmission and reflection scenarios. Moreover, methods for calculating the required collective polarizabilities in both scenarios are presented.

6.6.1. *Perfect transmitarrays*

Recently, it has been shown that for realizing perfect control of transmission in metasurfaces, an important condition that ensures complete suppression of the energy radiated in undesirable directions should be satisfied. Here we present this condition and an approach for calculating the collective polarizabilities that ensure this behaviour.

First, we define the electric and magnetic fields when a TE-polarized wave, with the amplitude $\mathbf{E}_\mathrm{i}$, illuminates the metasurface from medium 1 at a certain angle θ_i (see Fig. 6.8(a)). We refer to this case of incidence as *forward illumination*. Assuming ideally zero reflection, the tangential components of the electric and magnetic



(a) (b)

Fig. 6.8. Illustration of the performance desired in an ideal transmitarray. (a) Coordinate system and field definitions. (b) Schematic representation of the power conservation in the transmitarray.

fields on both sides of the metasurface can be written as

$$\mathbf{E}_{t1} = \mathbf{E}_{\mathrm{i}} e^{-jk_1 \sin\theta_{\mathrm{i}} z}, \quad \mathbf{n} \times \mathbf{H}_{t1} = \frac{\cos\theta_{\mathrm{i}}}{\eta_1} \mathbf{E}_{\mathrm{i}} e^{-jk_1 \sin\theta_{\mathrm{i}} z} \tag{6.71}$$

$$\mathbf{E}_{t2} = t_{\mathrm{TE}}\, \mathbf{E}_{\mathrm{i}} e^{-jk_1 \sin\theta_{\mathrm{i}} z}, \quad \mathbf{n} \times \mathbf{H}_{t2} = t_{\mathrm{TE}}\, \frac{\cos\theta_{\mathrm{t}}}{\eta_2} \mathbf{E}_{\mathrm{i}} e^{-jk_1 \sin\theta_{\mathrm{i}} z} \tag{6.72}$$

where t_{TE} is the complex transmission coefficient

$$t_{\mathrm{TE}} = |t_{\mathrm{TE}}| e^{j\Phi_{\mathrm{t}}}, \quad \Phi_{\mathrm{t}} = (k_1 \sin\theta_{\mathrm{i}} - k_2 \sin\theta_{\mathrm{t}}) z + \phi_{\mathrm{t}} \tag{6.73}$$

Perfect transmission can be obtained by ensuring that all the incident energy goes into the desirable direction, thus, the normal component of the Poynting vector has to be the same at both sides of the metasurface:

$$\frac{1}{2}\mathrm{Re}(\mathbf{E}_{\mathrm{t1}} \times \mathbf{H}^*_{\mathrm{t1}}) = \frac{1}{2}\mathrm{Re}(\mathbf{E}_{\mathrm{t2}} \times \mathbf{H}^*_{\mathrm{t2}}) \tag{6.74}$$

This condition defines a relation between the incident and the reflected fields, which reads

$$\mathbf{E}_{\mathrm{t}} = \mathbf{E}_{\mathrm{i}} \sqrt{\frac{\cos\theta_{\mathrm{i}}}{\cos\theta_{\mathrm{t}}}} \sqrt{\frac{\eta_1}{\eta_2}} = \mathbf{E}_{\mathrm{i}} |t_{\mathrm{TE}}| \tag{6.75}$$

Equation (6.75) shows that the amplitudes of the incident and the transmitted waves are different. Intuitively, this condition can be easily understood by considering the example of a metasurface which steers a beam of a finite width W_{i}. Figure 6.8(b) conceptually illustrates this idea, where the incident wave impinges normally into the metasurface for simplicity. We can see that the transmitted beam is tilted by θ_{t} and it has a smaller width W_{t}. Using simple trigonometrical analysis we can obtain the relation between both widths, $W_{\mathrm{i}} = W_{\mathrm{t}}/\cos\theta_{\mathrm{t}}$, supporting the conclusion extracted from Eq. (6.75).

To find the required collective polarizabilities, we need to study the reciprocal case when the metasurface is illuminated by the incident wave with the electric field $\mathbf{E}_{\mathrm{t}}$ from medium 2 at the angle θ_{t}. We refer to this incidence scenario as *backward illumination*. In this case the tangential components of the electric and magnetic fields read

$$\mathbf{E}_{t1} = \mathbf{E}_{\mathrm{i}} e^{jk_1 \sin\theta_{\mathrm{i}} z}, \quad \mathbf{n} \times \mathbf{H}_{t1} = -\frac{\cos\theta_{\mathrm{i}}}{\eta_1} \mathbf{E}_{\mathrm{i}} e^{jk_1 \sin\theta_{\mathrm{i}} z} \tag{6.76}$$

$$\mathbf{E}_{t2} = t^*_{\mathrm{TE}}\ \mathbf{E}_{\mathrm{i}} e^{jk_1 \sin\theta_{\mathrm{i}} z}, \quad \mathbf{n} \times \mathbf{H}_{t2} = -t^*_{\mathrm{TE}}\ \frac{\cos\theta_{\mathrm{t}}}{\eta_2} \mathbf{E}_{\mathrm{i}} e^{jk_1 \sin\theta_{\mathrm{i}} z} \tag{6.77}$$

where symbol "$*$" denotes the complex conjugate operator.

Once we have defined the functionality of the metasurface ensuring the reciprocal response, the next step in the design methodology is to obtain analytical formulas for the collective polarizabilities of the ideal refractive metasurface. Using equations (6.15) and (6.16) and the definition of the fields in the forward illumination (f) scenario, we can find an expression for the surface polarization

densities $\mathbf{p}_t^{\mathrm{f}}$ and $\mathbf{m}_t^{\mathrm{f}}$:

$$\begin{aligned} \mathbf{p}_t^{\mathrm{f}} &= \left(\widehat{\alpha}_{\mathrm{ee}}^{yy} + \widehat{\alpha}_{\mathrm{em}}^{yz}\frac{\cos\theta_{\mathrm{i}}}{\eta_1}\right)\mathbf{E}_{\mathrm{i}}e^{-jk_1\sin\theta_{\mathrm{i}}z}, \\ \mathbf{n}\times\mathbf{m}_t^{\mathrm{f}} &= \left(\widehat{\alpha}_{\mathrm{me}}^{zy} + \widehat{\alpha}_{\mathrm{mm}}^{zz}\frac{\cos\theta_{\mathrm{i}}}{\eta_1}\right)\mathbf{E}_{\mathrm{i}}e^{-jk_1\sin\theta_{\mathrm{i}}z} \end{aligned} \tag{6.78}$$

Following the same procedure for the case of backward illumination (b), the same polarizabilities relate another set of surface polarization densities $\mathbf{p}_t^{\mathrm{b}}$ and $\mathbf{m}_t^{\mathrm{b}}$ to the incident fields from medium 2 (that is, to $t_{\mathrm{TE}}^{*}\mathbf{E}_{\mathrm{i}}e^{jk_1\sin\theta_{\mathrm{i}}z}$):

$$\begin{aligned} \mathbf{p}_t^{\mathrm{b}} &= \left(\widehat{\alpha}_{\mathrm{ee}}^{yy} - \widehat{\alpha}_{\mathrm{em}}^{yz}\frac{\cos\theta_{\mathrm{t}}}{\eta_2}\right)t_{\mathrm{TE}}^{*}\mathbf{E}_{\mathrm{i}}e^{jk_1\sin\theta_{\mathrm{i}}z}, \\ \mathbf{n}\times\mathbf{m}_t^{\mathrm{b}} &= \left(\widehat{\alpha}_{\mathrm{me}}^{zy} - \widehat{\alpha}_{\mathrm{mm}}^{zz}\frac{\cos\theta_{\mathrm{t}}}{\eta_2}\right)t_{\mathrm{TE}}^{*}\mathbf{E}_{\mathrm{i}}e^{jk_1\sin\theta_{\mathrm{i}}z} \end{aligned} \tag{6.79}$$

In the next step, we apply the boundary conditions that link the tangential fields on the two sides of the metasurface to the electric $\mathbf{p}_t$ and magnetic $\mathbf{m}_t$ surface polarization densities induced in the metasurface. These classical boundary conditions simply express the fact that the jumps of the tangential field components at the interface equal to the corresponding surface current densities:

$$\begin{aligned} \mathbf{E}_{t1} - \mathbf{E}_{t2} &= \frac{j\omega\mu_0}{S}\mathbf{n}\times\mathbf{m}_t^{\mathrm{f,b}}, \\ \mathbf{n}\times\mathbf{H}_{t1} - \mathbf{n}\times\mathbf{H}_{t2} &= \frac{j\omega}{S}\mathbf{p}_t^{\mathrm{f,b}} \end{aligned} \tag{6.80}$$

We substitute (6.71)–(6.77) and (6.78)–(6.79) in (6.80) for the two cases of illumination, into the boundary conditions one can obtain

the following system of linear equations

$$\begin{aligned}
&\frac{\cos\theta_\mathrm{i}}{\eta_1} - \frac{\cos\theta_\mathrm{t}}{\eta_2} t_\mathrm{TE} = \frac{j\omega}{S}\left(\widehat{\alpha}_\mathrm{ee}^{yy} + \widehat{\alpha}_\mathrm{em}^{yz}\frac{\cos\theta_\mathrm{i}}{\eta_1}\right),\\
&1 - t_\mathrm{TE} = \frac{j\omega\mu_0}{S}\left(\widehat{\alpha}_\mathrm{me}^{zy} + \widehat{\alpha}_\mathrm{mm}^{zz}\frac{\cos\theta_\mathrm{i}}{\eta_1}\right)\\
&\qquad -\frac{\cos\theta_\mathrm{i}}{\eta_1} + \frac{\cos\theta_\mathrm{t}}{\eta_2} t_\mathrm{TE}^* = \frac{j\omega}{S}\left(\widehat{\alpha}_\mathrm{ee}^{yy} - \widehat{\alpha}_\mathrm{em}^{yz}\frac{\cos\theta_\mathrm{t}}{\eta_2}\right) t_\mathrm{TE}^*,\\
&1 - t_\mathrm{TE}^* = \frac{j\omega\mu_0}{S}\left(\widehat{\alpha}_\mathrm{me}^{zy} - \widehat{\alpha}_\mathrm{mm}^{zz}\frac{\cos\theta_\mathrm{t}}{\eta_2}\right) t_\mathrm{TE}^*
\end{aligned} \tag{6.81}$$

Solving this system, we find the required polarizabilities

$$\begin{aligned}
\widehat{\alpha}_\mathrm{ee}^{yy} &= \frac{S}{j\omega}\frac{\cos\theta_\mathrm{i}\cos\theta_\mathrm{t}}{\eta_1\cos\theta_\mathrm{t} + \eta_2\cos\theta_\mathrm{i}}\\
&\quad\times\left[2 - \left(\sqrt{\frac{\eta_1\cos\theta_\mathrm{t}}{\eta_2\cos\theta_\mathrm{i}}} + \sqrt{\frac{\eta_2\cos\theta_\mathrm{i}}{\eta_1\cos\theta_\mathrm{t}}}\right) e^{j\Phi_\mathrm{t}(z)}\right]
\end{aligned} \tag{6.82}$$

$$\begin{aligned}
\widehat{\alpha}_\mathrm{mm}^{zz} &= \frac{S}{j\omega\mu_0}\frac{\eta_1\eta_2}{\eta_1\cos\theta_\mathrm{t} + \eta_2\cos\theta_\mathrm{i}}\\
&\quad\times\left[2 - \left(\sqrt{\frac{\eta_1\cos\theta_\mathrm{t}}{\eta_2\cos\theta_\mathrm{i}}} + \sqrt{\frac{\eta_2\cos\theta_\mathrm{i}}{\eta_1\cos\theta_\mathrm{t}}}\right) e^{j\Phi_\mathrm{t}(z)}\right]
\end{aligned} \tag{6.83}$$

$$\widehat{\alpha}_\mathrm{em}^{yz} = -\mu_0\widehat{\alpha}_\mathrm{me}^{zy} = \frac{S}{j\omega}\frac{\eta_2\cos\theta_\mathrm{i} - \eta_1\cos\theta_\mathrm{t}}{\eta_1\cos\theta_\mathrm{t} + \eta_2\cos\theta_\mathrm{i}} \tag{6.84}$$

Equations (6.82) and (6.83) show that the electric and magnetic polarizabilities depend on the coordinate z. In other words, zero reflection at any point of the metasurface is required for obtaining a perfect refractive metasurface, which demands the balance of the induced electric and magnetic surface currents at any point (the Huygens condition). On the other hand, the omega coupling coefficient in (6.84) is constant with respect to z and depends only on the impedances and angles. This result reflects the fact that bianisotropic coupling of the omega type is necessary to ensure that the

waves incident on both sides of the metasurface see the same surface impedance, so that reciprocal full transmission is realized. It is important to notice that when the impedances of the incident and transmitted waves are the same, that is, $\frac{\eta_1}{\cos\theta_\mathrm{i}} = \frac{\eta_2}{\cos\theta_\mathrm{t}}$, the required coupling coefficient vanishes.

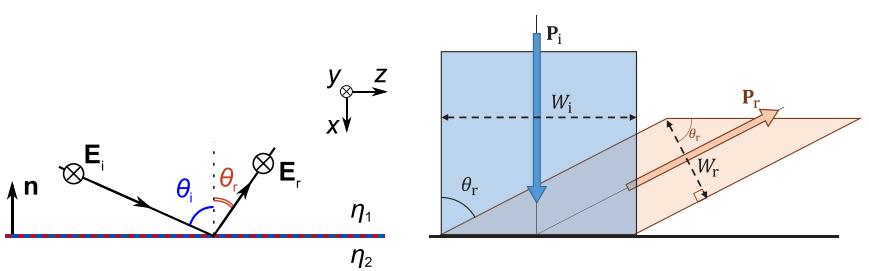
6.6.2. *Perfect reflectarrays*

Similar conclusions can be made for the perfect reflection control when an incident plane wave with the amplitude $\mathbf{E}_\mathrm{i}$ and the incident angle θ_i is reflected into another plane wave with the amplitude $\mathbf{E}_\mathrm{r}$ and the direction defined by the reflection angle θ_r (see Fig. 6.9(a)). In our analysis, we study reflective metasurfaces assuming that the tangential fields behind the metasurface are zero,

$$\mathbf{E}_{t2} = 0, \quad \mathbf{n} \times \mathbf{H}_{t2} = 0 \tag{6.85}$$

Tangential components of the electric and magnetic fields in medium 1 can be written as

$$\begin{aligned} \mathbf{E}_{t1} &= \mathbf{E}_\mathrm{i} e^{-jk_1 \sin\theta_\mathrm{i} z} + \mathbf{E}_\mathrm{r} e^{-jk_1 \sin\theta_\mathrm{r} z + j\phi_\mathrm{r}}, \\ \mathbf{n} \times \mathbf{H}_{t1} &= \mathbf{E}_\mathrm{i} \frac{1}{\eta_1} \cos\theta_\mathrm{i} e^{-jk_1 \sin\theta_\mathrm{i} z} - \mathbf{E}_\mathrm{r} \frac{1}{\eta_1} \cos\theta_\mathrm{r} e^{-jk_1 \sin\theta_\mathrm{r} z + j\phi_\mathrm{r}} \end{aligned} \tag{6.86}$$



(a) (b)

Fig. 6.9. Illustration of the performance desired in a perfect reflectarray. (a) Coordinate system and definitions of the fields. (b) Schematic representation of the power conservation in the reflectarray.

Using this definition of the fields, we define the complex reflection coefficient for the TE incidence as

$$r_{\mathrm{TE}} = \frac{|\mathbf{E}_{\mathrm{r}}|e^{-jk_1 \sin\theta_{\mathrm{r}} z + j\phi_{\mathrm{r}}}}{|\mathbf{E}_{\mathrm{i}}|e^{-jk_1 \sin\theta_{\mathrm{i}} z}} = |r_{\mathrm{TE}}|e^{j\Phi_{\mathrm{r}}}, \quad \Phi_{\mathrm{r}} = k_1(\sin\theta_{\mathrm{i}} - \sin\theta_{\mathrm{r}})z + \phi_{\mathrm{r}} \tag{6.87}$$

The condition for perfect reflection can be found, similarly to the transmission case, analysing the normal component of the Poynting vector. In this case, we have to ensure that all the energy illuminating the metasurface is reflected only into the desired direction. The power carried by the incident plan wave is $P_{\mathrm{i}} = \frac{|\mathbf{E}_{\mathrm{i}}|^2}{2\eta_1}\cos\theta_{\mathrm{i}}$, while the power carried by the reflected wave is $P_{\mathrm{r}} = \frac{|\mathbf{E}_{\mathrm{r}}|^2}{2\eta_1}\cos\theta_{\mathrm{r}}$. Equating these two expressions, we can find the relation between the incident and reflected field amplitudes. This relation reads

$$\mathbf{E}_{\mathrm{r}} = \mathbf{E}_{\mathrm{i}}\sqrt{\frac{\cos\theta_{\mathrm{i}}}{\cos\theta_{\mathrm{r}}}}\sqrt{\frac{\eta_1}{\eta_2}} = \mathbf{E}_{\mathrm{i}}|r_{\mathrm{TE}}| \tag{6.88}$$

Next, we use the same boundary conditions as in the case of perfect refractive metasurfaces, i.e., boundary conditions (6.80), however, only for the forward illumination direction (since the metamirror metasurface behaves as a boundary disconnecting medium 1 from medium 2). Substituting the fields from (6.86) and (6.85) and the surface polarizations (6.78) in the boundary conditions (6.80), one can obtain the following system of equations:

$$\begin{aligned} 1 + r_{\mathrm{TE}} &= \frac{j\omega}{S}\left(\widehat{\alpha}_{\mathrm{ee}}^{yy} + \widehat{\alpha}_{\mathrm{em}}^{yz}\frac{\cos\theta_{\mathrm{i}}}{\eta_1}\right), \\ \frac{\cos\theta_{\mathrm{i}}}{\eta_1} - \frac{\cos\theta_{\mathrm{r}}}{\eta_1}r_{\mathrm{TE}} &= \frac{j\omega\mu_0}{S}\left(\widehat{\alpha}_{\mathrm{me}}^{zy} + \widehat{\alpha}_{\mathrm{mm}}^{zz}\frac{\cos\theta_{\mathrm{i}}}{\eta_1}\right) \end{aligned} \tag{6.89}$$

Obviously, these equations have infinitely many solutions for polarizabilities which realize the desired response. The metasurface can be either bianisotropic (omega coupling) or it can be a non-bianisotropic pair of electric and magnetic current sheets. For the

non-bianisotropic realization we set

$$\widehat{\alpha}_{\mathrm{em}}^{yz} = \widehat{\alpha}_{\mathrm{me}}^{zy} = 0 \tag{6.90}$$

and find the unique solution

$$\widehat{\alpha}_{\mathrm{ee}}^{yy} = \frac{S}{j\omega}\left(1 + \sqrt{\frac{\cos\theta_{\mathrm{i}}}{\cos\theta_{\mathrm{r}}}}\sqrt{\frac{\eta_1}{\eta_2}}e^{j\Phi_{\mathrm{r}}}\right) \tag{6.91}$$

$$\widehat{\alpha}_{\mathrm{mm}}^{zz} = \frac{S}{j\omega\mu_0}\left(1 - \sqrt{\frac{\cos\theta_{\mathrm{r}}}{\cos\theta_{\mathrm{i}}}}\sqrt{\frac{\eta_1}{\eta_2}}e^{j\Phi_{\mathrm{r}}}\right) \tag{6.92}$$

These expressions show the collective polarizabilities required for the desired performance when the bianisotropy is neglected and we can immediately see what are the appropriate topologies of unit cells for implementing them. Since we need both electric and magnetic polarizations, the physical thickness of the reflecting layer must be different from zero, to allow formation of tangential magnetic moments in unit cells. For example, it is not possible to realize the desired performance by any patterning of a single, infinitesimally thin sheet of a perfect conductor.

If we relax the requirement of non-bianisotropic particles, the equations suggest the use of a single array of small particles which are polarizable both electrically and magnetically, such as small metal spirals as in Ref. 40. A typical realization based on the bianisotropic route is a high-impedance surface with a PEC ground plane (such as "mushroom layers"[61]). These implementations based on bianisotropic particles present the advantage of stronger magnetic excitation via bianisotropic coupling. Especially for optical applications, it is easier to realize strong bianisotropy (which is a first-order dispersion effect) as compared with the artificial magnetism (which is a weaker, second-order effect).[37] An experimental realization of a perfect reflectarray as an inhomogeneous high-impedance surface has been reported in Ref. 62.

6.7. Conclusion

We hope that after reading this chapter the reader understands and appreciates the power of the metasurface concept, which brings possibilities to realize extremely thin layers with rather general and powerful functionalities. The Huygens principle tells that action of volumetric currents can be replaced by equivalent action of surface-bound currents, suggesting that volumetric, 3D metamaterial devices often can be replaced by very thin, effectively 2D metasurfaces. For example, instead of bulky lenses, just a thin sheet can be used. A typical example of practical advantages of metasurfaces is the mantle cloak,[24,25] which can provide the same or better functionalities than a volumetric scattering-cancellation cloak,[23] while being extremely compact, easy to manufacture and low-cost. However, not all functionalities which are available with volumetric metamaterials can be realized with metasurfaces. It is obvious that if one wants full control over the distribution of electromagnetic fields in a given volume, the volume must contain sources (or secondary sources in form of meta-atoms). On the other hand, for vast majority of applications, field-controlling devices should control fields in space which is outside of the device itself (lenses, frequency-selective surfaces, cloaks, polarization converters, etc.), and in these applications metasurfaces show their advantageous properties.

The general synthesis approach based on an appropriate homogenization model and collective polarizabilities of unit cells has been explained, and we have seen how we can reveal all possible physical mechanisms which can be used to realize the desired electromagnetic response of metasurfaces. Starting from the desired reflection and transmission coefficients, we find the required collective polarizilities, and this result immediately tells what properties of the unit cells are needed (non-reciprocal or reciprocal, chiral or non-chiral, etc.). Based on this knowledge, we can finally determine suitable unit-cell topologies and realize devices with required functionalities.

Alternative synthesis methods are known and are actively developed by several research groups, as reviewed in Section 6.3. The choice of the most suitable model and design methods is not an

easy one, and it is based on the required metasurface functionalities, limitations in practical realizations, etc.

References

1. Glybovski, S. B., Tretyakov, S. A., Belov, P. A., Kivshar, Y. S., Simovski, C. R. (2016). Metasurfaces: From microwaves to visible, *Phys. Rep.* **634**, pp. 1–72.
2. Holloway, C. L., Kuester, E. F., Gordon, J. A., O'Hara, J., Booth, J. and Smith, D. R. (2012). An overview of the theory and applications of metasurfaces: The two-dimensional equivalents of metamaterials, *IEEE Antenn. Propag. Magaz.* **54**(2), pp. 10–35.
3. Kildishev, A. V., Boltasseva, A. and Shalaev, V. M. (2013). Planar photonics with metasurfaces. *Science* **339**(6125), pp. 1–6.
4. Yu, N. and Capasso, F. (2014). Flat optics with designer metasurfaces. *Nat. Mater.* **13**(2), pp. 139–150.
5. Harringthon, R. F. (2001). *Time-Harmonic Electromagnetic Fields,* (IEEE Press).
6. Lindell, I. V. (1992). *Methods for Electromagnetic Field Analysis* Wiley.
7. Lindell, I. V., Tretyakov, S. A. and Nikoskinen, K. (2000). Extended electromagnetic continuity condition and generalized Huygens' principle, *Electromagnetics* **20**(3), pp. 233–242.
8. Tretyakov, S. A., Nefedov, I. S. and Alitalo, P. (2008). Generalized field-transforming metamaterials, *New J. Phys.* **10**(11), p. 115028.
9. Kuester, E. F., Mohamed, M. A., Piket-May, M. and Holloway, C. L. (2003). Averaged transition conditions for electromagnetic fields at a metafilm, *IEEE Trans. Antenn. Propag.* **51**(10), pp. 2641–2651.
10. Holloway, C. L., Dienstfrey, A., Kuester, E. F., O'Hara, J. F., Azad, A. K. and Taylor, A. J. (2009). A discussion on the interpretation and characterization of metafilms/metasurfaces: The two-dimensional equivalent of metamaterials, *Metamaterials* **3**(2), pp. 100–112.
11. Simovski, C. R., Kondratjev, M. S., Belov, P. A. and Tretyakov, S. A. (1999). Interaction effects in two-dimensional bianisotropic arrays, *IEEE Trans. Antenn. Propag.* **47**(9), pp. 1429–1439.
12. Maslovski, S. I. and Tretyakov, S. A. (1999). Full-wave interaction field in two-dimensional arrays of dipole scatterers *Int. J. Electron. Comunic. Arch. Elek. Übertragungstech. (AEÜ)* **53**(3), pp. 135–139.
13. Tretyakov, S. A., Viitanen, A. J., Maslovski, S. I. and Saarela, I. E. (2003). Impedance boundary conditions for regular dense arrays of dipole scatterers, *IEEE Trans. Antenn. Propag.* **51**(8), pp. 2073–2078.
14. Yatsenko, V. V., Maslovski, S. I., Tretyakov, S. A., Prosvirnin, S. L. and Zouhdi, S. (2003). Plane-wave reflection from double arrays of small magnetoelectric scatterers, *IEEE Trans. Antenn. Propag.* **51**(1), pp. 2–11.

15. Pozar, D. M. (1996). Flat lens antenna concept using aperture coupled microstrip patches, *Electron. Lett.* **32**(23), pp. 2109–2111.
16. Gagnon, N., Petosa, A. and McNamara, D. A. (2013). Research and development on phase-shifting surfaces (PSSs), *IEEE Antenn. Propag. Magaz.* **55**(2), pp. 29–48.
17. Rahmati, B. and Hassani, H. R. (2015). High-efficient wideband slot transmitarray antenna, *IEEE Trans. Antenn. Propag.* **63**(11), pp. 5149–5155.
18. Abdelrahman, A. H., Elsherbeni, A. Z. and Yang, F. (2014). Transmission phase limit of multilayer frequency-selective surfaces for transmitarray designs, *IEEE Trans. Antenn. Propag.* **62**(2), pp. 690–697.
19. Ryan, C. G. M., Chaharmir, M. R., Shaker, J., Bray, J. R., Antar, Y. M. M. and Ittipiboon, A. (2010). A wideband transmitarray using dual-resonant double square rings, *IEEE Trans. Antenn. Propag.* **58**(5), pp. 1486–1493.
20. Pfeiffer C. and Grbic, A. (2013). Millimeter-wave transmitarrays for wavefront and polarization control, *IEEE Trans. Antenn. Propag.* **61**(12), pp. 4407–4417.
21. Monticone, F., Estakhri, N. M. and Alù, A. (2013). Full control of nanoscale optical transmission with a composite metascreen, *Phys. Rev. Lett.* **110**(20), p. 203903.
22. Milne, R. (1982). Dipole array lens antenna, *IEEE Trans. Antenn. Propag.* **30**(4), pp. 704–712.
23. Alù, A. and Engheta, N. Achieving transparency with plasmonic and metamaterial coatings. *Phys. Rev. E* **72**(1), p. 016623, [erratum *Phys. Rev. E* **73**(1), p. 019906 (2006)].
24. Alù, A. (2009). Mantle cloak: Invisibility induced by a surface. *Phys. Rev. B* **80**(24), p. 245115.
25. Soric, J. C., Monti, A., Toscano, A., Bilotti, F. and Alù, A. (2015). Dual-polarized reduction of dipole antenna blockage using mantle cloaks, *IEEE Trans. Antenn. Propag.* **63**(11), pp. 4827–4834.
26. Tretyakov, S. (2003). *Analytical modeling in applied electromagnetics,* Artech House.
27. Senior, T. B. A. and Volakis, J. L. (1995). *Approximate Boundary Conditions in Electromagnetics* (London: IEE Electromagnetic Waves Series).
28. Tretyakov, S. A. (1998). Generalized impedance boundary conditions for isotropic multilayers, *Microwave Opt. Technol. Lett.* **17**(4), pp. 262–265.
29. Vardaxoglou, Y. (1997). *Frequency Selective Surfaces: Analysis and Design* (John Wiley & Sons).
30. Munk, B. A. (2000). *Frequency Selective Surfaces: Theory and Design* (John Wiley & Sons).
31. Kuester, E. F., Mohamed, M. A., Piket-May, M. and Holloway, C. L. (2003). Averaged transition conditions for electromagnetic fields at a metafilm. *IEEE Trans. Ant. Prop.* **51**(10), pp. 2641–2651.
32. Niemi, T., Karilainen, A. and Tretyakov, S. (2013). Synthesis of polarization transformers. *IEEE Trans. Antenn. Propag.* **61**(6), pp. 3102–3111.

33. Ra'di, Y., Asadchy, V. S. and Tretyakov, S. A. (2014). Tailoring reflections from thin composite metamirrors. *IEEE Trans. Antenn. Propag.* **62**(7), pp. 3749–3760.
34. Asadchy, V. S., Ra'di, Y., Vehmas, J. and Tretyakov, S. A. (2015). Functional metamirrors using bianisotropic elements, *Phys. Rev. Lett.* **114**(9), p. 095503.
35. Tcvetkova, S. N., Asadchy, V. S. and Tretyakov, S. A. (2016). Scanning characteristics of metamirror antennas with sub-wavelength focal distance, *IEEE Trans. Antennas Propag.* **64**(8), pp. 3656–3660.
36. Landau, L. D. and Lifshitz, E. M. (1997). Statistical Physics. Oxford, U.K.: Butterworth-Heinemann vol. 5, sec. 125, Course of Theoretical Physics, pt. 1.
37. Serdyukov, A. N., Semchenko, I. V., Tretyakov, S. A. and Sihvola, A. (2001). Electromagnetics of bi-anisotropic materials: Theory and Applications, Amsterdam: Gordon and Breach Science Publishers.
38. Yatsenko, V. V., Maslovski, S. I., Tretyakov, S. A., Prosvirnin, S. L. and Zouhdi, S. (2003). Plane-wave reflection from double arrays of small magnetoelectric scatterers, *IEEE Trans. Antenn. Propag.* **51**(1), pp. 2–11.
39. Simovski, C. R., Tretyakov, S. A., Sochava, A. A., Sauviac, B., Mariotte, F. and Kharina, T. G. (1997). Antenna model for conductive omega particles, *J. Electromagnet. Waves Appl.* **11**(11), pp. 1509–1530.
40. Asadchy, V. S., Faniayeu, I. A., Radi, Y., Khakhomov, S. A., Semchenko, I. V. and Tretyakov, S. A. (2015). Broadband reflectionless metasheets: Frequency-selective transmission and perfect absorption, *Phys. Rev. X* **5**(3), p. 031005.
41. Radi, Y., Simovski, C. R. and Tretyakov, S. A. (2015). Thin perfect absorbers for electromagnetic waves: Theory, design, and realizations, *Phys. Rev. Appl.* **3**(3), p. 037001.
42. Kodera, T., Sounas, D. L. and Caloz, C. (2013). Magnetless nonreciprocal metamaterial (MNM) technology: Application to microwave components, *IEEE Trans. Microwave Theory Techn.* **61**(3), pp. 1030–1042.
43. Hadad, Y., Sounas, D. L. and Alu, A. (2015). Space-time gradient metasurfaces, *Phys. Rev. B* **92**(10), p. 100304(R).
44. Semchenko, I. V., Khakhomov, S. A. and Samofalov, A. L. (2010). Helices of optimal shape for nonreflecting covering, *Eur. Phys. J. Appl. Phys.*, **49**(3), p. 33002.
45. Saenz, E., Semchenko, I., Khakhomov, S., Guven, K., Gonzalo, R., Ozbay, E. and Tretyakov, S. (2008). Modeling of spirals with equal dielectric, magnetic, and chiral susceptibilities, *Electromagnetics,* **28**(7), p. 476.
46. Semchenko, I. V., Khakhomov, S. A. and Samofalov, A. L. (2009). Optimal helix shape: Equality of dielectric, magnetic, and chiral susceptibilities, *Russ. Phys. J* **52**(5), p. 472.
47. Alitalo, P., Karilainen, A. O., Niemi, T., Simovski, C. R. and Tretyakov, S. A. (2011). Design and realisation of an electrically small Huygens source for circular polarisation, *IET Microw. Antenn. Propag.* **5**(7), pp. 783–789.

48. Tretyakov, S. A., Mariotte, F., Simovski, C. R., Kharina, T. G. and Heliot, J.-P. (1996). Analytical antenna model for chiral scatterers: Comparison with numerical and experimental data, *IEEE Trans. Antenn. Propag.* **44**(7), pp. 1006–1014.
49. Asadchy, V., Albooyeh, M. and Tretyakov, S. (2016). Optical metamirror: All-dielectric frequency-selective mirror with fully controllable reflection phase, *J. Opt. Soc. America B* **33**(2), pp. A16–A20.
50. Zhaoa, Y., Engheta, N. and Alù, A. (2011). Homogenization of plasmonic metasurfaces modeled as transmission-line loads, *Metamaterials,* **5**(2), pp. 90–96.
51. Albooyeh, M., Tretyakov, S. and Simovski, C. (2016). Electromagnetic characterization of bianisotropic metasurfaces on refractive substrates: General theoretical framework, *Annalen der Physik,* **528**(9–10), pp. 721–737.
52. Achouri, K., Khan, B. A., Gupta, S., Lavigne, G., Salem, M. A. and Caloz, C. (2015). Synthesis of electromagnetic metasurfaces: Principles and illustrations, *EPJ Appl. Metamat.* **2**(12), pp. 1–11.
53. Pfeiffer, C. and Grbic, A. (2014). Bianisotropic metasurfaces for optimal polarization control: Analysis and synthesis, *Phys. Rev. Appl.* **2**(4), p. 044011.
54. Pfeiffer C. and Grbic, A. (2013). Cascaded metasurfaces for complete phase and polarization control, *Appl. Phys. Lett.* **102**(23), p. 231116.
55. Wong, J. P. S., Selvanayagam, M. and Eleftheriades, G. V. (2014). Design of unit cells and demonstration of methods for synthesizing Huygens metasurfaces, *Photonics and Nanostructures — Fundament. Appl.* **12**(4), pp. 360–375.
56. Epstein A. and Eleftheriades, G. V. (2014). Passive lossless Huygens metasurfaces for conversion of arbitrary source field to directive radiation, *IEEE Trans. Antenn. Propag.* **62**(11), pp. 5680–5695.
57. Asadchy, V. S., Albooyeh, M., Tcvetkova, S. N., Díaz-Rubio, A., Ra'di, Y. and Tretyakov, S. A. (2016). Perfect control of reflection and refraction using spatially dispersive metasurfaces, *Phys. Rev. B* **94**(7), p. 075142.
58. Wong, J., Epstein, A. and Eleftheriades, G. (2016). Reflectionless wide-angle refracting metasurfaces, *IEEE Antenn. Wireless Propag. Lett.* **15**, p. 1293.
59. Estakhri N. M. and Alù, A. (2016). Wavefront transformation with gradient metasurfaces, *Phys. Rev. X* **6**(4), 041008.
60. Epstein, A. and Eleftheriades, G. V. (2016). Arbitrary power-conserving field transformations with passive lossless Omega-type bianisotropic metasurfaces, *IEEE Trans. Antenn. Propag.* **64**(9), pp. 3880–3895.
61. Sievenpiper, D., Zhang, L., Broas, R. F. J., Alexópolous, N. G. and Yablonovitch, E. (1999). High-impedance electromagnetic surfaces with a forbidden frequency band, *IEEE Trans. Microw. Theory Tech.* **47**(11), p. 2059.
62. Díaz-Rubio, A., Asadchy, V., Elsakka, A. and Tretyakov, S. (2016). From the generalized reflection law to the realization of perfect anomalous reflectors, Preprint arxiv.org/abs/1609.08041.

CHAPTER 7

Scattering at the Extreme with Metamaterials and Plasmonics

FRANCESCO MONTICONE

Cornell University, USA

ANDREA ALÙ*

The University of Texas at Austin, USA

The ability to engineer and control the electromagnetic scattering from material bodies is of great importance in modern science and technology. In this chapter, we review some of the most exciting recent advances in this topic, enabled or inspired by metamaterials and plasmonics. We discuss, from a fundamental perspective, how to drastically suppress or enhance the scattering cross-section of a given object, as well as how to increase the lifetime of scattering resonances and the directivity of the scattering pattern. The possibility to control scattering processes "at the extreme" with metamaterials may find application in many diverse practical scenarios, including cloaking and invisibility, light trapping, energy harvesting, biochemical sensing and enhanced light–matter interaction at the micro- and nanoscale.

*Corresponding author: alu@mail.utexas.edu

7.1. Introduction

Scattering processes underpin several phenomena we are familiar with from our everyday experience, such as the distortion of sea waves washing against rocks, the dynamics of billiard balls and the color of the sky. Our very ability to see objects is largely determined by the light scattered from them that reaches our eyes. We therefore have a strong intuitive notion of what natural, or conventional, scattering looks like. Wouldn't it be surprising if an object became invisible before our eyes, or if the setting sun appeared blue instead of red?

Interestingly, with the advent of metamaterials in the past 15 years,[1] several anomalous scattering effects have indeed been demonstrated, including invisibility, superscattering, unidirectional scattering, light trapping, etc., which are not commonly observed in nature and are not part of our everyday experience. Metamaterials have therefore opened a wide range of new possibilities for manipulating and controlling the scattering of various types of waves (acoustic, electromagnetic, matter waves, etc.) in different frequency regimes. In this chapter, we will focus on classical scattering of electromagnetic waves without change of frequency during the scattering process. However, given the generality of elastic scattering processes in wave physics, some of the results discussed here may also be extended to other domains, e.g. acoustic and particle scattering.

The main goal of the chapter is to discuss to what extent we can control and engineer the electromagnetic scattering of material bodies, with particular emphasis on *extreme* scattering phenomena. In the spirit of a "handbook" on metamaterials, we will provide the reader with the fundamental physical insights into the different techniques available today to suppress, enhance and generally manipulate the scattering from a given object. In this context, particular attention will be devoted to fundamental limitations of different scattering processes, such as invisibility or scattering resonances. Such theoretical limitations are clearly of large significance from both the fundamental and practical standpoint, as they provide deep physical insights into the considered scattering effects and allow for quantitatively assessing their potential for practical applications.

In Section 1.2, we will briefly discuss some general properties of the scattering problem with particular emphasis on fundamental passivity and causality considerations. Sections 1.3 and 1.4 will be devoted, respectively, to the important topics of scattering suppression (cloaking and invisibility) and enhancement (scattering resonances and superscattering). The intriguing possibility of ideally trapping light in an open scatterer will be discussed in detail in Section 1.5. Finally, we will consider in Section 1.6 how to increase the directivity of the scattering pattern, and the related problem of enhancing the magnetic response at optical frequencies, by exploiting plasmonic and dielectric nanoparticles.

7.2. General Properties of Scattering Systems: Scattering Coefficients, Power Conservation and Causality

The typical benchmark example for scattering engineering is represented by spherically symmetric objects illuminated by a propagating plane wave. For this case, the scattering problem can be solved exactly using Mie theory.[2] The fields everywhere are expanded as a weighted sum of vector spherical harmonics, each one associated with a different angular momentum. Due to Noether's theorem of conservation laws,[3] the spherical symmetry directly implies that the angular momentum is conserved during the scattering process, which allows treating each spherical harmonic independently. To solve the scattering problem, we impose electromagnetic boundary conditions on the different spherical interfaces and solve the resulting system of equations, obtaining the weight coefficients of the scattered spherical waves for both transverse electric (TE) and transverse magnetic (TM) polarization (transverse with respect to the radial direction). These "scattering coefficients" can conveniently be written as[a]

$$c_n = -\frac{U_n}{U_n + iV_n}, \tag{7.1}$$

[a]Throughout this chapter, we assume a time-harmonic convention $e^{-i\omega t}$.

where the quantities U_n and V_n, purely real in the lossless case, can be written as matrix determinants associated to the boundary-condition system of equations (see, e.g. Ref. 4).

A one-to-one correspondence exists between the TE and TM scattering coefficients and the multipole moments in a multipole expansion of the current distribution induced in a given scatterer.[5,6] For example, the first TM spherical harmonic corresponds to the field radiated by the induced electric dipole moment $\mathbf{p}$, which is of particular importance as it represents the dominant scattering contribution for electrically small objects. If we compare the power carried by the first TM spherical harmonic with the power radiated by an electric dipole in free space, and we relate the electric dipole moment to the local electric field $\mathbf{E}_{\text{loc}}$ (in the absence of the scatterer) through a scalar polarizability α_e, i.e. $\mathbf{p} = \alpha_e \mathbf{E}_{\text{loc}}$, then we can easily find a relation between the scattering coefficient c_1^{TM} and the polarizability α_e

$$c_1^{\text{TM}} = \frac{ik_b^3 \alpha_e}{6\pi\varepsilon_b}, \tag{7.2}$$

where k_b and ε_b are the wavenumber and permittivity of the surrounding medium, respectively. Similar considerations apply to all the scattering coefficients: for example, the first TE coefficient corresponds to the magnetic dipole moment and the second TM coefficient to the electric quadrupole moment. However, the definition of higher-order polarizabilities becomes more difficult (and less useful), since the higher-order moments are not proportional to the local field $\mathbf{E}_{\text{loc}}$ but to the field gradients (e.g. magnetic dipole and electric quadrupole are proportional, respectively, to the antisymmetric and symmetric parts of the field gradient tensor).

By combining Eqs. (7.1) and (7.2), we can also write the electric polarizability in a particularly convenient form

$$\alpha_e = \left(\frac{V_1^{\text{TM}}}{U_1^{\text{TM}}} \frac{k_b^3}{6\pi\varepsilon_b} - i \frac{k_b^3}{6\pi\varepsilon_b} \right)^{-1}, \tag{7.3}$$

which expresses, in a compact way, all the available degrees of freedom to modify the induced dipole moment, and thereby the dipolar

scattering. In particular, in the case of a lossless scatterer, for which V_1^{TM} and U_1^{TM} are real, the imaginary part of α_e^{-1} uniquely depends on the properties of the surrounding medium. In fact, this term corresponds to the imaginary part of the free-space Green's function at the location of the particle or, in other words, to the local photonic density of states (LDOS).[7]

In conventional, quasi-static approximations of the polarizability, this imaginary term is often neglected, since it tends to be much smaller than the real part. This approximation, however, needs to be performed with care, since it implies neglecting the small energy associated with scattering. A lossless scatterer would not be able to radiate power without extracting it from the incident field, which requires an imaginary term in the polarizability. In fact, from Poynting's theorem, the time-averaged power extracted from the incident field by a material body with induced current $\mathbf{J}$ is[6] (see Ref. 6)

$$P_{\mathrm{ext}} = \frac{1}{2}\mathrm{Re}\left[\int_V \mathbf{J}^* \cdot \mathbf{E} dV\right], \tag{7.4}$$

where the integral is taken over the volume V of the body. If the scatterer can be approximated as a point electric dipole, then the extracted power can be written as

$$P_{\mathrm{ext}} = \frac{\omega}{2}\mathrm{Im}[\alpha_e]|\mathbf{E}_{\mathrm{loc}}|^2. \tag{7.5}$$

Passivity requires that the power extracted by the scatterer is positive (namely, the scatterer is not an independent power source), which directly implies that

$$\mathrm{Im}[\alpha_e] \geq 0. \tag{7.6}$$

In addition, by requiring that the power extracted from the incident field (7.5) is larger than the power re-radiated by the induced dipole, a more stringent passivity condition can be established on the *inverse* polarizability, known as the Sipe–Kranendonk condition[8]:

$$\mathrm{Im}[\alpha_e^{-1}] \leq -\frac{k_b^3}{6\pi\varepsilon_b}, \tag{7.7}$$

in which the equality sign holds for lossless scatterers (i.e. no material absorption).

Because of these considerations, only the first term in Eq. (7.3) can be actually manipulated to engineer the dipolar scattering of an object, whereas the second term can only be modified by changing the surrounding environment. In particular, by modifying the ratio $V_1^{\mathrm{TM}}/U_1^{\mathrm{TM}}$, which contains the geometrical and material properties of the scatterer, it is possible to either suppress or enhance the scattered energy. It should be stressed, however, that although the ratio $V_1^{\mathrm{TM}}/U_1^{\mathrm{TM}}$ can vary from zero to infinity, if the passivity condition (7.7) is respected, the range of possible values that the polarizability can assume is always finite, namely,

$$|\alpha_e| \in \left[0, \frac{6\pi\varepsilon_b}{k_b^3}\right], \tag{7.8}$$

where the upper bound is reached in the case of a lossless dipolar resonance (maximum power is extracted by the impinging wave), as we will further discuss in Section 1.4. Analogous considerations also apply to higher-order multipole moments, for which a passivity condition similar to (7.7) can be derived (see, e.g. Ref. 9 for the electric quadrupole moment). In general, passivity implies that the scattering coefficients, defined as in (7.1), have amplitude bound between zero and one, similar to the case of the reflection coefficient of a passive structure.

Scattering and reflection coefficients indeed have several similarities, as they both represent the coupling to "radiation channels" that extract power from the incident wave. However, it is important to recognize that the scattering coefficients defined above are not equal to conventional reflection coefficients in a spherical transmission-line model of the scatterer.[10] In particular, for a given spherical wave impinging over a lossless, spherically symmetric object, the spherical reflection coefficient s_n always has unitary amplitude, since all the energy of the spherical wave necessarily "bounces back" from the origin of the spherical coordinate system. A scattering coefficient c_n can be related to the corresponding spherical reflection

coefficient s_n by

$$c_n = (s_n e^{-2ik_0 a} - 1)/2, \tag{7.9}$$

where a indicates the maximum radius of the spherical scatterer. The scattering coefficient given by (7.9) indeed has magnitude between zero and one, related to the amount of phase delay that the scatterer introduces on the reflected spherical wave (c_n is equal to zero if the outgoing spherical wave has exactly the same phase as without scatterer). The two definitions of scattering coefficients, (7.1) and (1.9), are perfectly equivalent.

It is also worth mentioning that the conventional definition of scattering coefficients, although physically meaningful, inevitably leads to important non-causality issues, as discussed in Refs. 11 and 12. In fact, in the presence of a scatterer, the interaction of the incident wave with the object is expected to start when the wave hits its surface, not its center. On the other hand, each spherical harmonic is defined with respect to the center of the scatterer, and a causal response for each scattering coefficient would require the scattered wave to emerge only when the incident wave reaches the origin. As a result, individual scattering coefficients and polarizabilities are inherently non-causal functions characterized by a transient response starting at a negative time instant, as it can easily be verified by inverse-Fourier transforming them.[11,b] Only the sum of all scattering coefficients can be rigorously associated with a causal function, namely, the total forward scattering.[2] Although these considerations do not affect our following discussion about scattering suppression and enhancement, they become particularly important when trying to derive Kramers–Kronig relations and sum rules on the dynamic polarizabilities, or in the context of metamaterial homogenization, as discussed in Ref. 11.

After having briefly summarized here some general properties of scattering systems, in the next sections we will discuss a few specific

[b]Note also that it is not possible to enforce causality by simply introducing a time delay, i.e. by phase shifting c_n, as this would lead to a violation of the Sipe–Kranendonk passivity condition (7.7).

directions for *scattering engineering*, such as scattering suppression and cloaking, resonant scattering, superscattering, light trapping, directional scattering, etc.

7.3. Scattering Suppression, Cloaking and Invisibility

Invisibility devices, or cloaks, are arguably the quintessential example of how metamaterials can control and engineer scattering in unconventional ways. Several techniques are today available to suppress the scattering of given objects, each one of them with its own benefits and limitations, and we refer the interested reader to Ref. 13 for a broad literature review of different cloaking schemes. Here, instead, we focus on cloaking from the point of view of scattering cancellation, as in *plasmonic* or *mantle cloaking*, and, from this perspective, we discuss a few general results of relevance for any cloaking device.

We consider here a quantitative definition of cloaking, as the ability to restore the impinging field distribution, *in amplitude and phase*, all around the object to be concealed. In other words, an object is cloaked if the difference between the total field $\mathbf{E}_{\text{tot}}$ and the incident field $\mathbf{E}_{\text{inc}}$, namely, the scattered field, is minimized:

$$\int_{\infty/\text{object}} |\mathbf{E}_{\text{inc}} - \mathbf{E}_{\text{tot}}| \mathrm{d}V \approx 0, \tag{7.10}$$

where ∞/object indicates the entire space except the cloaked object. This corresponds to minimizing the total scattering cross-section of the object under consideration. We therefore exclude from this definition of cloaking other invisibility techniques that aim at restoring only the amplitude of the incident field, e.g. Refs. 14 and 15, or that works only for specific directions of incidence and observation, e.g. Refs. 16–18.

The total scattering cross-section is defined as a weighted sum of the amplitude of the scattering coefficients, each one being proportional to the power carried by the corresponding TE or TM spherical harmonic[2]:

$$\text{SCS} = \frac{\lambda_0^2}{2\pi} \sum_{n=1}^{N} (2n+1)(|c_n^{\text{TM}}|^2 + |c_n^{\text{TE}}|^2), \tag{7.11}$$

where N is the number of non-negligible scattering coefficients at the free-space wavelength λ_0 (this number is typically of the same order as $k_0 a$). In addition, for a lossless scatterer, the total scattering cross-section is equal to the total extinction cross-section, which represents the total amount of power extracted from the incident wave, either absorbed or re-radiated by the scatterer. Thanks to a fundamental theorem of scattering theory, known as the *optical theorem*,[c] the scattering cross-section is therefore directly proportional to the total forward scattering $s(0)$ (i.e. the "shadow" casted by the object)[6]:

$$\mathrm{SCS} = \frac{4\pi}{k_0^2}\mathrm{Im}[s(0)]. \tag{7.12}$$

According to the above definition of cloaking, a lossless object becomes invisible when all the non-negligible scattering coefficients are suppressed. If, for example, the scatterer is small compared to the incident wavelength, the dominant scattering mechanism is the net separation of (bound or free) charges that determines an induced electric dipole moment re-radiating energy into free space. In this quasi-static scenario, therefore, only the coefficient c_1^{TM} needs to be minimized, which can be done, for example, with the well-established cloaking techniques based on scattering-cancellation.[19–27] Such techniques are essentially based on surrounding the object to be concealed with a suitably designed shell, whose scattering compensates the original scattering of the object. In other words, as sketched in Fig. 7.1(a), both the shell and the object scatter when considered individually; however, since their induced dipole moments are exactly opposite, when they are combined the total dipolar scattering gets efficiently suppressed.[d] According to Eq. (7.1), the condition for cloaking is given by $U_1^{\mathrm{TM}} = 0$, which is a transcendental equation

[c]Equation (7.12) is a particular form of the optical theorem. According to how the forward scattering is defined, the optical theorem may be written in slightly different forms (for example, in Ref. 2, the real part of the forward scattering is taken, instead of the imaginary part).

[d]This heuristic explanation of the cloaking mechanism is consistent with the first-order Born approximation for low-scattering systems.[3,30]

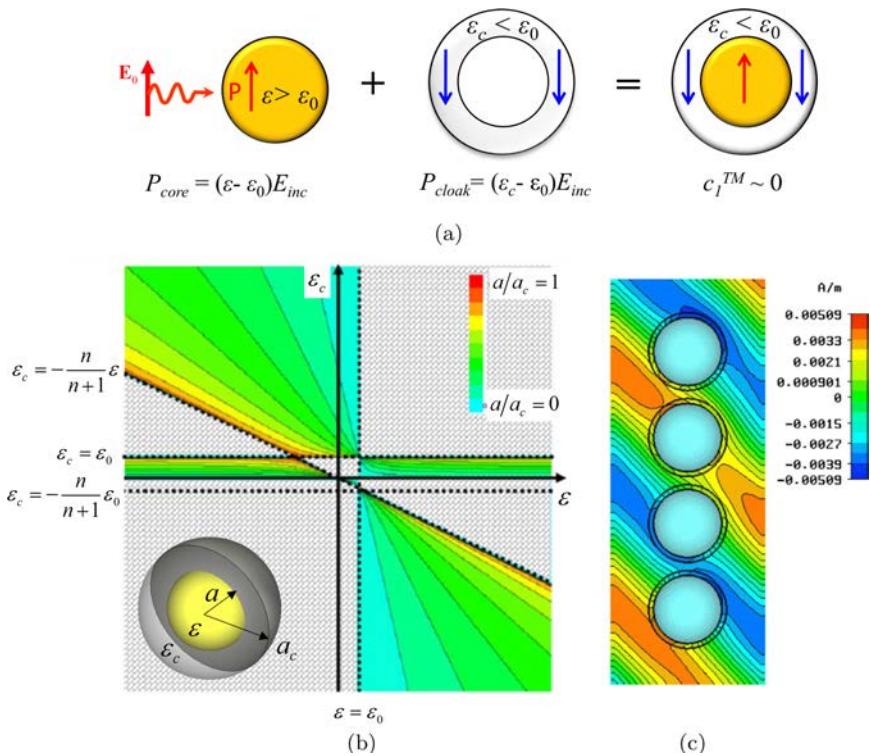


Fig. 7.1. (a) Working principle of plasmonic cloaking: The induced dipole moment in the cloak compensates the dipole moment in the object to be concealed, hence minimizing the overall dipolar scattering of the cloaked object. (b) Design chart for plasmonic cloaking of subwavelength objects: the map shows the range of permittivities for which the cloaking condition $U_n^{\mathrm{TM}} = 0$ admits physical solutions, and the corresponding aspect ratio a/a. ©2005 APS. Adapted with permission from Ref. 19. (c) Example of plasmonic cloaking of a collection of impenetrable objects. ©2007 OSA Publishing. Adapted with permission from Ref. 20.

without closed-form solution in the general case. However, for an isotropic core–shell spherical geometry in quasi-static regime, this condition can be greatly simplified,[19] obtaining

$$\left(\frac{a}{a_c}\right)^3 = \frac{(\varepsilon_c - \varepsilon_0)(2\varepsilon_c + \varepsilon)}{(\varepsilon_c - \varepsilon)(2\varepsilon_c + \varepsilon_0)}, \tag{7.13}$$

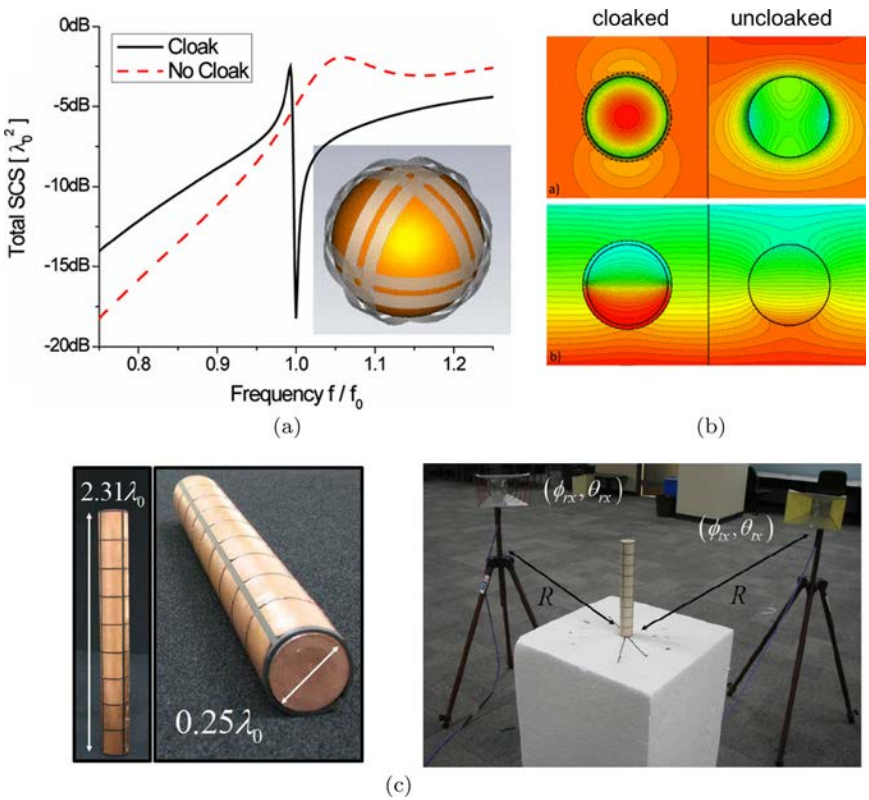


Fig. 7.2. Mantle cloaking. (a) Scattering performance of a mantle cloak applied to a dielectric sphere (in this case the cloak impedance is inductive). At the design frequency, the total SCS is strongly suppressed. (b) Near-field distribution of the amplitude of the electric field (top row) and phase of the magnetic field (bottom), for a cloaked (left) and uncloaked (right) object. ©2009 APS. Adapted with permission from Ref. 24. (c) Examples of fabricated cylindrical mantle cloak for operation at microwaves. ©2013 IOP Science. Adapted with permission from Refs. 34, and 35.

where a and a_c are the radii of the core and shell, respectively, and ε and ε_c are their permittivities. The design map in Fig. 7.1(b) shows the range of permittivity values for which Eq. (7.13) admits physical solutions (i.e. with aspect ratio $0 < a/a_c < 1$). By inspecting Fig. 7.2(b), we see that, if the core is dielectric ($\varepsilon > \varepsilon_0$), the cloaking condition (7.13) requires the permittivity of the shell to be

smaller than that of free space, i.e. $\varepsilon_c < \varepsilon_0$. This way, the polarization density of the core is compensated by the opposite polarization density of the cloak, resulting in a reduced overall polarizability of the cloaked object and, thereby, reduced dipolar scattering, consistent with Fig. 7.1(a). Note that, different from other cloaking techniques, here the original object is still polarized by the incident field and the internal fields are non-zero. The induced fields in the cloaked object can actually be very high, while the object is still invisible, an ideal condition to design cloaked sensors and low-invasive near-field probes.[28,29] Plasmonic cloaking can also be used to reduce the scattering of impenetrable objects and of large collections of scatterers, as seen in Fig. 7.1(c).

A permittivity smaller than that of free-space is naturally available in plasmonic materials at optical frequencies, such as gold and silver in the visible range, aluminum in the near UV and doped semiconductors in the infrared.[31] At lower frequencies, in the radio frequency (RF) and microwave range, a plasmonic cloak can be effectively emulated with engineered metamaterials.[21–23] In many practical cases, however, plasmonic cloaking may be challenging because the required material properties may be difficult to achieve at the frequency of interest. Besides, the effect of losses is often significant since this cloaking technique is based on the volumetric properties of the shell. An interesting alternative is represented by the technique of mantle cloaking, which is based on the wave interaction over a metasurface cloak with negligible thickness,[24–27] as shown in Fig. 7.2(a). The physical mechanism of mantle cloaking is also based on the concept of scattering cancellation: the currents induced on a patterned metallic shell re-radiate dipolar fields that compensate the scattering of the original subwavelength object. In particular, the cloaking effect is controlled by the surface impedance of the cloak, $Z_s = R_s - iX_s$, which relates the induced current to the tangential electric field, i.e. $\mathbf{E}_{\tan} = Z_s\mathbf{J}$.[e] As mentioned above, the condition for

[e]This can readily be extended to the case of anisotropic surfaces with tensorial surface impedance.

cloaking, $U_1^{\mathrm{TM}} = 0$, can be written in closed form in the quasi-static limit, obtaining[24]

$$Z_s = -i\frac{2[2+\varepsilon-\gamma^3(\varepsilon-1)]}{3\gamma^3\omega a\varepsilon_0(\varepsilon-1)}, \tag{7.14}$$

where $\gamma = a/a_c$. For a small dielectric sphere, the required surface impedance given by (7.14) is typically inductive ($X_s > 0$), whereas for a metallic sphere the cloak needs to be capacitive ($X_s < 0$). Interestingly, this can be interpreted in light of the *optical nanocircuit paradigm*, put forward in Ref. 32: a small dielectric sphere corresponds to an optical capacitance, which therefore can be compensated by introducing an inductance. As shown in Fig. 7.2(b), despite its ultrathin profile, a mantle cloak is indeed able to restore the incident field distribution, in amplitude and phase, all around the cloaked object (besides, notice that the field intensity inside the cloaked object is actually enhanced with respect to the uncloaked sphere).

The required surface impedance can be practically implemented by patterning a thin metallic shell, realizing arrays of patches, strips, crosses, etc., based on the well-established techniques of frequency-selective surfaces at RF and microwaves,[33] which make mantle cloaking particularly appealing in this frequency range. An example of fabricated cylindrical mantle cloak for operation at RF is shown in Fig. 7.2(c).[34,35] In addition, mantle cloaks made of graphene have been proposed for operation in the THz range, realizing the thinnest possible cloak,[36] and superconducting media have been considered for the possibility of realizing homogeneous mantle cloaks at microwaves, in which the required inductive surface impedance is directly provided by the high kinetic inductance of the supercurrent, without the need of patterning the surface.[37]

Interestingly, although scattering cancellation techniques are usually applied to subwavelength objects, moderate scattering suppression can also be achieved for larger objects (in the order of a few free-space wavelengths) by suitably optimizing the plasmonic shell.[38] Besides, by considering multiple concentric shells it may

be possible to realize multiband cloaks,[39] or moderately broadband cloaks in analogy with multisectors impedance-matching filters (e.g. Chebyshev, or Butterworth filters[40]).

In principle, there is no fundamental limit on the number of scattering coefficients that can be simultaneously suppressed at a given frequency. In fact, an ideal transformation-optics cloak[41,42] automatically aligns, at the desired frequency, the conditions $U_n = 0$ for any non-negligible scattering order. This cloaking technique is based on a completely different mechanism than scattering-cancellation cloaking, as it requires the cloak to be designed with strongly anisotropic and inhomogeneous permittivity and permeability, such that the incident wave is guided around a region of space, where the scatterer is hidden, without interacting with it. For further details on this and other cloaking techniques, we refer the reader to Ref. 13. The price to be paid for having an ideal invisible state is the high complexity of the cloaking design and its high sensitivity to losses and imperfections. In addition, there are other, more fundamental problems that limit the possibility of suppressing the scattering of a material body. In fact, relevant physical bounds on cloaking and invisibility can be directly deduced from passivity and causality considerations on the generic scattering system.[44,127] For example, since the total forward scattering $s(0)$ is a causal function as mentioned above, we can derive Kramers–Kronig relations on its real and imaginary parts (actually on a slightly modified version of $s(0)$, as discussed in Ref. 2). By further invoking power conservation, represented by the optical theorem (7.12), and after some manipulation, we obtain an important sum rule that relates the SCS of an object, integrated over the entire electromagnetic spectrum, to its electric and magnetic static polarizability tensors, $\boldsymbol{\alpha}_{\mathbf{e,s}}$ and $\boldsymbol{\alpha}_{m,\mathbf{s}}$. In the most general form, the sum rule reads[43]

$$\int_0^\infty C_s(\lambda)\mathrm{d}\lambda = \pi^2(\hat{p}_e^* \cdot \boldsymbol{\alpha}_{\mathbf{e,s}} \cdot \hat{p}_e + \hat{p}_m^* \cdot \boldsymbol{\alpha}_{\mathbf{m,s}} \cdot \hat{p}_m), \tag{7.15}$$

where $\hat{p}_e$ denotes the polarization and $\hat{p}_m$ the cross-polarization unit vectors. By further noting that the static polarizabilities are always

monotonically related to the static permittivity and permeability[44] (if the static permittivity/permeability grows in a region of space, also the electric/magnetic polarizability of the object is bound to grow), we can conclude that whenever a linear and passive cloak is wrapped around a given object, the static polarizability of the cloaked object necessarily increases, which also determines, according to (7.15), an increase of the integrated scattering. As a result, scattering suppression and invisibility in a given frequency window is always compensated by enhanced scattering in a different region of the electromagnetic spectrum (an interesting exception to this rule is represented by cloaks that include diamagnetic or superconducting media, as discussed in detail in Refs. 37, and 44). In this sense, cloaks that introduce less additional matter, such as mantle cloaks, are less prone to increase the integrated scattering. This is confirmed by comparing the typical performance of plasmonic, mantle and transformation-optics cloaks, applied to the same object, as shown in Fig. 7.3.

We also stress that due to causality considerations,[45] ideal cloaking (identically zero scattering) is only possible at discrete frequency points, namely, over a zero-measure continuous cloaking bandwidth. Instead, if larger scattering is tolerated (i.e. imperfect cloaking), the cloaking bandwidth can indeed be widened,[46,47,127] in strong analogy with the problem of impedance matching and reflection suppression in antenna and microwave-network theory.[40]

7.4. Scattering Resonances, Fano Resonances, Superscattering

In the previous section, we have considered the case of scattering suppression, namely, we investigated the condition $U_n = 0$ in Eq. (7.1). The opposite scenario is represented by *resonant scattering*, i.e. maximum scattering enhancement for a given spherical harmonic. This case corresponds to the condition $V_n = 0$ in (7.1), which implies that $|c_n| = 1$ and maximum energy is coupled from the incident wave into the n-th scattering channel.

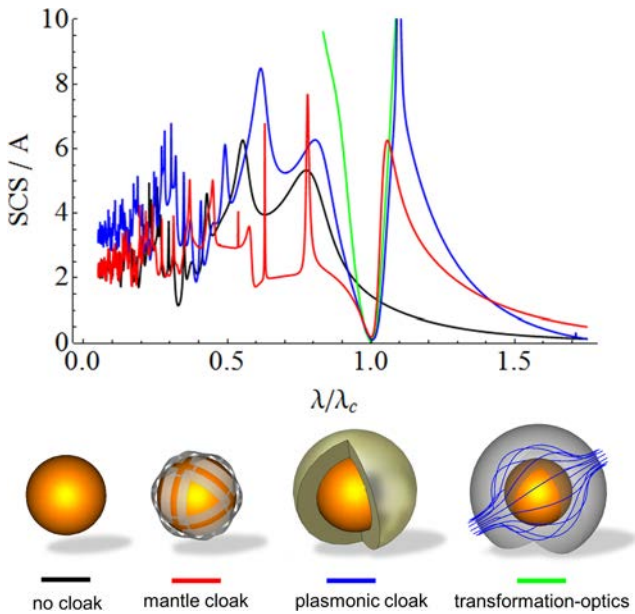


Fig. 7.3. Comparison between the typical performance (in terms of normalized scattering cross-section) of mantle, plasmonic and transformation-optics cloaks over a broad range of wavelengths. All the cloaks are designed to achieve invisibility at the wavelength λ_c and are applied to the same dielectric sphere with permittivity $\varepsilon_r = 5$ and diameter $d = \lambda_c/3$. ©2013 APS. Adapted with permission from Ref. 44.

In the case of electric dipolar scattering, the condition $V_1^{\mathrm{TM}} = 0$ implies that $\mathrm{Re}[\alpha_e^{-1}] = 0$ in Eq. (7.3). As a result, the electric dipole polarizability is purely imaginary at resonance, and its amplitude is at maximum, which guarantees maximal power extracted from the incident wave according to Eq. (7.5). As mentioned in Section 1.2, the polarizability and the radiated power at resonance are finite due to the Sipe–Kranendonk condition (7.7). Neglecting the radiation correction term in Eq. (7.3) would, in fact, lead to diverging scattered power, which clearly violates energy conservation. Besides, in the lossless limit, the value of the electric dipole polarizability and scattered power *at resonance* are fixed, no matter the shape and composition of the scatterer. The only way to enhance or suppress them is by modifying the surrounding environment (i.e. the optical LDOS

at the location of the scatterer), a phenomenon known as *Purcell effect* in the context of quantum emitters (the spontaneous emission rate, just as the classical radiation/scattering, is proportional to the LDOS[7]).

Since at resonance the electric dipole polarizability is purely imaginary, the local excitation field oscillates exactly in quadrature with the induced dipole moment $\mathbf{p}$, and exactly in phase with the polarization current $\mathbf{J} = -i\omega\mathbf{p}\delta(\mathbf{r})$ (where $\delta(\mathbf{r})$ is a delta function centered at the location of the object), in direct analogy with the operation of an RLC circuit, in which the voltage oscillates in phase with the current at resonance. A subwavelength dielectric scatterer is typically largely off-resonance, as the induced polarization current is out of phase with the excitation field, and thereby little power can be extracted from the incident field according to Eq. (7.4). This corresponds to the case of a non-resonant RC circuit, in which the voltage lags the current. Building on this analogy, in order to make the scatterer resonate, an inductive element should be introduced, which can be done by surrounding the small dielectric scatterer by a plasmonic shell, consistent with optical nanocircuit concepts.[35] Similar to the cloaking case, the condition for scattering resonance, $V_1^{\mathrm{TM}} = 0$, can be greatly simplified for an isotropic core–shell geometry in quasi-static regime, obtaining

$$\left(\frac{a}{a_c}\right)^3 = \frac{(2\varepsilon_0 + \varepsilon_c)(2\varepsilon_c + \varepsilon)}{2(\varepsilon_c - \varepsilon_0)(\varepsilon_c - \varepsilon)}, \tag{7.16}$$

where, again, a and a_c are the radii of the core and shell, respectively, and ε and ε_c are their permittivities. As in the cloaking case, a design map for scattering resonances is shown in Fig. 7.4, which indicates the range of permittivity values for which Eq. (7.16) admits physical solutions. In particular, if the core is dielectric ($\varepsilon > \varepsilon_0$), the scattering condition (7.16) requires the permittivity of the shell to be negative, i.e. $\varepsilon_c < 0$. This confirms that plasmonic materials offer an ideal solution to achieve scattering resonance over deeply subdiffractive footprints.

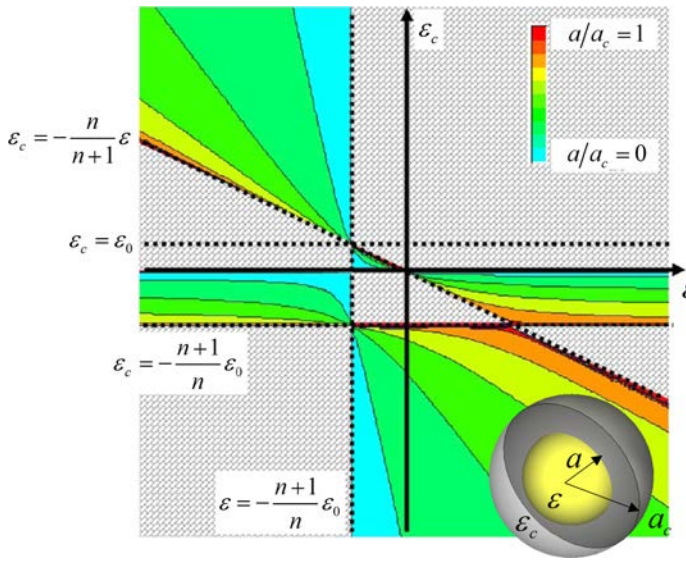


Fig. 7.4. Similar to Fig. 7.1(b), but for the case of subdiffractive scattering resonances. The map shows the range of permittivities for which the resonance condition $V_n^{\mathrm{TM}} = 0$ admits physical solutions, and the corresponding aspect ratio a/a_c. ©2005 AIP Publishing. Adapted with permission from Ref. 48.

At a scattering resonance, not only is the radiation at maximum, but also the fields induced inside the scatterer are enhanced. In the lossless limit, the ratio between stored energy in the scatterer (and in its reactive near field) and the radiated power is proportional to the resonance Q factor, and determines the linewidth of the resonance peak in the scattering spectrum. Interestingly, the Q factor of a dipolar scattering resonance has a lower bound, known as *Chu limit* in antenna theory[49]:

$$Q \geq \frac{1}{k_0 a} + \frac{1}{(k_0 a)^3}, \tag{7.17}$$

where a is the radius of the smallest spherical surface that completely encloses the radiator/scatterer. It is clear that the bound (7.17) becomes particularly stringent for small scatterers. In addition, the Q factor of any higher-order resonance is generally higher than that of a dipolar resonance. For example, for a small scatterer/radiator,

i.e. $k_0 a < 1$, the Q factor of the n-th TE or TM resonance can be estimated as[50]

$$Q \approx \frac{n\,(2n+1)!!^2}{(k_b a)^{2n+1}}, \tag{7.18}$$

where $(\cdot)!!$ indicates the double-factorial function.[51] It is therefore clear that the Q factor rapidly increases with the order n (note also that Eq. (7.18) is consistent with (7.17) for $n = 1$). In general, broadband resonant scattering/radiation, exceeding the Chu limit, can only be achieved by juxtaposing, at nearby frequencies, multiple resonances from different radiation modes,[52,53] or by using active and non-Foster structures, which allow achieving anomalously flat frequency dispersion.[54,55] Interestingly, while the Q factor of a scattering resonance is bounded from below by (7.17), an upper bound cannot be defined, meaning that the amount of stored energy can, in principle, diverge in the lossless limit, as we will see in Section 1.5.

We have discussed earlier that, when a certain scattering order resonates [i.e. $V_n = 0$ in Eq. (7.1)], maximum power is coupled to the corresponding radiation channel. In the lossless limit, this maximum power is a fixed quantity for each order n, and, according to Eq. (7.11), it contributes to the total SCS as $\lambda_0^2(2n+1)/2\pi$. Although this contribution is fixed for each n, there is no fundamental limit on the number of resonating scattering orders that can be "aligned" at the same frequency, hence achieving arbitrarily large scattering cross-sections. This "superscattering" effect has been demonstrated even in the quasi-static regime, by considering multi-layered nanoparticles with optimized plasmonic and dielectric layers.[56,57] In Fig. 7.5, we show an example of such a structure, whose SCS is indeed characterized by several scattering orders resonating at the same frequency, so that their resonant peaks overlap and "pile up" (Fig. 7.5(a)), determining much larger scattering than typically achieved with subwavelength particles, as seen in the field distribution in Fig. 7.5(b).

In this section and in Section 1.3, we have discussed how the scattering coefficients can be brought to zero or one by imposing

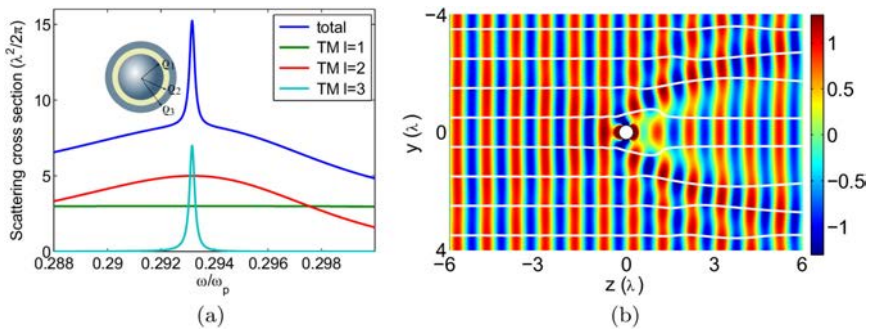


Fig. 7.5. Superscattering from subwavelength particles. (a) Normalized scattering cross-section from a layered nanosphere (shown in the inset, details in Ref. 57), in the lossless limit, and individual contributions from different scattering orders. (b) Time-snapshot of the electric field distribution, around the particle at resonance, under plane-wave illumination. The power flow (white stream lines) is strongly perturbed, particularly in the forward direction, as the scattering cross-section of the sphere is much larger than its geometrical cross-section. ©2013 AIP Publishing. Adapted with permission from Ref. 57.

$U_n = 0$ (cloaking) or $V_n = 0$ (scattering resonance) in Eq. (7.1). The dispersion of these two conditions determines therefore the distribution of scattering zeros and resonances along the real frequency axis. Interestingly, a few recent works have pointed out that, for subwavelength scatterers with purely electric dipolar response, the corresponding scattering coefficient c_1^{TM} exhibits a strict alternation of scattering zeros and resonances, indicating that the dispersion of $U_1^{\mathrm{TM}} = 0$ and $V_1^{\mathrm{TM}} = 0$ are necessarily interleaved.[58–60] In other words, two scattering zeros cannot exist without a scattering resonance between them, and vice-versa. Figure 7.6 shows an example of this behavior, for a subwavelength core–shell particle.[59] The dispersion of the cloaking (blue) and resonance (red) conditions is shown in Fig. 7.6(b) as a function of the cloak permittivity (or function of frequency, since the permittivity is assumed to follow Drude dispersion) and aspect ratio $\eta_c = a/a_c$, confirming that the different branches never cross. This intriguing phenomenon can be understood, at least qualitatively, from the point of view of circuit

theory. In fact, a subwavelength lossless particle is approximately equivalent to a *reactive* lumped-element circuit[32] (however, a small "radiation resistance" should always be added near a resonance to comply with passivity, consistent with the discussion above). Due to a relevant theorem of circuit theory known as *Foster's reactance theorem* (which is essentially a form of Kramers–Kronig relations for impedances), the reactive nanocircuit impedance monotonically increases with frequency and exhibits a strict alternation of zeros and poles (corresponding either to scattering zeros or resonances according to the circuit representation of the scatterer). Since in the quasi-static regime the scattering is proportional to this nanocircuit impedance, such an alternation is directly reflected in the scattering cross-section, hence explaining the distribution of cloaking and resonant states in Fig. 7.6(b). This "Foster's theorem for scattering" does not hold if the particle supports multiple scattering orders, or if it is not subwavelength.

The alternation of scattering zeros and resonances in subwavelength particles can be exploited to realize some intriguing scattering effects. In fact, as seen in Figs. 7.6(c) and 7.6(d), a particle could be designed such that a scattering zero lands close to a scattering resonance, hence obtaining an asymmetric scattering line-shape, known as Fano resonance,[61,62] which is usually interpreted as arising from the close coupling and interference between different modes supported by the scatterer. As seen in Fig. 7.2(a) and Figs. 7.6(b) and 7.6(c), the scattering spectrum of cloaked objects often exhibits this asymmetric shape, as the dipolar oscillations in the core and shell interfere destructively or constructively, determining a cloaking dip (dark state) or resonant peak (bright state), respectively. Based on this mechanism, multilayered plasmonic-dielectric particles can be designed to achieve anomalous scattering signatures,[58–60,63,64] including multifrequency cloaking,[58] electromagnetically induced transparency effects,[59] and comb-like scattering responses characterized by superscattering states alternated by invisible states.[60] Furthermore, the abrupt scattering variation of these Fano resonances,

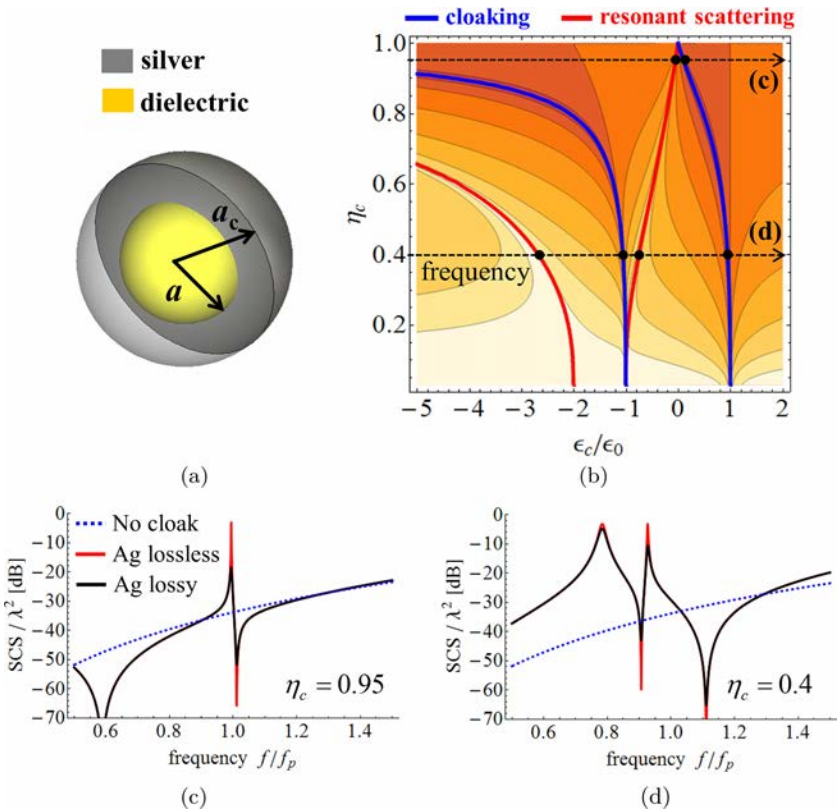


Fig. 7.6. Alternation of scattering zeros and resonances in subwavelength scatterers. (a) Spherical core–shell scatterer of subwavelength size (details in Ref. 59). (b) Dispersion of the conditions $U_1^{\mathrm{TM}} = 0$ (blue/darker curves, cloaking) and $V_1^{\mathrm{TM}} = 0$ (red/lighter curves, resonant scattering) for the particle in (a), as a function of geometrical aspect ratio $\eta_c = a/a_c$ and cloak permittivity ε_c. Since ε_c follows a Drude dispersion model, the horizontal axis also corresponds to frequency. The brighter (darker) colors of the background indicate regions of higher (lower) scattering cross-section. (c–d) Normalized SCS, as a function of frequency, for two different aspect ratios, corresponding to the two horizontal dashed lines in (b). ©2012 NPG. Adapted with permission from Ref. 59.

combined with the large field intensity induced inside the nanoparticle, represents an ideal condition to enhance and exploit weak nonlinear processes in order to realize, for example, all-optical scattering switches and nanomemories.[65]

7.5. Light Trapping in Open Structures

When the entire denominator of Eq. (7.1) goes to zero, i.e. $U_n + iV_n = 0$, the corresponding scattering coefficient and polarizability diverge. These *scattering poles*, which correspond to the eigenmodes, or self-sustained oscillations, of the scatterer, can occur only at complex frequencies, as required by passivity. The imaginary part of the complex eigenfrequency, in fact, is related to the oscillation damping of the corresponding eigenmode, which, in the limit of zero material absorption, is purely determined by radiation loss. The Q factor, or lifetime, of any scattering resonance is therefore determined by the "distance" of the corresponding pole from the real frequency axis. Note also that, if the radiation correction in Eq. (7.3), i.e. $-ik_b^3/6\pi\varepsilon_b$, was neglected, the scattering poles would be on the real frequency axis, exactly at the location of the scattering resonances ($V_n = 0$), which would violate passivity. Rather strikingly, however, it has been recently shown that an eigenmode can be "embedded" along the real frequency axis, leading to diverging oscillation lifetime and stored energy, without violating power conservation.[67,68] This intriguing and counterintuitive scattering effect is the topic of this section.

Consider the case of a layered sphere composed of a dielectric core, surrounded by two plasmonic shells with slightly different plasma frequencies, as it has been studied in.[68] Figure 7.7(a) shows the SCS of the particle as a function of wavelength and geometrical aspect ratio, similar to Fig. 7.6(b). The dark and bright bands follow the dispersion of the conditions $U_1^{\mathrm{TM}} = 0$ and $V_1^{\mathrm{TM}} = 0$ in Eq. (7.1), respectively, and, consistent with the discussion in the previous section, they strictly alternate since the scatterer is subwavelength. However, we see that in some regions two opposite bands merge, leading to a "degenerate state" in which $U_1^{\mathrm{TM}} = V_1^{\mathrm{TM}} = 0$. Near these points, therefore, a scattering pole moves arbitrarily close to the real axis, where it eventually "lands" exactly at the location of a zero, which cancels the pole, hence respecting passivity. This shows that, in the lossless limit, the Q factor and lifetime of a scattering resonance can be arbitrarily large, as seen in Fig. 7.7(b). Strikingly, this effect

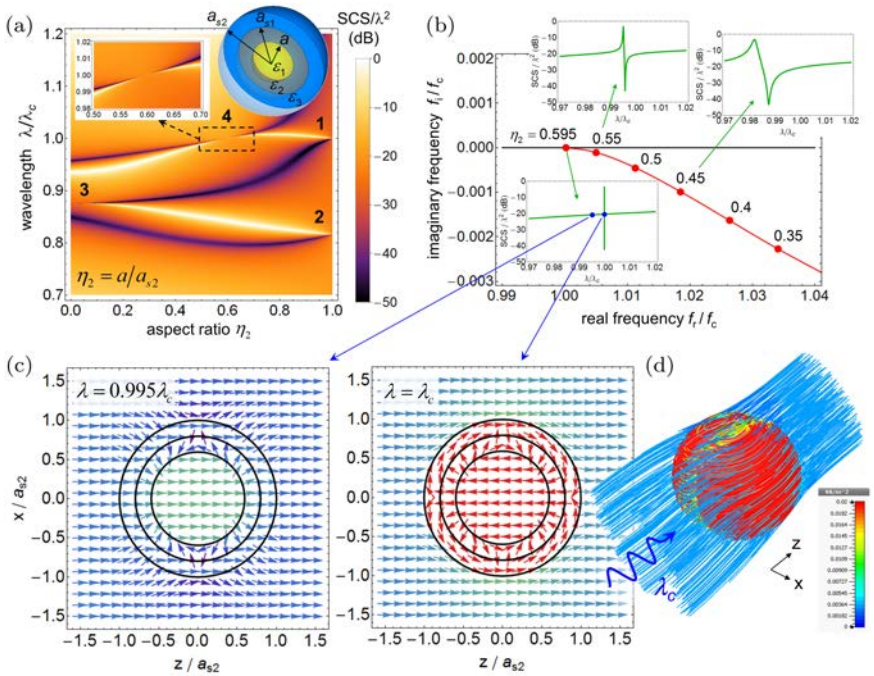


Fig. 7.7. Embedded scattering eigenstates in open 3D structures. (a) Scattering cross-section of a layered sphere, as shown in the inset (details in Ref. 68), as a function of wavelength and geometrical aspect ratio. Numbers denote the regions in which the Q factor of a scattering resonance diverges, as the dark and bright branches merge. (b) Evolution of a scattering pole on the complex frequency plane, as a function of the aspect ratio, and corresponding evolution of the SCS under plane-wave illumination. The energy leakage from the scatterer vanishes, as the pole gets closer and closer to the real axis. (c) Power flow distribution in the near field, for plane-wave excitation at a slightly off-resonance wavelength (left panel) and at the wavelength of the embedded eigenstate (right). (d) Corresponding 3D view of the power flow distribution at resonance. ©2015 APS. Adapted with permission from Ref. 68.

allows creating an ideally bound optical state, without radiation loss, supported by a three-dimensional (3D) *open* cavity. In fact, different from conventional light-confinement methods, which involve closing the cavity with metals or photonic crystals, such that the coupling to radiation modes is directly forbidden, here the cavity remains electromagnetically open, namely, radiation is in principle allowed, as the

photonic LDOS of the surrounding environment has not been modified (the radiation-correction term in the electric polarizability (7.3) is unaltered). This intriguing effect is the optical analog of so-called "embedded eigenstates" in certain quantum potentials[69] and different photonic implementations have been recently demonstrated, considering both unbounded two-dimensional structures[70,71,128] and bounded 3D geometries,[65,67] as in Fig. 7.7.

In the quasi-static limit, the equation $U_1^{\mathrm{TM}} = V_1^{\mathrm{TM}} = 0$ can be simplified, leading to different closed-form conditions to realize embedded photonic eigenstates in a spherical open resonator.[68] For example, the degenerate state "4" in Fig. 7.7(a) corresponds to the following conditions:

$$\begin{cases} \dfrac{{\eta_2}^3}{{\eta_1}^3} = \dfrac{2\varepsilon_2 + \varepsilon_1}{2\left(\varepsilon_2 - \varepsilon_1\right)} \\ \varepsilon_3 \to 0 \end{cases}, \tag{7.19}$$

where the ratios $\eta_1 = a_{c1}/a_{c2}$ and $\eta_2 = a/a_{c2}$ define the geometry, as indicated in the inset of Fig. 7.7(a). Equation (7.19) implies that a bound state with infinite lifetime is achieved when the permittivity of the outer shell ε_3 goes to zero and, at the same time, the two inner layers support a subdiffractive magnetic resonance (first condition in (7.19)), as discussed in Ref. 68. It is important to stress that the system is indeed electromagnetically open, as the fields inside the particle, under external excitation, are non-zero when Eq. (7.19) are ideally met, but radiation vanishes for a specific eigenmode of the scatterer.[f] In other words, conditions (7.19) correspond to a non-radiating current distribution, of purely dipolar nature, supported by the open structure.[67,68] This nicely connects the idea of embedded photonic eigenstates with the intriguing concept of non-radiating

[f]Note that the non-radiating states discussed here are fundamentally different from scattering dark states and cloaking states, which also imply a reduction of the re-radiated energy, but do not correspond to eigenmodes of the scattering system.

sources, which was important, for example, in the development of early classical atomic models.[72–74]

From an external excitation, we can pump energy into the embedded eigenstate, working arbitrarily close to the non-radiating condition (of course, due to reciprocity we cannot couple energy into an *ideally* non-radiating mode from external excitation). In the lossless limit, therefore, an arbitrarily large amount of energy can be stored in the particle. As seen in Figs. 7.7(c) and 7.7(d), the energy is indeed "trapped" in a self-sustained power flow, realizing a sort of optical vortex or whirlpool. If a small amount of ohmic losses is introduced, it can be shown that the stored energy rapidly drops when conditions (7.19) are exactly met,[68] because the optical vortices in the particle lead to enhanced absorption even for very small level of losses.

The structures discussed in this section can be implemented at optical frequencies with low-loss plasmonic materials, and at microwaves and RF with suitably designed metamaterials (e.g. Refs. 21–23), and may have applications for enhanced energy harvesting and absorption/emission. In the spirit of this chapter, the topic of embedded photonic eigenstates indeed represents a particularly illustrative case of "scattering at the extreme" enabled by plasmonics and metamaterials.

7.6. Huygens Sources, Directional Scattering, and Artificial Magnetism

In the previous sections, we have discussed electromagnetic scattering processes mostly from the point of view of the total scattering cross-section, for which all the different scattering orders contribute *incoherently*, i.e. only the amplitude, not the phase, of the individual scattering coefficients matters, consistent with Eq. (7.11). In addition, large emphasis has been placed on subwavelength particles with a dominant dipolar response, as they represent a fundamental building block of metamaterials and metasurfaces. However, scatterers supporting a single, electric or magnetic, dipolar mode have a mostly

isotropic scattering pattern (omnidirectional in the azimuthal direction, and proportional to $\cos(\theta)$ with respect to the elevation angle θ) with low directivity.[50] In order to make the scattering pattern more directive, higher-order multipolar scattering can be considered, as the directivity grows with the scattering order n as[50]

$$D = \frac{2n+1}{2}, \tag{7.20}$$

The multipolar radiation patterns for the first four orders at resonance are reported in Fig. 7.8(a), showing how the number of radiation beams increases, and the beam-width decreases, with the order n. Furthermore, the scattering in specific directions can be controlled by suitably combining different scattering orders. In fact, different from the total scattering cross-section, in specific directions the scattering depends on the phase of the different scattering coefficients, namely, the different scattering orders contribute *coherently*. For example, the backward scattering cross-section is given by[2]

$$\sigma_{\mathrm{bw}} = \frac{\lambda_0^2}{4\pi}\left|\sum_{n=1}^{N}(-1)^n(2n+1)(c_n^{\mathrm{TM}} - c_n^{\mathrm{TE}})\right|^2. \tag{7.21}$$

Therefore, by controlling the interference between TE and TM spherical harmonics, it is possible to largely suppress or enhance the backward scattering, hence modifying the shape and directivity of the radiation pattern. Of particular interest is the case of exactly zero backward scattering, obtained with $c_n^{\mathrm{TM}} = c_n^{\mathrm{TE}}$, so that TM and TE spherical waves interfere constructively in the forward direction and destructively in the backward direction, as predicted in a seminal paper by Kerker.[75] Such a radiating structure known as a *Huygens source*, is of particular importance in several practical scenarios, such as in antenna theory as it allows realizing electrically small antennas with moderately high directivity (see, e.g. Refs. 76, and 77, and the following discussion about optical nanoantennas), as well as in stealth technology, in which the ability to minimize the backward scattering (or "monostatic radar cross-section") is typically more important

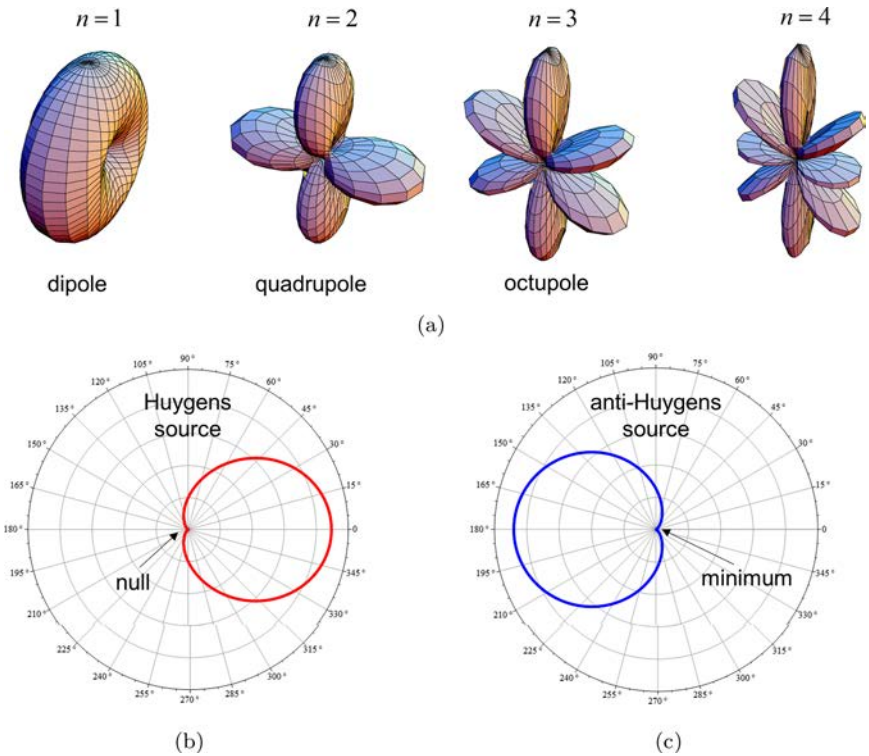


Fig. 7.8. (a) Scattering patterns of the first four (TE or TM) scattering orders at resonance (the electric or magnetic origin of these patterns influences their orientation, but not their shape). ©2007 IEEE. Reprinted with permission from Ref. 50. (b–c) Normalized scattering patterns of a subwavelength magneto-electric sphere (radius $a = \lambda_0/10$) with relative permittivity and permeability: (a) $\varepsilon_r = \mu_r = 3$, to fulfill the condition for a Huygens source, and (b) $\varepsilon_r = 3$ and $\mu_r = 0.14$ to fulfill condition (7.23) for an anti-Huygens source. The patterns are rotationally symmetric with respect to the horizontal axis. Backward scattering in (b) is identically zero, whereas the forward scattering in (c) has a residual imaginary component, consistent with the discussion in the text.

than suppressing the total scattering.[78] Figure 7.8(b) shows the typical radiation pattern of a Huygens sources with a marked null in the backward direction.

The concept of Huygens sources is also strongly related to the problem of *impedance matching*. To better appreciate this

connection, consider the case of a magneto-electric subwavelength sphere supporting only electric and magnetic dipolar responses. In this case, it is easy to show that the condition $c_1^{\mathrm{TM}} = c_1^{\mathrm{TE}}$ directly implies that the relative permittivity and permeability of the sphere need to become equal, i.e. $\varepsilon_r = \mu_r$.[75] As a result, the impedance of the material, $\eta_0\sqrt{\mu_r/\varepsilon_r}$, matches that of free space, $\eta_0 \approx 377\,\Omega$, and the wave will propagate through the interface with zero back reflections (whereas the total SCS will generally be non-zero). In general, it can be shown that any axially symmetric body (of any size) with $\varepsilon_r = \mu_r$ exhibits zero backscattering if illuminated along its axis of symmetry, a result known as *Weston's theorem*.[79–81]

The possibility of having zero backscattering by suitably combining electric and magnetic responses is also very important in the context of metasurfaces operating in transmission.[82–87] In fact, early metasurface designs were based on subwavelength scattering inclusions (e.g. V-shaped nanoantennas) supporting a purely electric dipolar response,[82,83] which therefore exhibit a symmetric scattering pattern, as discussed above. As a result, such metasurfaces suffered from very low transmission efficiency, since half of the incident power was inevitably reflected. It was soon realized that, even by using ideally optimized inclusions, the maximum efficiency of these metasurfaces was fundamentally limited.[84] These limitations, however, can be easily overcome by breaking the radiation symmetries of the metasurface, based on inclusions that radiate both electric and magnetic dipolar fields.[84,85] Consistent with the above discussion, in fact, a metasurface composed of a dense array of Huygens sources will be perfectly matched to free space, thereby producing no reflections. Then, if the inclusions are locally designed to control the phase of the forward scattering, while keeping the Huygens condition, the metasurface will be able to fully control the pattern of the transmitted wave with 100% efficiency and no reflection.[84–87]

The possibility of controlling the backscattering by engineering the interaction of electric and magnetic modes may suggest that a similar degree of control may be achieved for the forward scattering as well. This idea seems supported by the fact that the expression

for the forward scattering cross-section is rather similar to that of the backward scattering[2]:

$$\sigma_{\mathrm{fw}} = \frac{\lambda_0^2}{4\pi}\left|\sum_{n=1}^{N}(2n+1)(c_n^{\mathrm{TM}}+c_n^{\mathrm{TE}})\right|^2, \tag{7.22}$$

In particular, a few works have investigated the possibility of realizing *anti-Huygens sources*, whose radiation pattern would have a null in the forward direction, by imposing the condition $c_n^{\mathrm{TM}} = -c_n^{\mathrm{TE}}$.[75,88] According to Kerker,[75] this could be achieved with a magneto-electric sphere in quasi-static regime, provided that its relative electric permittivity and magnetic permeability respect the following condition:

$$\varepsilon_r = \frac{4-\mu_r}{2\mu_r+1}. \tag{7.23}$$

However, as extensively discussed in Ref. 88, this has led to some confusion, and misinterpretation of experimental results.[89,90] In fact, different from the backward scattering considered above, the forward scattering is directly related to power conservation by the optical theorem (7.12), implying that if the forward scattering goes identically to zero, so does the total scattering cross-section, i.e. the object does not scatter at all! Following the discussion in Ref. 88, the seemingly paradoxical nature of anti-Huygens sources can be explained by recognizing that if the radiation-correction term is not neglected in the polarizabilities, Eq. (7.23) implies that *only the real part* of the forward scattering is identically zero, whereas its imaginary part is small but finite. It is this residual imaginary component of the forward scattering that sustains a finite total scattering cross-section, consistent with the optical theorem (7.12). Therefore, although the condition $c_n^{\mathrm{TM}} = -c_n^{\mathrm{TE}}$ can never be exactly fulfilled with passive media, Kerker's condition (7.23) leads to a *minimum* of forward scattering (not a null), hence realizing an anti-Huygens source that complies with power conservation. An example of radiation/scattering pattern of this kind is shown in Fig. 7.8(c).

The above discussion shows that in order to control the scattering in specific directions, design directive nanoantennas, realize efficient

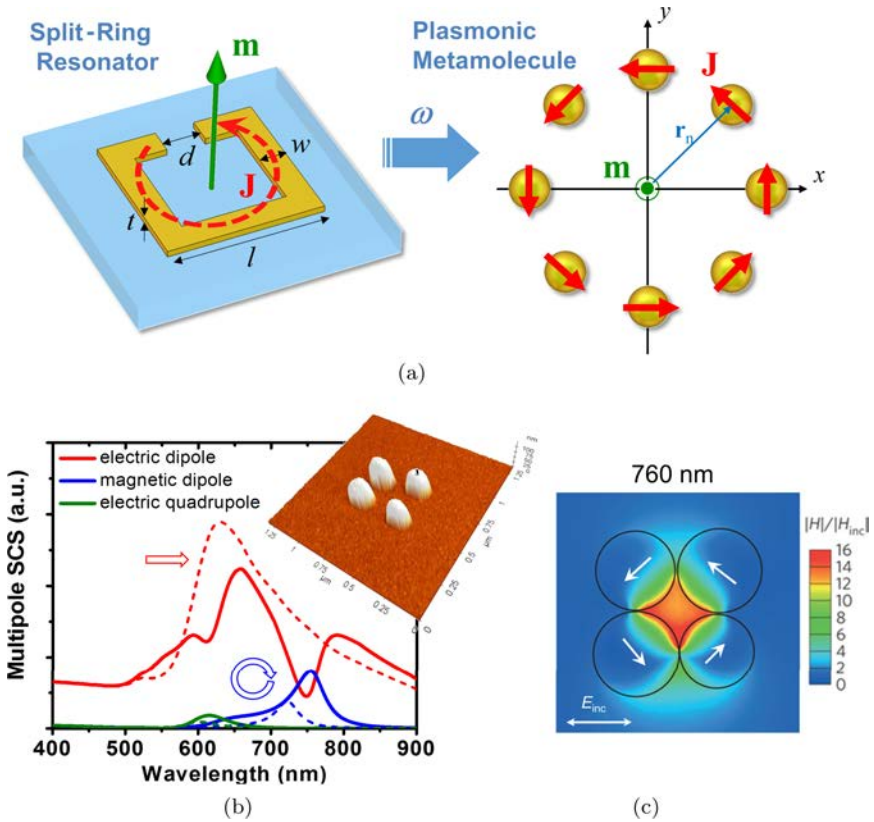


Fig. 7.9. Optical magnetism based on rings of plasmonic nanoparticles. (a) A metallic split-ring resonator at microwaves realizes a resonant circulation of conduction current over a subwavelength footprint, supporting a strong magnetic response without magnetic materials. This idea can be transplanted at optical frequencies based on rings of plasmonic nanoparticles, in which the displacement current takes the role of the conduction current. ©2014 RSC. Reprinted with permission from Ref. 92. (b) Multipolar contributions to the total SCS, for a four-particle nanoring assembled with AFM nanomanipulation (AFM image in the inset). Dashed and solid lines correspond to the case of symmetric and slightly asymmetric nanorings, respectively. Small structural asymmetries efficiently boost the magnetic response (blue line). (c) Magnetic field intensity distribution at the magneto-electric Fano resonance in (b). White arrows indicate the induced dipole moments in the particles. ©2013 NPG. Adapted with permission from Ref. 96.

metasurfaces, etc., it is necessary to efficiently excite magnetic and higher-order multipolar scattering responses. This becomes particularly challenging if the scatterer needs to be subwavelength, as in the case of metamaterial and metasurface inclusions, since the excitation of higher-order multipoles requires some field retardation within the particle. The resonant frequency of higher-order spherical modes can be lowered by loading the structure with reactive elements, as typically done in the case of split-ring resonators at RF to realize strong magnetic resonances within a subwavelength footprint.[91] At optical frequencies, an analogous effect can be achieved with plasmonic materials. For example, it has been shown that subwavelength rings of nanoparticles, as sketched in Fig. 7.9(a), can be designed to support a resonant circulation of displacement current, realizing strong magnetic response at optical frequencies, where natural magnetism is typically weak.[92–95] Within the assumption of uniform magnetic field across the loop, we can also define a magnetic polarizability of the metamolecule, relating the local magnetic field to the induced magnetic dipole moment, i.e. $\mathbf{m} = \alpha_m \mathbf{H_{loc}}$ (here, the polarizability is scalar assuming magnetic field orthogonal to the nanoring plane), and its closed form expression is given in Ref. 93. Consistent with the optical nanocircuit paradigm,[32] this nanoring is the optical equivalent of split-ring resonators at RF, as the plasmonic nanoparticles correspond to inductors interleaved by capacitors (the insulating gaps) in a ring configuration. The magnetic response can be further boosted, for a given excitation, by suitably breaking the symmetry of the nanoring, as demonstrated in Ref. 96. In fact, when the symmetry of the metamolecule is broken, a magneto-electric coupling α_{em} is introduced, which modifies the induced dipole moments as:

$$\begin{aligned} \mathbf{p} &= \alpha_e \mathbf{E}_{\text{loc}} - \alpha_{\text{em}} \eta_0 \hat{\mathbf{n}} \times \mathbf{H}_{\text{loc}} \\ \mathbf{m} &= \alpha_m \mathbf{H}_{\text{loc}} - \alpha_{\text{em}} \frac{\hat{\mathbf{n}} \times \mathbf{E}_{\text{loc}}}{\eta_0} \end{aligned} , \tag{7.24}$$

where $\hat{\mathbf{n}}$ is the unit vector normal to $\mathbf{E}_{\text{loc}}$ and $\mathbf{H}_{\text{loc}}$. The scattering from such a bi-anisotropic object may exhibit particularly strong Fano resonances, due to the near-field interference between electric

and magnetic modes of the scatterer. As seen in Fig. 7.9(b), this may lead to strongly boosted magnetic response, which becomes the dominant scattering contribution of the metamolecule, as well as intense magnetic "hot-spots" (Fig. 7.9(c)), which may be particularly useful to probe and enhance the direct interaction of matter with the optical magnetic field.

The possibility to enhance the optical magnetic response represents an important scientific breakthrough, as it overcomes the long-held belief that magnetism is inherently weak at optical frequencies,[92–97] and it is a relevant example of "extreme scattering engineering" inspired by metamaterial concepts and enabled by plasmonics.

In general, systems of closely coupled plasmonic nanoparticles represent a particularly appealing platform to realize advanced scattering responses at optical frequencies, based on suitably tailoring the collective plasmon modes of these metamolecules. A vast literature exists on this topic, and we refer the interested readers to Ref. 98–100 for further details. Particularly interesting examples include the realization of directive optical scatterers, nanoantennas and nanoemitters,[100–102] optical nanocircuits and nanofilters,[103,104] Fano resonances for enhanced energy concentration and sensing,[62,105–107] etc.

An interesting alternative to realize higher-order scattering responses is represented by high-index dielectric nanoparticles, which have recently attracted large attention, since materials with relatively high permittivity (and low losses) are readily available at optical frequencies (e.g. silicon, gallium arsenide, germanium, etc.).[108–112] For a dielectric sphere with refractive index larger than about two, in fact, the lowest-frequency Mie resonance is magnetic, corresponding to the first TE scattering coefficient c_1^{TE}. This magnetic resonance occurs when the diameter of the particle is approximately equal to the wavelength in the dielectric material, i.e.

$$2a \approx \frac{\lambda_0}{\sqrt{\varepsilon_r}}. \tag{7.25}$$

Therefore, a magnetic resonance can be obtained within a subwavelength footprint if the refractive index of the particle is sufficiently

high. Different from plasmonic scatterers, which can, in principle, support a resonance for arbitrarily small dimensions, beating diffraction thanks to the excitation of surface plasmons, a dielectric scatterer is inherently limited by diffraction and, therefore, it can support a resonance for smaller and smaller dimensions (at a given frequency) only if its refractive index is also proportionally increased according to (7.25). Therefore, while the scaling of a plasmonic resonator is limited by the unavoidable losses of the plasmonic material (i.e. the imaginary part of ε_r), the scaling of a dielectric resonator is limited by the availability of materials with high refractive index at the frequency of interest (i.e. it is limited by the real part of ε_r).

High-index dielectric scatterers can be tailored to exhibit electric and magnetic dipolar responses in the same frequency range, an ideal condition to realize robust Huygens or anti-Huygens optical nanoantennas,[113–115] as shown in Fig. 7.10, as well as Huygens metasurfaces at optical frequencies[116–119] and antireflection coatings.[120] Furthermore, as in the case of plasmonic nanoparticle clusters, the combination of multiple dielectric particles allows realizing more advanced scattering responses, such as different forms of electric and magnetic Fano resonances.[121,122] Rings of dielectric particles can also be used to realize a so-called "toroidal" response at optical frequencies,[123] namely, the generation of an electric dipole moment, orthogonal to the plane of the particle, from the circulation of a magnetic current. This represents the dual of the case of artificial magnetism realized in rings of plasmonic nanoparticles discussed above. A toroidal current distribution radiates just as an electric dipole,[g] with typically lower radiation efficiency, which results in scattering resonances with higher Q factor. Recently, these ideas have also been applied to design a scatterer in which the toroidal contribution cancels the total electric dipole, leading to a so-called

[g]The so-called toroidal moment, in fact, is a second-order contribution to the electric dipole moment $\mathbf{p}$, supported by a circulating magnetization current (see, e.g. Ref. 124).

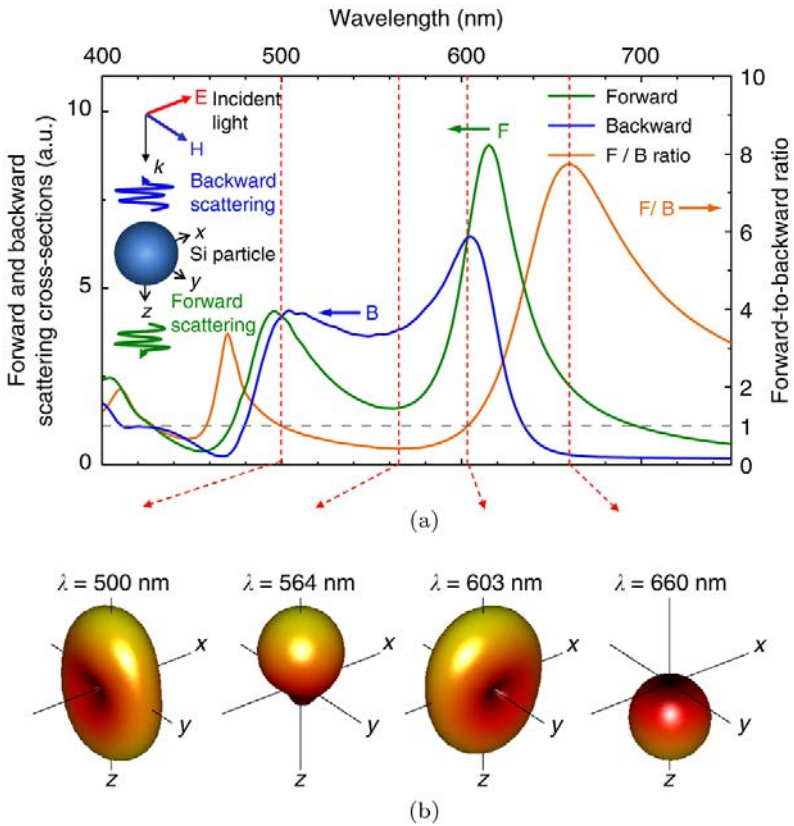


Fig. 7.10. Scattering properties of a silicon nanoparticle in the visible range. (a) Forward (green curve) and backward (blue) scattering cross-sections, and forward-to-backward ratio (orange) of a spherical silicon nanoparticles with radius of 75 nm, illuminated by a plane wave, as shown in the inset. (b) At different frequencies, the combination of electric and magnetic dipolar responses results in different scattering patterns. From left to right: scattering pattern of a resonating electric dipole, anti-Huygens source (notice the non-zero forward scattering), resonating magnetic dipole, and Huygens source. ©2013 NPG. Reprinted with permission from Ref. 115.

"anapole" current distribution with reduced radiation damping.[125] Interestingly, this has also been connected to the concept of non-radiating sources discussed above and to speculative models of dark matter.[125,126] However, some fundamental theorems[67] imply that the radiation loss can never be ideally zero in a 3D open resonator,

unless materials with zero permittivity or permeability are used, or the geometry has extreme features (infinitesimal or diverging dimensions). The only way to achieve an ideally bound optical state supported by an open resonator seems to be through the concept of embedded eigenstates,[67,68] as discussed in Section 1.5. Nevertheless, research on quasi-non-radiating current distributions, and their application for extreme scattering, light trapping, Aharonov–Bohm-like optical effects, etc., is an active and exciting area of research.

7.7. Conclusions

In this chapter, we have discussed some of the most exciting and intriguing scattering effects based on, or inspired by, metamaterial concepts and plasmonics. After a brief introduction on the general scattering properties of linear and passive structures, we have reviewed some of the different techniques available today to suppress, enhance and manipulate the electromagnetic scattering. Large emphasis has also been placed on fundamental limits of passive scatterers, suggesting how to approach or exceed them. In this context, we have discussed some extreme and counterintuitive scattering phenomena, which clearly illustrate the power of metamaterial concepts applied to scattering systems.

We believe that the reach of these concepts can be further extended, in the near future, by introducing active, non-local, non-reciprocal, and nonlinear materials and structures, which may allow realizing even more intriguing and useful scattering effects. We expect that the ability to largely control scattering processes, enabled by metamaterials and plasmonics, will have disruptive impact for many practical applications in the coming years.

Acknowledgments

This work was supported by the NSF CAREER award with Grant No. ECCS-0953311, the AFOSR with Grant No. FA9550-13-1-0204, the Defense Threat Reduction Agency with Grant No. HDTRA1-12-1-0022 and the ONR MURI Grant No. N00014-10-1-0942.

References

1. Engheta, N. and Ziolkowski, R. W. (2006). *Electromagnetic Metamaterials: Physics and Engineering Explorations* (Wiley-IEEE Press, New York).
2. Bohren, C. F. and Huffman, D. R. (2008). *Absorption and Scattering of Light by Small Particles* (John Wiley & Sons, New York).
3. Weinberg, S. (2012). *Lectures on Quantum Mechanics* (Cambridge University Press, Cambridge).
4. Alú, A. and Engheta, N. (2005). *J. Appl. Phys.* **97**, p. 094310.
5. Papas, C. H. (2013). *Theory of Electromagnetic Wave Propagation* (Courier Dover Publications, Mineola).
6. Jackson, J. D. (1998). *Classical Electrodynamics* (Wiley, Hoboken).
7. de Vries, P., van Coevorden, D. and Lagendijk, A. (1998). *Rev. Mod. Phys.* **70**, p. 447.
8. Sipe, J. and Kranendonk, J. (1974). *Phys. Rev. A* **9**, p. 1806.
9. Alù, A. and Engheta, N. (2009). *Phys. Rev. B* **79**, p. 235412.
10. Felsen, L. B. and Marcuvitz, N. (1994). *Radiation and Scattering of Waves* (Wiley-IEEE Press, Hoboken).
11. Alù, A., Yaghjian, A. D., Shore, R. A. and Silveirinha, M. G. (2011). *Phys. Rev. B* **84**, p. 054305.
12. Nussenzveig, H. M. (1972). *Causality and Dispersion Relations* (Academic Press, New York).
13. Fleury, R., Monticone, F. and Alù, A. (2015). *Phys. Rev. Appl.* **4**, p. 037001.
14. Leonhardt, U. (2006). *Science* **312**, p. 1777.
15. Perczel, J., Tyc, T. and Leonhardt, U. (2011). *New J. Phys.* **13**, p. 083007.
16. Chen, H., Zheng, B., Shen, L., Wang, H., Zhang, X., Zheludev, N. I. and Zhang, B. (2013). *Nat. Commun.* **4**, p. 2652.
17. Chen H. and Zheng, B. (2012). *Sci. Rep.* **2**, p. 255.
18. Choi, J. S. and Howell, J. C. (2014). *Opt. Express* **22**, p. 29465.
19. Alù, A. and Engheta, N. (2005). *Phys. Rev. E* **72**, p. 16623.
20. Alu, A. and Engheta, N. (2007). *Opt. Express* **15**, p. 7578.
21. Silveirinha, M. G., Alù, A. and Engheta, N. *Phys. Rev. E* **75**, p. 036603.
22. Edwards, B., Alù, A., Silveirinha, M. G. and Engheta, N. (2009). *Phys. Rev. Lett.* **103**, p. 153901.
23. Rainwater, D., Kerkhoff, A., Melin, K., Soric, J. C., Moreno, G. and Alù, A. (2012). *New J. Phys.* **14**, p. 013054.
24. Alù, A. (2009). *Phys. Rev. B* **80**.
25. Chen, P.-Y., Monticone, F. and Alu, A. (2011). *IEEE Antennas Wirel. Propag. Lett.* **10**, p. 1598.
26. Chen, P.-Y. and Alù, A. (2011). *Phys. Rev. B* **84**.
27. Soric, J. C., Chen, P. Y., Kerkhoff, A., Rainwater, D., Melin, K. and Alù, A. (2013). *New J. Phys.* **15**, p. 033037.
28. Alù, A. and Engheta, N. (2009). *Phys. Rev. Lett.* **102**.
29. Alù, A. and Engheta, N. (2010). *Phys. Rev. Lett.* **105**, p. 263906.

30. Born, M. and Wolf, E. (2002). *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Cambridge University Press, Cambridge).
31. Boltasseva, A. and Atwater, H. A. (2011). *Science* **331**, p. 290.
32. Engheta, N., Salandrino, A. and Alù, A. (2005). *Phys. Rev. Lett.* **95**, p. 095504.
33. Munk, B. A. (2000). *Frequency Selective Surfaces: Theory and Design* (John Wiley & Sons Hoboken).
34. Soric, J. C., Chen, P. Y., Kerkhoff, A., Rainwater, D., Melin, K. and Alù, A. (2013). *New J. Phys.* **15**, p. 033037.
35. Schofield, R. S., Soric, J. C., Rainwater, D., Kerkhoff, A. and Alù, A. (2014). *New J. Phys.* **16**, p. 063063.
36. Chen, P.-Y. and Alù, A. (2011). *ACS Nano* **5**, p. 5855.
37. Monticone, F. and Alù, A. (2014). *Photonics Nanostruct.* **12**, p. 330.
38. Alù, A. and Engheta, N. (2008). *Phys. Rev. E* **78**, p. 045602.
39. Alù, A. and Engheta, N. (2008). *Phys. Rev. Lett.* **100**, p. 113901.
40. Pozar, D. M. (2011). *Microwave Engineering*, 3rd edn. (Wiley).
41. Pendry, J. B., Schurig, D. and Smith, D. R. (2006). *Science* **312**, p. 1780.
42. Schurig, D., Mock, J. J., Justice, B. J., Cummer, S. A., Pendry, J. B., Starr, A. F. and Smith, D. R. (2006). *Science* **314**, p. 977.
43. Gustafsson, M., Sohl, C. and Kristensson, G. (2007). *Proc. R. Soc. A Math. Phys. Eng. Sci.* **463**, p. 2589.
44. Monticone, F. and Alù, A. (2013). *Phys. Rev. X* **3**, p. 041005.
45. Miller, D. A. B. (2006). *Opt. Express* **14**, p. 12457.
46. Chen, H., Liang, Z., Yao, P., Jiang, X., Ma, H. and Chan, C. T. (2007). *Phys. Rev. B* **76**.
47. Craeye, C. and Bhattacharya, A. (2012). *IEEE Trans. Antennas Propag.* **60**, p. 3516.
48. Alú, A. and Engheta, N. (2005). *J. Appl. Phys.* **97**, p. 094310.
49. Chu, L. J. (1948). *J. Appl. Phys.* **19**, p. 1163.
50. Alu, A. and Engheta, N. (2007). *IEEE Trans. Antennas Propag.* **55**, p. 3027.
51. Olver, F. W., Lozier, D. W., Boisvert, R. F. and Clark, C. W. (2010). *NIST Handbook of Mathematical Functions* (Cambridge University Press, Cambridge).
52. Pozar, D. M. (2009). "New results for minimum Q, maximum gain, and polarization properties of electrically small arbitrary antennas," in EuCAP 2009. 3rd European Conference on Antennas and Propagation 2009, 1993–1996.
53. Sievenpiper, D. F., Dawson, D. C., Jacob, M. M., Kanar, T., Kim, S., Long, J. and Quarfoth, R. G. (2012). *IEEE Trans. Antennas Propag.* **60**, p. 8.
54. Ziolkowski, R. W. and Erentok, A. (2007). *IET Microwaves, Antennas Propag.* **1**, p. 116.
55. Ziolkowski, R. W. (2010). *IEEE Trans. Antennas Propag.* **58**, p. 318.
56. Ruan, Z. and Fan, S. (2010). *Phys. Rev. Lett.* **105**, p. 013901.

57. Ruan, Z. and Fan, S. (2011). *Appl. Phys. Lett.* **98**, p. 043101.
58. Alù, A. and Engheta, N. (2008). *New J. Phys.* **10**, p. 115036.
59. Monticone, F., Argyropoulos, C. and Alù, A. (2012). *Sci. Rep.* **2**, p. 912.
60. Monticone, F., Argyropoulos, C. and Alù, A. *Phys. Rev. Lett.* **110**, p. 113901.
61. Miroshnichenko, A. E., Flach, S. and Kivshar, Y. S. (2010). *Rev. Mod. Phys.* **82**, p. 2257.
62. Luk'yanchuk, B., Zheludev, N. I., Maier, S. A., Halas, N. J., Nordlander, P., Giessen, H. and Chong, C. T. (2010). *Nat. Mater.* **9**, p. 707.
63. Argyropoulos, C., Monticone, F., D'Aguanno, G. and Alù, A. (2013). *Appl. Phys. Lett.* **103**, p. 143113.
64. Hsu, C. W., DeLacy, B. G., Johnson, S. G., Joannopoulos, J. D. and Soljačić, M. (2014). *Nano Lett.* **14**, p. 2783.
65. Argyropoulos, C., Chen, P.-Y., Monticone, F., D'Aguanno, G. and Alù, A. *Phys. Rev. Lett.* **108**.
66. Monticone, F. and Alù, A. (2014). Forum for Electromagnetic Research Methods and Application Technologies (FERMAT) **6**, Available at: http://www.e-fermat.org/files/news_views/1546b8ee8cf6e5.pdf.
67. Silveirinha, M. G. (2014). *Phys. Rev. A* **89**, p. 023813.
68. Monticone, F. and Alù, A. (2014). *Phys. Rev. Lett.* **112**, p. 213903.
69. von Neuman, J. and Wigner, E. (1929). *Zhurnal Phys.* **30**, p. 467.
70. Marinica, D., Borisov, A. and Shabanov, S. (2008). *Phys. Rev. Lett.* **100**, p. 183902.
71. Hsu, C. W., Zhen, B., Lee, J., Chua, S.-L., Johnson, S. G., Joannopoulos, J. D. and Soljačić, M. (2013). *Nature* **499**, p. 188.
72. Ehrenfest, P. (1910). *Phys. Z* **11**, pp. 708–709.
73. Schott, G. A. (1933). *Philos. Mag.* **15**, pp. 752–761.
74. Devaney, A. J. and Wolf, E. (1973). *Phys. Rev. D* **8**, p. 1044.
75. Kerker, M., Wang, D.-S. and Giles, C. L. (1983). *J. Opt. Soc. Am.* **73**, p. 765.
76. Ziolkowski, R. W. (2010). *IEEE Antennas Wirel. Propag. Lett.* **9**, p. 501.
77. Niemi, T., Alitalo, P., Karilainen, A. O. and Tretyakov, S. A. (2012). *IET Microwaves, Antennas Propag.* **6**, p. 735.
78. Knott, E. F., Shaeffer, J. and Tuley, M. (2004). *Radar Cross Section, 2nd edn.* (SciTech Publishing Stevenage).
79. Weston, V. (1963). *IEEE Trans. Antennas Propag.* **11**, p. 578.
80. Monzon, C., Medgyesi-Mitschang, L. N. and Forester, D. W. (2005). *J. Opt. Soc. Am. A* **22**, p. 1035.
81. Vacus, O. and Ziolkowski, R. W. (2015), *IEEE Trans. Antennas Propag.* **63**, p. 4418.
82. Yu, N., Genevet, P., Kats, M. A., Aieta, F., Tetienne, J.-P., Capasso, F. and Gaburro, Z. (2011). *Science* **334**, p. 333.
83. Ni, X., Emani, N. K., Kildishev, A. V., Boltasseva, A. and Shalaev, V. M. (2012). *Science* **335**, p. 427.
84. Monticone, F., Estakhri, N. M. and Alù, A. (2013). *Phys. Rev. Lett.* **110**, p. 203903.

85. Alù, A. (2013). *Physics* **6**, p. 53.
86. Pfeiffer, C. and Grbic, A. (2013). *Phys. Rev. Lett.* **110**, p. 197401.
87. Selvanayagam, M. and Eleftheriades, G. V. (2013). *Opt. Express* **21**, p. 14409.
88. Alu, A. and Engheta, N. (2010). *J. Nanophotonics* **4**, p. 041590.
89. Mehta, R. V., Patel, R., Desai, R., Upadhyay, R. V. and Parekh, K. (2006). *Phys. Rev. Lett.* **96**, p. 127402.
90. Ramachandran, H. and Kumar, N. (2008). *Phys. Rev. Lett.* **100**, p. 229703.
91. Pendry, J. B., Holden, A. J., Robbins, D. J. and Stewart, W. J. (1999). *IEEE Trans. Microw. Theory Tech.* **47**, p. 2075.
92. Monticone, F. and Alu, A. (2014). *J. Mater. Chem. C.*
93. Alù, A., Salandrino, A. and Engheta, N. (2006). *Opt. Express* **14**, p. 1557.
94. Urzhumov, Y. A., Shvets, G., Fan, J. A., Capasso, F., Brandl, D. and Nordlander, P. (2007). *Opt. Express* **15**, p. 14129.
95. Simovski, C. R. and Tretyakov, S. A. (2009). *Phys. Rev. B* **79**, p. 045111.
96. Shafiei, F., Monticone, F., Le, K. Q., Liu, X.-X., Hartsfield, T., Alù, A. and Li, X. (2013). *Nat. Nanotechnol.* **8**, p. 95.
97. Landau, L. D., Pitaevskii, L. P. and Lifshitz, E. M. (1984). *Electrodynamics of Continuous Media* (Butterworth-Heinemann, Oxford).
98. Halas, N. J., Lal, S., Chang, W.-S., Link, S. and Nordlander, P. (2011). *Chem. Rev.* **111**, p. 3913.
99. Monticone, F. and Alù, A. (2014). *Chinese Phys. B* **23**, p. 047809.
100. Mario, A. and Alú, A. (2013). *Optical Antennas* (Cambridge University Press, Cambridge).
101. Li, J., Salandrino, A. and Engheta, N. (2007). *Phys. Rev. B* **76**, p. 245403.
102. Curto, A. G., Volpe, G., Taminiau, T. H., Kreuzer, M. P., Quidant, R. and van Hulst, N. F. (2010). *Science* **329**, p. 930.
103. Shi, J., Monticone, F., Elias, S., Wu, Y., Ratchford, D., Li, X. and Alù, A. (2014). *Nat. Commun.* **5**.
104. Liu, N., Wen, F., Zhao, Y., Wang, Y., Nordlander, P., Halas, N. J. and Alù, A. (2013). *Nano Lett.* **13**, p. 142.
105. Fan, J. A., Wu, C., Bao, K., Bao, J., Bardhan, R., Halas, N. J., Manoharan, V. N., Nordlander, P., Shvets, G. and Capasso, F. (2010). *Science* **328**, p. 1135.
106. Dregely, D., Hentschel, M. and Giessen, H. (2011). *ACS Nano* **5**, p. 8202.
107. Stockman, M. I. (2010). *Nature* **467**, p. 541.
108. Schuller, J. A., Zia, R., Taubner, T. and Brongersma, M. L. (2007). *Phys. Rev. Lett.* **99**, p. 107401.
109. Popa, B.-I. and Cummer, S. A. (2008). *Phys. Rev. Lett.* **100**, p. 207401.
110. Zhao, Q., Zhou, J., Zhang, F. and Lippens, D. (2009). *Mater. Today* **12**, p. 60.
111. Ginn, J. C., Brener, I., Peters, D. W., Wendt, J. R., Stevens, J. O., Hines, P. F., Basilio, L. I., Warne, L. K., Ihlefeld, J. F., Clem, P. G. and Sinclair, M. B. (2012). *Phys. Rev. Lett.* **108**, p. 097402.

112. Kuznetsov, A. I., Miroshnichenko, A. E., Fu, Y. H., Zhang, J. and Luk'yanchuk, B. (2012). *Sci. Rep.* **2**, p. 492.
113. Rolly, B., Stout, B. and Bonod, N. (2012). *Opt. Express* **20**, p. 20376.
114. Person, S., Jain, M., Lapin, Z., Sáenz, J. J., Wicks, G. and Novotny, L. (2013). *Nano Lett.* **13**, p. 1806.
115. Fu, Y. H., Kuznetsov, A. I., Miroshnichenko, A. E., Yu, Y. F. and Luk'yanchuk, B. (2013). *Nat. Commun.* **4**, p. 1527.
116. Lin, D., Fan, P., Hasman, E. and Brongersma, M. L. (2014). *Science* **345**, p. 298.
117. Arbabi, A., Horie, Y., Bagheri, M. and Faraon, A. (2015). "Dielectric metasurfaces for complete control of phase and polarization with subwavelength spatial resolution and high transmission," Nat. Nanotechnol., in press.
118. Decker, M., Staude, I., Falkner, M., Dominguez, J., Neshev, D. N., Brener, I., Pertsch, T. and Kivshar, Y. S. (2015). *Adv. Opt. Mater.* **3**, p. 813.
119. Campione, S., Basilio, L. I., Warne, L. K. and Sinclair, M. B. (2015). *Opt. Express* **23**, p. 2293.
120. Spinelli, P., Verschuuren, M. A. and Polman, A. (2012). *Nat. Commun.* **3**, p. 692.
121. Chong, K. E., Hopkins, B., Staude, I., Miroshnichenko, A. E., Dominguez, J., Decker, M., Neshev, D. N., Brener, I. and Kivshar, Y. S. (2014). *Small* **10**, p. 1985.
122. Hopkins, B., Filonov, D. S., Miroshnichenko, A. E., Monticone, F., Alù, A. and Kivshar, Y. S. (2015). *ACS Photonics* **2**, p. 724.
123. Basharin, A. A., Kafesaki, M., Economou, E. N., Soukoulis, C. M., Fedotov, V. A., Savinov, V. and Zheludev, N. I. (2015). *Phys. Rev. X* **5**, p. 011036.
124. Van Bladel, J. G. (2007). *Electromagnetic Fields* (John Wiley & Sons, Hoboken).
125. Miroshnichenko, A. E., Evlyukhin, A. B., Yu, Y. F., Bakker, R. M., Chipouline, A., Kuznetsov, A. I., Luk'yanchuk, B., Chichkov, B. N. and Kivshar, Y. S. (2015). *Nat. Commun.* **6**, p. 8069.
126. Fedotov, V. A., Rogacheva, A. V., Savinov, V., Tsai, D. P. and Zheludev, N. I. (2013). *Sci. Rep.* **3**, p. 2967.
127. Monticone, F. and Alù, A. (2016). *Optica* **3**, p. 718.
128. Hsu, C. W., Zhen, B., Stone, A. D., Joannopoulos, J. D. and Soljačić, M. Bound states in the continuum. Nat. Rev. Mater. 1, 16048 (2016).

CHAPTER 8

All-Dielectric Nanophotonics: Fundamentals, Fabrication, and Applications

ALEXANDER KRASNOK[*,†], ROMAN SAVELEV[†], DENIS BARANOV[‡,§] and PAVEL BELOV[†]

[*]ITMO University, Russia
[†]The University of Texas at Austin, USA
[‡]Moscow Institute of Physics and Technology, Russia
[§]Chalmers University of Technology, Sweden

This chapter reviews a novel, rapidly developing field of modern light science named all-dielectric nanophotonics. This branch of nanophotonics is based on the properties of high-index dielectric nanoparticles which allow for controlling both magnetic and electric responses of a nanostructured matter. Here, we discuss optical properties of high-index dielectric nanoparticles, methods of their fabrication, and recent advances in practical applications, including the quantum source emission engineering, Fano resonances in all-dielectric nanoclusters, surface enhanced spectroscopy and sensing, coupled-resonator optical waveguides, metamaterials and metasurfaces, and nonlinear nanophotonics.

[*]Corresponding author: akrasnok@utexas.edu

8.1. Introduction

Since the modern technologies largely depend on the rapidly growing demands for powerful computational capacities and efficient information processing, the development of conceptually new approaches and methods is extremely valuable. One of these approached is based on replacing electrons with photons as the main information carriers.[1] The advantages of light for fast computing are obvious: the parallel transfer and processing of signals using polarization and orbital momentum of photons as additional degrees of freedom,[2] the possibility of multi-frequency operations, and the high operating frequency around 500 THz (wavelength of 600 nm). However, photons as alternative information carriers have relatively large "size" determined by their wavelength. This leads to a weak interaction of photons with nanoscale objects including quantum emitters, subwavelength waveguides, and others, whereas the effective light-matter coupling is extremely important for all-optical information processing.

The efficient light manipulation implies simultaneous control of its electric and magnetic components. However, the magnetic response of natural materials at optical frequencies is usually weak, as was originally posted by Landau and Lifshitz.[3,4] This is the reason why photonic devices operate mainly with the electric part of a light wave.[5] At the same time, magnetic dipoles are very common sources of the magnetic field in nature. A common example of a magnetic dipole radiation is an electromagnetic wave produced by an excited metal split-ring resonator (SRR), which is a basic constituting element of *metamaterials*.[6–8] Currents excited by the external electromagnetic radiation and running inside the SRR produce a transverse magnetic field in the center of the ring oscillating up and down, which simulates an oscillating magnetic dipole. The major interest in such artificial systems is due to their ability to response to the magnetic component of incident light and thus to have a non-unity or even negative magnetic permeability (μ) at optical frequencies. This provides the possibilities to design materials with highly unusual properties such as negative refraction,[6,9–11] cloaking,[12] or superlensing.[13] The

SRR concept works very well for gigahertz,[8,14] terahertz[15] and even near-infrared[16] frequencies. However, this approach fails for shorter wavelengths and, in particular, in the visible spectral range due to increasing losses and technological difficulties in fabrication of smaller and smaller constituting split-ring elements.[17] Several other designs based on metal nanostructures have been proposed to shift the magnetic resonance wavelength to the visible spectral range.[9,10] However, all of them are suffering from losses inherent to metals at visible frequencies.

An alternative approach to achieve strong magnetic response with low losses is to use nanoparticles made of high-refractive index dielectric materials (e.g. Si, Ge).[18–31] As it follows from the exact Mie theory of light scattering,[32] a strong magnetic dipole response can be achieved in spherical dielectric particles. Remarkably, for the refractive indices larger than a certain value, there is a well-established hierarchy of the magnetic and electric resonances. In contrast to plasmonic particles, the fundamental resonance of high-index spherical nanoparticles is the *magnetic dipole* one, occurring when the wavelength of light inside the particle approximately equals its diameter $\lambda/n \simeq 2R$, where λ is the wavelength in free space, R and n are the radius and the refractive index of spherical particle. Under this condition, the polarization of the electric field is anti-parallel at opposite boundaries of the sphere, which gives rise to strong coupling to circulation displacement currents, while the magnetic field oscillates up and down in the middle. Today, such high-index dielectric nanoparticles are the base of so-called *all-dielectric nanophotonics*.

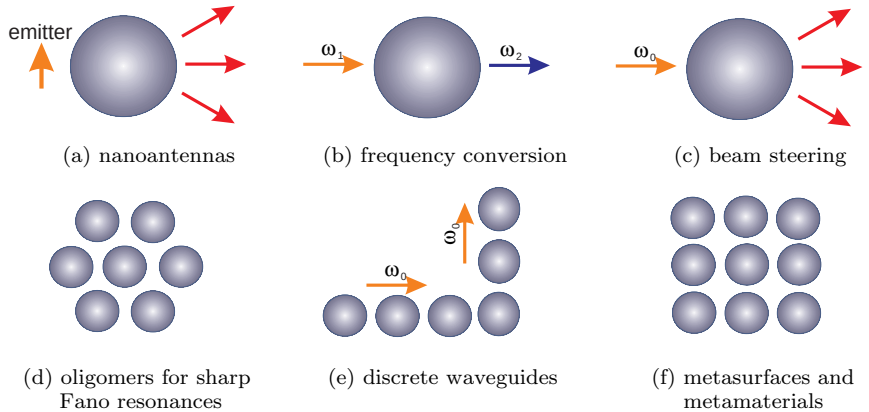
It should be noted that the all-dielectric nanophotonics is not a part of so-called *silicon photonics*[33,34] since it does not limited by silicon as a material. Moreover, the silicon photonics is based on silicon waveguides and circuits with dimensions of few hundred nanometers, whereas the heart of all-dielectric nanophotonics is high-index nanoparticles exhibiting a strong magnetic optical response.

The prediction of the resonant electromagnetic response of small dielectric particles is not an absolutely new discovery. These resonances have been recognized since the original work of G. Mie.[35] Later, in 1930 J. A. Stratton has highlighted the effect of magnetic and electric dipole resonances of water drops in the atmosphere (rain, fog, or clouds) on the propagation of short radio waves.[36] In the following studies (see, e.g., Ref. 37) the theoretical prediction of man-made artificial dielectrics based on high-index particles was made. Then, the metamaterials based on the dielectric particles have been proposed in the microwave and mid-IR frequency ranges.[18,19,23]

In 2010 the possibility of enhanced optical magnetic response of high-index nanoparticles in the visible range was discussed theoretically in Ref. 38. It should be noted that before this work, a number of papers have been published, where the spectra of light scattering by high-index dielectric cylinders including magnetic resonances were measured.[39–42] However, an attention to the magnetic nature of these resonances has not been paid. In 2011 the scattering properties of silicon (Si) nanoparticles have been studied in details.[43] Shortly after these theoretical works, the concept of "magnetic light" has been experimentally realized in visible,[44–46] infrared,[47] and even microwave[48] frequency ranges.

This magnetic light concept has paved the way for many fascinating applications in the areas of quantum source emission engineering[49–57] (Fig. 8.1(a)), frequency conversion[58] (Fig. 8.1(b)), tunable routers and switchers[59–61] (Fig. 8.1(c)), sensors[62–64] (Fig. 8.1(d)), dielectric waveguides[65–67] (Fig. 8.1(e)), and all-dielectric metasurfaces,[68,69] and metamaterials[22–24,47,70] (Fig. 8.1(f)). All these applications are discussed in this Review (see Sec. 8.4).

This Chapter is intended to review the all-dielectric nanophotonics as a part of the modern light-science. First, in Sec. 8.2 we discuss the optical properties of high-index dielectric nanoparticles with an example of silicon nanoparticles. Then, in Sec. 8.3 we describe the various existing methods of such nanoparticles fabrication, including the chemical deposition, the thin film dewetting, the femtosecond



(a) nanoantennas (b) frequency conversion (c) beam steering

(d) oligomers for sharp Fano resonances (e) discrete waveguides (f) metasurfaces and metamaterials

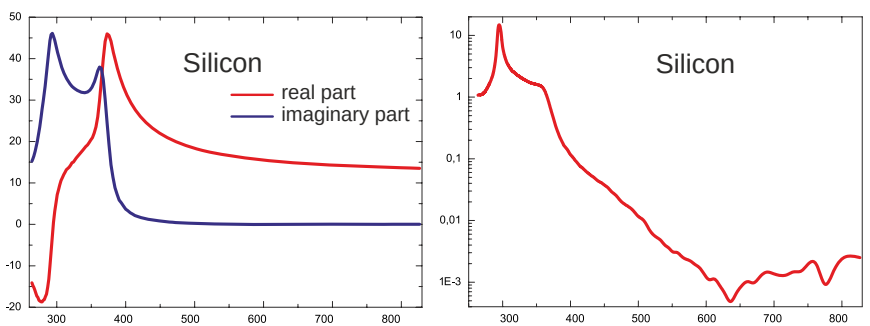
Fig. 8.1. The variety of photonics structures based on high-indexed dielectric nanoparticles are discussed in this Review: (a) Nanoantennas for enhanced emission and focusing; (b) Nanoantennas for enhanced frequency conversion effects; (c) Nanoantennas for steering of light; (d) Fano-resonant nanostructures; (e) Waveguides composed of dielectric nanoparticles; (f) All-dielectric metasurfaces and metamaterials.

laser ablation of bulk and thin Si films, and the reactive-ion-etching approach. The examples of the produced structures are presented. In Sec. 8.4.1 the applications of all-dielectric nanoantennas for the quantum source emission engineering are presented. In Sec. 8.4.2 the all-dielectric oligomers and their Fano resonances are discussed. Then in Sec. 8.4.3 the applications of dielectric nanoantennas for the surface enhanced spectroscopy and sensing are provided. The all-dielectric nanoantennas and plasmonic ones are compared in the context of the sensing applications. In Sec. 8.4.4 the properties of coupled-resonator nanoparticle waveguides are discussed. The optical solitons and bound-states-in-continuum in the dielectric waveguides in form of chains of dielectric particles are also reviewed. Sec. 8.4.5 is devoted to the brief review of the all-dielectric metamaterials and metasurfaces. The references to the more comprehensive reviews of this topic are proposed. In Sec. 8.4.6 the nonlinear properties of silicon nanoparticles with electric and magnetic dipole responses, including higher harmonics generation and electron-hole plasma photoexcitation, are

discussed. In Conclusions the main statements of the Chapter are summarized and the outlook of this area is given.

8.2. Optical Properties of High-index Dielectric Nanoparticles

Two types of electric currents appear as the sources of electromagnetic fields in Maxwell equations: the conductivity current and the displacement current.[72] In dielectrics and semiconductors, the displacement currents strongly exceed the conductivity currents in the spectral range far from their main absorption band. As an example, the spectral dependencies of real and imaginary parts of the refractive index of crystalline Si, measured at room temperatures, are shown in Fig. 8.2(a).[71] In Fig. 8.2(b) the ratio of the conductivity and the displacement current is shown. One can see that in silicon the displacement currents strongly exceed the conductivity currents at wavelengths larger 400 nm. Therefore, from the electromagnetic point of view, the pure and crystalline silicon can be considered as a dielectric in the visible spectral range.

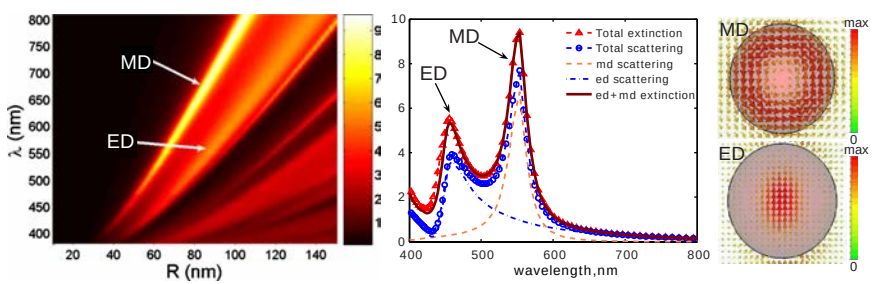


(a) dielectric permittivity (b) conduction current/polarization current

Fig. 8.2. (a) Real and imaginary parts of the permittivity of crystalline silicon.[71] (b) Ratio of the conductivity and displacement currents in crystalline silicon. The displacement current strongly exceeds the conductivity current at wavelengths above 400 nm — in this spectral range the pure and crystalline Si can be considered as dielectric.

A dielectric nanoparticle can be treated as an open resonator supporting a series of electromagnetic resonances — eigenmodes. Exact analysis of the plane wave diffraction by a spherical particle (Mie scattering) shows that multipole modes of different orders can be excited in the dielectric nanoparticle.[32] The coupling strengths between the incident wave and eigenmodes depend on the size parameter $x = k_0 nR$, where n is the refractive index of the particle, k_0 is the free space wavenumber of the incident radiation, and R is the radius of the particle. If $x \ll 1$, the particle is optically small and its diffraction can be described by the Rayleigh approximation. With increasing x, the fundamental magnetic dipole (MD) resonance appears in the particle response. The electric field lines at this resonance are shown in Fig. 8.3(c). The scattered field by the particle at the MD resonance corresponds to the radiation field of a magnetic dipole. With further increase of x, the first electric dipole (ED) resonance is formed. For even larger values of the size parameter, the higher order (quadrupole, octupole, etc.) multipole modes are excited.

The frequencies of the Mie resonances for a spherical particle in the dipole approximation can be determined from the following



(a) (b) (c)

Fig. 8.3. (a) Scattering efficiency spectra of Si spherical particles with the radius R located in air. (b) Extinction and scattering spectra of a Si particle (R = 65 nm). The arrows indicate the electric dipole (ED) and magnetic dipole (MD) contributions to the total efficiencies. From the Ref. 38. (c) Distributions of the electric field within the nanoparticle at magnetic dipole (MD) and electric dipole (ED) resonances.

conditions:

$$\mathrm{Re}(\alpha_e^{-1}) = \mathrm{Re}\left[\left(i\frac{3\varepsilon_h}{2k_h^3}a_1\right)^{-1}\right] = 0,$$
$$\mathrm{Re}(\alpha_m^{-1}) = \mathrm{Re}\left[\left(i\frac{3}{2k_h^3}b_1\right)^{-1}\right] = 0, \quad (8.1)$$

where α_e and α_m are electric and magnetic polarizabilities, respectively, ε_h is the host permittivity, $k_h = \sqrt{\varepsilon_h}\omega/c$ is the wavenumber of light in host medium, ω is the angular frequency, c is the speed of light in a vacuum, a_1 and b_1 are the scattering Mie coefficients.[32] In Ref. 38 it was shown, that for a Si spherical nanoparticle, the conditions of the lowest order multipole (dipole) resonances are fulfilled for the radius of $\approx$ 70 nm. Figure 8.3 shows the numerically calculated scattering efficiency spectra (a), as well as the extinction and scattering spectra (b) of such particle. We would like to note that the resonance frequencies of the particle can be shifted not only by changing its size, by also its shape.[73,74]

Almost complete absence of conductivity currents in silicon in the optical frequency range leads to low dissipative losses, in contrast to plasmonic structures where the strong field localization is always accompanied by high dissipation. Therefore, by exploiting dielectric particles with the magnetic response one can design different low-loss nanostructures, composite materials and metasurfaces with unique functionalities.

8.3. Methods of High-index Dielectric Nanoparticles Fabrication

Silicon is the most frequently used high-index dielectric in optical and IR ranges owing to its relatively low cost and low imaginary part of the refractive index. Moreover, the technology of fabrication of Si nanoparticles with Mie resonances has been developing intensively during the last several years, resulting in the emerging of various techniques. The proposed methods of Si nano- and

microparticles fabrication can be classified on the level of the particles size and location controllability. Here, we describe the various existing methods of the high-index nanoparticles fabrication, including the *chemical deposition*, the *thin film dewetting*, the *femtosecond laser ablation* of bulk and thin Si films, and the *reactive-ion-etching* approach.

Chemical deposition. Fabrication method of Si nanoparticles with different sizes can be carried out by means of chemical vapor deposition technique, in which disilane gas (Si_2H_6) decomposes at high temperatures into solid Si and hydrogen gas by the following chemical reaction: $Si_2H_6 \rightarrow 2Si(s) + 3H_2(g)$. Spherical polycrystalline Si nanoparticles were produced by this method in Ref. 47. Further, fabrication of monodispersed Si colloid was achieved via decomposition of trisilane (Si_3H_8) in supercritical n-hexane at high temperature in Ref. 75. In this advanced method, the particles size can be controlled by changing of trisilane concentration and temperature of the reaction. This relatively simple method allows one to obtain plenty of similar Si nanoparticles with the size dispersion of several percents, which can be ordered into hexagonal lattice via a self-assembly process [Fig. 8.4(b)]. The main disadvantage of this method is the porosity and high hydrogen content in each nanoparticle as well as the necessity of their additional ordering to fabricate functional structures.

Thin film dewetting. Disordered Si nanoparticles of different sizes can be also produced via dewetting of thin Si film after its heating [Fig. 8.4(b)].[76] In this case, the nanoparticles can be crystalline and their sides are aligned along crystallographic facets. The main controlling parameters in this method are the heating temperature and the film conditions (defects and initial pattern).[76] In thin film dewetting technique, the control over the nanoparticles size and location can be achieved only by using additional lithographical methods, which is even more complicated in comparison with the chemical deposition techniques. Indeed, the chemical deposition and thin film dewetting methods are more suitable for the high-throughput and low-cost nanoparticles fabrication.

Femtosecond laser ablation. In order to improve the control over the location of fabricated nanoparticles, the laser ablation by focused beam can be used. Indeed, an ultrashort laser pulse focused on the Si surface can heat the material up to the critical point, leading to the material fragmentation into spherical nanoparticles and their deposition nearby the heated area[25,44] [Fig. 8.4(c)]. It worth noting that the colloids of chemically pure nanoparticles can be obtained by means of the laser ablation as well as the chemical deposition. The main advantages of the ablation approach are the high-productivity and the lack of harmful chemical waste.

Fabrication of Si nanoparticles, demonstrating Mie-resonances in the visible range, with accurate control over their positions was

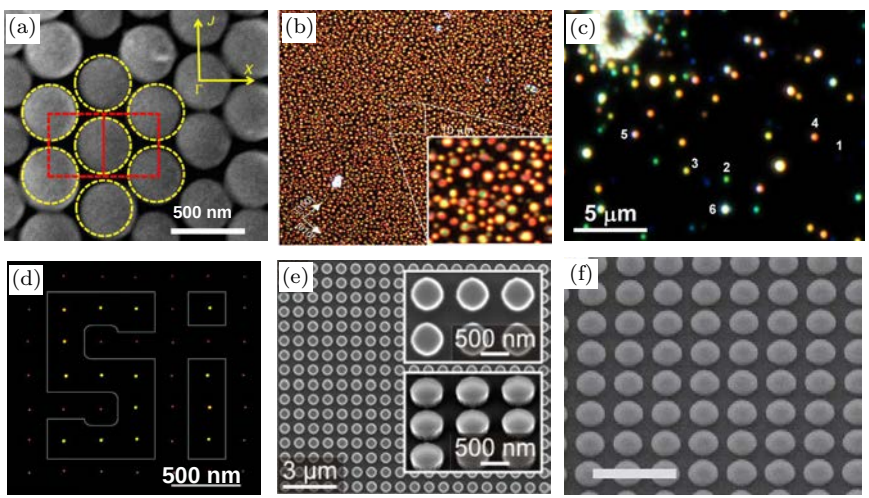


Fig. 8.4. (a) Electron microscopy image of self-aligned Si nanoparticles obtained by chemical deposition.[75] (b) Dark-field optical image of Si nanoparticles obtained by thin film dewetting.[76] (c) Dark-field optical image of Si nanoparticles obtained via femtosecond laser ablation of bulk Si.[44] (d) Dark-field optical image of Si nanoparticles obtained via femtosecond laser ablation of thin Si film.[77] In the picture (d) red nanoparticles are amorphized, while yellow ones are annealed and crystalline. (e) Electron microscopy image of Si nanoparticles obtained by means of reactive-ion-etching through a mask;[26] (f) the same but with additionally deposited Si_3N_4 thin film.[78]

carried out by focusing the femtosecond laser onto the Si surface, from which nanoparticles were emitted to the transparent receiver substrate[46,77,79] [Fig. 8.4(d)]. There are three main parameters that affect ablated Si nanoparticles: laser intensity, beam spatial distribution and sample thickness. For instance, a single Si nanoparticle with a certain size can be formed from bulk Si after irradiation by a single laser pulse with ring-type spatial distribution,[79] or after irradiation of thin Si film by conventional Gaussian beam.[77] The ultrashort laser can be used not only for fabrication but also for Si nanoparticles post-processing. In particular, the well-known effect of the laser annealing was applied for Si nanoparticles in order to controllably change them from the initially amorphized state to crystalline one, thus tailoring their optical properties [Fig. 8.4(d)].[77]

Reactive-ion-etching. The most controllable fabrication of Si nanoparticles was achieved by a multi-stage method, including electron-beam lithography on Si-on-insulator wafers (formation of a mask from resist) and the reactive-ion-etching process with following removing of the remaining electron-beam resist mask. This advanced technology enables the formation of Si nanocylinders [Fig. 8.4(e)], in which Mie-resonances can be precisely tuned by varying the basic geometrical parameters (diameter and height). Various types of structures based on the Si nanocylinders have been designed in order to show unique properties of the all-dielectric nanophotonics devices.[26,58,78,80] To achieve higher absorption of the fabricated Si metasurface, this method was supplemented by a deposition of Si_3N_4 thin film[78] [Fig. 8.4(f)]. Note that lithography-based methods have such serious disadvantages as high-cost and low-productivity of technological process in comparison with above mentioned lithography-free methods.

In order to summarize the section on methods of high-index dielectric nanoparticles fabrication, we would like to emphasize that currently developing approaches allow one to create various types of all-dielectric functional structures with given optical properties, which will be discussed below.

8.4. All-Dielectric Nanophotonics Applications

8.4.1. *Quantum source emission engineering*

The recently emerged field of optical nanoantennas is promising for its potential applications in various areas of nanotechnology. The ability to effectively emit light in a desired direction, redirect propagating radiation, and transfer it into localized subwavelength modes at the nanoscale[81] makes *optical nanoantennas* highly desirable for many applications. Originally, antennas were suggested as sources of electromagnetic radiation at radio frequencies and microwaves, emitting the radiation via oscillating currents. Different types of antennas were suggested and demonstrated for the effective manipulation of the electromagnetic radiation.[82] The recent progress in the fabrication of nanoscale elements allows bringing the concept of the radio frequency antennas to optics, leading to the development of optical nanoantennas consisting of subwavelength elements.[81] Currently, nanoantennas are used mainly for near-field microscopy,[83] high resolution biomedical sensors,[84] photovoltaics,[85] and medicine.[86] This section is devoted to the review of nanoantennas based on all-dielectric nanoparticles.

Apparently, the first work on the all-dielectric nanoantennas was Ref. 87. However, in this article, specific implementations of such nanoantennas have not been proposed. The first layout of *all-dielectric nanoantennas* has been proposed in Ref. 88 in 2011. In this work it has been shown that one Si nanoparticle can have the properties of a *Huygens element* in the optical wavelength range (Fig. 8.5(a)). It has been demonstrated, that such nanoantennas are able to switch the radiation pattern between forward and backward directions due to the presence of the electric and magnetic resonant modes. Then, in the Ref. 49 it has been shown that Si nanoparticles can be arranged in the *Yagi-Uda geometry* for creating highly efficient and directive optical nanoantennas (see Fig. 8.5(b)). The radii of the directors and the reflector have been chosen to achieve the maximal constructive interference in the forward direction along the directors array. The optimal performance of the Yagi-Uda nanoantenna has been achieved

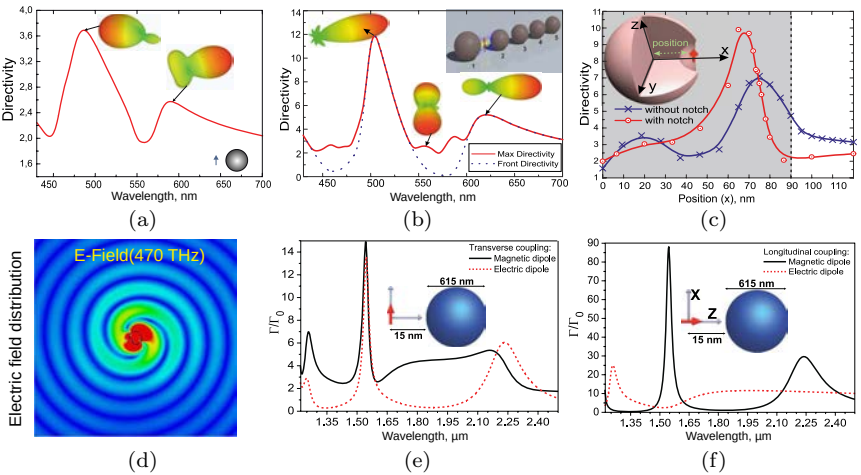


Fig. 8.5. (a) Directivity of the all-dielectric nanoantenna *Huygens element*, consisting of a Si nanoparticle and poind dipole source vs. the radiation wavelength.[49,88] (b) Directivity of the all-dielectric *Yagi-Uda nanoantenna*, consisting of the reflector with the radius of $R_r = 75$ nm, and smaller directors with the radii of $R_d = 70$ nm vs. the radiation wavelength.[49] Insert: 3D radiation patterns at particular wavelengths. (c) Maximum of the directivity of the *superdirective nanoantenna* depending on the position of the dipole source at the wavelength of 455 nm in the case of the Si nanosphere with (red curve) and without (blue curve) notch.[101] Inset: the geometry of the superdirective nanoantenna and the dipole source location in the notch. (d) Near field of the all-dielectric nanoantenna for chiral near-field formation and unidirectional excitation of electromagnetic guided modes.[55] (e) and (f) Normalized decay rates (Γ/Γ_0) as a function of the wavelength, for an emitter placed at 15 nm from the surface of a Si sphere of diameter 615 nm for transverse and longitudinal orientation of the dipole, respectively; the surrounding medium is air.[89]

for the radii of the directors correspond to the magnetic resonance ($R_d = 70$ nm), and the radius of the reflector, which corresponds to the electric resonance ($R_r = 75$ nm) at a given frequency. In Fig. 8.5(b) the directivity of the all-dielectric Yagi-Uda nanoantenna as a function of wavelength for the separation distance between the particles of $D = 70$ nm is presented. Inserts demonstrate the 3D radiation patterns at particular wavelengths. A strong maximum of the nanoantenna directivity at $\lambda = 500$ nm has been achieved. In this resonant regime the nanoantenna has the directivity about 12 and the

extremely narrow main lobe. The maximum of the dependence does not correspond exactly to either magnetic or electric resonances of a single dielectric sphere, which implies the importance of the interaction between constitutive nanoparticles. By comparing plasmonic and the all-dielectric Yagi-Uda nanoantennas, it has been demonstrated that the all-dielectric ones may exhibit better radiation efficiency also allowing more compact design.

These ideas were developed in the series of subsequent works.[89–97] It has been demonstrated that the unique optical properties and low dissipative losses make dielectric nanoparticles perfect candidates for design of high-performance nanoantennas, low-loss metamaterials, and other novel all-dielectric nanophotonic devices. The key to such novel functionalities of high-index dielectric nanophotonic elements is the ability of subwavelength dielectric nanoparticles to support simultaneously both electric and magnetic resonances, which can be controlled independently for particles of non-spherical forms.[26] Moreover, the magnetic field localization in a vicinity of all-dielectric nanoantennas has been theoretically predicted[98] and experimentally realized in the microwave[99] and in the visible[100] ranges.

Generally speaking, achieving of a high radiation directivity is a very important issue for the field of optical nanoantennas.[81] One of the most straightforward ways to achieve high directivity is to combine the nanoantenna constituent elements in the described above Yagi-Uda geometry. However, there is one more general approach consisting in the excitation of high-order multipole modes in the nanoantenna. In Refs. 53, and 102, the concept of the so-called *superdirective nanoantennas* has been proposed (see Fig. 8.5(c)). The superdirective regime has been achieved by placing an emitter (e.g. a quantum dot) within a small notch created on the nanosphere surface. The notch has the shape of a hemisphere. The emitter is shown in Fig. 8.5(c) by a red arrow. It turns out that such a small modification of the sphere allows for the efficient excitation of higher-order spherical multipole modes. Figure 8.5(c) shows the dependence of the maximum directivity $D_{\max}$ on the position of the emitting dipole

in the case of a sphere $R_s = 90$ nm without a notch at the wavelength $\lambda = 455$ nm (blue curve with crosses). This function achieves the maximum value ($D_{max} = 7.1$) when the emitter is placed inside the particle at the distance 20 nm from its surface. The multipole decomposition analysis has shown that, in this case, the electric field distribution inside a particle corresponds to the noticeable excitation of higher-order multipole modes. This becomes possible due to a strong inhomogeneity of the external field produced by the emitter. Furthermore, it has been demonstrated that the excitation efficiency of higher-order multipoles can be significantly improved by making a small notch on the Si nanoparticle surface and placing the emitter inside that notch, as shown in Fig. 8.5(c). The notch has the form of a hemisphere with the center at the dielectric nanoparticle surface. The optimal radius of the notch $R_n = 40$ nm has been found by means of numerical optimization. The red curve with circles in Fig. 8.5(c) shows the directivity maximum dependence on the dipole source location for this geometry. The maximal directivity at the wavelength of 455 nm is $D_{max} = 10$. Thus, Fig. 8.5(c) demonstrates the importance of the notch presence for achieving the higher directivity of the source radiation. Note that the proposed superdirectivity effect is not associated with high dissipative losses, because of the generally magnetic nature of the nanoantenna operation.

Recently, study of nanoantennas for formation of chiral distributions of the near-field has gained considerable interest.[103–105] In particular, in Ref. 105 it was shown that the chiral near-field can be produced by a symmetric non-chiral nanoantenna. In work[103] the chiral distribution of the near-field was investigated in the context of trapping and rotation of nanoparticles. In paper[106] it was demonstrated that excitation of the chiral near-field leads to the emergence of lateral optomechanical force acting on a chiral particle. Moreover, such nanostructures enable the generation of light beams with orbital angular momentum.[107] These effects are the consequence of the fundamental coupling between the spin angular momentum

of an evanescent wave and the direction of its propagation, known in the literature as the *spin-orbit coupling*.[108–110] The asymmetric excitation of high-index dielectric subwavelength nanoantenna by a point source, located in the notch at the nanoantenna surface has been studied in Ref. 55. The generation of the chiral near-field similar to that of a circularly polarized dipole or quadrupole depending on the frequency of the driving source has been demonstrated (see Fig. 8.5(d)). Using numerical simulations, it has been shown that this effect is the result of the higher multipole modes excitation within the nanoantenna. In this work it has been demonstrated that this effect can be applied for the unidirectional launching of waveguide modes in the dielectric and plasmonic waveguides. Contrary to the strategy employed in Refs. 111, 112, the directional launching of the guided modes achieved without a rotating or a circularly polarized point dipole source, but due to the violation of the rotational symmetry of the system.

One more important feature of the optical nanoantennas is their ability to exhibit strong *Purcell effect*. The Purcell effect is manifested in a modification of the spontaneous emission rate (Γ) of a quantum emitter induced by its interaction with inhomogeneous environment and is quantitatively expressed by the Purcell factor.[113–118] This modification is significant if the environment is a resonator tuned to the emission frequency. Open nanoscale resonators such as plasmonic nanoantennas can change the spontaneous emission lifetime of a single quantum emitter, that is very useful in microscopy of single NV centers in nanodiamonds,[119] Eu^{3+}-doped nanocrystals,[120] plasmon-enhanced optical sensing,[114] and the visualization of biological processes with large molecules.[121]

Mie resonances in dielectric particles can also increase the Purcell factor associated with either electric or magnetic transition rates in nearby quantum emitters. Their large quality factors compensate their low field confinement as compared to the plasmon resonances of metallic nanostructures for which nonradiative decay channels dominate. In Ref. 89 it has been shown theoretically that near-infrared quadrupolar magnetic resonances in Si nanoparticles can

preferentially promote magnetic versus electric radiative deexcitation in trivalent erbium ions at 1.54 μm (see Fig. 8.5(e,f)). The distance-dependent interaction between magnetic (electric) dipole emitters and induced magnetic or electric dipoles and quadrupoles has been derived analytically and compared to quasiexact full-field calculations based on Mie theory. The detailed analysis of the Purcell effect in the plasmonic and all-dielectric nanoantennas is presented in Sec. 8.4.3.

8.4.2. *Fano resonances in all-dielectric oligomers*

The Fano resonance[50,122,123] is known to originate from the interference of two scattering channels, one of which is nonresonant, while the other is strongly resonant. Fano resonance was observed in different areas of physics, including photonics, plasmonics, and metamaterials.[124] It has been demonstrated that the Fano resonance is highly sensitive to the optical properties of the background medium, which makes it perspective for the design of sensors. In the last few years, there is a growing interest in studying the Fano resonances in the so-called plasmonic *oligomer structures*, that consist of several symmetrically positioned metallic nanoparticles.[125,126] In such structures, the Fano resonance appears as a resonant suppression of the scattering cross-section of the structure, and it is accompanied by a strong near-field enhancement and, consequently, absorption.

Recently, it was shown that the oligomers composed of high-index dielectric nanoparticles are also able to exhibit the Fano resonance.[50–52,127–130] The important feature of such dielectric oligomers, comparing to their metallic counterparts, is the localization of the electromagnetic field inside the dielectric nanoparticles. Another important property of such oligomers is that the fundamental mode of the high-index spherical nanoparticle is magnetic dipole mode.[44,46] Formation of this magnetic mode, as was mentioned above, is due to excitation of a circular displacement current. It occurs when the diameter of the particle is comparable to the wavelength inside the nanoparticle.

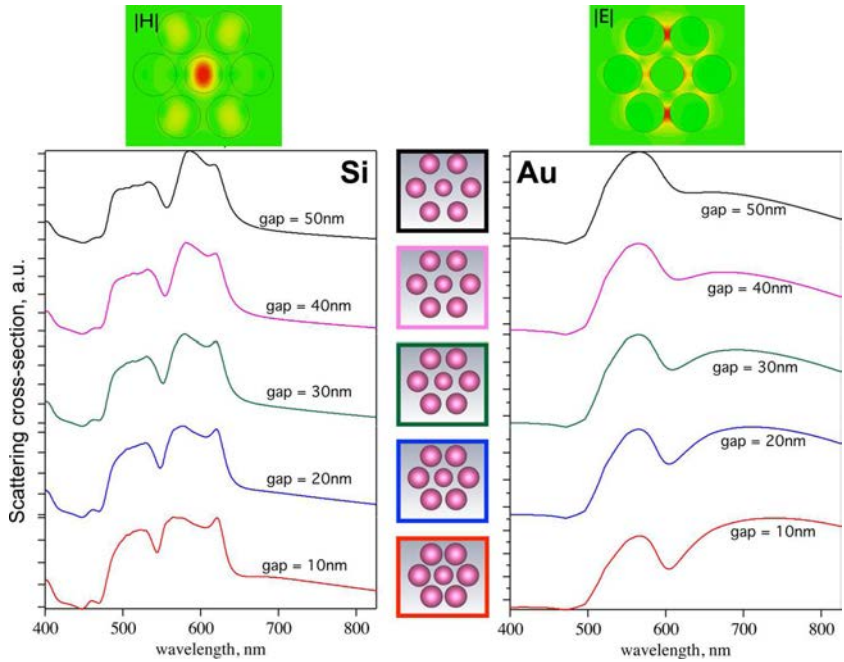


Fig. 8.6. Scattering cross-section of oligomer of silicon (left) and gold (right) nanoparticles for various separation between the particles. The radius of the central particle is $R_1 = 65$ nm and outer particles $R_2 = 75$ nm. The Fano resonance depends weakly on the separation between particles for all-dielectric oligomers, while for plasmonic analogue this dependence is very strong. It demonstrates a difference in the coupling mechanism in both situations. From Ref. 50.

Authors of Ref. 50 have shown that the structure arranged of six identical dielectric nanoparticles, positioned in the vertices of the regular hexagon, and the particle of another radius in the center [see Fig. 8.6], exhibits the Fano resonance at the resonance frequency of the central particle, while six other particles are not resonant at this frequency and they form a non-resonant mode of the whole structure. The near-field interference of this two modes leads to the suppression of the whole structure scattering and formation of the Fano resonance.[50] In Fig. 8.6 the scattering cross-section spectra of oligomer of silicon (left) and gold (right) nanoparticles for various separation between the particles are presented. It has been demonstrated that

the Fano resonance depends weakly on the separations between particles for all-dielectric oligomers, while for plasmonic analogue this dependence is very strong (see Fig. 8.6).

In Ref. 52 the existence of the Fano resonances in dielectric oligomers has been demonstrated for the first time. Due to the scalability of Maxwell equations, the authors used microwave ceramic spheres with sizes of several centimeters (instead of Si nanoparticles). Such particles exhibit the magnetic response in the microwave frequency range. The authors measured the near magnetic field in the vicinity of the dielectric oligomer with high accuracy, which allowed to verify the origin of the Fano resonance, predicted in the theoretical study.[50]

8.4.3. *Surface enhanced spectroscopy and sensing*

Resonant nanoparticles and nanostructures are proven to be powerful tools for sensing applications due to their ability to enhance and localize the optical energy in the near-field, whereas the positions of resonances depend on ambient media properties.[132–135] Moreover, nanoscale resonators provide so-called Purcell effect,[97,118,136,137] when the power radiated by a quantum light source (atom, molecule, quantum dot, etc.) is enhanced due to the increase of the local density of states (LDOS). There are numerous sensing techniques based on all these effects: surface enhanced Raman scattering (SERS),[138–140] surface enhanced fluorescence (SEF),[141] Forster resonance energy transfer,[142–144] refractive index sensing,[145–148] and thermometry.[149] Despite plasmonics has demonstrated tremendous success in the sensing applications, the concept of all-dielectric nanophotonics can also serve as a platform for high-effective sensors. First of all, low Ohmic losses in all-dielectric resonant nanostructures prevent parasitic heating of the analyzed objects,[98,131,150] and second, high radiative part of Purcell factor and directivity improve signal extraction.[54]

Figure 8.7 shows the enhancement of the radiative decay rate and quantum efficiency of an electric dipolar emitter positioned between two silicon (a–b) and gold (c–d) nanospheres, and the results of temperature measurement in such nanoantennas (e–g). Figure 8.7

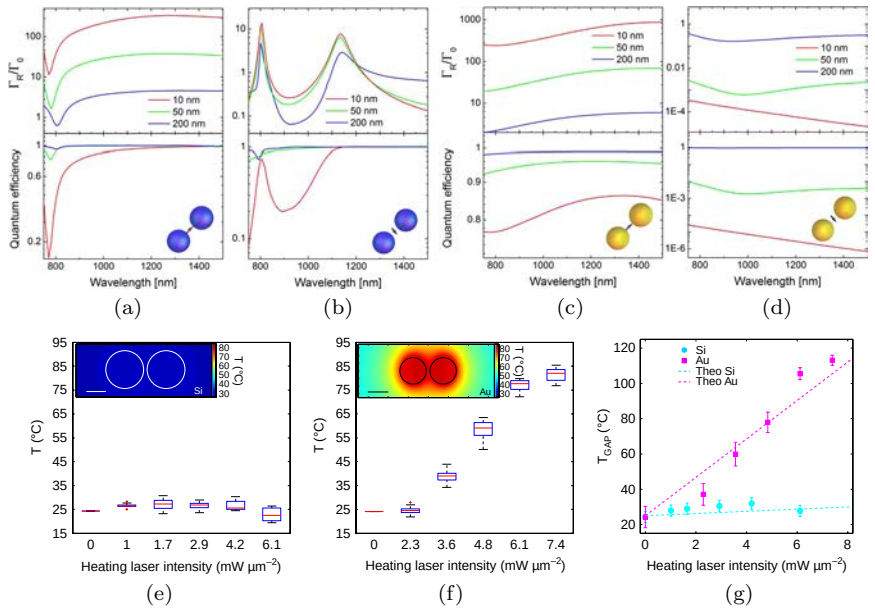


Fig. 8.7. Enhancement of the radiative decay rate and quantum efficiency of an electric dipolar emitter positioned in between two silicon ((a),(b)) and gold ((c),(d)) nanospheres of 150 nm radius. Orientations of the emitters positioned at the centers of the systems are shown in the schematics. Gap widths are given in the legends. (e), (f), (g) — Temperature measurement in nanoantennas. Box plot for the average temperature T, measured for (e) silicon and (f) gold nanoantennas, excited at resonance. The inset in each figure shows the calculated temperature map around the disks for the heating laser intensity of 5 mW μm^{-2} in both cases. Scale bar is 100 nm. (g) Extracted temperature in the gap for selected silicon (cyan) and gold (magenta) nanoantennas as a function of the heating laser intensity at 860 nm. The dashed lines show the numerical calculations for the temperature at the gap, presenting good agreement with the experimental data. The error bars show the s.d. of the temperature measurements, obtained from error propagation from the fluorescence measurements. From Refs. 98,and 131.

(a–d) shows that the all-dielectric dimer nanoantennas have an ability to strongly enhance both electric and magnetic LDOS, whereas the plasmonic nanoantennas work only with electric dipole sources. It can be seen that the all-dielectric nanoantennas have the quantum efficiency exceeding that of the plasmonic nanoantennas. Moreover,

it also can be seen that the Au nanoantennas significantly increase their temperature when the heating laser intensity increases while the Si temperature remains nearly constant and does not affect the molecules under study.

From the perspective of enhancement by individual nanoparticles, the absence of cut-off frequency for dipole plasmon resonance results in much higher field enhancement near plasmonic nanoparticles as compared to dielectric ones with same sizes in the sub-100-nm range. However, larger Si nanoparticles, possessing a magnetic Mie-type resonance at the optical frequencies, yield comparable or even larger near-field enhancement.[98,131] This effect was proved in SERS experiments, where Si resonant nanoparticles produce larger SERS effect as compared to gold ones of the same sizes.[151]

In order to get a huge local field enhancement, plasmonic dimers[152] or oligomers are used.[133,153–155] The same approach is also possible for dielectrics, when local field enhancement factor in the gap of a Si dimer could be more than one order of magnitude.[100] Such enhancement was applied to achieve high SERS and SEF effects.[98,131,150] Therefore, all-dielectric nanostructures also provide the field enhancement, which is high enough for detection of small amount of organic materials.

Figure 8.8 shows the experimental results of all-dielectric nanoantennas application for surface enhanced Raman scattering (a) and for surface enhanced fluorescence (b). It has been demonstrated that the Si-dimer nanoantennas exhibit high near-field enhancement within a 20 nm gap at the near IR wavelengths. This all-dielectric nanoantenna is able to enhance the Raman scattering of a polymer thin film by a factor of 10^3 (a) and also allow surface enhanced fluorescence by a factor of 2×10^3, avoiding the well-known fluorescence quenching effects observed for metallic structures when no spacer layers are used. Moreover, the molecular thermometry measurements have demonstrated that the dielectric nanoantennas produce ultra-low heating when illuminating at their resonance wavelength, thus overcoming one of the main drawbacks of traditional plasmonic materials such as gold.

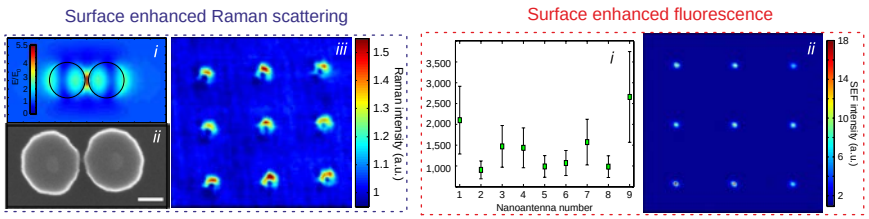


Fig. 8.8. (a) Surface enhanced Raman scattering (SERS) via all-dielectric dimer nanoantenna. (i) Near-field distribution map for the silicon structure excited at resonance, showing good confinement of the electric field in the gap; the maximum enhancement value is 5.5. (ii) The SEM image of the Si nanoantenna fabricated on a silicon-on-insulator substrate. The nanoantenna consists of two identical disks with a diameter of 220 nm, a height of 200 nm, and a 20 nm gap in between. (iii) Experimental 2D normalized Raman map, showing enhanced signal coming from the molecules close to the nanoantennas. The SERS enhancement factor reached in experiments is about 1500. (b) Surface enhanced fluorescence via all-dielectric dimer nanoantenna. (ii) Experimental SEF map obtained for the Si antennas. It can be clearly seen that fluorescence is enhanced over the nanoantennas. (i) SEF enhancement factor (FSEF) obtained from the maximum values over each nanoantenna in (ii). The error bars show half the difference between the minimum and the maximum value in each nanoantenna. From Ref. 131.

It is worth noticing that placing a detected nanoobject in a "hot spot" of a resonant nanostructure could be less effective for sensing than positioning it in a place with the highest Purcell factor. Moreover, high Purcell factor can be achieved in "cold spots" of the nanostructure.[118] So, the conceptually different approach to extract more signal from quantum emitters is to enhance Purcell factor within the nanostructure. In case of all-dielectric nanostructures, a periodical chain of resonant nanoparticles, supporting magnetic dipole resonance, looks a promising device to achieve huge radiative part of Purcell factor for an electric dipole source,[118] whereas dielectric material does not lead to any quenching effects.

Another important feature of the dielectric nanoparticles is the ability to shift the incident light frequency via the *Raman scattering* process. The Raman scattering is inherent for dielectric materials, but almost does not exist in metals. Moreover, an intensity of the Raman scattering strongly depends on the resonant properties of

dielectric nanoparticle. In particular, magnetic types of low-order Mie resonances provide much larger enhancement factors as compared to corresponding electric types of Mie resonances.[30] The enhancement factor dependence on excitation wavelength demonstrates sharp peak at a magnetic dipole resonance, demonstrating more than two orders of magnitude variation in its vicinity and narrower width (∼10 nm) as compared the width of the Mie resonance (∼30 nm). This effect could also be promising for a number of applications related to sensitivity of Raman signal to thermal and refractive index variations of ambient medium.

8.4.4. *Coupled-resonator optical waveguides*

A design of highly efficient integrated circuits with combined optical and electronic components for the subwavelength guiding of the electromagnetic energy is one of the main trends of the optical communications technologies in the beginning of the 21st century.[156] In order to achieve high integration densities, optical waveguides with subwavelength light localization have been proposed. Among those are conventional Si (or other dielectric/semiconductor) nanowires, photonics crystal waveguides and plasmonic waveguides. Silicon nanowire waveguides have very small cross-section sizes, and they can be manufactured being of a high quality.[157] However, such waveguides do not provide low-loss propagation of optical signals through sharp bends and require rather large bending geometries thus increasing the overall size of an optical chip.[158] Photonic crystals have been viewed as a possible alternative, and it has been already demonstrated that light can be guided by a waveguide composed of defects, and that such waveguides can have sharp bends.[159] However, due to the different mechanism of waveguiding, namely Bragg reflection, the overall transverse size of the photonic crystal waveguide is usually about several wavelengths. Besides, the nice property of photonic crystals to propagate light through sharp bends was found to depend strongly on the bend geometry being also linked to the strict resonant conditions associated with the Fano resonance where the waveguide bend plays a role of a specific localized defect,[159] thus

demonstrating narrowband functionalities. Conventional plasmonic waveguides allow for truly subwavelength localization of light, but it is always accompanied with severe Ohmic losses, which makes the propagation lengths of surface plasmons impractically short.[160]

Another candidate for the efficient subwavelength guiding is a coupled-resonator optical waveguide, where guided modes are formed by coupled resonances of the single elements.[162] The most recent realization of such type of waveguide was suggested in Ref. 163 in the form of the chain of high-index low-loss dielectric nanoparticles. Such waveguide was fabricated and its properties were measured several years later.[164] In Ref. 161 it was shown that guiding of the electromagnetic energy in dielectric discrete waveguides is achieved due to coupled Mie resonances. In addition, since dielectric nanoparticles support both MD and ED resonances simultaneously, waveguides composed of such nanoparticles support several modes of different types. For the case of spherical particles, first several modes (at low frequencies) calculated within the framework of the dipole approximation and via full-wave numerical simulations are shown in Fig. 8.9(a). Two longitudinally (transversely) polarized modes marked LE and LM (TE and TM) are formed by coupling between MD and ED dipoles, oriented along (perpendicular to) the axis of the chain. Note that in the case of transverse modes there is a coupling between EDs and MDs induced in different particles, while longitudinal modes are independent of each other. However in spherical particles ED and MD resonances are well separated in frequency and, consequently, the coupling between EDs and MDs is quite small. Therefore coupled resonances of nanospheres form quasi-independent separate pass bands in different spectral ranges, due to the different dipole-dipole interaction strength.

In Fig. 8.9(b) the results of modelling of light transmission through the chain of 6 particles are presented. One can observe a transmission band around $ka/\pi = 0.5$ formed by excited TM and LM modes [Figs. 8.9(f,g)]. Transmission band around $ka/\pi = 0.7$ is formed by multipole modes. The most high-frequency peak corresponds to the longitudinal magnetic quadrupole mode with $\beta = 0$

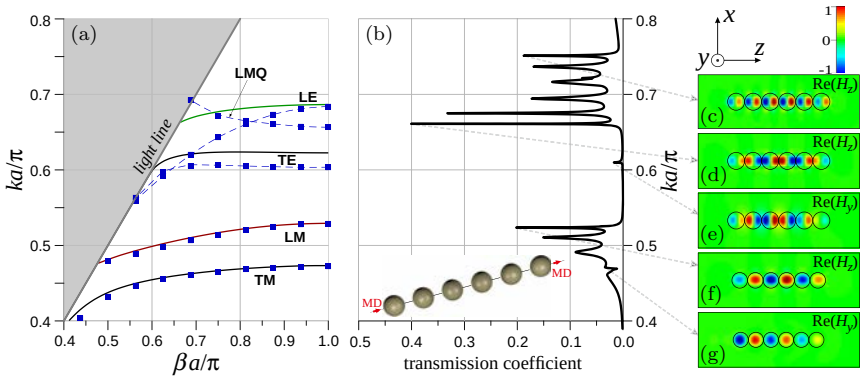
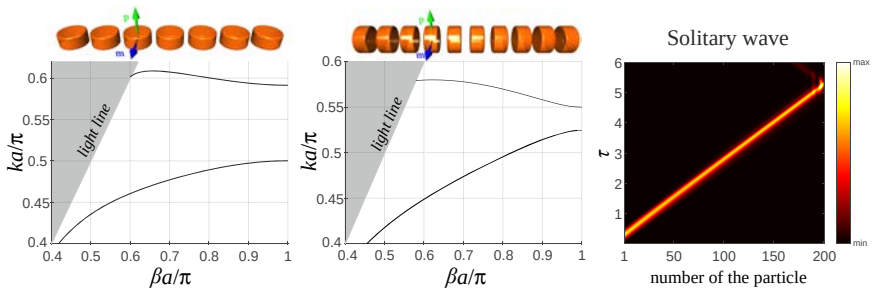


Fig. 8.9. (a) Dispersion diagram of an infinite chain of lossless spherical Si nanoparticles with radius 70 nm and period $a = 140$ nm. Only waveguide modes under the light line $\beta = k$ are shown. Numerically calculated eigenmodes are shown with blue squares (and thin dashed blue lines). Oblique grey line is the light line. (b) Numerically calculated transmission spectrum of a chain of 6 Si spheres. (c–g) Electric field distributions in the corresponding modes. The operational range of normalized frequenices ka/π lies within an optical spectral range for the chosen parameters. Adopted from the Ref. 161.

[Fig. 8.9(c)]. This mode crosses the light line (i.e. it is a radiating leaky wave), and therefore it is not shown in Fig. 8.9(a), where only unattenuated modes are present. Numerically found frequency for $\beta = 0$ is $ka/\pi \approx 0.76$, which coincide with the value in transmission spectrum at the upper edge of the longitudinal magnetic quadrupole band. One can also see a transmission peak at $ka/\pi \approx 0.61$ [Fig. 8.9(e)] corresponding to the TE mode, which is also excited due to the inhomogeneity of current in the probes.

For nonspherical particles, resonance frequencies depend on the orientation of the dipole oscillations, and the corresponding passbands can be shifted by changing the particle parameters. Besides the shifting of the operational frequency range, change of the resonance frequencies can also substantially increase the coupling between EDs and MDs of the neighboring particles. This happens for transversely polarized modes when resonance frequencies of EDs and MDs oriented perpendicular both to the axis of the chain and to each other get closer. In Figs. 8.10(a,b) it is shown that second branch of the



(a) (b) (c)

Fig. 8.10. (a,b) Numerically calculated dispersion curves for the transversely polarized modes of an infinite chain of dielectric nanocylinders (a) with radius $R = 90$ nm and height $h = 180$ nm and (b) with radius $R = 130$ nm and height $h = 120$ nm; period is $a = 200$ nm in all cases. $\beta a/\pi$ is the normalized Bloch wave number and ka/π is the normalized frequency. (c) Calculated nonlinear dynamics of the pulse propagating in the chain of dielectric nanocylinders with dispersion properties shown in (b) in the nonlinear regime, excited with a Gaussian 100 fs pulse with normalized center frequency 0.553. Adopted from Ref. 67.

chain of cylindrical nanoparticles with close ED and MD resonance frequencies changes the sign of group velocity and group velocity dispersion in a certain frequency range. One can see, that while it is not possible for the chain of spherical nanoparticles, the interplay of the ED and MD resonances in the cylindrical nanoparticles with certain parameters can induce the anomalous dispersion regime in the discrete waveguide. This feature was employed in the study of the nonlinear regimes of femtosecond optical pulse propagation through all-dielectric waveguides.[67] It has been shown that for the chains of cylindrical particles the broadening of the propagating pulse can be compensated by the nonlinear Kerr effect, thus making possible the formation and propagation of *solitary waves*.

In Ref. 65, the transmission efficiency of the discrete waveguides composed of arrays of high-index dielectric nanodisks with and without sharp bends has been studied. The appropriate period of the chain has been chosen so that longitudinal and transverse pass bands of the nanodisk chain, formed by coupled MD resonances of nanodisks, overlap (which cannot be done with spherical particles). This condition allows to realize an efficient transmission through sharp

90° bends. In the straight chain of nanodisks two passbands formed by Fabry-Perot resonances of mixed LM-TM modes and TE modes of the finite chain were observed in transmission efficiency spectrum. In the waveguide composed of 30 disks with a 90° bend in the middle the maximum value of the transmission efficiency of about 0.6 was achieved. Also it has been shown that the LM mode in the horizontal branch can transform into the TM mode propagating in the vertical branch and vice versa. These theoretical conclusions have been supported by presenting the experimental results for the microwave frequencies for an efficient guiding through 90° bend in a microwave dielectric waveguide.

Recently, several studies proposed the chains of high-index dielectric nanoparticles as a simple 1D platform for studying *bound states in the continuum* (BICs).[165,166] The BICs are the localized states that exist within the continuum spectrum of radiative waves.[167,168] In the Ref. 165 TE and TM BICs in a linear periodic array of dielectric spheres were demonstrated. Field distribution and force lines of the simplest TM BIC in a zero diffraction channel with zero azimuthal wavenumber and zero Bloch wavenumber are shown in Figs. 8.11(a,b), respectively. Different TE- and TM-polarized BICs with non-zero azimuthal numbers and non-zero Bloch wavenumbers were also predicted. Such peculiar states manifest themselves in the spectra of scattering cross-section of a plane wave by infinite array of nanoparticles. For the parameters of the system close to the existence of a BIC, Fano-type resonance occurs due to the interference between the plane wave and the localized mode of the array [see dashed green and solid red curves in Figs. 8.11(c,d)]. For the parameters exactly corresponding to the BIC state, quality factor of the localized state tends to infinity which results in the collapse of the Fano resonance [see dashed blue curves in Figs. 8.11(c,d)].

In the Ref. 166 light guiding above the light line in the chain of dielectric nanospheres was also demonstrated. It was shown that at the frequencies close to the BICs light can propagate to the large distances in both stationary and pulse regimes. In the stationary regime the possibility of waveguiding at different frequencies can

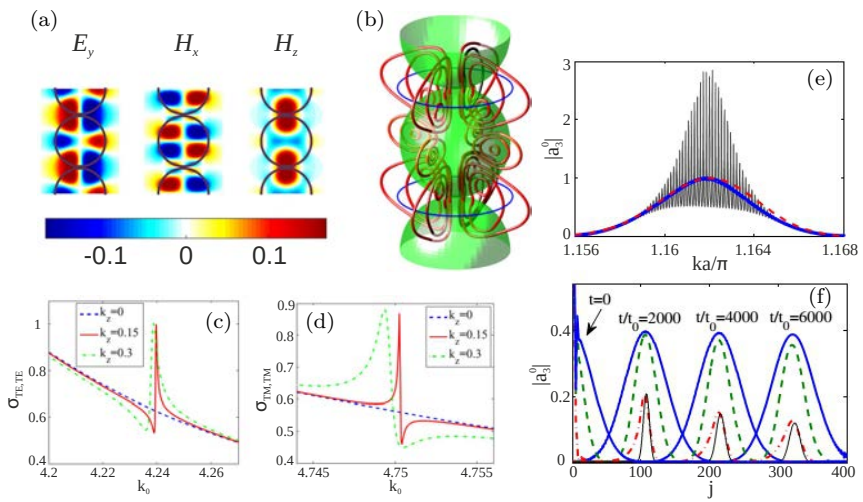


Fig. 8.11. (a) The real parts of the EM field components and (b) the electric force lines (red) and the magnetic force lines (blue) of the symmetry-protected bound state embedded into the TM continuum with zero azimuthal number and zero Bloch wavenumber. (c,d) Total scattering cross section of the array illuminated by a plane wave. Scattering of the (c) TE and (d) TM plane waves is strongly affected by the presence of the symmetry-protected BIC. (e,f) Light propagation in the array of 400 nanoparticles. Absolute value of the leading Mie coefficient versus (e) wave number of the stationary wave injected into the array and (f) the particle number j for a light pulse with different pulse widths. Adopted from the Refs. 165, and 166.

be determined from the standing waves formed in finite chain of 400 nanoparticles [Fig. 8.11(e)]. In Fig. 8.11(f) moments induced in nanoparticles during the pulse propagation are shown for different pulse widths. Such calculations revealed that pulses with a certain width and the central frequency tuned exactly to the frequency of the BIC can propagate in the chains of dielectric nanospheres for the distances of tens and hundreds wavelengths.

Thus, chains of dielectric nanoparticles provides one with the simple and efficient platform, allowing guiding and localization of light in linear and nonlinear regimes. The discrete waveguides based on high-index dielectric nanoparticles may exceed its currently existing analogs: plasmonic waveguides, dielectric photonic crystals, and

homogeneous Si waveguides, offering a large number of customizable options and negligible energy dissipation. Such waveguides can be used in photonic components responsible for the transmission of information in the optical and optoelectronic integrated circuits.

8.4.5. *Metamaterials and metasurfaces*

Future technologies will push for a steep increase in photonic integration and energy efficiency, far surpassing that of bulk optical components, Si photonics, and plasmonic circuits. Such level of integration can be achieved by embedding the data processing and waveguiding functionalities at the level of material rather than a chip, and the only possible solution to meet those challenges is to employ the recently-emerged concept of *metamaterials and metasurfaces*. Metamaterials are artificial media with exotic electromagnetic properties not available in natural media which are specially created in order to reach functionalities required for particular applications.[11,169] Metasurfaces are their two-dimensional implementations that are much simpler for fabrication.[170] Metamaterials have been studied since 2000 and revealed such effects as negative refraction, backward waves, beating of diffraction limit (subwavelength imaging), and became a paradigm for engineering electromagnetic space and controlling propagation of waves by means of transformation optics.[169–171] The research agenda is now focusing on achieving tunable, switchable, nonlinear and sensing functionalities of metamaterials. Since 2010 the studies have been shifted to the stage of practical implementation and development of real metadevices. As a result, a novel concept of metadevices, that can be defined as metamaterial-based devices with novel and useful functionalities achieved by the structuring of functional matter on the subwavelength scale, has been developed.[11] The metadevices practical implementation is the general trend in the area of metamaterials.

The area of metamaterials has opened a broad range of technologically important capabilities ranging from the subwavelength focusing to "stopped light", including their ability to control magnetic response of novel subwavelength structured materials. This is

important because the magnetic response of natural materials at optical frequencies is very weak due to the diminishing of electronic spin states at high frequencies.[3] That is why only the electric component of light is directly controlled in conventional photonic devices. However, effective control of light at the nanoscale requires the presence of the electric and magnetic responses, simultaneously. A vast majority of the current metamaterial structures exhibiting magnetic response contain metallic elements with high conductive losses at optical frequencies, which limits their performance. As it was mentioned above, one of the canonical examples is a split-ring resonator that is an inductive metallic ring with a gap that is a building block of many metamaterials. This concept works very well for gigahertz,[8,14] terahertz[15] and even near-infrared (few hundreds THz)[16] frequencies. However, for shorter wavelengths and in particular for the visible spectral range, this concept fails due to the dissipative losses and fabrication difficulties.[17]

In order to overcome these fundamental problems, an alternative approach of all-dielectric metamaterials has been proposed.[18–25,27,70,172,175,176] In this case, a high-index dielectric particle, exhibiting magnetic and electric Mie resonances plays a role of a single *meta-atom*. Such high-index dielectric particles replace their metallic counterparts in metamaterials and metasurfaces. For example, it has been shown that the 3D dielectric composite of high-index dielectric particles [as shown in Fig. 8.12(a)] exhibits the negative permeability near the first Mie resonance.[20,172] Even more complex geometry of such all-dielectric metamaterials have been proposed in Refs. 173, 174. For theoretical treatment of such composites consisting of high-index dielectric spherical particles embedded in a low-index dielectric matrix the Levins model can be used.[172,177]

Conventional optical components rely on gradual phase shifts accumulated during light propagation to shape light beams. The nanostructured design can introduce new degrees of freedom by making abrupt phase changes over the scale of the wavelength. A two-dimensional lattice of optical resonators or nanoantennas on a planar surface, with the spatially varying phase response and subwavelength

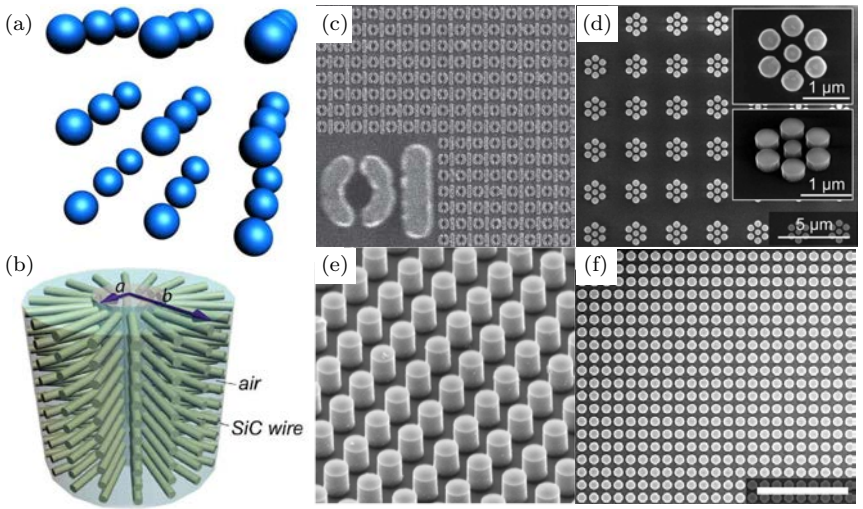


Fig. 8.12. (a,b) All-dielectric metamaterials based on spherical and cylindrical particles;[22,172–174] (c-f) all-dielectric metasurfaces.[68,70,78]

separation, can imprint such phase patterns and discontinuities on propagating light as it traverses the interface between two media. In this regime, anomalous reflection and refraction phenomena can be observed in optically thin metamaterial layers, or optical metasurfaces (see Fig. 8.12(c–f)), creating surfaces with unique functionalities and engineered reflection and transmission laws. The first example of such metasurface — a lattice of metallic nanoantennas on Si with a linear phase variation along the interface, was demonstrated recently.[178–182] The concept of metasurfaces with phase discontinuities allows introducing generalized laws for light reflection, and such surfaces provide great flexibility in the control of light beams, being also associated with the generation of optical vortices. Metasurfaces can also be used for the implementation of important applications such as light bending[178] and specific lenses.[183]

The phase gradient metasurfaces created by high-index nanoparticles of varying shape have been recently proposed.[68,184] In Ref. 68, for the first time, highly efficient all-dielectric metasurfaces for near-infrared frequencies using arrays of Si nanodisks as meta-atoms

have been proposed and realized [see Fig. 8.12(f)]. The authors have employed the main features of Huygens' sources, namely spectrally overlapping electric and magnetic dipole resonances of equal strength, to demonstrate Huygens' metasurfaces with a full transmission-phase coverage of 360 degrees and near-unity transmission, and confirmed experimentally full phase coverage combined with high efficiency in transmission. Based on these key properties, the authors show that all-dielectric Huygens' metasurfaces could become a new paradigm for flat optical devices, including beam-steering, beam-shaping, and focusing, as well as holography and dispersion control.

We also should note here that the studies in the field of dielectric metamaterials in microwave frequency range performed until 2009 are summarized in the review paper.[185] The current state of research in this area given in Refs. 29, 186–188 and especially in the excellent Review paper by S. Jahani and Zubin Jacob.[176]

8.4.6. *Nonlinear nanophotonics applications*

An enhancement of nonlinear optical response at the nanoscale is also challenging area of nanooptics, where dielectric materials have already been implemented for various micro-devices. In particular, Raman lasing,[189,190] supercontinuum generation,[190] and all-optical switching[59,171,190] are the bright examples of nonlinear photonics based on Si micro-devices (waveguides, ring-resonators, photonic crystals etc.). Indeed, inherent nonlinear response of many dielectrics (especially, semiconductors) is very high in the optical and IR range, being comparable with metals or even much larger due to non-centrosymmetrical crystalline lattice of some dielectric materials (GaAs, GaP, Te, etc.). On the other hand, plasmonic nanodevices paved the way to creation of deeply subwavelength nonlinear devices.[171,191] Therefore, implementation of plasmonic principles for developing of all-dielectric nonlinear nanodevices looks tempting. Recently, the enhancement of optical nonlinearities in resonant Si nanostructures has been demonstrated theoretically and experimentally at the scale of single nanoparticles.[58–61,192–195]

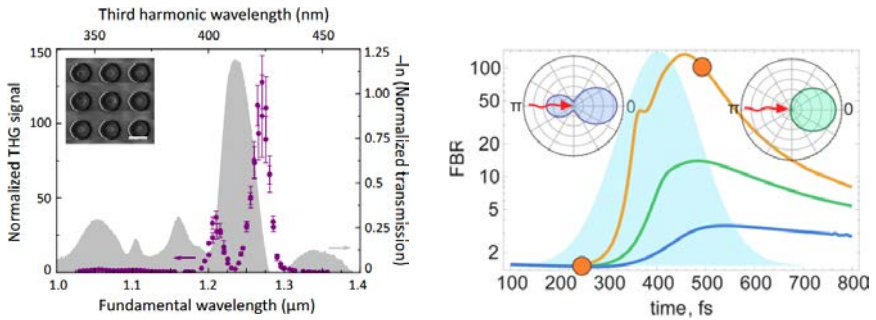


Fig. 8.13. (a) Purple dots: third harmonic generation spectrum of the nanodisks array (shown in the inset). The gray area represents the linear transmission spectrum. (b) Dynamical reconfiguration of a Si nanoantenna directivity via photoexcitation: Front-to-Back ratio of a nanoparticle during the action of a 200 fs pulse with peak intensities of 10 GW/cm^2 (blue), 20 GW/cm^2 (green) and 40 GW/cm^2 (orange). Scattering diagrams of the incident beam at the largest intensity are shown in the two insets.

One of the most attractive applications of all-dielectric nanostructures is the efficient *frequency conversion*. In the pioneering work on this topic[58] enhancement of third-harmonic generation from Si nanoparticles (in form of nanodisks) exhibiting both electric and magnetic dipolar resonances has been demonstrated, Fig. 8.13(a). The efficiency of IR-to-visible conversion by 2 orders of magnitude in the vicinity of the magnetic dipole resonance with respect to the unstructured bulk Si slab was achieved. The idea of the conversion enhancement at the magnetic resonance has been developed in subsequent works with regard to the generation of higher optical harmonics[127,194,196,197] and Raman scattering.[30] Dielectric oligomers[196] and nanoparticles supporting the anapole mode excitation[198] have also been employed for third harmonic generation enhancement.

Another intriguing non-linear effect arising in resonant Si nanoparticles is the electron-hole plasma photoexcitation. Silicon is a semiconductor, and at normal conditions, its conduction band is almost empty. However, optical absorption causes the electrons to fill the conduction band thus altering its permittivity and optical response.[199] Recently, the photoexcitation of plasma was employed

for tuning of Si nanoantenna optical properties in the IR and visible regions.[60,61,193,194] It has been shown that the plasma photoexcitation allows for manipulating of electric and magnetic nanoparticle responses, resulting in dramatic changes of both scattering diagram and scattering cross section, Fig. 8.13(b). The 20% tuning of reflectance of a single Si nanoparticle by femtosecond laser pulses with the wavelength in the vicinity of the magnetic dipole resonance has been demonstrated. In the recent work[61] this effect has been utilized for achieving a pronounced *beam steering effect* in an all-dielectric dimer nanoantenna.

Last but not least, resonant dielectric nanoparticles demonstrate *much higher damage threshold*. For comparison, the typical values of damage threshold of metallic nanostructures are: gold nanorods ($\sim$70 GW/cm^2 or $\sim$10 mJ/cm^2 at 130 fs[200]), gold G-shaped nanostructures ($\sim$100 GW/cm^2 or $\sim$3 mJ/cm^2 at 30 fs[201]), and gold nanocylinders ($\sim$200 GW/cm^2 or $\sim$20 mJ/cm^2 at 100 fs[202]). According to the known data from the literature, low-loss Si nanoparticles have significantly higher damage threshold: $\sim$400 GW/cm^2 or $\sim$100 mJ/cm^2 at 250 fs;[194] and $\sim$1000 GW/cm^2 or $\sim$100 mJ/cm^2 at 100 fs.[193] Such considerable difference in damage thresholds for plasmonic (e.g. gold) and all-dielectric (e.g. Si) materials originates form difference in their melting temperatures (Tm(Au)=1337 K and Tm(Si)=1687 K), and enthalpies of fusion (H(Au)=12.5 kJ/mol and H(Si)=50.2 kJ/mol). Therefore, silicon-based nanostructures are more stable than plasmonic ones upon intense laser irradiation, which makes them very attractive for nonlinear applications.

8.5. Conclusions and Outlook

In this Chapter, we have reviewed some of the recent developments in the field of all-dielectric nanophotonics. This area of optical science studies the light interaction with high-index dielectric nanoparticles supporting optically-induced electric and magnetic Mie resonances. We have described several advances in this field which demonstrate that dielectric structures allow to control both magnetic and electric components of light in a desirable way, and also discuss properties

of high-index nanoparticles along with their fabrication methods. We have reviewed practical applications of all-dielectric nanophotonics, including the nanoantennas for the quantum source emission engineering, the all-dielectric oligomers and their Fano resonances, the surface enhanced spectroscopy and sensing, coupled-resonator optical waveguides, optical solitons and bound-states-in-continuum, all-dielectric metamaterials and metasurfaces, and the nonlinear nanophotonics.

Future technologies will demand a huge increase in photonic integration and energy efficiency far surpassing that of bulk optical components and silicon photonics. Such an integration can be achieved by embedding the data-processing and waveguiding functionalities at the material level, creating the new paradigm of metadevices. It is now believed that robust and reliable metadevices will allow photonics to compete with electronics not only in telecommunication systems but also at the level of consumer products. The main challenges in achieving this goal will be in developing cost-efficient fabrication and device integration technologies. All-dielectric nanophotonics is seen as a practical way to implement many of the important concepts of nanophotonics allowing high functionalities and low-loss performance of metadevices.

References

1. Caulfield, H. J. and Dolev, S. (2010). Why future supercomputing require optics?, *Nat. Photon.* **4**, pp. 261–263.
2. Yan, Y., Xie, G., Lavery, M. P. J., Huang, H., Ahmed, N., Bao, C., Ren, Y., Cao, Y., Li, L., Zhao, Z., Molisch, A. F., Tur, M., Padgett, M. J. and Willner, A. E. (2014). High-capacity millimetre-wave communications with orbital angular momentum multiplexing, *Nature Communications* **5**, p. 4876.
3. Landau, L. D., Lifshitz, E. M. and Pitaevskii, L. P. (1984). *Electrodynamics of Continuous Media*. (Butterworth-Heinemann, London).
4. Merlin, R. (2009). Metamaterials and the landaulifshitz permeability argument: Large permittivity begets high-frequency magnetism, *Proc. Natl. Acad. Sci. USA* **106**, p. 1693.
5. Burresi, M., van Oosten, D., Kampfrath, T., Schoenmaker, H., Heideman, R., Leinse, A. and Kuipers, L. (2009). Probing the magnetic field of light at optical frequencies, *Science* **326**, p. 550.

6. Soukoulis, C. M. and Wegener, M. (2010). Optical metamaterialsmore bulky and less lossy, *Science* **330**, pp. 1633–1634.
7. Pendry, J. B., Holden, A. J., Robbins, D. J. and Stewart, W. J. (1999). Magnetism from conductors and enhanced nonlinear phenomena, *IEEE Trans. Microwave Theory Tech.* **47**, pp. 2075–2084.
8. Shelby, R. A., Smith, D. R. and Schultz, S. (2001). Experimental verification of a negative index of refraction, *Science* **292**, pp. 77–79.
9. Shalaev, V. M. (2007). Optical negative-index metamaterials, *Nature Photon* **1**, pp. 41–47.
10. Zheludev, N. I. (2010). The road ahead for metamaterials, *Science* **328**, pp. 582–583.
11. Kivshar, Y. and Zheludev, N. (2012). From metamaterials to metadevices, *Nature Materials* **11**, pp. 917–924.
12. Leonhardt, U. (2006). Optical conformal mapping, *Science* **312**, pp. 1777–1780.
13. Pendry, J. B. (2000). Negative refraction makes a perfect lens, *Phys. Rev. Lett.* **85**, pp. 3966–3969.
14. Smith, D. R., Padilla, W. J., Vier, D. C., Nemat-Nasser, S. C. and Schultz, S. (2000). Composite medium with simultaneously negative permeability and permittivity, *Phys. Rev. Lett.* **84**, pp. 4184–4187.
15. Padilla, W. J., Taylor, A. J., Highstrete, C., Lee, M. and Averitt, R. D. (2006). Dynamical electric and magnetic metamaterial response at terahertz frequencies, *Phys. Rev. Lett.* **96**, p. 107401.
16. Liu, N., Guo, H., Fu, L., Kaiser, S., Schweizer, H. and Giessen, H. (2008). Three-dimensional photonic metamaterials at optical frequencies, *Nature Mater.* **7**, pp. 31–37.
17. Soukoulis, C. M., Koschny, T., Zhou, J., Kafesaki, M. and Economou, E. N. (2007). Magnetic response of split ring resonators at terahertz frequencies, *Phys. Stat. Sol. B* **244**, pp. 1181–1187.
18. Peng, L., Ran, L., Chen, H., Zhang, H., Kong, J. A. and Grzegorczyk, T. M. (2007). Experimental observation of left-handed behavior in an array of standard dielectric resonators, *Phys. Rev. Lett.* **98**(15), pp. 157403.
19. Schuller, J. A., Zia, R., Taubner, T. and Brongersma, M. L. (2007). Dielectric metamaterials based on electric and magnetic resonances of silicon carbide particles, *Phys. Rev. Lett.* **99**, p. 107401.
20. Zhao, Q., Du, B., Kang, L., Zhao, H., Xie, Q., Li, B., Zhang, X., Zhou, J., Li, L. and Meng, Y. (2008). Tunable negative permeability in an isotropic dielectric composite, *Applied Physics Letters* **92**, p. 051106.
21. Vynck, K., Felbacq, D., Centeno, E., Cabuz, A. I., Cassagne, D. and Guizal, B. (2009). All-dielectric rod-type metamaterials at optical frequencies, *Phys. Rev. Lett.* **102**, p. 133901.
22. Zhao, Q., Zhou, J., Zhang, F. and Lippens, D. (2009). Mie resonance-based dielectric metamaterials, *Mater. Today* **12**, pp. 60–69.

23. Popa, B.-I. and Cummer, S. A. (2008). Compact dielectric particles as a building block for low-loss magnetic metamaterials, *Physical Review Letters* **100**, p. 207401.
24. Ginn, J. C. and Brener, I. (2012). Realizing optical magnetism from dielectric metamaterials, *Physical Review Letters* **108**, p. 097402.
25. Fu, Y. H., Kuznetsov, A. I., Miroshnichenko, A. E., Yu, Y. F. and Lukyanchuk, B. (2013). Directional visible light scattering by silicon nanoparticles, *Nature Communication* **4**, p. 1527.
26. Staude, I., Miroshnichenko, A. E., Decker, M., Fofang, N. T., Liu, S., Gonzales, E., Dominguez, J., Luk, T. S., Neshev, D. N., Brener, I. and Kivshar, Y. (2013). Tailoring directional scattering through magnetic and electric resonances in subwavelength silicon nanodisks, *ACS Nano* **7**(9), pp. 7824–7832.
27. Bi, K., Guo, Y., Liu, X., Zhao, Q., Xiao, J., Lei, M. and Zhou, J. (2014). Magnetically tunable mie resonance-based dielectric metamaterials, *Scientific Reports* **4**, p. 7001.
28. Habteyes, T. G., Staude, I., Chong, K. E., Dominguez, J., Decker, M., Miroshnichenko, A., Kivshar, Y. S. and Brener, I. (2014). Near-field mapping of optical modes on all-dielectric silicon nanodisks, *ACS Photonics* **1**, pp. 794–798.
29. Savelev, R. S., Makarov, S. V., Krasnok, A. E. and Belov, P. A. (2015). From optical magnetic resonance to dielectric nanophotonics (a review), *Optics and Spectroscopy* **119**(4), pp. 551–568.
30. Dmitriev, P. A., Baranov, D. G., Milichko, V. A., Makarov, S. V., Mukhin, I. S., Samusev, A. K., Krasnok, A. E., Belov, P. A. and Kivshar, Y. S. (2016). Resonant raman scattering from silicon nanoparticles enhanced by magnetic response, *Nanoscale* **8**, p. 9721.
31. Dmitriev, P., Makarov, S., Milichko, V., Mukhin, I., Gudovskikh, A., Sitnikova, A., Samusev, A., Krasnok, A. and Belov, P. (2016). Laser fabrication of crystalline silicon nanoresonators from an amorphous film for low-loss all-dielectric nanophotonics, *Nanoscale* **8**, pp. 5043–5048.
32. Bohren, C. and Huffman, D. (1998). *Absorption and scattering of light by small particles* (New York: Wiley).
33. Soref, R., Fellow, L. and Paper, I. (2006), The past, present and future of silicon photonics, *IEEE journal of selected topics in quantum electronics* **12**(6), pp. 1678–1687.
34. Daldosso, N. and Pavesi, L. (2009). Nanosilicon photonics, *Laser Photonics Rev.* **3**(6), pp. 508–534.
35. Mie, G. (1908). Beitrage zur optik truber medien, speziell kolloidaler metallosungen, *Annalen der Physik* **330**, pp. 377–445.
36. Stratton, J. A. (1930). The effect of rain and fog on the propagation of very short radio waves, *Proceedings of the Institute of Radio Engineers* **18**(6), pp. 1064–1074.

37. Lewin, L. (1947). The electrical constants of a material loaded with spherical particles, *Electrical Engineers — Part III: Radio and Communication Engineering, Journal of the Institution of* **94**(27), pp. 65–68.
38. Evlyukhin, A. B., Reinhardt, C., Seidel, A., Lukyanchuk, B. S. and Chichkov, B. N. (2010). Optical response features of si-nanoparticle arrays, *Phys. Rev. B* **82**, p. 045404.
39. Schuller, J. A., Taubner, T. and Brongersma, M. L. (2009). Optical antenna thermal emitters, *Nat. Photonics* **3**(11), pp. 658–661.
40. Cao, L., Fan, P., Barnard, E. S., Brown, A. M. and Brongersma, M. L. (2010). Tuning the color of silicon nanostructures, *Nano Lett.* **10**(7), pp. 2649–2654.
41. Cao, L., Fan, P., Vasudev, A. P., White, J. S., Fan, S. and Brongersma, M. L. (2010). Semiconductor nanowire optical antenna solar absorbers, *Nano Lett.* **10**, pp. 439–445.
42. Cao, L., Park, J.-S., Fan, P., Clemens, B. and Brongersma, M. L. (2010). Resonant germanium nanoantenna photodetectors, *Nano Lett.* **10**, pp. 1229–1233.
43. García-Etxarri, A., Gómez-Medina, R., Froufe-Pérez, L. S., López, C., Chantada, L., Scheffold, F., Aizpurua, J., Nieto-Vesperinas, M. and Sáenz, J. J. (2011). Strong magnetic response of submicron silicon particles in the infrared, *Opt. Express* **19**(6), pp. 4815–4826.
44. Kuznetsov, A. I., Miroshnichenko, A. E., Fu, Y. H., Zhang, J. and Lukyanchuk, B. (2012). Magnetic light, *Sci. Rep.* **2**, p. 492.
45. Vesseur, E. J. R., Coenen, T., Caglayan, H., Engheta, N., Polman, A., Aizpurua, J., Reyes-Coronado, A., Batson, P. E., Schoen, D. T., García De Abajo, F. J., Brongersma, M. L., Sapienza, R., Renger, J., Kuttge, M., van Hulst, N. F., Frimmer, M., Koenderink, A. F., Van De Groep, J., Koenderink, F., Bernal Arango, F., Femius Koenderink, A., Barnard, E. S., Acar, H. and Kuipers, L. K. (2012). Resonant modes of single silicon nanocavities excited by electron irradiation, *ACS Nano.* **11**(9), pp. 4265–4269.
46. Evlyukhin, A. B., Novikov, S. M., Zywietz, U., Eriksen, R. L., Reinhardt, C., Bozhevolnyi, S. I. and Chichkov, B. N. (2012). Demonstration of magnetic dipole resonances of dielectric nanospheres in the visible region, *Nano Lett.* **12**, p. 3749.
47. Shi, L., Tuzer, T. U., Fenollosa, R. and Meseguer, F. (2012). A new dielectric metamaterial building block with a strong magnetic response in the sub-1.5-micrometer region: Silicon colloid nanocavities, *Advanced Materials* **24**, pp. 5934–5938.
48. Geffrin, J. M., García-Etxarri, B., Gómez-Medina, R., Albella, P., Froufe-Pérez, L. S., Eyraud, C., Litman, A., Vaillon, R., González, F., Nieto-Vesperinas, M., Sáenz, J. J. and Moreno, F. (2012). Magnetic and electric coherence in forward- and back-scattered electromagnetic waves by a single dielectric subwavelength sphere, *Nat. Commun.* **3**, p. 1171.

49. Krasnok, A. E., Miroshnichenko, A. E., Belov, P. A. and Kivshar, Y. S. (2012). All-dielectric optical nanoantennas, *Optics Express* **20**, pp. 20599–20604.
50. Miroshnichenko, A. E. and Kivshar, Y. S. (2012). Fano resonances in all-dielectric oligomers, *Nano Letters* **12**(12), pp. 6459–6463.
51. Chong, K. E., Hopkins, B., Staude, I., Miroshnichenko, A. E., Dominguez, J., Decker, M., Neshev, D. N., Brener, I. and Kivshar, Y. S. (2014). Observation of fano resonances in all-dielectric nanoparticle oligomers, *Small* **10**, pp. 1985–1990.
52. Filonov, D. S., Slobozhanyuk, A. P., Krasnok, A. E., Belov, P. A., Nenasheva, E. A., Hopkins, B., Miroshnichenko, A. E. and Kivshar, Y. S. (2014). Near-field mapping of fano resonances in all-dielectric oligomers, *Appl. Phys. Lett.* **104**, p. 021104.
53. Krasnok, A. E., Simovski, C. R., Belov, P. A. and Kivshar, Y. S. (2014). Superdirective dielectric nanoantenna, *Nanoscale* **6**, pp. 7354–7361.
54. Krasnok, A. E., Maloshtan, A., Chigrin, D. N., Kivshar, Y. S. and Belov, P. A. (2015). Enhanced emission extraction and selective excitation of nv centers with alldielectric nanoantennas, *Laser and Photonics Reviews* **9**, pp. 385–391.
55. Li, S. V., Baranov, D. G., Krasnok, A. E. and Belov, P. A. (Oct., 2015). All-dielectric nanoantennas for unidirectional excitation of electromagnetic guided modes, *Applied Physics Letters* **107**(17), p. 171101.
56. Polubavkina, Y. S., Kryzhanovskaya, N. V., Moiseev, E. I., Kulagina, M. M., Mukhin, I. S., Komissarenko, F. E., Zadiranov, Y. M., Maximov, M. V., Krasnok, A. E. and Bogdanov, A. A. (2016). Improved emission outcoupling from microdisk laser by si nanospheres, *Journal of Physics: Conference Series* **741**, p. 012158.
57. Markovich, D., Baryshnikova, K., Shalin, A., Samusev, A., Krasnok, A. and Belov, P. (Mar., 2016). Enhancement of artificial magnetism via resonant bianisotropy, *Scientific Reports* **6**, p. 22546.
58. Shcherbakov, M. R., Neshev, D. N., Hopkins, B., Shorokhov, A. S., Staude, I., Melik-Gaykazyan, E. V., Decker, M., Ezhov, A. A., Miroshnichenko, A. E., Brener, I., Fedyanin, A. A. and Kivshar, Y. S. (2014). Enhanced third-harmonic generation in silicon nanoparticles driven by magnetic response, *Nano Letters* **14**(11), pp. 6488–6492.
59. Noskov, R. E., Krasnok, A. E. and Kivshar, Y. S. (2012). Nonlinear metal-dielectric nanoantennas for light switching and routing, *New Journal of Physics* **14**.
60. Baranov, D. G., Makarov, S. V., Milichko, V. A., Kudryashov, S. I., Krasnok, A. E. and Belov, P. A. (2016). Nonlinear transient dynamics of photoexcited resonant silicon nanostructures, *ACS Photonics* **3**, p. 1546.
61. Baranov, D. G., Makarov, S. V., Krasnok, A. E., Belov, P. A. and Alu, A. (2016). Tuning of near- and far-field properties of all-dielectric dimer nanoantennas via ultrafast electron-hole plasma photoexcitation, *Laser Photonics Rev.* **10**, pp. 1009–1015.

62. Granzow, N., Schmidt, M. A. Chang, W., Wang, L., Coulombier, Q., Troles, J., Toupin, P., Hartl, I., Lee, K. F., Fermann, M. E., Wondraczek, L. and Russell, P. S. J. (2013). Mid-infrared supercontinuum generation in as2 s3 — silica nano-spike step-index waveguide, *Opt. Express* **21**, pp. 1596–1603.
63. Guo, Q., Zhu, H., Liu, F., Zhu, A. Y., Reed, J. C., Yi, F. and Cubukcu, E. (2014). Silicon-on-glass graphene-functionalized leaky cavity mode nanophotonic biosensor, *ACS Photonics* **1**(3), pp. 221–227.
64. Huang, Z., Wang, J., Liu, Z., Xu, G., Fan, Y., Zhong, H., Cao, B., Wang, C. and Xu, K. (2015). Strong-field-enhanced spectroscopy in silicon nanoparticle electric and magnetic dipole resonance near a metal surface, *Journal of Physical Chemistry C* **119**(50), pp. 28127–28135.
65. Savelev, R. S., Filonov, D. S., Kapitanova, P. V., Krasnok, A. E., Miroshnichenko, A. E., Belov, P. A. and Kivshar, Y. S. (2014). Bending of electromagnetic waves in all-dielectric particle array waveguides, *Applied Physics Letters* **105**, p. 181116.
66. Savelev, R. S., Filonov, D. S., Petrov, M. I., Krasnok, A. E., Belov, P. A. and Kivshar, Y. S. (2015). Resonant transmission of light in chains of high-index dielectric particles, *Phys. Rev. B* **92**, p. 155415.
67. Savelev, R., Yulin, A., Krasnok, A. and Kivshar, Y. (2016). Solitary waves in chains of high-index dielectric nanoparticles, *ACS Photonics* **3**, pp. 1869–1876.
68. Decker, M., Staude, I., Falkner, M., Dominguez, J., Neshev, D. N., Brener, I., Pertsch, T. and Kivshar, Y. S. (2015). High-efficiency dielectric huygens surfaces, *Advanced Optical Materials* **3**, pp. 813–820.
69. Makarov, S. V., Tsypkin, A., Voytova, T., Milichko, V., Mukhin, I., Yulin, A., Putilin, S., Baranov, M., Krasnok, A., Morozov, I. A. and Belov, P. (2016). Self-adjusted all-dielectric metasurfaces for deep ultraviolet femtosecond pulse generation, *Nanoscale* **8**, pp. 17809–17814.
70. Moitra, P., Slovick, B. A., Yu, Z. G., Krishnamurthy, S. and Valentine, J. (2014). Experimental demonstration of a broadband all-dielectric metamaterial perfect reflector, *Applied Physics Letters* **104**, p. 171102.
71. Vuye, G., Fisson, S., Van, V. N., Wang, Y., Rivory, J. and Abeles, F. (1993). Temperature dependence of the dielectric function of silicon using in situ spectroscopic ellipsometry, *Thin Solid Films* **233**, pp. 166–170.
72. Novotny, L. and Hecht, B. (2006). *Principles of Nano-Optics* (Cambridge University Press).
73. Evlyukhin, A., Reinhardt, C. and Chichkov, B. (2011). Multipole light scattering by nonspherical nanoparticles in the discrete dipole approximation, *Physical Review B* **23**, p. 235429.
74. Evlyukhin, A. B., Eriksen, R. L., Cheng, W., Beermann, J., Reinhardt, C., Petrov, A., Prorok, S., Eich, M., Chichkov, B. N. and Bozhevolnyi, S. I. (2014). Optical spectroscopy of single si nanocylinders with magnetic and electric resonances, *Scientific Reports* **4**, p. 4126.

75. Shi, L., Harris, J. T., Fenollosa, R., Rodriguez, I., Lu, X., Korgel, B. A. and Meseguer, F. (2013). Monodisperse silicon nanocavities and photonic crystals with magnetic response in the optical region, *Nat. Commun.* **4**, p. 1904.
76. Abbarchi, M., Naffouti, M., Vial, B., Benkouider, A., Lermusiaux, L., Favre, L., Ronda, A., Bidault, S., Berbezier, I. and Bonod, N. (2014). Wafer scale formation of monocrystalline silicon-based mie resonators via silicon-on-insulator dewetting, *ACS Nano* **8**(11), pp. 11181–11190.
77. Zywietz, U., Evlyukhin, A., Reinhardt, C. and Chichkov, B. (2014). Laser printing of silicon nanoparticles with resonant optical electric and magnetic responses, *Nat. Commun.* **5**, p. 3402.
78. Spinelli, P., Verschuuren, M. and Polman, A. (2012). Broadband omnidirectional antireflection coating based on subwavelength surface mie resonators, *Nat. Commun.* **3**, p. 692.
79. Zywietz, U., Reinhardt, C., Evlyukhin, A. B., Birr, T. and Chichkov, B. N. (2014). Generation and patterning of si nanoparticles by femtosecond laser pulses, *Appl. Phys. A* **114**, p. 45.
80. Person, S., Jain, M., Lapin, Z., Saenz, J. J., Wicks, G. and Novotny, L. (2013). Demonstration of zero optical backscattering from single nanoparticles, *Nano Letters* **13**(4), pp. 1806–1809.
81. Novotny, L. and van Hulst, N. (2011). Antennas for light, *Nat. Photon.* **5**, pp. 83–90.
82. Balanis, C. (1982). *Antenna Theory: Analysis and Design*. (New York; Brisbane: J. Wiley).
83. Fan, J. A., Bao, K., Lassiter, J. B., Bao, J., Halas, N. J., Nordlander, P. and Capasso, F. (2012). Near-normal incidence dark-field microscopy: Applications to nanoplasmonic spectroscopy, *Nano Letters* **12**(6), pp. 2817–2821.
84. Zhang, S., Bao, K., Halas, N. J., Xu, H. and Nordlander, P. (2011). Substrate-induced fano resonances of a plasmonic nanocube: A route to increased-sensitivity localized surface plasmon resonance sensors revealed, *Nano Letters* **11**(4), pp. 1657–1663.
85. Spinelli, P., Ferry, V. E., van de Groep, J., van Lare, M., Verschuuren, M. A., Schropp, R. E. I., Atwater, H. A. and Polman, A. (2012). Plasmonic light trapping in thin-film si solar cells, *Journal of Optics* **14**, p. 024002.
86. Kim, K., Kim, J. H., Kim, H. P., Park, K., Nam, H., Lee, S., Kim, S., Choi, K., Kim, S. Y. and Kwon, C. (2010). Tumor-homing multifunctional nanoparticles for cancer theragnosis: Simultaneous diagnosis, drug delivery, and therapeutic monitoring, *Journal of Controlled Release* **146**.
87. Schuller, J. A. and Brongersma, M. L. (2009). General properties of dielectric optical antennas, *Opt. Express* **17**, pp. 24084–24095.
88. Krasnok, A. E., Miroshnichenko, A. E., Belov, P. A. and Kivshar, Y. S. (2011). Huygens optical elements and yagi–uda nanoantennas based on dielectric nanoparticles, *JETP Letters* **94**, pp. 635–640.

89. Rolly, B., Bebey, B., Bidault, S., Stout, B. and Bonod, N. (2012). Promoting magnetic dipolar transition in trivalent lanthanide ions with lossless mie resonances, *Phys. Rev. B* **85**(24), pp. 245432.
90. Filonov, D. S., Krasnok, A. E., Slobozhanyuk, A. P., Kapitanova, P. V., Nenasheva, E. A., Kivshar, Y. S. and Belov, P. A. (2012). Experimental verification of the concept of all-dielectric nanoantennas, *Appl. Phys. Lett.* **100**, p. 201113.
91. Rolly, B., Stout, B. and Bonod, N. (2012). Boosting the directivity of optical antennas with magnetic and electric dipolar resonant particles, *Opics Express* **20**, p. 20376.
92. Liu, Y. G., Choy, W. C. H., Sha, W. E. I. and Chew, W. C. (2012). Unidirectional and wavelength-selective photonic sphere-array nanoantennas, *Opt. Lett.* **37**, p. 2112.
93. Schmidt, M. K., Esteban, R., Sáenz, J. J., Suárez-Lacalle, I., Mackowski, S. and Aizpurua, J. (2012). Dielectric antennas — a suitable platform for controlling magnetic dipolar emission: Errata, *Optics Express* **20**(17), p. 18609.
94. Maksymov, I. S., Staude, I., Miroshnichenko, A. E. and Kivshar, Y. S. (2012). Optical yagi-uda nanoantennas, *Nanophotonics* **1**(1), pp. 65–81.
95. Zou, L., Withayachumnankul, W., Shah, C. M., Mitchell, A., Bhaskaran, M., Sriram, S. and Fumeaux, C. (2013). Dielectric resonator nanoantennas at visible frequencies, *Opt. Express* **21**(1), pp. 83–90.
96. Liu, W., Miroshnichenko, A. E. and Kivshar, Y. S. (2014). Control of light scattering by nanoparticles with optically-induced magnetic responses, *Chinese Physics B* **23**(4), p. 047806.
97. Krasnok, A. E., Slobozhanyuk, A. P., Simovski, C. R., Tretyakov, S. A., Poddubny, A. N., Miroshnichenko, A. E., Kivshar, Y. S. and Belov, P. A. (2015). An antenna model for the purcell effect, *Scientific Reports* **5**, p. 12956.
98. Albella, P., Poyli, M. A., Schmidt, M. K., Maier, S. A., Moreno, F., Saenz, J. J. and Aizpurua, J. (2013). Low-loss electric and magnetic field-enhanced spectroscopy with subwavelength silicon dimers, *Journal of Physical Chemistry C* **117**(26), pp. 13573–13584.
99. Boudarham, G., Abdeddaim, R. and Bonod, N. (2014). Enhancing magnetic nearfield intensities with dielectric resonators, *Applied Physics Letters* **104**, p. 021117.
100. Bakker, R. M., Permyakov, D., Yu, Y. F., Markovich, D., Paniagua-Domínguez, R., Gonzaga, L., Samusev, A., Kivshar, Y. S., Luk'yanchuk, B. and Kuznetsov, A. I. (2015). Magnetic and electric hotspots with silicon nanodimers, *Nano Letters* **15**(33) pp. 2137–2142.
101. Krasnok, A. E., Simovski, C. R., Belov, P. a. and Kivshar, Y. S. (2014). Superdirective dielectric nanoantennas, *Nanoscale* **6**(13), p. 7354.
102. Krasnok, A. E., Filonov, D. S., Simovski, C. R., Kivshar, Y. S. and Belov, P. A. (2014). Experimental demonstration of superdirective dielectric antenna, *Appl. Phys. Lett.* **104**, p. 133502.

103. Tsai, W.-Y., Huang, J.-S. and Huang, C.-B. (2014). Selective trapping or rotation of isotropic dielectric microparticles by optical near field in a plasmonic archimedes spiral, *Nano Lett.* **14**, pp. 547–552.
104. Chen, W., Abeysinghe, D. C., Nelson, R. L. and Zhan, Q. (2010). Experimental confirmation of miniature spiral plasmonic lens as a circular polarization analyzer, *Nano Lett.* **10**, pp. 2075–2079.
105. Schaferling, M., Yin, X. and Giessen, H. (2012). Formation of chiral fields in a symmetric environment, *Optics Exp.* **20**, p. 26326.
106. Wang, S. and Chan, C. (2014). Lateral optical force on chiral particles near a surface, *Nature Comm.* **5**, p. 3307.
107. Yu, N. and Capasso, F. (2014). Flat optics with designer metasurfaces, *Nature Materials* **13**, pp. 139–150.
108. Bliokh, K. Y. and Nori, F. (2012). Transverse spin of a surface polariton, *Physical Review A* **85**(6), pp. 1–5.
109. Bliokh, K. Y., Bekshaev, A. Y. and Nori, F. (2014). Extraordinary momentum and spin in evanescent waves, *Nature communications* **5**, p. 3300.
110. Van Mechelen, T. and Jacob, Z. (2016). Universal spin-momentum locking of evanescent waves, *Optica.* **3**(2), p. 118.
111. Rodriguez-Fortuno, F. J., Marino, G., Ginzburg, P., O′Connor, D., Martinez, A., Wurtz, G. A. and Zayats, A. V. (2013). Near-field interference for the unidirectional excitation of electromagnetic guided modes, *Science* **340**, pp. 328–330.
112. Mueller, J. P. B. and Capasso, F. (2013). Asymmetric surface plasmon polariton emission by a dipole emitter near a metal surface, *Phys. Rev. B* **88**, p. 121410(R).
113. Purcell, E. M. (1946). Spontaneous emission probabilities at radio frequencies, *Phys. Rev.* **69**, p. 681.
114. Sauvan, C., Hugonin, J. P., Maksymov, I. S. and Lalanne, P. (2013). Theory of the spontaneous optical emission of nanosize photonic and plasmon resonators, *Physical Review Letters* **110**, p. 237401.
115. Barthes, J., des Francs, G. C., Bouhelier, A., Weeber, J.-C. and Dereux, A. (2011). Purcell factor for a point-like dipolar emitter coupled to a two-dimensional plasmonic waveguide, *Phys. Rev. B* **84**(7), p. 73403.
116. Poddubny, A. N., Belov, P. A., Ginzburg, P., Zayats, A. V. and Kivshar, Y. S. (2012). Microscopic model of purcell enhancement in hyperbolic metamaterials, *Phys. Rev. B* **86**, p. 035148.
117. Krasnok, A. E., Slobozhanyuk, A. P., Simovski, C. R., Tretyakov, S. A., Poddubny, A. N., Miroshnichenko, A. E., Kivshar, Y. S. and Belov, P. A. (2015). An antenna model for the purcell effect, *Sci. Rep.* **5**, p. 12956.
118. Krasnok, A., Glybovski, S., Petrov, M., Makarov, S., Savelev, R., Belov, P., Simovski, C. and Kivshar, Y. (2016). Demonstration of the enhanced purcell factor in all-dielectric structures, *Appl. Phys. Lett.* **108**(21), p. 211105.
119. Beams, R., Smith, D., Johnson, T. W., Oh, S.-H., Novotny, L. and Vamivakas, A. N. (2013). Nanoscale fluorescence lifetime imaging of an optical antenna with a single diamond nv center, *Nano Letters* **13**, pp. 3807–3811.

120. Aigouy, L., Caze, A., Gredin, P., Mortier, M. and Carminati, R. (2014). Mapping and quantifying electric and magnetic dipole luminescence at the nanoscale, *Physical Review Letters* **113**, p. 076101.
121. Acuna, G. P., Moller, F. M., Holzmeister, P., Beater, S., Lalkens, B. and Tinnefeld, P. (2012). Fluorescence enhancement at docking sites of dna-directed self-assembled nanoantennas, *Science* **338**, pp. 506–510.
122. Fano, U. (1961). Effects of configuration interaction on intensities and phase shifts, *Phys. Rev.* **124**, pp. 1866–1878.
123. Cai, D.-J., Huang, Y.-H., Wang, W.-J., Ji, W.-B., Chen, J.-D., Chen, Z.-H. and Liu, S.-D. (2015). Fano resonances generated in a single dielectric homogeneous nanoparticle with high structural symmetry, *J. Phys. Chem. C* **119**, pp. 4252–4260.
124. Luk'yanchuk, B., Zheludev, N. I., Maier, S. A., Halas, N. J., Nordlander, P., Giessen, H. and Chong, C. T. (2010). The fano resonance in plasmonic nanostructures and metamaterials, *Nature Materials* **9**, pp. 707–715.
125. Rahmani, M., Lei, D. Y., Giannini, V., Lukiyanchuk, B., Ranjbar, M., Liew, T. Y. F., Hong, M. and Maier, S. A. (2012). Subgroup decomposition of plasmonic resonances in hybrid oligomers: Modeling the resonance lineshape, *Nano Letters* **12**(4), pp. 2101–2106.
126. Zhang, Y., Wen, F., Zhen, Y.-R., Nordlander, P. and Halas, N. J. (2013). Coherent fano resonances in a plasmonic nanocluster enhance optical four-wave mixing, *Proceedings of the National Academy of Sciences* **110**(23), pp. 9215–9219.
127. Shcherbakov, M. R., Shorokhov, A. S., Neshev, D. N., Hopkins, B., Staude, I., Melik-Gaykazyan, E. V., Ezhov, A. a., Miroshnichenko, A. E., Brener, I., Fedyanin, A. A. and Kivshar, Y. S. (2015). Nonlinear interference and tailorable third-harmonic generation from dielectric oligomers, *ACS Photonics* **2**, pp. 578–582.
128. Hopkins, B., Poddubny, A. N., Miroshnichenko, A. E. and Kivshar, Y. S. (2013). Revisiting the physics of fano resonances for nanoparticle oligomers, *Phys. Rev. A* **88**, p. 053819.
129. Hopkins, B., Liu, W., Miroshnichenko, A. E. and Kivshar, Y. S. (2013). Optically isotropic responses induced by discrete rotational symmetry of nanoparticle clusters, *Nanoscale* **5**, pp. 6395–6403.
130. Hopkins, B., Filonov, D. S., Miroshnichenko, A. E., Monticone, F., Alù, A. and Kivshar, Y. S. (2015). Interplay of magnetic responses in all-dielectric oligomers to realize magnetic fano resonances, *ACS Photonics* **2**, pp. 724–729.
131. Caldarola, M., Albella, P., Cortes, E., Rahmani, M., Roschuk, T., Grinblat, G., Oulton, R. F., Bragas, A. V. and Maier, S. A. (2015). Non-plasmonic nanoantennas for surface enhanced spectroscopies with ultra-low heat conversion, *Nature Communications* **6**, p. 7915.
132. Lal, S., Link, S. and Halas, N. J. (2007). Nano-optics from sensing to waveguiding, *Nat. Photonics* **1**(11), pp. 641–648.

133. Deng, H.-D., Chen, X.-Y., Xu, Y. and Miroshnichenko, A. E. (2015). Single protein sensing with asymmetric plasmonic hexamer via fano resonance enhanced two-photon luminescence, *Nanoscale* **7**, pp. 20405–20413.
134. Lei, D. Y., Wondraczek, L., Nazabal, V., Maier, S. A. and Schmidt, M. A. (2012). Extended remote-sensing ability, *Nat. Commun.* **3**, p. 1108.
135. Saha, K., Agasti, S., Kim, C., Li, X. and Rotello, V. (2012). Gold nanoparticles in chemical and biological sensing, *Chem. Rev.* **112**(5), pp. 2739–2779.
136. Slobozhanyuk, A. P., Poddubny, A. N., Krasnok, A. E. and Belov, P. A. (2014). Magnetic purcell factor in wire metamaterials, *Applied Physics Letters* **104**, p. 161105.
137. Baranov, D. G., Savelev, R. S., Li, S. V., Krasnok, A. E. and Alu, A. (2016). Modifying magnetic dipole spontaneous emission with nanophotonic structures, *Laser Photonics Rev.* **11**, p. 1600268.
138. Nie, S. and Emory, S. R. (1997). Probing single molecules and single nanoparticles by surface-enhanced raman scattering, *Science* **275**, pp. 1102–1106.
139. Maier, S. A. (2007). *Plasmonics: Fundamentals and Applications*. (Springer Science and Business Media).
140. Rui Tan, J. M., Ruan, J. J., Lee, H. K., Phang, I. Y. and Ling, X. Y. (2014). A large-scale superhydrophobic surface-enhanced Raman scattering (SERS) platform fabricated via capillary force lithography and assembly of Ag nanocubes for ultratrace molecular sensing, *Phys. Chem. Chem. Phys.* **16**(48), pp. 26983–26990.
141. Giannini, V., Fernandez-Dominguez, A. I., Heck, S. C. and Maier, S. A. (2011). Plasmonic nanoantennas: Fundamentals and their use in controlling the radiative properties of nanoemitters, *Chemical Reviews* **111**(6), pp. 3888–3912.
142. Forster, T. (1948). Zwischenmolekulare energiewanderung und fluoreszenz, *Annalen der Physik* **437**, pp. 55–75.
143. de Torres, J., Mivelle, M., Moparthi, S. B., Rigneault, H., Van Hulst, N. F., García-Parajó, M. F., Margeat, E. and Wenger, J. (2016). Plasmonic Nanoantennas Enable Forbidden Förster DipoleDipole Energy Transfer and Enhance the FRET Efficiency, *Nano Lett.* **16**, pp. 6222–6230.
144. Krainer, G., Hartmann, A. and Schlierf, M. (2015). FarFRET: Extending the Range in Single-Molecule FRET Experiments beyond 10 nm, *Nano Lett.* **15**(9), pp. 5826–5829.
145. Shen, Y., Zhou, J., Liu, T., Tao, Y., Jiang, R., Liu, M., Xiao, G., Zhu, J., Zhou, Z.-K., Wang, X., Jin, C. and Wang, J. (2013). Plasmonic gold mushroom arrays with refractive index sensing figures of merit approaching the theoretical limit, *Nat. Commun.* **4**, p. 2381.
146. Yang, J., Giessen, H. and Lalanne, P. (2015). Simple analytical expression for the peak-frequency shifts of plasmonic resonances for sensing, *Nano Lett.* **15**(5), pp. 3439–3444.

147. Yu, R., Cox, J. D. and de Abajo, F. J. G. (2016). Nonlinear plasmonic sensing with nanographene, *Phys. Rev. Lett.* **117**, p. 123904.
148. Mesch, M., Metzger, B., Hentschel, M. and Giessen, H. (2016). Nonlinear Plasmonic Sensing, *Nano Lett.* **16**, pp. 3155–3159.
149. Kucsko, G., Maurer, P. C., Yao, N. Y., Kubo, M., Noh, H. J., Lo, P. K., Park, H. and Lukin, M. D. (2013). Nanometre-scale thermometry in a living cell, *Nature* **500**, pp. 54–58.
150. Albella, P., Alcaraz de la Osa, R., Moreno, F. and Maier, S. A. (2014). Electric and magnetic field enhancement with ultra-low heat radiation dielectric nanoantennas: Considerations for surface enhanced spectroscopies, *ACS Photonics* **1**, pp. 524–529.
151. Rodriguez, I., Shi, L., Lu, X., Korgel, B. A., Alvarez-Puebla, R. A. and Meseguer, F. (2014). Silicon nanoparticles as raman scattering enhancers, *Nanoscale* **6**, p. 5666.
152. Punj, D., Regmi, R., Devilez, A., Plauchu, R., Moparthi, S. B., Stout, B., Bonod, N., Rigneault, H. and Wenger, J. (2015). Self-assembled nanoparticle dimer antennas for plasmonic-enhanced single-molecule fluorescence detection at micromolar concentrations, *ACS Photonics* **2**(8), pp. 1099–1107.
153. Rahmani, M., Luk'yanchuk, B. and Hong, M. (2013). Fano resonance in novel plasmonic nanostructures, *Laser Photonics Rev.* **7**(3), pp. 329–349.
154. Hentschel, M., Saliba, M., Vogelgesang, R., Giessen, H., Alivisatos, A. P. and Liu, N. (2010). Transition from isolated to collective modes in plasmonic oligomers, *Nano Lett.* **10**, pp. 2721–2726.
155. Verre, R., Yang, Z. J., Shegai, T. and Kall, M. (2015). Optical magnetism and plasmonic fano resonances in metal-insulator-metal oligomers, *Nano Lett.* **15**, pp. 1952–1958.
156. Chen, Y., Li, H. and Mo, L. (2012). Flexible and tunable silicon photonic circuits on plastic substrates, *Sci. Rep.* **2**, 622.
157. Law, M., Sirbuly, D. J., Johnson, J. C., Goldberger, J., Saykally, R. J. and Yang, P. (2004). Nanoribbon waveguides for subwavelength photonics integration, *Science* **305**(5688), pp. 1269–1273.
158. Espinola, R., Ahmad, R., Pizzuto, F., Steel, M. and Osgood, R. (2001). A study of high-index-contrast 90 degree waveguide bend structures, *Opt. Express* **8**(9), pp. 517–528.
159. Miroshnichenko, A. and Kivshar, Y. (2005). Sharp bends in photonic crystal waveguides as nonlinear fano resonators, *Opt. Express* **13**(11), pp. 3969–3976.
160. Khurgin, J. B. (2015). How to deal with the loss in plasmonics and metamaterials, *Nature Nanotechnology* **10**, pp. 2–6.
161. Savelev, R. S., Slobozhanyuk, A. P., Miroshnichenko, A. E., Kivshar, Y. S. and Belov, P. A. (2014). Subwavelength waveguides composed of dielectric nanoparticles, *Phys. Rev. B* **89**, p. 035435.
162. Yariv, A., Xu, Y., Lee, R. K. and Scherer, A. (1999). Coupled-resonator optical waveguide: A proposal and analysis, *Opt. Lett.* **24**(11), pp. 711–713.

163. Du, J., Liu, S., Lin, Z., Zi, J. and Chui, S. T. (2009). Guiding electromagnetic energy below the diffraction limit with dielectric particle arrays, *Phys. Rev. A* **79**, p. 051801.
164. Bakker, R. M., Yu, Y. F., Paniagua-Domínguez, R., Luk'yanchuk, B. and Kuznetsov, A. (2015). Silicon nanoparticles for waveguiding. In *Frontiers in Optics 2015*, p. FM1B.2. Optical Society of America.
165. Bulgakov, E. N. and Sadreev, A. F. (2015). Light trapping above the light cone in a one-dimensional array of dielectric spheres, *Phys. Rev. A* **92**, p. 023816.
166. Bulgakov, E. N. and Maksimov, D. N. (2016). Light guiding above the light line in arrays of dielectric nanospheres, *Opt. Lett.* **41**(16), pp. 3888–3891.
167. Marinica, D. C., Borisov, A. G. and Shabanov, S. V. (2008). Bound states in the continuum in photonics, *Phys. Rev. Lett.* **100**, p. 183902.
168. Hsu, C. W., Zhen, B., Stone, A. D., Joannopoulos, J. D. and Soljacic, M. (2016). Bound states in the continuum, *Nature Reviews Materials* **1**, p. 16048.
169. Silva, A., Monticone, F., Castaldi, G., Galdi, V. and Engheta, A. A. N. (2014). Performing mathematical operations with metamaterials, *Science* **343**, pp. 160–163.
170. Ni, X., Emani, N. K., Kildishev, A. V., Boltasseva, A. and Shalaev, V. M. (2011). Broadband light bending with plasmonic nanoantennas, *Science* **335**, p. 427.
171. Kauranen, M. and Zayats, A. V. (2012). Nonlinear plasmonics, *Nature Photonics* **6**, pp. 737–748.
172. Ahmadi, A. and Mosallaei, H. (2008). Physical configuration and performance modeling of all-dielectric metamaterials, *Physical Review A* **77**, p. 045104.
173. Vendik, I., Odit, M. and Kozlov, D. (2009). 3d metamaterial based on a regular array of resonant dielectric inclusions, *Radioengineering* **18**, p. 111.
174. Jylha, L., Kolmakov, I., Maslovski, S. and Tretyakov, S. (2006). Modeling of isotropic backward-wave materials composed of resonant spheres, *Journal of Applied Physics* **99**, p. 043102.
175. Rybin, M. V., Filonov, D. S., Belov, P. A., Kivshar, Y. S. and Limonov, M. F. (2015). Switching from visibility to invisibility via fano resonances: Theory and experiment, *Scientific Reports* **5**, p. 8774.
176. Jahani, S. and Jacob, Z. (2016). All-dielectric metamaterials, *Nature Nanotechnology* **11**, pp. 23–36.
177. Lewin, L. (1947). The electrical constants of a material loaded with spherical particles, *Proc. Inst. Electr. Eng.* **94**, p. 65.
178. Yu, N., Genevet, P., Kats, M., Aieta, F., Tetienne, J.-P., Capasso, F. and Gaburro, Z. (2011). Light propagation with phase discontinuities: Generalized laws of reflection and refraction, *Science* **334**, p. 333.

179. Huang, L., Chen, X., Muhlenberhd, H., Li, G., Bai, B., Tan, Q., Jin, G., Zentgraf, T. and Zhang, S. (2012). Dispersionless phase discontinuities for controlling light propagation, *Nano Lett.* **12**, p. 5750.
180. Sun, S., Yang, K.-Y., Wang, C.-M., Juan, T.-K., Chen, W.-T., Liao, C.-Y., He, Q., Xiao, S., Kung, W.-T., Guo, G.-Y., Zhou, L. and Tsai, D. (2012). High-efficiency broadband anomalous reflection by gradient metasurfaces, *Nano Lett.* **12**, p. 6223.
181. Monticone, F., Estakhri, N. and Alu, A. (2013). Full control of nanoscale optical transition with a composite metascreen, *Phys. Rev. Lett.* **110**, p. 203903.
182. Pfeiffer, C. and Grbic, A. (2013). Metamaterial huygens' surfaces: Tailoring wavefront with reflectionless sheets, *Phys. Rev. Lett.* **110**, p. 19740.
183. Aieta, F., Genevet, P., Kats, M., Yu, N., Blanchard, R., Gaburro, Z. and Capasso, F. (2012). Aberration-free ultrathin flat lenses and axicons at telecom wavelengths based on plasmonic metasurfaces, *Nano Lett.* **12**, p. 4932.
184. Aieta, F., Kats, M. A., Genevet, P. and Capasso, F. (2015). Multiwavelength achromatic metasurfaces by dispersive phase compensation, *Science* **347**, pp. 1342–1345.
185. Zhao, Q., Zhou, J., Zhang, F. and Lippens, D. (2009). Mie resonance-based dielectric metamaterials, *Mater. Today* **12**(12), pp. 60–69. ISSN 13697021.
186. Jahani, S. and Jacob, Z. (2014). Transparent subdiffraction optics: Nanoscale light confinement without metal, *Optica*. **1**, pp. 96–100.
187. Corbitt, S. J., Francoeur, M. and Raeymaekers, B. (2014). Implementation of optical dielectric metamaterials: A review, *Journal of Quantitative Spectroscopy and Radiative Transfer* **158**, pp. 3–16.
188. Krasnok, A., Makarov, S., Petrov, M., Savelev, R., Belov, P. and Kivshar, Y. (2015). Towards all-dielectric metamaterials and nanophotonics, *Proc. SPIE.* **9502**, p. 950203.
189. Spillane, S. M., Kippenberg, T. J. and Vahala, K. J. (2002). Ultralow-threshold raman laser using a spherical dielectric microcavity, *Nature* **415**, pp. 621–623.
190. Leuthold, J., Koos, C. and Freude, W. (2010). Nonlinear silicon photonics, *Nature Photon.* **4**, pp. 535–544.
191. Krasavin, A. V., Ginzburg, P., Wurtz, G. A. and Zayats, A. V. (2016). Nonlocality-driven supercontinuum white light generation in plasmonic nanostructures, *Nature Communications* **7**, p. 11497.
192. Shcherbakov, M. R., Vabishchevich, P. P., Shorokhov, A. S., Chong, K. E., Choi, D.-Y., Staude, I., Miroshnichenko, A. E., Neshev, D. N., Fedyanin, A. A. and Kivshar, Y. S. (2015). Ultrafast all-optical switching with magnetic resonances in nonlinear dielectric nanostructures, *Nano Letters* **15**, p. 6985. ISSN 1530-6984.
193. Makarov, S., Kudryashov, S., Mukhin, I., Mozharov, A., Milichko, V., Krasnok, A. and Belov, P. (2015). Tuning of magnetic optical response in a dielectric nanoparticle by ultrafast photoexcitation of dense electron-hole plasma, *Nano Letters* **15**(9), pp. 6187–6192.

194. Yang, Y., Wang, W., Boulesbaa, A., Kravchenko, I. I., Briggs, D. P., Puretzky, A., Geohegan, D. and Valentine, J. (2015). Nonlinear fano-resonant dielectric metasurfaces, *Nano Lett.* **15**, pp. 7388–7393.
195. Iyer, P. P., Butakov, N. A. and Schuller, J. A. (2015). Reconfigurable semiconductor phased-array metasurfaces, *ACS Photonics* **2**, pp. 1077–1084. ISSN 2330-4022.
196. Shorokhov, A. S., Melik-Gaykazyan, E. V., Smirnova, D. A., Hopkins, B., Chong, K. E., Choi, D.-Y., Shcherbakov, M. R., Miroshnichenko, A. E., Neshev, D. N., Fedyanin, A. A. and Kivshar, Y. S. (2016). Multifold enhancement of third-harmonic generation in dielectric nanoparticles driven by magnetic fano resonances, *Nano Letters* **16**(8), pp. 4857–4861.
197. Smirnova, D. A., Khanikaev, A. B., Smirnov, L. A. and Kivshar, Y. S. (2016). Multipolar third-harmonic generation driven by optically induced magnetic resonances, *ACS Photonics* **3**, p. 1468.
198. Grinblat, G., Li, Y., Nielsen, M. P., Oulton, R. F. and Maier, S. A. (2016). Enhanced third harmonic generation in single germanium nanodisks excited at the anapole mode, *Nano Letters* **16**(7), pp. 4635–4640.
199. Sokolowski-Tinten, K. and von der Linde, D. (2000). Generation of dense electron-hole plasmas in silicon, *Phys. Rev. B* **61**(4), pp. 2643–2650.
200. Wurtz, G. A., Pollard, R., Hendren, W., Wiederrecht, G. P., Gosztola, D. J., Podolskiy, V. A. and Zayats, A. V. (2011). Designed ultrafast optical nonlinearity in a plasmonic nanorod metamaterial enhanced by nonlocality, *Nature Nanotechnology* **6**(2), pp. 107–111.
201. Valev, V. K., Denkova, D., Zheng, X., Kuznetsov, A. I., Reinhardt, C., Chichkov, B. N., Tsutsumanova, G., Osley, E. J., Petkov, V., Clercq, B. D., Silhanek, A. V., Jeyaram, Y., Volskiy, V., Warburton, P. A., Vandenbosch, G. A. E., Russev, S., Aktsipetrov, O. A., Ameloot, M., Moshchalkov, V. V. and Verbiest, T. (2012). Plasmon-enhanced sub-wavelength laser ablation: Plasmonic nanojets, *Advanced Materials* **24**, pp. OP29–OP35.
202. Zuev, D. A., Makarov, S. V., Mukhin, I. S., Milichko, V. A., Starikov, S. V., Morozov, I. A., Shishkin, I. I., Krasnok, A. E. and Belov, P. A. (2016). Fabrication of hybrid nanostructures via nanoscale laser-induced reshaping for advanced light manipulation, *Adv. Mater.* **28**, pp. 3087–3093.

CHAPTER 9

Tunable Metamaterials

ILYA V. SHADRIVOV* **and**
DRAGOMIR N. NESHEV†

Nonlinear Physics Centre,
Australian National University, Australia
*Ilya.Shadrivov@anu.edu.au
†Dragomir.Neshev@anu.edu.au

In this chapter, we give an overview of the various approaches for making tunable metamaterials and focus in more detail on several particular examples.

9.1. Introduction

Metamaterials and metasurfaces are designed to exhibit various exotic properties and this is often done by utilizing resonant phenomena. As a result, the electromagnetic response of metamaterials is usually dispersive, and desired properties occur only in a very narrow frequency range. This is why it was understood from the very beginning of metamaterial research that one will need to be able to *dynamically tune* the response of the composite structures in order to achieve practically usable metamaterials. Alternatively, when the broadband performance is not important, one may want to change the properties of materials at a given frequency so that the electromagnetic waves can be dynamically manipulated. This can be

used, for example, in switching and modulation applications, making the ability to tune material properties highly desirable. There is a substantial amount of work done on tunable metamaterials to date, with several reviews published[1–5] and different journals having made this a topic of special issues.[6] Table 9.1 summarizes several tuning methods, with their applicability in various frequency ranges as well as their advantages and limitations.

First, we would like to define what we call tunable metamaterials. Some authors say that they "tune" the properties of materials by fabricating a new sample with different parameters, or by changing the operating frequency, or by changing the incidence angle of the electromagnetic waves. Clearly, in these examples the properties of each given material are not changed, with either a totally new material fabricated or by changing the excitation conditions. Here, we will call metamaterials tunable if their properties can be changed by external influence, e.g. by control voltage, or by temperature, or by magnetic field, etc. Overall, we can distinguish three distinct tuning mechanisms that can be applied for changing the material properties:

- *Tuning by changing the structural geometry:* Achieved by mechanically deforming the shape of the constituent elements or their mutual arrangement in the metamaterial, which affects the overall properties due to mutual coupling between these elements,[7] e.g. elastic deformation of the structure.
- *Tuning by changing the constituent materials:* Achieved by changing the properties of the materials composing the individual meta-atoms, e.g. actively changing the conductivity of semiconductors by injecting free electrons into them.
- *Tuning by changing of the surrounding environment:* Achieved by immersing the metamaterial in an environment, which properties can change, e.g. by a liquid crystal (LC).

Each of these tuning mechanisms can be realized in a different way, which is also specific to the frequency range where this metamaterial operates. Table 9.1 summarizes the most common techniques

Table 9.1. Mechanisms for achieving tunability in various frequency ranges.

	Tuning mechanism	μW	THz	Optics	Advantages	Limitations
Changing geometry	Mechanical displacement of components, including micro-electromechanical systems (MEMS)	+	+	+	Strong tunability	Slow, hard to implement at higher frequencies
	Pneumatic actuation	+	+	+/−	Gas pressure sensitive	Complex fabrication and control
Changing meta-atoms	Lumped electronic components	+	−	−	Strong tunability, commercially available components	Only available for lower frequencies
	Superconductors	+	−	−	Strong tunability, link to quantum effects in Josephson junctions	Cryogenic temperatures
	Photo-excitation	+	+	+	Fast	Currently requires large light intensities
	Phase change materials	+	+	+	Controlled by temperature	Usually add losses
	Ferrimagnetics, Ferromagnetics	+	−	−	Higher microwave frequency bands	Require large magnetic fields
	Ferroelectrics	+	−	−	Control by relatively small electric field	Relatively high losses in ferroelectrics
Changing environment	LCs	+	+	+	Available in most frequency regions	Large volumes required in microwaves. Difficult pre-alignment
	Semiconductor substrates	+/−	+	−	Can be tuned by voltage or by light	

(+) indicates that the tunability was demonstrated, (±) shows that the tunability is, in principle, possible, but was not demonstrated yet, while (−) indicates that the tunability in a given frequency range is probably impossible by using the corresponding method.

for achieving such tunability and the corresponding frequency range of applicability.

Since various tunability concepts work differently in different frequency bands, we split our chapter in three sections reflecting three broad frequency ranges: microwave, terahertz (THz) and optics. In Section 9.2, we briefly review the tunability mechanisms used for microwave metamaterials. In the following sections, we review (THz) and optical tunable metamaterials with a bigger focus on several selected works in this area.

9.2. Tunable Microwave Metamaterials

There is a substantial amount of work done on tunable metamaterials and composite right and left-handed transmission lines in the microwave frequency range, and we see three main reasons for this. Firstly, we have a large number of materials and commercially available components whose properties change under external influence. Secondly, the fabrication of metamaterials is much easier at microwave frequencies, where constituent components are quite large. Thirdly, immediate applications of metamaterials seem to be more feasible in this frequency range with several products currently being developed.

The most obvious way to tune metamaterials or transmission lines in microwaves is to use semiconductor components, such as diodes. This was demonstrated, e.g. in Refs. 8–10. The tunability obtained by this approach is quite large, since the parameters of, e.g. varactor diodes can be changed by almost an order of magnitude. An extension of this approach is based on the use of photodiodes for biasing varactor diodes.[11] When such structures are illuminated by light, the photodiodes generate voltage, and microwave response of the metamaterial changes. This may be convenient for creating light-controlled tunable metasurfaces, as it was demonstrated in Ref. 11. Moreover, such tunability not only controls linear properties of the metamaterial, but also its *nonlinear properties*.[12]

Another way of making tunable metamaterials at low frequencies is by utilizing superconductors. As an example, the properties of

superconductor metamaterials in Ref. 13 were controlled by three different means in order to make tunable metamaterials, where authors used temperature, dc magnetic field, and rf magnetic field to change the properties of metamaterials. Superconductors in metamaterials can further be used in Josephson junction configuration, and this opens the path for the applications of metamaterials for solving quantum problems. As an example, in Ref. 14, the authors used metamaterial-type superconducting quantum interference devices (SQUID) in order to achieve a tunable amplifier that can be used for amplifying quantum signals without adding noise.

Further, tunable metamaterials in microwave frequency range were achieved by using ferroelectrics,[15] ferrimagnetics[16] and ferromagnetics.[17] LCs are a good platform for achieving tunability in nearly all frequency ranges, and for microwaves the tunable metamaterials were demonstrated by using either the electric[18] or magnetic[19] fields. This method is, however, often impractical for bulk microwave metamaterial use due to large amounts of expensive LC required.

A completely new class of tunable (and nonlinear) metamaterials was developed after detailed studies of the near-field interaction between metamaterial elements have shown strong potential and importance of this method.[7] Mechanical displacements of the metamaterial elements lead to change in the coupling of these elements, and this tunes electromagnetic response of metamaterials. This was demonstrated both in microwave and THz frequency ranges, and in principle possible in optics, though required tolerances make it technologically difficult. An extension of this concept involves the use of MEMS and pneumatic metamaterials, where the displacement of metamaterial elements is achieved either by activating the MEMS mechanisms[1] or by changing pressure.[20]

9.3. Tunable Terahertz Metamaterials

Metamaterials are well placed for applications in manipulating THz waves, since there is a small number of natural materials which can be used for these purposes. The first demonstration of tunable THz metamaterials was done in 2006, when the

Schottky-type semiconductor structure with metamaterial-shaped electrodes exhibited strong tunability by voltage.[21] In such a structure, the conductivity of the substrate was changing, resulting in the modification of the strength of the metamaterial resonance. Further advances in this field were done by creating a patterned semiconductor structure that would have conductivity increased only in certain areas of metamaterial, leading to the frequency shift of the metamaterial resonance.[22,23] In these structures, the change in the semiconductor excitation was achieved via photo-doping by intense laser beam.

Ferroelectrics[24] as well as vanadium dioxide[25] can be used for creating temperature-tunable metamaterials. At the same time, it was shown that the using temperature for tuning LC-infiltrated THz metamaterials is not efficient.[26]

The general concept of tuning metamaterial response by displacing meta-atoms can be readily applied in the THz frequency range. This was achieved by direct mechanical motion of the metamaterial layers,[27] by using polymer actuation,[28] by pneumatic actuation[29] or by stretching.[30] We will focus on the first two of these methods along with the description of the liquid metamaterial, whose properties change when we apply dc electric field[31] that rotates the metamolecules.

9.3.1. *Mechanical offset of metamaterial layers*

To demonstrate the practically strong modification of the transmission spectrum in a double-layer metamaterial, we consider metamaterial layers containing electric field-coupled resonators.[27] As the resonators in neighboring cells are electrically connected, the structure exhibits substantial transmission only near its resonances. The metamaterial layers are fabricated using conventional optical lithography followed by a lift-off process. The thickness of the gold is 150 nm, and the resonators are arranged in a square lattice with a lattice constant of 135 μm. The overall structure occupies an area of 5 mm by 5 mm. After two metamaterial samples are fabricated, a small amount of solidified heat-sensitive glue is put on top of each sample and then

heated until it reaches the melting point. The two metamaterial layers are face-to-face bonded and aligned with each other under a microscope. The lateral displacement can be varied by reheating the sample and moving the top layer with respect to the bottom one. The spacing between the two layers can be controlled by moving the top layer back and forth to redistribute the glue. Tunability is achieved by controlling the coupling between the two adjacent metamaterial structures: the frequency shift is controlled by moving the structures laterally, while the maximum frequency shift is controlled by the thickness of the glue. The lateral displacements can be changed from $S = 0\,\mu$m to half unit cell $S = 67.5\,\mu$m.

Numerical simulations are performed using commercial electromagnetic solver CST Microwave Studio. In our simulations, we have used normally incident waves polarized across the gap of the resonator. It is assumed that the structures are periodic in the x–y plane and the computational area is terminated by perfectly matching boundary layers in the z direction. The spacing between the metamaterial layers is determined by measuring the optical transmission at the edges of the fabricated samples outside the metamaterial region. The frequency spacing between Fabry–Perot transmission fringes depend on the distance between the plates. The average spacing for six positions measured around the metamaterial region is $10.4\,\mu$m with variations below $2.5\,\mu$m.

Transmission through the double-layered metamaterials is measured by using a THz time-domain spectroscopy system. The polarization of the incident field is aligned across the gaps of the resonators. The system is excited by a femtosecond laser pulse, and the transmission response is sampled in the time domain. A Fourier transform is used to obtain the amplitude and phase of the transmitted wave, and further signal processing in the frequency domain allows the calculation of the experimental transmission spectrum. Fabry–Perot interference fringes in the transmission spectrum are reduced by employing time gating in the time-domain signal.

The experimental and simulation results for the double-layered structure with different lateral displacements are shown in Fig. 9.1.

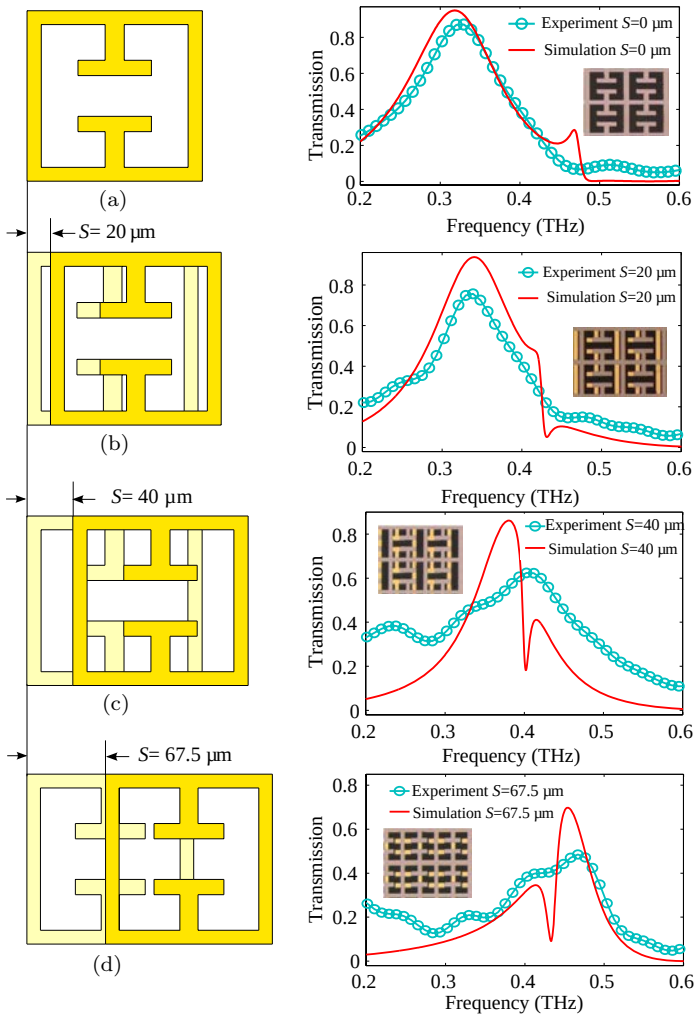


Fig. 9.1. Schematics, measurement versus simulation results, and alignment taken under microscope (inset) of the double-layered structure when (a) $S = 0\,\mu$m, (b) $S = 20\,\mu$m, (c) $S = 40\,\mu$m, (d) $S = 67.5\,\mu$m.

The alignment is also shown for different lateral displacements. The alignment achieved by manual adjustment of alignment marks under the microscope is reasonably good, demonstrating the feasibility of the proposed post-processing approach. The shift of the transmission

band is quite substantial, and it is accompanied by reshaping of the spectrum. We also simulated the resonance shift for larger spacing between the layers of 110 μm. For such large spacing, the vertical coupling between two layers is significantly reduced. The resonance of stacked layers in this case closely resembles isolated layers of metamaterials due to the absence of strong interactions between them. Hence, in this case the lateral displacement will induce negligible resonance tuning.

This post-processing approach may find applications in building tunable THz devices without involving re-fabrication processes and can lead to a large tunability in a pair of layers. One can imagine a THz filter made of two metamaterial layers, where one layer is fixed, while another one can be moved by a translation stage. Various spectral features can be designed by using other types of metamaterial patterns.

9.3.2. *Conductive polymers for metamaterial tuning*

In order to achieve electric control of the metamaterial layers offset, we proposed to use conductive polymers.[28] The applications of π-conjugated polymers, or conducting polymers, have been widely studied to develop various types of electronic devices ranging from organic light-emitting diodes, organic thin film transistors (TFT), organic photovoltaic cells and so on. They have also been used to develop soft actuators. The basic working principle typical in π-conjugated polymer actuators is an electrochemical reduction/oxidation reaction, in which doped ions move into or out of the π-conjugated polymer by applying and reversing voltages, giving mechanical deformation of the polymer in a reversible manner. It was found that heavily doped polypyrrole (PPy) film could be contracted by applying only a few Volts, which controls the desorption of water molecules. Such PPy films are known to behave as metals in the THz and lower frequency ranges, and can exhibit extraordinary transmission in a perforated film.[32]

Here we demonstrate electrically tunable device working in the THz frequency range, however our approach may, in principle, be

extendable to near infrared or visible wavelengths. We have fabricated thin films of PPy doped with hexafluorophosphate (PF6) ions, PPy(PF6), by electrochemical polymerization. They are used as linear actuators to laterally shift the relative position of two stacked metamaterial arrays designed to work as electro-active THz metadevices. In this experiment, we used the same design of metamaterials as in previous subsection with one resonator array fixed while the adjacent array is free to move laterally, see Fig. 9.2. The PPy(PF6) linear actuators are attached to both sides of the freely movable metamaterial array to allow one-dimensional (1D) linear lateral shift along the capacitive gap of the SRR and also to recover the initial position, which is required owing to the elastic nature of the film. It has been reported that PPy(PF6) films contract only a few percent, therefore

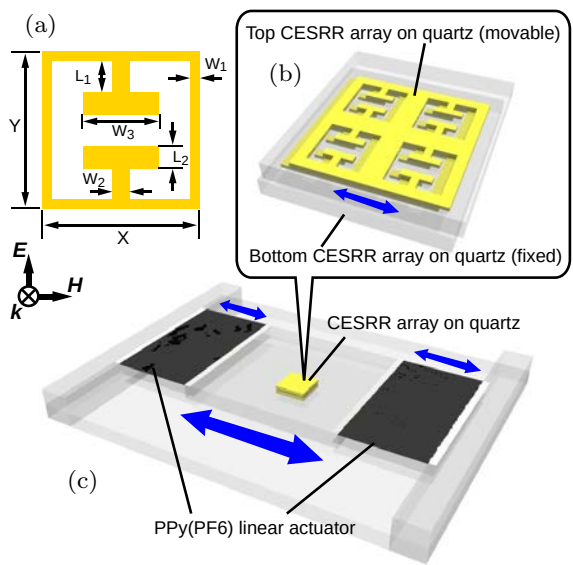
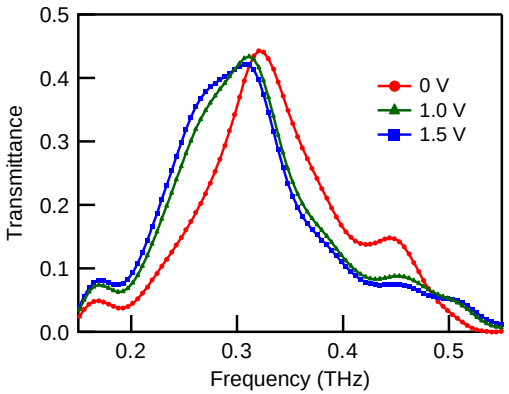


Fig. 9.2. Schematics of the tunable THz device based on a PPy(PF6) linear actuator. (a) Geometry of the unit cell along with the polarization directions of the incident THz wave (top view). (b) Schematics of the double-layered metamaterial arrays with fixed substrate and movable superstrate. (c) Schematics of the tunable THz device based on double-layered metamaterial with fixed substrate and movable superstrate suspended with PPy(PF6) linear actuator. The top layer can be laterally moved by applying voltage.

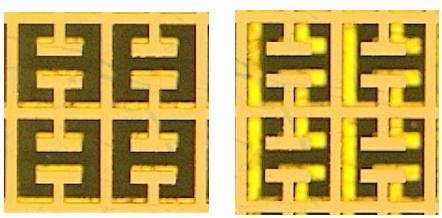
we cut the PPy(PF6) film into 10 mm by 10 mm square pieces in order to achieve tens of μm actuation range that will allow us to obtain significant frequency shift in our THz device. The repeating unit in each metamaterial layer is an electric field-coupled resonator, and all of them are connected, therefore the single-layer structure shows resonant transmission only near its resonances around 0.375 THz. It has been confirmed that by stacking these CESRR arrays, two resonant modes were observed at 0.318 THz and 0.468 THz. These two modes are symmetric and anti-symmetric modes corresponding to the in-phase and out-of-phase loop currents, respectively.[27] The well-aligned device shows resonance around 0.320 THz due to the symmetric mode (Fig. 9.3 (a)). We have successfully obtained a frequency shift of resonant THz transmission by applying a low bias voltage of only a few Volts to one of the polymer actuators to pull the SRR array in one direction. Applying voltage leads to contraction of the polymer film, and this laterally shifts the freely movable layer. By turning off the applied voltage, the system recovers its initial state and transmission spectrum. The optical microscope images also confirm that an applied voltage of 1.5 V is enough to have significant lateral shift of the SRR array to obtain frequency shift in resonant transmission in our device, as shown in Fig. 9.3(b). This approach can be a cheap alternative to a more complex and expensive fabrication process for MEMS type devices.

9.3.3. *Meta liquid crystals*

Another structure utilizing the concept of mechanical motion of the metamolecules for tuning its properties is *meta liquid crystals* (MLCs).[31] To achieve tunability, the meta-atoms are encapsulated in elongated dielectric bars which are dispersed in a host liquid, and are rotated by electro phoretic forces when a biasing electric field is applied. When all such particles are aligned along the field, the structure will exhibit strongly anisotropic properties. The axis of anisotropy can be rotated by changing the direction of the biasing field, tuning the electromagnetic properties of the metamaterial in a manner similar to the mesogens of liquid crystals. Due to this



(a)



(b)

Fig. 9.3. (a) Measured THz transmission spectra of the tunable THz device based on double-layered CESRR arrays with PPy(PF6) linear actuator without and with (1.5 V) applied voltage. (b) Optical microscope images of double-layered CESRR arrays without and with (1.5 V) applied voltage. Scale bar is 100 μm.

strong analogy, we call our metamaterial an MLC, and we call the constituent elements meta-mesogens. The elements of MLCs are complex meta-atoms with sub-wavelength features, and their properties can be engineered in a wide range. Moreover, the *electrostatic* and *electromagnetic* properties of our meta-mesogens are weakly correlated, enabling us to construct MLCs with tunable properties not available with conventional LCs, such as larger anisotropy, strong magnetic and/or chiral (and in general bianisotropic) responses, or well-designed spectral features such as electromagnetically induced transparency. The liquid nature of such metamaterials allows them to

flow and fill in spaces or cover surfaces of arbitrary 3D shapes, which significantly widens the field where metamaterials can be employed.

While the concept could be realized over a wide range of frequencies, here we experimentally demonstrate the idea in THz frequency range. Metamaterials play a vital role in controlling THz waves and creating passive and tunable devices, due to the lack of intrinsic response from natural materials.

We chose two types of meta-atoms, electric split-ring resonators (ESRRs) and I-beam resonators (IBRs). Each meta-mesogen consists of a short array of identical metallic units encapsulated within 10-μm thick polyimide, so that the whole meta-mesogen functions as an anisotropic dipole in an external bias field. The response of the chosen meta-atoms to the high frequency field is strongly anisotropic, therefore by rotating the whole meta-mesogen, we can tune the overall response of the composite medium.

Here, we study three designs — meta-mesogens made of three ESRRs (3-ESRR), five ESRRs (5-ESRR) and two IBRs (2-IBR). Each type of meta-mesogen is mixed in a non-polar liquid so that they are free to rotate in response to an external bias electric field. Paraffin oil is used in the experimental work as it is a readily available liquid which is safe to handle and reasonably transparent in the THz range.

The electrostatic torque of a meta-mesogen in a uniform static electric field is well described within the dipole approximation and is primarily determined by the length and aspect ratio of the meta-mesogens. From the torque, we can calculate the response time of meta-mesogens in different liquids. In paraffin oil, all three designs show estimated response time around 100 ms for a bias field strength of 1.6×10^5V m^{-1}. This time can be further improved by using less viscous liquids. In addition, paraffin oil is not ideal for realizing MLC, since it is less dense than the mesogens, and they slowly sink. Further search for more suitable host liquids is required, and as an example, we simulated MLC with trichloroethylene. It is less viscous and more dense than paraffin oil, making it more suitable for a faster performing MLC. Our calculations predict that this can provide at

least one order of magnitude enhancement of the operation speed. Trichlorethylene was not used in our experiments due to its high toxicity. The dynamics can further be optimized by choosing the shape and size of the meta-mesogens. Since the sign of the voltage output from our source cannot be changed, there is a net dc contribution leading to charge build up, which would result in the aggregation of meta-mesogens, and therefore bursts of pulses are used to maintain the alignment. To enable practical application of this idea, further technical optimization is required so that the charging effect and sinking problem can be addressed.

Figure 9.4 shows microscope photographs of three different MLCs before and after application of the static electric field. To quantify the anisotropy induced in the THz spectral response, the transmission is measured with a commercial THz time-domain spectrometer. In the measurement, we rotate the sample with respect to the polarization of the incident wave to characterize parallel and perpendicular effective refractive indices as well as the modulation of transmission. Figures 9.5(a)–(c) show transmittance spectra $|t|^2$ for

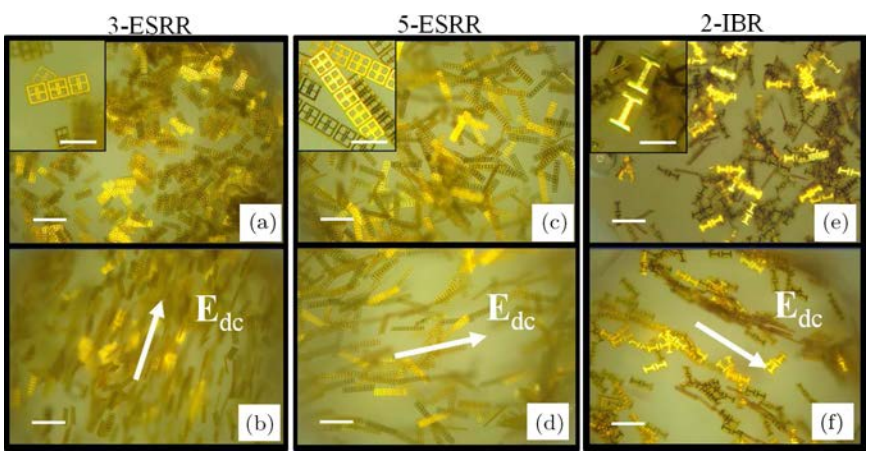


Fig. 9.4. Structural change of the three different MLCs before (top row) and 5 s after (bottom row) application of the bias electric field. (a) and (b) 3-ESRR, (c) and (d) 5-ESRR, (e) and (f) 2-IBR. The scale bars in the main figures and insets are 300 μm and 100 μm, respectively.

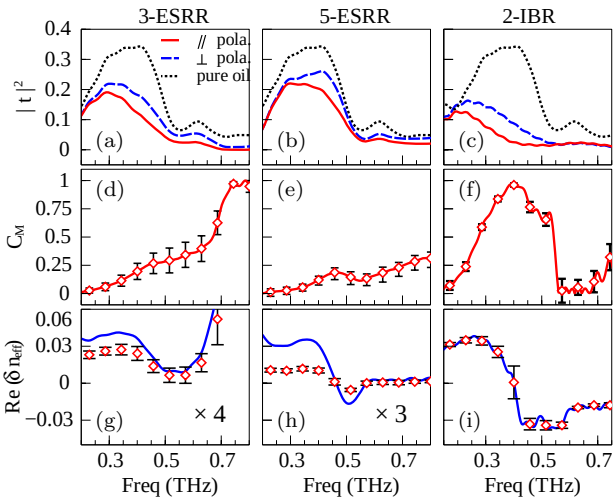


Fig. 9.5. (a)–(c) Measured transmittance spectra of three different types of meta-mesogens for orthogonal polarizations. The corresponding particle numbers for 3-ESRR, 5-ESRR and 2-IBR are around 800, 370 and 1140, respectively. Spectra of pure paraffin oil are also given as a reference. (d)–(f) Modulation contrast of transmission for different number density. (g)–(i) Effective index difference calculated from the relative transmission phase difference for orthogonal polarizations.

three different samples. We note that in our experiments the transmission curves are normalized to the free space transmission of the THz signal. Thus, the transmission shown in the figures is affected both by losses in the structure and by reflection. The most profound effect is due to reflection at the air/quartz interface at the input side and the PTFE/air interface at the output side, with only around 1/3 of the energy left when the signal transmits through an empty cell. Another contribution is due to absorption of energy by the paraffin oil. The oil is quite transparent below 0.5 THz, and the loss increases as frequency goes up. We believe the transmission of the device can be improved by reducing reflection at the input interface with a proper antireflection coating, and by finding a more appropriate liquid with lower absorption. The modulation contrast, defined as $C_{\mathrm{M}} = |t_{||}^2 - t_{\perp}^2|/|t_{||}^2 + t_{\perp}^2|$, shows that the modulation effect is quite noticeable (see Figures 9.5(d)–(f)). It is interesting to note

that the trends in modulation contrast are different for the designs with ESRRs and IBRs. For the two samples with ESRRs, there is an increasing trend as frequency increases, and this is consistent with the growing trend predicted in the calculated extinction cross-section.[31] As a comparison, we also calculated the modulation properties of the proposed MLCs based on the scattering properties of the meta-mesogens. In the calculation, we assume that the meta-mesogens form a cubic lattice and use Clausius–Mossotti formula to estimate the effective permittivity for different number density. The results of calculations showed overall good agreement with the experiment apart from red shift of the resonances.

For sample 3-ESRR, we used a larger concentration ($N \sim 800$, $\rho_{\mathrm{N}} \sim 32\,\mu\mathrm{L}^{-1}$) due to its smaller number of effective meta-atoms. Yet, due to the inherently smaller anisotropy and the partially non-uniform orientation of meta-mesogens, the resonant peak around 0.5 THz is less noticeable (see Figure 9.5(d)). To compare the modulation contrast of IBR and ESRR, we first used similar numbers of meta-mesogens for sample 5-ESRR ($N \sim 370$) and sample 2-IBR ($N \sim 1140$), and we found that the modulation contrast of sample 2-IBR is approximately twice the value of 5-ESRR (see Figures 9.5(e) and (f)), which is in agreement with the predictions from simulations.[31]

From the measured transmission phase, we calculate the effective linear birefringence $\delta n_{\mathrm{eff}} \approx \arg(t_{||}/t_{\perp})/(k_0 h)$ produced by the three different designs, where $k_0 = \omega/c$ and h is the thickness of the MLC sample (see Figures 9.5(g)–(i)). To have a fair comparison, we normalize the index to the same mesogen number of $N = 1200, \rho_{\mathrm{N}} = 48\,\mu\mathrm{L}^{-1}$. This normalization is based on the linear approximation of the relation between effective index and number density of mesogens $n_{\mathrm{eff}} \propto N$, and should be valid when the concentration is low. The resonant features for all three samples are clearly observed, and the effective index difference of sample 2-IBR is about three times the value of 5-ESRR.

To summarize this section, we presented an experimental study of 3D tunable THz MLCs, where engineered meta-mesogens play a similar role to the mesogens of a natural LC. The most important feature

of MLCs is that the electromagnetic and electrostatic response can be independently engineered, which provides a much greater flexibility in constructing tunable metamaterials and metadevices. Different from conventional LCs, the MLCs studied here can provide a direct and pronounced amplitude modulation of transmission, whereas conventional LCs change the phase of the waves only. The meta-atoms can be fabricated in a wide variety of shapes, and could even be made from high index dielectric resonators, which are appealing due to their low losses at optical wavelengths. We expect that with advances in nanofabrication, MLCs could be fabricated on a length scale suitable for applications in optics.

9.4. Tunable Optical Metamaterials

In the early days of the development of optical metamaterials, the research was mainly focused on the realization of bulk 3D materials,[33] motivated primarily by effects such as negative refraction and cloaking. However, the fabrication of 3D metamaterials still remains challenging, and the huge propagation losses at optical frequencies in such materials make their use impractical. As such, most research was done on 2D metamaterial structures, which were considered only as a planar realization of the more desirable bulk structures. A paradigm shift in this research happened after Capasso's group introduced the generalized law of refraction based on spatially varying optical surfaces[34] (see also review papers [35, 36]). The term *metasurfaces* was established as 2D nanostructured materials with unusual properties derived from the structure of their constituent elements, called meta-atoms. Based on such ultra-thin metasurfaces, a large number of important functionalities have been demonstrated, including anomalous beam-deflection, beam-shaping and even carpet cloaking.

However, similar to all other metamaterial structures, the properties of the optical metasurface are fixed by the design and arrangement of their meta-atoms. However, for many applications it is important to dynamically control the optical properties of such designer metasurfaces. Following the need of such dynamic tunability, a number of groups have demonstrated different tuning

approaches, see Ref. 37. To a large extent, these optical approaches overlap with the approaches reviewed above for microwave and THz metamaterials. However, the 2D arrangements of such ultra-thin optical metasurfaces enable greater flexibility. The field of optical tunable metasurfaces has therefore been able to borrow a lot from a number of established optical technologies, including nanoelectromechanical systems (NEMS), LC displays, etc. Below, we review some of the key works in this development.

9.4.1. *Tuning by changing the structural geometry*

The tuning derived from changes of the structural geometry relies on mechanical variations of the shapes of individual meta-atoms or their mutual arrangement. Within this approach, there are two possible scenarios that have been explored: (*i*) Mechanical changes of the meta-atoms through exploration of nanomechanical forces, e.g. acousto-optical effects and (*ii*) changes of the mutual arrangements of meta-atoms through deformation of the substrate, e.g. metasurfaces on flexible substrates.

The use of nanomechanical forces for dynamically reconfiguring the properties of optical metasurfaces has been pioneered by the group of Zheludev, see Ref. 5. These include thermal,[38] electrical,[39] magnetic[40,41] and optical[42,43] actuation of the metasurface structure for changing its structural geometry. While enabling opportunities for strong tunability, the fabrication of such nanomechanical structures is complicated and the long-term reliability is low.

The fabrication of metamaterials on flexible substrates offers a lot easier fabrication. Importantly, this approach also offers good tunability due to stretching and flexing of the substrate, which alters the geometry of mutual arrangement of meta-atoms. Following this idea, flexible plasmonic[44,45] as well as dielectric metamaterials[46] have been demonstrated. In the latter (dielectric) case, the sensitivity of the metasurface properties to the lattice periodicity is very strong,[47] which allows for highly tunable elastic metamaterials. However, despite the simplicity of their fabrication, the tuning speed of metamaterials on flexible substrates is substantially lower than in the case

of nanomechanically actuated metamaterials. Therefore, the choice of the particular technology will depend on the specificity of the applications, including the cost-effectiveness and tunability speed.

9.4.2. *Tuning by changing the constituent materials*

This approach relies on changing of the constituent meta-atoms through varying the properties of materials they are made of. For example, changing the Fermi level of the electrons in the metallic components can result in an overall change of the properties of the meta-atoms. However, in contrast to the case of metamaterials in the THz frequency domain, for optics changing the density of electrons is much more difficult. Nevertheless, several important tunability schemes have been devised and demonstrated as a proof of principle. These include the use of *phase change materials, electron excitation of carriers in semiconductors*, and the use of *graphene*.

9.4.2.1. *Phase change materials*

Phase change materials are materials that change their structure with temperature. Due to presence of a phase transition, the properties of the material can change from crystalline to amorphous or from dielectric or metallic. When such materials are used to build meta-atoms, then the overall optical properties of the meta-atoms can be dramatically altered through the phase transition.

An important example of such a material is vanadium dioxide (VO_2). At room temperature, VO_2 exhibits dielectric properties, but at $\sim 67°C$ it undergoes a change into a conducting state. Using this phase transition of VO_2, Kats *et al.*[48] demonstrated thermally tunable metasurfaces. A metasurface designed from Y-shaped elements was fabricated on top of VO_2 and its reflectivity was investigated at different temperature. The reflection spectrum was changing dramatically as the sample was heated up, with reflection peak corresponding to antenna's resonance shifting to the position of a dip in the mid-infrared range.

Another important example of phase change materials are some chalcogenides, often used in rewritable DVDs. The use of such

materials is a well-established technology; however, their application in reconfigurable metasurfaces has only been recently demonstrated by Wang *et al.*.[49] The authors combined germanium–antimony–tellurium-based films with a diffraction-limited resolution optical writing process to demonstrate a variety of devices, including visible-range reconfigurable bichromatic and multifocus Fresnel zone plates, a super-oscillatory lens with subwavelength focus, a greyscale hologram, and a dielectric metasurfaces with on-demand reflection and transmission resonances. Due to the reversible phase change transition in the chalcogenide films, such components can be written, erased and rewritten as 2D binary or greyscale patterns into the thin film.

9.4.2.2. *Graphene*

The inclusion of graphene flakes as a part of the constituent meta-atoms is a promising candidate for strongly tunable response of optical metasurfaces. Because of the presence of Dirac cones in the electronic dispersion of graphene, the electron density increases linearly with the energy. This is in stark contrast to plasmonic metals, where the electron density increases quadratically with the energy. Therefore, in the case of graphene, it is easier to change the electron density by external factors, such us temperature or biasing electric field.

Yao *et al.*[50] demonstrated electrical tuning of graphene-loaded antennas over a broad wavelength range of 10% of the resonance frequency in the mid-infrared region. Similar ideas have been applied at shorter infrared wavelengths ($2\,\mu$m) by Emani *et al.*,[51] where Fano resonant plasmonic nanostructures were fabricated on top of a graphene sheet. The use of Fano resonant elements enhances the interaction of incident radiation with the graphene sheet and enables efficient electrical modulation of the plasmonic resonance.

9.4.2.3. *Excitation of free carriers*

The process of excitation of free charges in semiconductor materials has also been explored for tuning the properties of optical metasurfaces. This is especially important in the case of

dielectric metasurfaces made out of silicon (or other semiconductor) nanoparticles. When exciting such nanostructures with high photon flux, free carriers can be generated in the silicon due to one- or two-photon absorption processes. The excited free carriers can cause changes in the conductivity of the semiconductor and therefore result in significant shift of the metasurface resonances.[52] The use of this effect has been recently employed to achieve strong modulation of transmission[52,53] through silicon nanodisc arrays. However, the observed tuning of the parameters of such metasurfaces was relatively weak, due to the limited density of the photo-excited carriers.

To overcome this problem, Huang *et al.*[54] have recently demonstrated metasurfaces tunable by the field-effect modulation of the complex refractive index of conducting oxide layers incorporated into the metasurface antenna elements. The conductive oxide, in their case — indium tin oxide, was operated in the ε near zero regime, which enabled strong modulation with relative small bias electric field. As such, they have demonstrated modulation of the metasurface reflectance at frequencies exceeding 10 MHz and electrical switching of $\pm$ first-order diffracted beams. Importantly, the modulation was realized by electrical control, which is a basic requirement for electrically tunable beam-steering by gradient metasurfaces.

9.4.3. *Tuning by changing of the surrounding environment*

While the above two approaches offer some unique advantages, their implementation often requires sophisticated fabrication techniques and the use of various exotic materials. In many cases, however, the structure fabrication has to be as simple as possible, realized by standard fabrication techniques. In this case, their structural geometry or constituent material properties cannot be easily altered. Therefore, for post-fabrication tuning of such optical metasurfaces, the only remaining option is to change the properties of the surrounding environment. For example, liquids with different refractive indices can be applied on top of the metasurfaces to alter their background refractive index and hence their optical response. However, probably

the most flexible technique is the infiltration of the metasurface with LCs. LCs have a common use in several matured technologies, most important of which are the LC displays for monitors and flat screen television sets.

The LCs consist of highly elongated molecules with strong electric dipole moment. Due to the dipole–dipole interactions between the LC molecules, they can align in a crystal-like fashion. Because of this regular orientation of the elongated LC molecules, the LCs exhibit the highest natural optical anisotropy of $\Delta n \approx 0.3$ and above. Importantly, the LCs orientation and refractive index can be tuned by temperature, electric or magnetic fields, as well as by the optical field itself. These tuning mechanisms provide unique opportunities for the use of LCs in tunable optical metasurfaces.

The first use of LC to tune the resonances of cut-wire pair metamaterials was demonstrated by Xiao *et al.*[57] The authors used the change of the LC refractive index from nematic to isotropic phase through increase of the temperature of the metamaterial above a critical value.

Shortly after, Minovich *et al.*[58] suggested how to control the optical properties of fishnet metasurfaces infiltrated with a nematic LC, as shown schematically in Fig. 9.6(a). The tuning was achieved due to the reorientation of LC molecules inside the fishnet holes, which led to frequency shift of the hole modes. Changing of the effective refractive index of the fishnet from negative to positive values was predicted to be achievable with this method. The same group experimentally demonstrated electrical and all-optical transmission control through the LC fishnet structure[55] (see Fig. 9.6(b)). In that case, the reorientation of LC molecules was caused by the strong light field of the incident laser beam. As seen in Fig. 9.6(b), the transmission through the structure was changing non-linearly with the increase of the laser power. Importantly, the transmission was further controlled by the application of bias electric field, applied across the substrate and the top layer of the metasurface, thus demonstrating electrically tunable nonlinear transmission through optical metasurfaces.

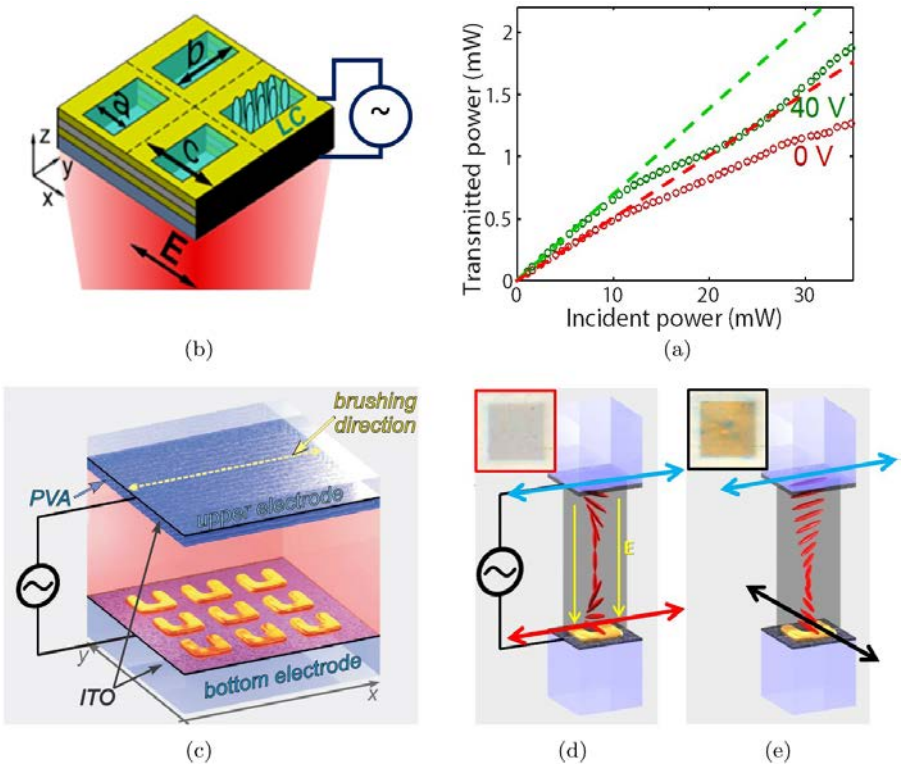


Fig. 9.6. (a) Schematic of the LC-infiltrated fishnet metasurface (from Ref. 55). (b) Tunable transmittance through the metasurface as a function of incident optical power and applied electric field. (c) Schematic illustration of magnetic metasurface attached to an LC cell (from Ref. 56). (d-e) Illustration of the LC alignment with and without applied bias voltage. The insets show the corresponding real-color photographs of the metasurface. The change of color is clearly visible.

Furthermore, Decker *et al.*[56] have demonstrated electro-optical tuning of split-ring resonator (SRR) metasurface utilizing a pre-oriented layer of LCs (see Fig. 9.6(c)). By applying external electric field across the LC cell, 90° polarization rotation was achieved from the top to the bottom surface of the LC cell (see Figs. 9.6(d) and 9.6(e)), which led to the switching of the excitation between the magnetic and the electric modes of the SRR metasurface. This switching

causes a dramatic change of spectral transmission through the metasurface, accompanied by a visual change of its color, as seen in the insets of Figs. 9.6(d) and 9.6(e). A similar effect was also reported by Buchnev *et al.*,[59] though for a metasurface consisting of plasmonic zig-zag strips.

While, significant changes in the transmission was detected in both works above, the actual spectral position of the resonances was not strongly affected. Indeed, the optical mode of plasmonic metasurfaces is strongly confined to the metal layer, thus extending only a few nanometers into the LC. On the other hand, due to strong surface anchoring of the LC molecules, the refractive index of the surrounding environment changes negligibly for these strongly confined plasmonic modes. Therefore, all observed changes in the transmission were due to the polarization rotation inside the bulk of the LC cell. Therefore, a practical scheme for spectral tuning of the optical resonances of metasurfaces remained to be found.

A new solution to this problem was possible with the invention of all-dielectric metasurfaces. Such metasurfaces are composed of resonant high-refractive index dielectric particles (see Fig. 9.7(a)) and exhibit negligible losses at optical wavelengths. In addition, the resonant modes of such dielectric particles extend further into the surrounding offering better opportunities for tuning of the resonant frequency.

LC tuning of dielectric metasurfaces has been recently demonstrated by Sautter *et al.*[60] A metasurface composed of silicon nanodiscs (see Figs. 9.7(a) and 9.7b)) was infiltrated with E7 LC and sandwiched on the top with a cover slip that imposed linear alignment of the LC at room temperature (see Fig. 9.7(c)). By heating the metasurface to temperatures near 60°C, the LC undergoes phase transition and transforms into an isotropic phase (see Fig. 9.7(d)). Through this phase transition, the resonance frequencies of both electric and magnetic dipolar resonances of the metasurface shifted significantly, with the strongest shift of about a full resonance width observed for the electric mode of the metasurface (see Fig. 9.7(e)). This spectral shift of the resonances also resulted in strong phase and intensity

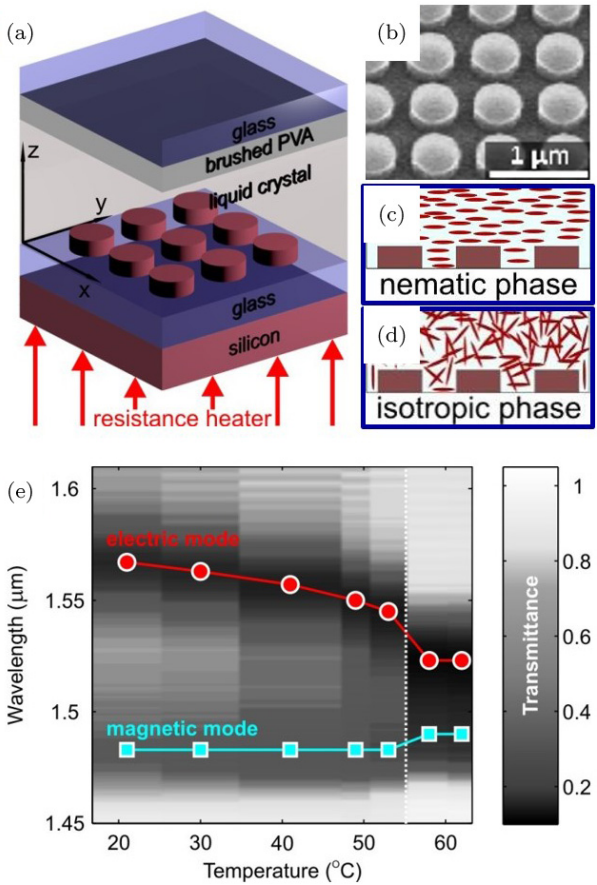


Fig. 9.7. (a) Schematic illustration of the LC cell with the dielectric metasurface. (b) Scanning electron micrograph of the silicon metasurface, before infiltration. (c) and (d) Illustration of the two different phases of the LC at temperatures before and above the phase transition, respectively. (e) Measured transmittance spectra *versus* temperature, demonstrating the strong tuning of both electric and magnetic dipolar resonances of the metasurface (from Ref. 60).

modulation. For wavelengths of about 1.53 μm, we observed intensity modulation of 300% (see Fig. 9.7(e)). Even higher transmission modulation of 500% could be achieved for properly designed optical resonances of the metasurface.

9.4.4. *Outlook*

The progress made in tunable optical metasurfaces in the last few years is truly outstanding. This is a fast developing field and many new exciting results remain in front of us. However, the biggest challenge in the field is to be able to design and demonstrate a tunable metasurface that can modulate the phase of the input beam in the full phase range of $0-2\pi$ without modulating the intensity of the beam. Increasing the spectral range of operation of the metasurfaces is also of enormous importance. Achieving these key functionalities will truly open the field of tunable metasurfaces to a plethora of practical applications, including tunable lenses, beam steering and even dynamic holograms, just to mention a few.

9.5. Conclusions

To date we have seen the extensive research into the field of tunable metamaterials. Various schemes have already been explored and some great advances in the field have been demonstrated. This gained knowledge opens up the doors for the practical applications of tunable metamaterials. We already see the first examples of commercialization of the technologies for tunable metamaterials. Starting from longer wavelengths in the microwave spectral domain, the commercialization efforts are moving to higher frequencies. We would like to mention one particular example of application of tunable metamaterials. It is the metamaterial-based antenna developed by Kymeta,[61] which is set to play a pivotal role in the development of global communications. It will be able to deliver satellite communications not only to stationary structures, but also to moving objects, such as cars, ships and airplanes. The tunability achieved by using TFT technology, allows to dynamically tune the directivity of the flat antenna so that it does not loose connection with a satellite as the antenna moves.[61] We expect that further developments will lead to the applications of tunable metamaterials at higher frequency ranges, including THz, infrared and visible.

References

1. Liu, A. Q., Zhu, W. M., Tsai, D. P. and Zheludev, N. I. (2012). Micromachined tunable metamaterials: A review, *J. Opt.* **14**(11), p. 114009. ISSN 2040-8986. doi: 10.1088/2040-8978/14/11/114009.
2. Zhong J., Huang, Y., Wen, G., Sun, H. and Zhu, W. (2012). The design and applications of tunable metamaterials, *Procedia Eng.* **29**, pp. 802–807. ISSN 1877-7058. doi: 10.1016/j.proeng.2012.01.045.
3. Zheludev, N. I. and Kivshar, Y. S. (2012). From metamaterials to metadevices, *Nat. Mater.* **11**(11), pp. 917–924. ISSN 1476-1122. doi: 10.1038/nmat3431.
4. Turpin, J. P., Bossard, J. A., Morgan, K. L., Werner, D. H. and Werner, P. L. (2014). Reconfigurable and tunable metamaterials: A review of the theory and applications, *Int. J. Antennas Propag.* **2014**, p. e429837. ISSN 1687-5869, 1687-5869. doi: 10.1155/2014/429837,\%002010.1155/2014/429837.
5. Zheludev, N. I. and Plum, E. (2016). Reconfigurable nanomechanical photonic metamaterials, *Nat. Nanotechnol.* **11**(1), pp. 16–22. ISSN 1748-3387. doi: 10.1038/nnano.2015.302.
6. Zheludev, N., Padilla, W. J. and Brener, I. (2012). Photonic materials on demand, *J. Opt.* **14**(11), p. 110201. ISSN 2040-8986. doi: 10.1088/0240-8978/14/11/110201.
7. Powell, D. A., Lapine, M., Gorkunov, M. V., Shadrivov, I. V. and Kivshar, Y. S. (2010). Metamaterial tuning by manipulation of near-field interaction, *Phys. Rev. B* **82**(15). ISSN 1098-0121, pp. 1550-235X. doi: 10.1103/PhysRevB.82.155128.
8. Lim, S., Caloz, C. and Itoh, T. (2004). Metamaterial-based electronically controlled transmission-line structure as a novel leaky-wave antenna with tunable radiation angle and beamwidth, *IEEE Trans. Microw. Theory Tech.* **52**(12), pp. 2678–2690. ISSN 0018-9480. doi: 10.1109/TMTT.2004.838302.
9. Gil, I., Bonache, J., Garcia-Garcia, J. and Martin, F. (2006). Tunable metamaterial transmission lines based on varactor-loaded split-ring resonators, *IEEE Trans. Microw. Theory Tech.* **54**(6), pp. 2665–2674. ISSN 0018-9480. doi: 10.1109/TMTT.2006.872949.
10. Shadrivov, I. V., Morrison, S. K. and Kivshar, Y. S. (2006). Tunable split-ring resonators for nonlinear negative-index metamaterials, *Opt. Express.* **14**(20), p. 9344. ISSN 1094-4087. doi: 10.1364/OE.14.009344.
11. Shadrivov, I. V., Kapitanova, P. V., Maslovski, S. I. and Kivshar, Y. S. (2012). Metamaterials controlled with light, *Phys. Rev. Lett.* **109**(8). ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.109.083902.
12. Kapitanova, P. V., Slobozhnanyuk, A. P., Shadrivov, I. V., Belov, P. A. and Kivshar, Y. S. (2012). Competing nonlinearities with metamaterials, *Appl. Phys. Lett.* **101**(23), p. 231904.
13. Ricci, M., Xu, H., Prozorov, R., Zhuravel, A., Ustinov, A. and Anlage, S. (2007). Tunability of superconducting metamaterials, *IEEE Trans. Appl.*

Supercond. **17**(2), pp. 918–921. ISSN 1051-8223. doi: 10.1109/TASC.2007.898535.

14. Castellanos-Beltran, M. A., Irwin, K. D., Hilton, G. C., Vale, L. R. and Lehnert, K. W. (2008). Amplification and squeezing of quantum noise with a tunable Josephson metamaterial, *Nat. Phys.* **4**(12), pp. 929–931. ISSN 1745-2473. doi: 10.1038/nphys1090.
15. Hand, T. H. and Cummer, S. A. (2008). Frequency tunable electromagnetic metamaterial using ferroelectric loaded split rings, *J. Appl. Phys.* **103**(6), p. 066105. ISSN 0021-8979, 1089-7550. doi: 10.1063/1.2898575.
16. He, Y., He, P., Dae Yoon, S., Parimi, P. V., Rachford, F. J., Harris, V. G. and Vittoria, C. (2007). Tunable negative index metamaterial using yttrium iron garnet, *J. Magn. Magn. Mater.* **313**(1), pp. 187–191. ISSN 0304-8853. doi: 10.1016/j.jmmm.2006.12.031.
17. Kang, L., Zhao, Q., Zhao, H. and Zhou, J. (2008). Magnetically tunable negative permeability metamaterial composed by split ring resonators and ferrite rods, *Opt. Express.* **16**(12), p. 8825. ISSN 1094-4087. doi: 10.1364/OE.16.008825.
18. Zhang, F., Zhang, W., Zhao, Q., Sun, J., Qiu, K., Zhou, J. and Lippens, D. (2011). Electrically controllable fishnet metamaterial based on nematic liquid crystal, *Opt. Express.* **19**(2), p. 1563. ISSN 1094-4087. doi: 10.1364/OE.19.001563.
19. Zhang, F., Kang, L., Zhao, Q., Zhou, J., Zhao, X. and Lippens, D. (2009). Magnetically tunable left handed metamaterials by liquid crystal orientation, *Opt. Express.* **17**(6), p. 4360. ISSN 1094-4087. doi: 10.1364/OE.17.004360.
20. Khodasevych, I. E., Shadrivov, I. V., Powell, D. A., Rowe, W. S. T. and Mitchell, A. (2013). Pneumatically switchable graded index metamaterial lens, *Appl. Phys. Lett.* **102**(3), p. 031904.
21. Chen, H.-T., Padilla, W. J., Zide, J. M. O., Gossard, A. C., Taylor, A. J. and Averitt, R. D. (2006). Active terahertz metamaterial devices, *Nature.* **444**(7119), 597–600. ISSN 0028-0836. doi: 10.1038/nature05343.
22. Chen, H.-T., O'Hara, J. F., Azad, A. K., Taylor, A. J., Averitt, R. D., Shrekenhamer, D. B. and Padilla, W. J. (2008). Experimental demonstration of frequency-agile terahertz metamaterials, *Nat. Photon.* **2**(5), pp. 295–298. ISSN 1749-4885, 1749-4893. doi: 10.1038/nphoton.2008.52.
23. Kafesaki, M., Shen, N. H., Tzortzakis, S. and Soukoulis, C. M. (2012). Optically switchable and tunable terahertz metamaterials through photoconductivity, *J. Opt.* **14**(11), p. 114008. ISSN 2040-8986. doi: 10.1088/2040-8978/14/11/114008.
24. Němec, H., Kužel, P., Kadlec, F., Kadlec, C., Yahiaoui, R. and Mounaix, P. (2009). Tunable terahertz metamaterials with negative permeability, *Phys. Rev. B* **79**(24), p. 241108. doi: 10.1103/PhysRevB.79.241108.
25. Driscoll, T., Kim, H.-T., Chae, B.-G., Kim, B.-J., Lee, Y.-W., Jokerst, N. M., Palit, S., Smith, D. R., Ventra, M. D. and Basov, D. N. (2009). Memory metamaterials, *Science.* **325**(5947), pp. 1518–1521. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1176580.

26. Liu, L., Shadrivov, I., Powell, D., Raihan, R., Hattori, H., Decker, M., Mironov, E. and Neshev, D. (2013). Temperature control of terahertz metamaterials with liquid crystals, *IEEE Trans. Terahertz Sci. Technol.* **3**(6), pp. 827–831. ISSN 2156-342X. doi: 10.1109/TTHZ.2013.2285570.
27. Liu, L., Chen, W.-c., Powell, D. A., Padilla, W. J., Karouta, F., Hattori, H. T., Neshev, D. N. and Shadrivov, I. V. (2014). Post-processing approach for tuning multi-layered metamaterials, *Appl. Phys. Lett.* **105**(15), p. 151102. ISSN 0003-6951, 1077-3118. doi: 10.1063/1.4897949.
28. Matsui, T., Inose, Y., Powell, D. A. and Shadrivov, I. V. (2015). Electroactive tuning of double-layered metamaterials based on π-conjugated polymer actuators, *Adv. Opt. Mater.* pp. 135–140. ISSN 2195-1071. doi: 10.1002/adom.201500276.
29. Kan, T., Isozaki, A., Kanda, N., Nemoto, N., Konishi, K., Takahashi, H., Kuwata-Gonokami, M., Matsumoto K. and Shimoyama, I. (2015). Enantiomeric switching of chiral metamaterial for terahertz polarization modulation employing vertically deformable MEMS spirals, *Nat. Commun.* **6**, p. 8422. doi: 10.1038/ncomms9422.
30. Li, J., Shah, C. M., Withayachumnankul, W., Ung, B. S.-Y., Mitchell, A., Sriram, S., Bhaskaran, M., Chang, S. and Abbott, D. (2013). Mechanically tunable terahertz metamaterials, *Appl. Phys. Lett.* **102**(12), p. 121101. ISSN 0003-6951, 1077-3118. doi: 10.1063/1.4773238.
31. Liu, M., Fan, K., Padilla, W., Powell, D. A., Zhang, X. and Shadrivov, I. V. (2015). Tunable meta-liquid crystals, *Adv. Mater.* pp. 1553–1558. ISSN 1521-4095. doi: 10.1002/adma.201504924.
32. Matsui, T., Vardeny, Z. V., Agrawal, A., Nahata, A. and Menon, R. (2006). Resonantly-enhanced transmission through a periodic array of subwavelength apertures in heavily-doped conducting polymer films, *Appl. Phys. Lett.* **88**(7), p. 071101. ISSN 0003-6951, 1077–3118. doi: 10.1063/1.2175482.
33. Soukoulis, C. M. and Wegener, M. (2011). Past achievements and future challenges in the development of three-dimensional photonic metamaterials, *Nat. Photon.* **5**, pp. 523–530.
34. Yu, N., Genevet, P., Kats, M. A., Aieta, F., Tetienne, J.-P., Capasso, F. and Gaburro, Z. (2011). Light propagation with phase discontinuities: Generalized laws of reflection and refraction, *Science.* **334**, pp. 333–337. doi: 10.1126/science.1210713. URL `http://www.sciencemag.org/content/334/6054/333.abstract`.
35. Yu, N. and Capasso, F. (2014). Flat optics with designer metasurfaces, *Nat. Mater.* **13**(2), pp. 139–150. URL `http://dx.doi.org/10.1038/nmat3839`.
36. Minovich, A. E., Miroshnichenko, A. E., Bykov, A. Y., Murzina, T. V., Neshev, D. N. and Kivshar, Y. S. (2015). Functional and nonlinear optical metasurfaces, *Laser Photon. Rev.* **9**(2), pp. 195–213. ISSN 1863-8899. doi: 10.1002/lpor.201400402. URL `http://dx.doi.org/10.1002/lpor.201400402`.
37. Zheludev N. I. and Kivshar, Y. S. (2012). From metamaterials to metadevices, *Nat. Mater.* **11**, pp. 917–924. URL `http://dx.doi.org/10.1038/nmat3431`.

38. Ou, J. Y., Plum, E., Jiang, L. and Zheludev, N. I. (2011). Reconfigurable photonic metamaterials, *Nano Lett.* **11**(5), pp. 2142–2144. doi: 10.1021/nl200791r. URL http://dx.doi.org/10.1021/nl200791r. PMID: 21480583.
39. Ou, J.-Y., Plum, E., Zhang, J. and Zheludev, N. I. (2013). An electromechanically reconfigurable plasmonic metamaterial operating in the near-infrared, *Nat. Nanotech.* **8**(4), pp. 252–255.
40. Valente, J. A., Ou, J.-Y., Plum, E., Youngs, I. J. and Zheludev, N. I. (2015). Reconfiguring photonic metamaterials with currents and magnetic fields, *Appl. Phys. Lett.* **106**(11) p. 111905. doi: http://dx.doi.org/10.1063/1.4913609. URL http://scitation.aip.org/content/aip/journal/apl/106/11/10.1063/1.4913609.
41. Valente, J., Ou, J.-Y., Plum, E., Youngs, I. J. and Zheludev, N. I. (2015). A magneto-electro-optical effect in a plasmonic nanowire material, *Nat. Commun.* **6**.
42. Zhang, J., MacDonald, K. F. and Zheludev, N. I. (2013). Nonlinear dielectric optomechanical metamaterials, *Light Sci. Appl.* **2**, p. e96. Supplementary information available for this article at URL http://www.nature.com/lsa/journal/v2/n8/suppinfo/lsa201352s1.html.
43. Ou, J.-Y., Plum, E., Zhang, J. and Zheludev, N. I. (2016). Giant nonlinearity of an optically reconfigurable plasmonic metamaterial, *Adv. Mater.* **28**(4), pp. 729–733. ISSN 1521-4095. doi: 10.1002/adma.201504467. URL http://dx.doi.org/10.1002/adma.201504467.
44. Di Falco, A., Ploschner, M. and Krauss, T. F. (2010). Flexible metamaterials at visible wavelengths, *New J. Phys.* **12**(11), p. 113006. URL http://stacks.iop.org/1367-2630/12/i=11/a=113006.
45. Di Falco, A., Zhao, Y. and Al, A. (2011). Optical metasurfaces with robust angular response on flexible substrates, *Appl. Phys. Lett.* **99**(16), p. 163110. doi: http://dx.doi.org/10.1063/1.3655332. URL http://scitation.aip.org/content/aip/journal/apl/99/16/10.1063/1.3655332.
46. Gutruf, P., Zou, C., Withayachumnankul, W., Bhaskaran, M., Sriram, S. and Fumeaux, C. (2016). Mechanically tunable dielectric resonator metasurfaces at visible frequencies, *ACS Nano.* **10**(1), pp. 133–141. doi: 10.1021/acsnano.5b05954. URL http://dx.doi.org/10.1021/acsnano.5b05954. PMID: 26617198.
47. Chong, K. E., Staude, I., James, A., Dominguez, J., Liu, S., Campione, S., Subramania, G. S., Luk, T. S., Decker, M., Neshev, D. N., Brener, I. and Kivshar, Y. S. (2015). Polarization-independent silicon metadevices for efficient optical wavefront control, *Nano Lett.* **15**(8), pp. 5369–5374. doi: 10.1021/acs.nanolett.5b01752. URL http://dx.doi.org/10.1021/acs.nanolett.5b01752. PMID: 26192100.
48. Kats, M. A., Blanchard, R., Genevet, P., Yang, Z., Qazilbash, M. M., Basov, D. N., Ramanathan, S. and Capasso, F. (2013). Thermal tuning of mid-infrared plasmonic antenna arrays using a phase change material, *Opt. Lett.* **38**, pp. 368–370. doi: 10.1364/OL.38.000368. URL http://ol.osa.org/abstract.cfm?URI=ol-38-3-368.

49. Wang, Q., Rogers, E. T. F., Gholipour, B., Wang, C.-M., Yuan, G., Teng, J. and Zheludev, N. I. (2016). Optically reconfigurable metasurfaces and photonic devices based on phase change materials, *Nat. Photon.* **10**(1), pp. 60–65.
50. Yao, Y., Kats, M. A., Genevet, P., Yu, N., Song, Y., Kong, J. and Capasso, F. (2013). Broad electrical tuning of graphene-loaded plasmonic antennas, *Nano Lett.* **13**, pp. 1257–1264.
51. Emani, N. K., Chung, T.-F., Kildishev, A. V., Shalaev, V. M., Chen, Y. P. and Boltasseva, A. (2014). Electrical modulation of fano resonance in plasmonic nanostructures using graphene, *Nano Lett.* **14**(1), pp. 78–82. doi: 10.1021/nl403253c. URL `http://dx.doi.org/10.1021/nl403253c`. PMID: 24303876.
52. Shcherbakov, M. R., Vabishchevich, P. P., Shorokhov, A. S., Chong, K. E., Choi, D.-Y., Staude, I., Miroshnichenko, A. E., Neshev, D. N., Fedyanin, A. A. and Kivshar, Y. S. (2015). Ultrafast all-optical switching with magnetic resonances in nonlinear dielectric nanostructures, *Nano Lett.* **15**(10), pp. 6985–6990. doi: 10.1021/acs.nanolett.5b02989. URL `http://dx.doi.org/10.1021/acs.nanolett.5b02989`. PMID: 26393983.
53. Makarov, S., Kudryashov, S., Mukhin, I., Mozharov, A., Milichko, V., Krasnok, A. and Belov, P. (2015). Tuning of magnetic optical response in a dielectric nanoparticle by ultrafast photoexcitation of dense electronhole plasma, *Nano Lett.* **15**(9), pp. 6187–6192. doi: 10.1021/acs.nanolett.5b02534. URL `http://dx.doi.org/10.1021/acs.nanolett.5b02534`. PMID: 26259100.
54. Huang, Y.-W., Lee, H. W. H., Sokhoyan, R., Pala, R., Thyagarajan, K., Han, S., Tsai, D. P. and Atwater, H. A. (2015). Gate-tunable conducting oxide metasurfaces, *Nano Lett.*, **16**, pp 5319–5325.
55. Minovich, A., Farnell, J., Neshev, D. N., McKerracher, I., Karouta, F., Tian, J., Powell, D. A., Shadrivov, I. V., Tan, H. H., Jagadish, C. and Kivshar, Y. S. (2012). Liquid crystal based nonlinear fishnet metamaterials, *Appl. Phys. Lett.* **100**, p. 121113.
56. Decker, M., Kremers, C., Minovich, A., Staude, I., Miroshnichenko, A. E., Chigrin, D., Neshev, D. N., Jagadish, C. and Kivshar, Y. S. (2013). Electro-optical switching by liquid-crystal controlled metasurfaces, *Opt. Express.* **21**, pp. 8879–8885. doi: 10.1364/OE.21.008879. URL `http://www.opticsexpress.org/abstract.cfm?URI=oe-21-7-8879`.
57. Xiao, S., Chettiar, U. K., Kildishev, A. V., Drachev, V., Khoo, I. C. and Shalaev, V. M. (2009). Tunable magnetic response of metamaterials, *Appl. Phys. Lett.* **95**(3) p. 033115. doi: http://dx.doi.org/10.1063/1.3182857. URL `http://scitation.aip.org/content/aip/journal/apl/95/3/10.1063/1.3182857`.
58. Minovich, A., Neshev, D. N., Powell, D. A., Shadrivov, I. V. and Kivshar, Y. S. (2010). Tunable fishnet metamaterials infiltrated by liquid crystals, *Appl. Phys. Lett.* **96**, p. 193103.
59. Buchnev, O., Ou, J. Y., Kaczmarek, M., Zheludev, N. I. and Fedotov, V. A. (2013). Electro-optical control in a plasmonic metamaterial hybridised with a liquid-crystal cell, *Opt. Express.* **21**(2), pp. 1633–1638. doi: 10.

1364/OE.21.001633. URL `http://www.opticsexpress.org/abstract.cfm?URI=oe-21-2-1633`.

60. Sautter, J., Staude, I., Decker, M., Rusak, E., Neshev, D. N., Brener, I. and Kivshar, Y. S. (2015). Active tuning of all-dielectric metasurfaces, *ACS Nano*. **9**(4), pp. 4308–4315.
61. Kymeta. URL `https://www.kymetacorp.com`.

CHAPTER 10

Spatial Solitonic and Nonlinear Plasmonic Aspects of Metamaterials

ALLAN D. BOARDMAN[*], **ALESANDRO ALBERUCCI**[†], **GAETANO ASSANTO**[‡], **YU. G. RAPOPORT**[§,¶], **VLADIMIR V. GRIMALSKY**[‖], **VASYL M. IVCHENKO**[§] **and EUGEN N. TKACHENKO**[¶]

[*]University of Salford, UK
[†]Tampere University of Technology, Finland
[‡]University "Roma Tre", Italy
[§]Space Research Institute NAS of Ukraine and SSA of Ukraine, Ukraine
[¶]University of Kyiv, Ukraine
[‖]Autonomous University of State Morelos (UAEM), Mexico

10.1. Introduction

Solitons are well-known globally with roots in hydrodynamics, dating back[1] to John Scott Russell in 1834. The investigation of the ability of new substances, like metamaterials, to permit soliton[2] propagation is

[†]Corresponding author: alessandro.alberucci@gmail.com

therefore a beautiful modern development. For a realistic discussion, however, it is convenient to focus upon only certain members of both the large soliton family and the growing metamaterial family. For this reason, without loss of generality, the attention of this chapter will be upon spatial solitons[3] and both double-negative and hyperbolic metamaterials.[4] Spatial solitons are beams of electromagnetic energy that rely upon balancing diffraction and nonlinearity in order to retain their shape. This property, contrasts nicely with temporal solitons, which are pulses that rely upon balancing phase changes across their width that arise from material dispersion and nonlinearity. Spatial solitons, like their temporal counterparts, are stable if they are the outcomes of what is known as the 1D cubic nonlinear Schrödinger equation.[1] This is because, even though a beam of electromagnetic energy in an unconfined, bulk, medium, possesses two degrees of freedom perpendicular to its propagation direction, and can balance diffraction with nonlinearity, the balance is unstable. Nevertheless, it was shown, some time ago, that placing a beam in a planar waveguide produces stability in an elegant fashion, and this is the basic model adopted here.[5] Within a planar waveguide, a stable soliton can be created by permitting the beam to diffract in the plane of the guide and the role of any diffraction-management that is present, naturally, or artificially created, will be important. In fact, for positive phase materials, diffraction-management, and also managing dispersion, has been investigated already,[6] not only for optical fibers, but also for spatial solitons controlled through the deployment of waveguide arrays.[7]

The fundamental 1D nonlinear Schrodinger equation is a satisfactory model to describe how electromagnetic beams behave in a bounded dielectric, but important additions to this basic equation have to be investigated as well. One of these is the appearance of nonparaxiality when the slowly varying amplitude approximation is partially relaxed. However, for this chapter, it is the nonlinearly-induced diffraction[2,8,9] that is the vital addition to the nonlinear Schrodinger equation. This type of diffraction dominates as the beams become very narrow and can dominate any non-paraxial terms and quintic

nonlinearity in its task of preventing beam collapse at high powers. The coupling of nonlinear diffraction and the possibility of managing the usual diffraction[10] is exciting, and is something that will be discussed below through the appearance of the negative phase behavior that is critically associated with metamaterials.[11] If the core nonlinear Schrodinger equation is modified, in order to take into account other effects, the solutions will still be referred to as 'solitons'. This is,[12] a convenient, and common, practice when dealing with excitations that look and behave like solitons throughout their range of applicability.

Investigating solitons in a nonlinear negative metamaterial gives rise to a generalized nonlinear Schrödinger equation,[13] which stresses that by using the dispersion brought in by the relative dielectric permittivity and the relative permeability a lot of new features emerge. For example, it soon becomes apparent that the sign of the self-steepening parameter can be changed through frequency management. This has a dramatic impact upon modulation instability and short pulse propagation.[14–17] Another example is that, near to the band edges of the relative dielectric permittivity and the relative permeability, slow solitons may be possible[18] and the intensities needed to launch such interesting beams can be assessed. The work that will now unfold does *not address* dispersion but *diffraction* leading to dimensionless equations that deploy coefficients calculated for a purely monochromatic beam. It is interesting, then, that the metamaterial properties emerge as an influence upon the nonlinearly induced diffraction.

The nonlinear Schrödinger equation developed below, concentrates upon diffraction-management and nonlinear diffraction. The absence of any description of loss is because, as recent literature shows, it is possible to minimize the role of loss,[19–22] and still introduce the fundamental double negative behavior that is such a globally attractive feature of negative phase metamaterials. At this stage, it is also worth mentioning that there is a perfect analogy between the general, multidiscipline, nonlinear electromagnetic Schrödinger work and the study of unstable waves on deep oceans. It is not a surprise

that such instability in the hydrodynamic area was originally met with considerable, and broadly based, skepticism. However, now it has been shown to be correct that the conditions for modulation instability are the conditions for the creation of solitons. This applies to both spatial and temporal solitons.[23,24]

Taking another route, investigations of plasma oscillations and waves in metallic and semiconductor nanostructures are of a great interest, however, due to using these structures as building blocks in metamaterials.[25] Plasmonic metamaterials have applications in nanocircuits, lasers, lenses with super-resolution and electromagnetic concentrators. An important direction of metamaterial research and applications for THz and optical transmissions lines, nanocircuits, lasers, sensors, subwavelength imaging, and electromagnetic concentrators[26] is metamaterial plasmonics. Resonances of surface plasmons can now be exploited for concentration of electromagnetic field near plasmonic nanoparticles, and, respectively, enhancement of effective (resonant) nonlinearity[27] and sensitivity of corresponding nonlinear devices. Because of the concentration of the electromagnetic energy in bounded plasmas, various nonlinear phenomena can be realized there, like harmonic generation and self-action.[28] Usually the nonlinearity is considered as moderate, when various perturbation methods are applied. Generation of plasmonic second-harmonics in the set of nanoparticles with proper asymmetry and plasmonic solitons in a "metal-dielectric" structure have been considered in Refs. 29 and 30, respectively. Surface nonlinearity of metal nanoparticles has been addressed, for example, in a classic publication of Ref. 28.

For wave propagation at the plasma –dielectric interface, oscillations of the surface charge occur. This surface charge has a high value and is localized near the surface. Therefore, when the nonlinearity in the volume is still moderate, or even very small, the values of electron concentrations can be quite large at the interfaces. Moreover, the electron concentration cannot be negative; this fact limits the use of the modern nonlinearity itself, because the amplitudes of oscillations of the perturbations of the electron concentration cannot be greater than the equilibrium value of the electron concentration.

Below, the nonlinear electron waves in the plasma layer are investigated. Within the approximation of a moderate surface nonlinearity, the resonant generation of the second harmonic of the surface electron plasma wave is investigated. A general method to take into account both surface and volume nonlinearity has been put forward.[31–35] It is shown that the new frequency regions of modulation instability can be realized. For the (nonlinear) wave processes in mesoscale layered structures,[26,36] where the wavelength λ is of the order of the scale L_{inh} of the medium inhomogeneity, $\lambda \sim L_{\text{inh}}$ in one part of the system and $\lambda \ll L_{\text{inh}}$ in the other part of the system, the combined method of "Complex geometrical optics-Full-Wave nonlinear electromagnetics" has been developed.[26,36] At the same time, the grading in electromagnetic field concentrator, at least in the linear approximation, is determined by the ideas of providing an ideal all-direction concentration,[37] similar to the transformational optics.[38] The methods for the derivation of the nonlinear evolution equations for (wave processes in) layered structures (NEELS)[26,31–35] are, in fact, the parts of the general metamaterial approach to the wave processes, applicable for both artificial[26] and natural[36] media.

The general case that includes the strong surface nonlinearity is simulated. The excitation of surface electron plasma waves in the plasma slab is investigated with an initial oscillatory distribution of perturbations of the volume electron concentration. The main result is the occurrence of the strong nonlinearity at the interfaces, when the amplitudes of initial perturbations exceed some threshold. The collapse of nonlinear plasma waves takes place. The nonlinearity cannot be considered as moderate there. The values of the electron concentration near the plasma boundaries reach extremely high values, and sharp peaks of concentration are formed that are localized both in longitudinal and transverse directions. A similar behavior occurs when the oscillatory distributions of the longitudinal electron velocity are given initially. Below the threshold for observing wave collapse, the surface nonlinearity dominates over the volume one but can be considered as moderate.

10.2. Solitons in Metamaterials

10.2.1. *Bright spatial solitons in double-negative metamaterials*

10.2.1.1. *General discussion*

Suppose that ω is the angular frequency of a guided wave in the planar structure shown in Fig. 10.1, then it can be associated with the Fourier transforms $\tilde{\mathbf{D}}$, $\tilde{\mathbf{B}}$,$\tilde{\mathbf{E}}$,$\tilde{\mathbf{H}}$, $\tilde{\mathbf{P}}_{\mathrm{NL}}$, $\tilde{\mathbf{M}}_{\mathrm{NL}}$, of the displacement vector, the magnetic flux density vector, the electric field, the magnetic field, the dielectric nonlinear polarization and the nonlinear magnetization, respectively. Bianisotropy, anisotropy or spatial dispersion, is not taken into account at this stage. Any nonlinearity is introduced by the polarizations that are labelled with the subscript NL. In the time domain, t, these fundamental quantities will be simply designated as $\mathbf{D}$, $\mathbf{B}$, $\mathbf{E}$, $\mathbf{H}$, $\mathbf{P}_{\mathrm{NL}}$, $\mathbf{M}_{\mathrm{NL}}$.

For the type of metamaterial capable of sustaining both nonlinear electric and magnetic polarizations, the frequency Fourier transforms of the constitutive relations are

$$\tilde{\mathbf{D}}(\mathbf{r},\omega) = \varepsilon_0\varepsilon(\omega)\tilde{\mathbf{E}}(\mathbf{r},\omega) + \tilde{\mathbf{P}}_{\mathrm{NL}}(\mathbf{r},\omega), \tag{10.1}$$

$$\tilde{\mathbf{B}}(\mathbf{r},\omega) = \mu_0\mu(\omega)\tilde{\mathbf{H}}(\mathbf{r},\omega) + \mu_0\tilde{\mathbf{M}}_{\mathrm{NL}}(\mathbf{r},\omega), \tag{10.2}$$

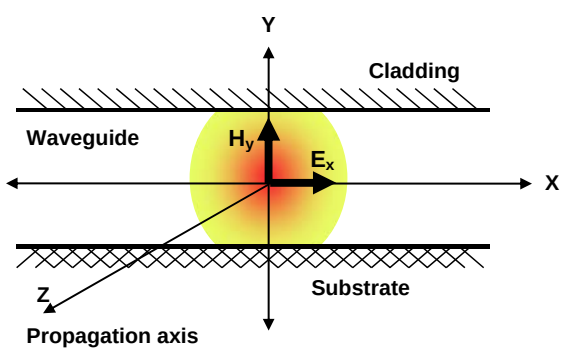


Fig. 10.1. Typical planar waveguide structure sustaining an optical beam characterized by an electric field $\mathbf{E}$ and a magnetic field $\mathbf{H}$. Propagation of the beam is along the z-axis and diffraction is only possible in the x-directions.

where $\varepsilon(\omega)$ and $\mu(\omega)$ are, respectively, relative permittivity and permeability functions. ε_0 and μ_0 are the standard free-space values of the permittivity and permeability. In order to observe spatial solitons, a monochromatic beam, with a frequency ω_0, should be launched into the kind of planar guide sketched in Fig. 10.1. Maxwell's equations then yield the fundamental equations

$$\nabla^2\mathbf{E} - \nabla(\nabla\cdot\mathbf{E}) + k_0^2(\omega_0)\mathbf{E} + \frac{\omega_0^2}{c^2}\mu(\omega_0)\frac{\mathbf{P}_{\mathrm{NL}}}{\varepsilon_0} + i\omega_0\mu_0\nabla\times\mathbf{M}_{\mathrm{NL}} = 0, \tag{10.3}$$

$$\nabla^2\mathbf{H} - \nabla(\nabla\cdot\mathbf{H}) + k_0^2(\omega_0)\mathbf{H} + \frac{\omega_0^2}{c^2}\varepsilon(\omega_0)\mathbf{M}_{\mathrm{NL}} - i\omega_0\nabla\times\mathbf{P}_{\mathrm{NL}} = 0, \tag{10.4}$$

in which $k_0^2(\omega_0) = \frac{\omega_0^2}{c^2}\varepsilon(\omega_0)\mu(\omega_0)$.

The notation can be made simpler by using the definitions $k \equiv k_0(\omega_0), \varepsilon \equiv \varepsilon(\omega_0)$ and $\mu \equiv \mu(\omega_0)$. For a TE-polarized beam, the field vectors are $\mathbf{E} = \hat{\mathbf{x}}E_x$ and $\mathbf{H} = \hat{\mathbf{y}}H_y$, where $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are standard unit vectors. Nonlinear polarizations, generated by Kerr-like material properties, are $P_{\mathrm{NL}} = \varepsilon_0\varepsilon_{NL}^{(3)}|E_x|^2E_x$ and $M_{\mathrm{NL}} = \mu_{NL}^{(3)}|H_y|^2H_y$, where $\varepsilon_{\mathrm{NL}}^{(3)}$ and $\mu_{\mathrm{NL}}^{(3)}$ are, respectively, the electric and magnetic cubic nonlinearity coefficients. Although this assumption avoids introducing any saturation, such a medium can be readily modelled, for example, by including quintic terms in this nonlinear model.

Any fast spatial variation can be accounted for by making the transformation $H_y \Rightarrow H_y(x,z)e^{ik_0z}$. Hence, with this procedure, H_y is replaced by the slowly varying function $H_y(x,z)$. In the same manner $E_x \Rightarrow E_x(x,z)e^{ikz}$. Retaining only the significant order terms, the equations for the TE-mode components of $\mathbf{E}$ and $\mathbf{H}$ are

$$2ik\frac{\partial E_x}{\partial z} + \frac{\partial^2}{\partial x^2}E_x + \frac{\omega_0^2}{c^2}[\mu\varepsilon_{\mathrm{NL}}^{(3)}|E_x|^2 + \varepsilon\mu_{\mathrm{NL}}^{(3)}|H_y|^2]E_x = 0, \tag{10.5}$$

$$2ik\frac{\partial H_y}{\partial z} + \frac{\partial^2}{\partial x^2}H_y + \frac{\omega_0^2}{c^2}[\mu\varepsilon_{\mathrm{NL}}^{(3)}|E_x|^2 + \varepsilon\mu_{\mathrm{NL}}^{(3)}|H_y|^2]H_y = 0. \tag{10.6}$$

At this stage, it should be noted that the nonlinear coefficients $\varepsilon_{\rm NL}^{(3)}$ and $\mu_{\rm NL}^{(3)}$ can be positive, or negative, and that a sign has not been given to the wave number k. Also

$$|H_y|^2 = \frac{\varepsilon_0|\varepsilon|}{\mu_0|\mu|}|E_x|^2. \tag{10.7}$$

Equations (10.5) and (10.6) can now be modified by including higher-order linear diffraction, nonlinear diffraction and non-paraxiality. Nonlinear diffraction is dominant,[11,12] and the quintic contribution to the nonlinearity can be taken as negligible. The final form of the nonlinear Schrodinger equations to be used in the numerical simulations, shown later, assumes that narrow beam formation requires the dominance of nonlinear diffraction over non-paraxial terms. Hence, the latter will not be included but higher-order linear diffraction will be retained because it may emerge into prominence during any diffraction-management stage. The next set of equations recognizes that the nonlinear diffraction of the TE-polarized beam equation depends upon the condition $\nabla \cdot \mathbf{E} \neq 0$ and that the TM-polarized beam equation depends upon the condition $\nabla \cdot \mathbf{H} \neq 0$.

The evolution equations for spatial solitons are, therefore:

(*a*) *TE-polarized beams*

$$i\frac{\partial E_x}{\partial z} + \frac{1}{2k}\frac{\partial^2 E_x}{\partial x^2} - \frac{1}{8k^3}\frac{\partial^4 E_x}{\partial x^4} + \frac{1}{2k}\frac{\omega_0^2}{c^2}\left[\mu\varepsilon_{\rm NL}^{(3)} + \varepsilon\mu_{\rm NL}^{(3)}\frac{\varepsilon_0|\varepsilon|}{\mu_0|\mu|}\right]|E_x|^2E_x$$
$$+ \frac{1}{2\varepsilon k}\varepsilon_{\rm NL}^{(3)}\frac{\partial^2}{\partial x^2}(|E_x|^2E_x) = 0, \tag{10.8}$$

which shows clearly that *only* the dielectric nonlinear polarization can produce the nonlinear diffraction. In the now classic double-negative metamaterial, $\mu = -|\mu|$ and $\varepsilon = -|\varepsilon|$ and backward waves exist, for which $k = -|k|$. For this scenario, bright spatial solitons only exist whenever $N_{TE} = \left(|\mu|\varepsilon_{\rm NL}^{(3)} + |\varepsilon|\mu_{\rm NL}^{(3)}\frac{\varepsilon_0|\varepsilon|}{\mu_0|\mu|}\right) < 0$. A sufficient condition to make $N_{\rm TE} < 0$ is to make $\varepsilon_{\rm NL}^{(3)} < 0$ and $\mu_{\rm NL}^{(3)} < 0$. However, it is not a necessary condition and it is possible to have either $\varepsilon_{\rm NL}^{(3)} > 0$, or $\mu_{\rm NL}^{(3)} > 0$, according to their magnitude and

it does not need to happen simultaneously. This is very interesting because, although the introduction of a metamaterial could be directed towards lenses,[39] the coefficient of the nonlinear diffraction term is entirely controllable with such artificial materials.[40] Gap solitons are also of interest.[41]

(*b*) *TM-polarized beams*

$$i\frac{\partial H_x}{\partial z}+\frac{1}{2k}\frac{\partial^2 H_x}{\partial x^2}-\frac{1}{8k^3}\frac{\partial^4 H_x}{\partial x^4}+\frac{1}{2k}\frac{\omega_0^2}{c^2}\left[\varepsilon\mu_{\rm NL}^{(3)}+\mu\varepsilon_{\rm NL}^{(3)}\frac{\mu_0|\mu|}{\varepsilon_0|\varepsilon|}\right]|H_x|^2H_x+\frac{1}{2\mu k}\mu_{\rm NL}^{(3)}\frac{\partial^2}{\partial x^2}(|H_x|^2H_x)=0. \tag{10.9}$$

Here, the vector $\mathbf{H}=(H_x,0,0)$ is used for the TM beam and *only* the magnetic polarization can produce the nonlinear diffraction. Once again, it is the signs of $k,\varepsilon_{\rm NL}^{(3)}$ and $\mu_{\rm NL}^{(3)}$ that determine the behavior of the critical coefficient $N_{\rm TM}=\left(|\varepsilon|\mu_{\rm NL}^{(3)}+|\mu|\varepsilon_{\rm NL}^{(3)}\frac{\mu_0|\mu|}{\varepsilon_0|\varepsilon|}\right)<0$.

For the propagation of bright spatial solitons in a double-negative metamaterial, each polarization, has a fast variation that is a backward wave, with a wave number $k=-|k|$. The beam propagation is along the z-axis, and this can be scaled in the familiar manner by setting $z=|k|w^2Z$ and $x=wX$, where w is the beam width that will be measured as an integral number of wavelengths. In fact, it will be set equal to $m\lambda$, where m is an integer and λ is the wavelength of the guided beam. At this stage, by legitimately ignoring higher-order linear diffraction term, *both* nonlinear Schrödinger equations become

$$i\frac{\partial\psi}{\partial Z}-\frac{1}{2}\frac{\partial^2\psi}{\partial X^2}-|\psi|^2\psi-\kappa\frac{\partial^2}{\partial X^2}(|\psi|^2\psi)=0. \tag{10.10}$$

It is correct to assert that the nonlinear coefficient, in the TE-polarized case, is $\varepsilon_{\rm NL}^{(3)}=-|\varepsilon_{\rm NL}^{(3)}|$ and that $\mu_{\rm NL}^{(3)}$ is negligible, whilst, in the TM-polarized case, it is correct to say that the nonlinear coefficient is $\mu_{\rm NL}^{(3)}=-|\mu_{\rm NL}^{(3)}|$, with $\varepsilon_{\rm NL}^{(3)}$ set as negligible. Physically, for TE beams, the nonlinear diffraction is controlled by $\varepsilon_{\rm NL}^{(3)}$ and for

TM beams, it is controlled by $\mu_{NL}^{(3)}$. In either case, the fundamental control parameter is κ, and this can be modelled for a particular metamaterial. The latter can be varied by changing the operating frequency of the beams, and ψ has been normalized to the following:

$$\textbf{TE: } \psi = \left(\sqrt{\frac{w^2\omega_0^2}{2c^2}|\mu||\varepsilon_{\mathrm{NL}}^{(3)}|} \right) E_x$$

$$\textbf{TM: } \psi = \left(\sqrt{\frac{w^2\omega_0^2}{2c^2}|\varepsilon||\mu_{\mathrm{NL}}^{(3)}|} \right) H_x. \tag{10.11}$$

It is important now to get some idea of how a metamaterial can influence spatial soliton propagation, through the role of κ, and then to go on to investigate whether the diffraction can be controlled in any kind of way. For the waveguide shown in Fig. 10.2

$$\kappa_{\mathrm{TE}} = \kappa_{\mathrm{TM}} = \frac{1}{k^2w^2} = \frac{c^2}{\omega^2w^2(\varepsilon(\omega)\mu(\omega))} \equiv \kappa, \tag{10.12}$$

where k is the wave number. If the well-known Drude model is adopted for the permittivity and the permeability, associated, respectively, with plasma frequencies ω_{pe} and ω_{pm}, then both the permittivity and permeability assume Drude models, where the ratio of the plasma frequencies is assumed to be $\omega_{\mathrm{pm}}/\omega_{\mathrm{pe}} = 0.6$. Positive phase material yields the red line, and the negative phase metamaterials

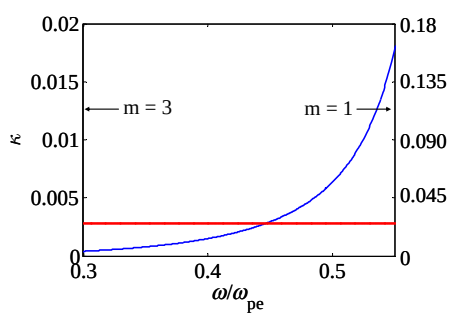


Fig. 10.2. Plot of the variation of the nonlinear diffraction coefficient, κ, with frequency normalized to the dielectric resonance (plasma) frequency, ω_{pe}. The permeability resonance frequency is ω_{pm}.

yields the blue curve. Left-hand scale, $m = 3$, for which the input beam width is 3λ and the right-hand scale, $m = 1$, is for an input beam width equal to λ.

Clearly, κ changes very rapidly with frequency compared to using a normal positive phase material. It is important, therefore, to investigate how nonlinear diffraction can be controlled with the use of metamaterials. One method of control can be called diffraction-management, which will now be discussed.

10.2.1.2. *Diffraction-management*

A form of diffraction-management can be arranged using the arrangement shown in Fig. 10.3, which is assumed to be impedance-matched so that any degradation due to inherent mismatch need not be considered. A nonlinear positive phase medium (PPM) competes with nonlinear negative phase medium (NPM) and creates an influence over spatial soliton beam propagation along the z-axis. The assessment of this process will now be presented in a way that will yield the final form of a slowly-varying envelope equation capable of being used for bright spatial soliton applications.

As shown in Fig. 10.3, the unit cell is split into two lengths, l_1L and l_2L, each containing a PPM and an NPM material, respectively. For a TE-polarized spatial soliton, the appropriate form of the nonlinear Schrödinger equation, in each part of the unit cell, is

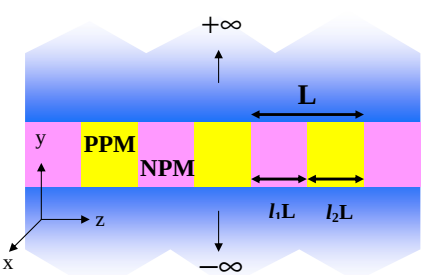


Fig. 10.3. A planar waveguide structure consisting of alternating layers of PPM and NPM. z is the propagation direction and diffraction that takes place along the x-axis. Note that the periodic structure has a unit cell of length L.

PPM:

$$2i\frac{\partial E_x}{\partial z}+\frac{1}{k_1}\frac{\partial^2 E_x}{\partial x^2}+\frac{\omega^2}{c^2 k_1}\varepsilon^{(3)}_{\mathrm{NL1}}|E_x|^2E_x - \frac{1}{4k_1^3}\frac{\partial^4 E_x}{\partial x^4}+\frac{\varepsilon^{(3)}_{\mathrm{NL1}}}{\varepsilon_1 k_1}\frac{\partial^2}{\partial x^2}(|E_x|^2E_x)=0, \qquad (10.13)$$

NPM:

$$2i\frac{\partial E_x}{\partial z}-\frac{1}{|k_2|}\frac{\partial^2 E_x}{\partial x^2}-\frac{\omega_0^2}{c^2|k_2|}|N_{\mathrm{TE}}||E_x|^2E_x + \frac{1}{4|k_2^3|}\frac{\partial^4 E_x}{\partial x^4}+\frac{\varepsilon^{(3)}_{\mathrm{NL2}}}{|\varepsilon_2||k_2|}\frac{\partial^2}{\partial x^2}(|E_x|^2E_x)=0, \qquad (10.14)$$

where

$$N_{\mathrm{TE}}=\left(|\mu|\varepsilon^{(3)}_{\mathrm{NL}}+|\varepsilon|\mu^{(3)}_{\mathrm{NL}}\frac{\varepsilon_0|\varepsilon|}{\mu_0|\mu|}\right)$$

and the subscripts 1 and 2 are used to label the regions of the unit cell. For a TM-polarized beam, N_{TE} is simply substituted by N_{TM}. For this polarization, the final term in Eq. (10.14) is absent because, in a PPM, there is no nonlinear magnetic polarization. In the NPM, however, the TM-polarization does permit a nonlinear diffraction term, for which the coefficient is $\mu^{(3)}_{\mathrm{NL2}}/(|\mu_2||k_2|)$. T is interesting. Note that *both* nonlinear diffraction and higher-order diffraction are included *at this stage*. Building upon the previous discussion, however, it can be assumed, now, that the nonlinear diffraction actually dominates over the other possible contributions from non-paraxiality and the impact of any possible quintic nonlinearity and any nonlinear saturation that could prevent beam collapse. The procedure is to consider the unit cell and introduce all the terms in the Schrödinger equations one at a time. The coefficients are adjusted according to whether they refer to the PPM, or the NPM, part of this cell. In other words, the averages are simply going to be integrations with respect to z over L, but it must be accepted that this type of unit cell is smaller in scale than a diffraction length measured as k_1w^2 or $|k_2|w^2$, where $1, 2$ label the unit cell regions and w is the width of the

spatial soliton. It is possible now to look at the unit cell in Fig. 10.3 and perform a straightforward average of each term in the envelope equation for TE modes, for example. The integration, for each term, is over the distance L and this is split into parts concerning the PPM and NPM, which are small enough to make the terms appear to be constant over each section. Useful outcomes of this strategy are

Linear diffraction:

$$\frac{1}{\mathrm{L}}\int_0^{\mathrm{L}} \frac{1}{k}\frac{\partial^2 E_x}{\partial x^2}dz \approx \frac{1}{k_1}\left(l_1 - \frac{k_1}{|k_2|}l_2\right)\frac{\partial^2 E_x}{\partial x^2}. \tag{10.15}$$

Nonlinearity:

$$\left(\left(\frac{\omega_0^2\varepsilon_{\mathrm{NL1}}^{(3)}}{c^2 k_1}\right)\frac{1}{\mathrm{L}}\int_0^{\mathrm{l_1L}} dz - \left(\frac{\omega_0^2}{c^2|k_2|}\right)\frac{1}{\mathrm{L}}\int_{l_1\mathrm{L}}^{\mathrm{L}} |N_{\mathrm{TE2}}|dz\right)(|E_x|^2E_x)$$
$$\approx \frac{\omega_0^2}{c^2 k_1}\left(l_1\varepsilon_{\mathrm{NL1}}^{(3)} - \frac{k_1}{|k_2|}l_2|N_{\mathrm{TE2}}|\right)(|E_x|^2E_x). \tag{10.16}$$

Nonlinear diffraction:

$$\frac{1}{L}\int_0^L \frac{\varepsilon_{\mathrm{NL}}^{(3)}}{\varepsilon k}\frac{\partial^2}{\partial x^2}(|E_x|^2E_x) \approx \frac{1}{k_1}\left(l_1\frac{\varepsilon_{\mathrm{NL1}}^{(3)}}{\varepsilon_1} + \frac{k_1}{|k_2|}l_2\frac{\varepsilon_{\mathrm{NL2}}^{(3)}}{|\varepsilon_2|}\right)\frac{\partial^2}{\partial x^2}(|E_x|^2E_x). \tag{10.17}$$

Higher-order diffraction:

$$\frac{1}{L}\int_0^L \frac{1}{4k^3}\frac{\partial^4 E_x}{\partial x^4} \approx \frac{1}{4k_1^3}\left(l_1 - l_2\frac{k_1^3}{|k_2^3|}\right)\frac{\partial^4 E_x}{\partial x^4}. \tag{10.18}$$

The net results are the following equations:
TE:

$$2ik_1\frac{\partial E_x}{\partial z} + D\frac{\partial^2 E_x}{\partial x^2} + \frac{\omega_0^2}{c^2}\left(l_1\varepsilon_{\mathrm{NL1}}^{(3)} - \frac{k_1}{|k_2|}l_2|N_{\mathrm{TE2}}|\right)|E_x|^2E_x$$
$$-\frac{F}{4k_1^2}\frac{\partial^4 E_x}{\partial x^4} + \left(l_1\frac{\varepsilon_{\mathrm{NL1}}^{(3)}}{\varepsilon_1} + \frac{k_1}{|k_2|}l_2\frac{\varepsilon_{\mathrm{NL2}}^{(3)}}{|\varepsilon_2|}\right)\frac{\partial^2}{\partial x^2}(|E_x|^2E_x) = 0. \tag{10.19}$$

TM:

$$2ik_1\frac{\partial H_x}{\partial z}+D\frac{\partial^2 H_x}{\partial x^2}+\frac{\omega_0^2}{c^2}\left(l_1\varepsilon_{\mathrm{NL1}}^{(3)}\frac{\mu_0|\mu|}{\varepsilon_0|\varepsilon|}-\frac{k_1}{|k_2|}l_2|N_{\mathrm{TM2}}|\right)|H_x|^2H_x$$
$$-\frac{F}{4k_1^2}\frac{\partial^4 H_x}{\partial x^4}+\left(\frac{k_1}{|k_2|}l_2\frac{\mu_{\mathrm{NL2}}^{(3)}}{|\mu_2|}\right)\frac{\partial^2}{\partial x^2}(|H_x|^2H_x)=0, \qquad (10.20)$$

where the linear diffraction and higher-order linear diffraction coefficients are

$$D=l_1-\frac{\sqrt{\varepsilon_1\mu_1}}{|\sqrt{\varepsilon_2\mu_2}|}l_2, \quad F=l_1-\left(\frac{\varepsilon_1\mu_1}{|\varepsilon_2\mu_2|}\right)^{\frac{3}{2}}l_2 \qquad (10.21)$$

and ε_i, μ_i are the relative permittivities and permeabilities of the parts of the unit cell shown in Fig. 10.3. The control process will now aim at reducing D so that any first-order linear diffraction can be minimized and enable nonlinear diffraction to become important. If the first-order linear diffraction is reduced in this way, however, it opens up the possibility, in principle, that the higher-order linear diffraction term, controlled by the parameter F, is not minimized. However, $\frac{k_1}{|k_2|}$ does not have to be unity, and higher-order linear diffraction can be changed by simply manipulating the structure, in order to control D *or* F. This means that for any choice of D, it is possible to arrange for the ratio $\frac{k_1}{|k_2|}$ to make the contribution of higher-order linear diffraction negligible. This good outcome shows that a reduction of D does not necessarily mean that higher-order linear diffraction has to be introduced. This conclusion means that it will not be regarded here as an important contribution to the envelope equation. Instead, the emphasis will be directed towards first-order diffraction-management. Also, even though non-paraxiality is reduced as D tends to zero, nonlinear diffraction is the dominant influence as beams become narrower. Indeed, it takes over the role of preventing beam collapse. Clearly, if a nonlinear material is selected that has a large quintic contribution, then the latter will compete with the nonlinear diffraction. There are two broad scenarios, namely, (a) no diffraction-management but the nonlinear diffraction becomes a very important influence as the beams narrow and (b) manipulating

D, to engage in some form of linear diffraction-management that can be arranged in a number of ways to make $D \rightarrow D(z)$ through the deployment of a metamaterial.

The scales $z = (kw^2)Z$ and $x = (w)X$ can still be used, where w is actually an arbitrary unit but can be physically interpreted as the beam width. Hence,

$$i\frac{\partial\psi}{\partial Z} + \frac{D}{2}\frac{\partial^2\psi}{\partial X^2} + |\psi|^2\psi + \kappa_{\mathrm{TE,TM}}\frac{\partial^2}{\partial X^2}(|\psi|^2\psi) = 0, \tag{10.22}$$

where

TE:

$$\psi = w\sqrt{G_1^{\mathrm{TE}}}E_x, \quad G_1^{\mathrm{TE}} = \frac{\omega_0^2}{c^2}\left(l_1\varepsilon_{\mathrm{NL1}}^{(3)} - \frac{k_1}{|k_2|}l_2|N_{\mathrm{TE2}}|\right),$$

$$G_2^{\mathrm{TE}} = \left(l_1\frac{\varepsilon_{\mathrm{NL1}}^{(3)}}{\varepsilon_1} + \frac{k_1}{|k_2|}l_2\frac{\varepsilon_{\mathrm{NL2}}^{(3)}}{|\varepsilon_2|}\right). \tag{10.23}$$

TM:

$$\psi = w\sqrt{G_1^{\mathrm{TM}}}H_x, \quad G_1^{\mathrm{TM}} = \frac{\omega_0^2}{c^2}\left(l_1\varepsilon_{\mathrm{NL1}}^{(3)}\frac{\mu_0|\mu|}{\varepsilon_0|\varepsilon|} - \frac{k_1}{|k_2|}l_2|N_{\mathrm{TM2}}|\right),$$

$$G_2^{\mathrm{TM}} = \left(\frac{k_1}{|k_2|}l_2\frac{\mu_{\mathrm{NL2}}^{(3)}}{|\mu_2|}\right) \tag{10.24}$$

and for each polarization $\kappa_{\mathrm{TE}} = \frac{G_2^{\mathrm{TE}}}{w^2G_1^{\mathrm{TE}}}$ and $\kappa_{\mathrm{TM}} = \frac{G_2^{\mathrm{TM}}}{w^2G_1^{\mathrm{TM}}}$.

Equations (10.22)–(10.24) reveal that a number of metamaterial types could support typical bright spatial solitons. However, the focus here is upon two quite realistic cases. PPM, for TE beams, contributes the major part of the nonlinearity. In this case, $\varepsilon_{\mathrm{NL2}}^{(3)} = 0$ and $\mu_{\mathrm{NL2}}^{(3)} = 0$ and, of course, $\mu_{\mathrm{NL1}}^{(3)} = 0$. Without nonlinear diffraction, it is not necessary for TM beams to use only magnetic nonlinearity. However, if κ_{TM} is required, then it will be assumed that this applies to the case when the nonlinearity is located within the NPM, and that $\varepsilon_{\mathrm{NL2}}^{(3)} = 0, \varepsilon_{\mathrm{NL1}}^{(3)} = 0$ and $\mu_{\mathrm{NL2}}^{(3)} \neq 0$. For the diffraction-managed

cases to be presented below, therefore, $\kappa_{\rm TE}$ and $\kappa_{\rm TM}$ reduce to the simple forms

$$\kappa_{\rm TE} = \frac{1}{k_1^2 w^2}, \quad \kappa_{\rm TM} = \frac{1}{k_2^2 w^2}. \tag{10.25}$$

10.2.1.3. *Simulations*

Nonlinear diffraction will prevent beam collapse as the power carried by a spatial soliton increases and results in a limit on how narrow the beams can be. It is a process that dominates non-paraxiality and even quintic nonlinearity, provided the latter is not too large. Nonlinear diffraction is still, nevertheless, a diffraction process so that as beams become narrow, and the power goes up, and it will still have to compete with the linear diffraction, unless the latter is managed to be of small influence.

The question of soliton control has been addressed[39,40] as has the properties of gap solitons.[41] The latter, however deals with quadratic nonlinear material that does not feature in this discussion. All of the issues addressed here involve only bright spatial solitons driven by third-order nonlinearity. Figure 10.4 shows the behavior of bright spatial solitons in a homogeneous metamaterial guide for which $l_2 = 0$ and both TE and TM modes have the same nonlinear diffraction coefficient

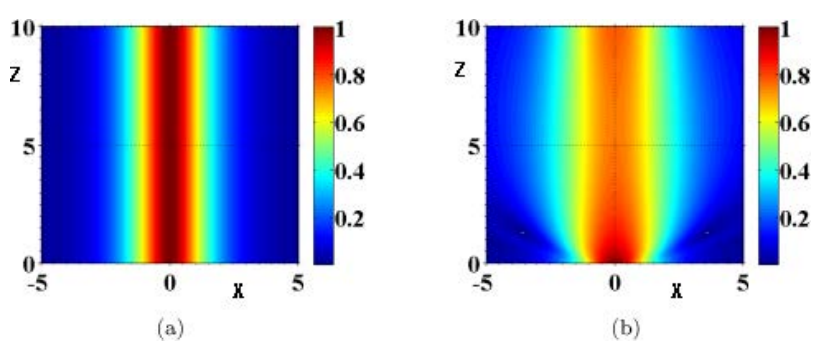
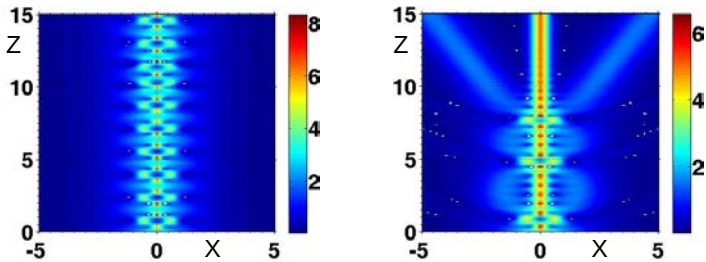


Fig. 10.4. Behavior of a bright soliton beams, with beam width the order of a wavelength ($m = 1$), ($m = 1$) in a homogeneous guide ($D = 1$). Initial (input) beam shape is sec$h(X)$. (a) $\kappa = 0$, (b) $\kappa = 0.17$.



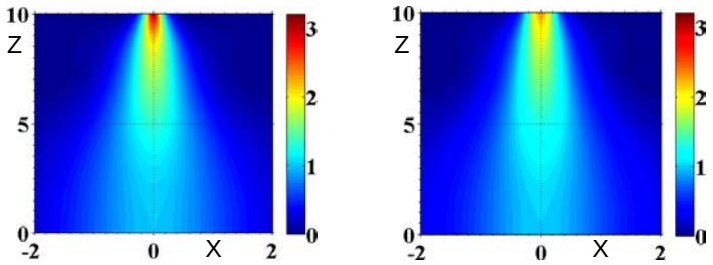
(a) (b)

Fig. 10.5. Initial input: 3 sec$h(X)$, $m = 3$. No linear diffraction management. (a) $\kappa = 0$, (b) $\kappa = 0.0028$.

κ. Figure 10.4(a) shows the behavior in the absence of nonlinear diffraction. There is no diffraction management so, in 10.4(b), the presence of nonlinear diffraction creates a limitation, with a lowering of intensity and the introduction of some broadening.

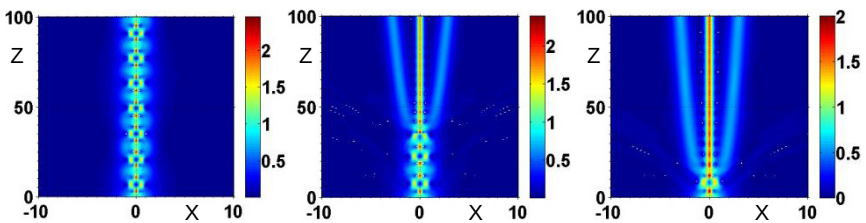
In the *absence* of diffraction-management, Fig. 10.5. shows what happens to a higher order, breathing soliton if nonlinear diffraction is introduced. The presence of 100% of the linear diffraction means that a first-order soliton, namely a sec$h(X)$ input, would need to be a beam with a width the order of a wavelength for the nonlinear diffraction to disturb its stability. Figure 10.5(a) shows, however, that a higher-order soliton, such as 3 sec$h(X)$ is, in fact, three first-order solitons held together with zero binding energy. Unperturbed, this soliton will propagate down the z-axis as a breather. Figure 10.5(b) shows that it is rapidly perturbed by the presence of nonlinear diffraction. The breathing is interrupted and a narrow, first-order soliton is created, plus some low-energy radiated beams. However, this behavior could be frequency-controlled. It can be seen that there is a special position on the z-axis at which the beam starts to split, where control by a metamaterial, an appropriate frequency, could be introduced.

In order to achieve the outcomes in Fig. 10.6, diffraction-management is created, by the deployment of a double-negative metamaterial. To begin with, Fig. 10.6. shows, in the absence of nonlinear diffraction, what happens when the linear diffraction is reduced from its starting value, on the input plane, to 10% on the



(a) (b)

Fig. 10.6. Illustrating the impact of linear diffraction management. Initial inputs: sec$h(X)$, $m = 3$. The diffraction-management factor is reduced from an input value of 100%–10% over a distance of 10 Rayleigh lengths setting (a) $\kappa = 0$ and (b) $\kappa = 0.0028$.



(a) (b) (c)

Fig. 10.7. Linear diffraction set at 10%. Inputs: $\psi = \text{sec}h(X)$. (a) $\kappa = 0$ (b) $\kappa = 0.00168$ (c) $\kappa = 0.005$. In (b) splitting occurs after 45 Rayleigh lengths.

output plane. This is illustrated here, just as an example, by using a linear reduction as progression down the z-axis takes place. It is interesting that the intensity scale is now much more compressed than in the previous figures. $\kappa = 0$ is an interesting case because it can arise simply because the linear diffraction is dominant. κ need not be exactly zero for this outcome to appear. In contrast, $\kappa = 0.0028$ is closer to the κ resonance region. In both cases, ultra-narrow beam formation occurs.

Figure 10.7 shows what happens when the diffraction-management is fixed to reduce the linear diffraction to 10%. For a first-order soliton, launched at $z = 0$, decreasing the linear diffraction to 10% creates a situation in which there is now too much

power available for the propagation of this first-order soliton. In optical fiber work, this kind of diffraction reduction would lead to a breathing soliton of order $\sqrt{10}$. Figure 10.7(a) shows that this is precisely what happens, and a third-order breathing soliton is created. In Fig. 10.7(b), nonlinear diffraction is introduced into the system and it acts as a perturbation that splits the breather. The outcome now is three low-power, pseudo-solitons. However, the power contained within each one is not enough to retain a soliton status if it is entered into a normal, 100% diffraction medium. Each one will simply radiate. In Fig. 10.7(c), a higher value of the nonlinear-diffraction coefficient has been used and, therefore, as a perturbation, it is larger and the breather is split quite quickly.

10.3. Bright Spatial Solitons in Hyperbolic Metamaterials

10.3.1. *Basic features of the metamaterial*

In this type of metamaterial, monochromatic electromagnetic fields, with complex amplitudes that vary with time as $e^{-i\omega t}$, propagate in a non-magnetic dielectric, anisotropic environment, which has a diagonal permeability tensor $\boldsymbol{\mu} = \mu_0 \mathbf{I}$, where $\mathbf{I}$ is the standard unit tensor and μ_0 is the permeability of free space. The dielectric permittivity, however, is characterized by a tensor $\boldsymbol{\epsilon}$, which has off-diagonal terms determined by its uniaxial nature. Two out of three eigenvalues of $\boldsymbol{\epsilon}$ coincide. The latter can defined as $\epsilon_\perp$ (double multiplicity), while the other eigenvalue is defined as $\epsilon_\parallel$ (single multiplicity). From an optical viewpoint, $\epsilon_\perp$ and $\epsilon_\parallel$ are the dielectric permittivities "sensed" by linearly polarized electric fields oscillating normal and parallel to the optic axis $\hat{n}$, respectively, corresponding to the symmetry axis of the medium. Without loss of generality, it can be assumed that $\hat{n}$ lies in the (yz) plane. The dielectric tensor $\boldsymbol{\epsilon}$ then assumes the form

$$\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_{xx} & 0 & 0 \\ 0 & \epsilon_{yy} & \epsilon_{yz} \\ 0 & \epsilon_{zy} & \epsilon_{zz} \end{pmatrix},$$

where, in general, all the elements are space-dependent functions. Cylindrical symmetry implies $\epsilon_{ij} = \epsilon_\perp \delta_{ij} + \epsilon_a n_i n_j (i, j = x, y, z)$, where δ_{ij} is the Kronecker's delta, n_i are the Cartesian components of the optic axis, and $\epsilon_a = \epsilon_\parallel - \epsilon_\perp$ is the optical anisotropy. Given the condition $\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{x}} = 0$, then setting θ to be the angle between $\hat{n}$ and the axis z, yields the tensor components $\epsilon_{xx} = \epsilon_\perp$, $\epsilon_{yy} = \epsilon_\perp + \epsilon_a \sin^2\theta$, $\epsilon_{zz} = \epsilon_\perp + \epsilon_a \cos^2\theta$, and $\epsilon_{yz} = \epsilon_a \sin\theta\cos\theta$. In the absence of sources, and using the label NL to denote nonlinear components, Maxwell's equations give

$$\nabla \times \boldsymbol{E} = i\omega\mu_0 \boldsymbol{H}, \tag{10.26}$$

$$\nabla \times \boldsymbol{H} = -i\omega \boldsymbol{\epsilon} \cdot \boldsymbol{E} - i\omega \boldsymbol{P}_{\mathrm{NL}}, \tag{10.27}$$

where, in Eq. (10.27), the linear and nonlinear contributions to the medium polarizability are separated. Also, in this equation, a local response is assumed in the linear regime that neglects spatial dispersive effects such as optical activity. An extra feature is that, without losing any generality, it is simpler, and perfectly accurate, to conduct investigations with 2D geometries by setting $\partial_x = 0$. In this limit, the ordinary and the extraordinary waves are decoupled, even in the nonparaxial limit. For the hyperbolic metamaterial being investigated here, the focus is upon the extraordinary wave, i.e. it is assumed that the electric field oscillates in the plane yz. In addition, $\boldsymbol{P}_{\mathrm{NL}} \cdot \hat{\boldsymbol{x}} = 0$, since the nonlinearity is assumed to not couple extraordinary and ordinary polarizations. If the spatial derivatives of ϵ_{ij} are neglected, Eqs. (10.26) and (10.27), using a Cartesian reference system, yield[42]

$$E_z = -\frac{i}{\omega\epsilon_{zz}}\frac{\partial H_x}{\partial y} - \frac{\epsilon_{zy}}{\epsilon_{zz}}E_y - \frac{1}{\epsilon_{zz}}P_{\mathrm{NL},z}, \tag{10.28}$$

$$\begin{aligned}\frac{\partial H_x}{\partial z} + \frac{\epsilon_{yz}}{\epsilon_{zz}}\frac{\partial H_x}{\partial y} = &-i\omega\left(\epsilon_{yy} - \frac{\epsilon_{yz}\epsilon_{zy}}{\epsilon_{zz}}\right)E_y \\ &-i\omega\left(P_{\mathrm{NL},y} + \frac{\epsilon_{yz}}{\epsilon_{zz}}P_{\mathrm{NL},z}\right),\end{aligned} \tag{10.29}$$

$$\frac{\partial E_y}{\partial z} + \frac{\epsilon_{zy}}{\epsilon_{zz}}\frac{\partial E_y}{\partial y} = -i\omega\mu_0 H_x - \frac{i}{\omega\epsilon_{zz}}\frac{\partial^2 H_x}{\partial y^2} - \frac{1}{\epsilon_{zz}}\frac{\partial P_{\mathrm{NL},z}}{\partial y}. \tag{10.30}$$

This system of equations governs light propagation in the paraxial and non-paraxial regimes, regardless of the nature of the nonlinear response, which, in principle, could include second- and third-order nonlinearities, and local and non-local responses.

10.3.1.1. *The walk-off angle*

In the absence of losses, or gains, the dielectric tensor $\boldsymbol{\epsilon}$ is Hermitian (i.e. $\epsilon_{ij} = \epsilon_{ij}^*$) and energy is conserved, as revealed by the Poynting theorem in the harmonic regime. Without losses, all the elements of $\boldsymbol{\epsilon}$ are real, yielding $\epsilon_{yz} = \epsilon_{zy}$. Also, in a uniaxial medium, the Poynting vector $\boldsymbol{S}$ of an electromagnetic wave packet, with a spatial spectrum centerd around $k_y = 0$, propagates at a walk-off angle $\delta = \arctan(\frac{\epsilon_{yz}}{\epsilon_{zz}})$ with respect to the axis z. This is sketched in Fig. 10.8. Another way of putting this is to note that, while the wave-fronts are normal to the axis z, the electromagnetic energy propagates at an angle δ with respect to it. The angle δ depends on the orientation of the average wavevector with respect to the optic axis, i.e. in this case on the angle θ. Finally, the walk-off angle does not depend on the spectrum of the input beam; thus, the approach remains valid for wave packets as narrow as a few wavelengths.[43]

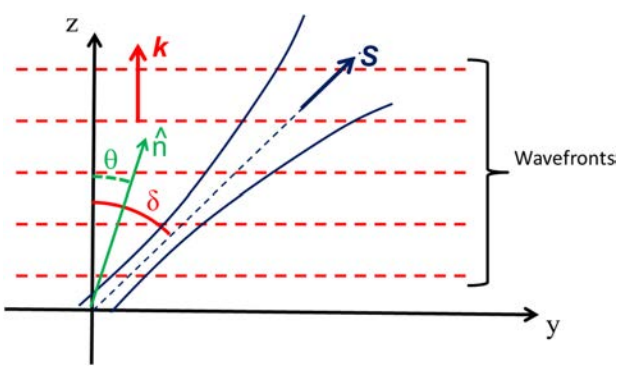


Fig. 10.8. Pictorial sketch of the beam walk-off. The Poynting vector $\boldsymbol{S}$, i.e. the transmission direction of the electromagnetic energy, forms an angle δ with respect to the wavevector $\boldsymbol{k}$, the latter depending on the angle θ between the optic axis and the wavevector.

An important point is that Eqs. (10.28)–(10.30) apply no matter what the values of ϵ_{ij} are: this implies that the tensor elements can be real, positive or negative (i.e. the case of hyperbolic dispersion), or even complex (when losses are present).

In order to consider the creation of bright spatial solitons, Eqs. (10.28)–(10.30) can be rewritten as a single Helmholtz equation. Equations (10.28) and (10.29) can be combined into a single equation using the coordinate transformation[43]

$$\begin{aligned} x' &= x \\ y' &= y - (\tan\delta)z \\ z' &= z. \end{aligned}$$

This new coordinate system, $x'y'z'$, is the frame of reference in which the ray is at rest. The equations are valid if the walk-off angle does not undergo substantial changes across the wave front (i.e. in planes $z =$ constant) and its longitudinal (i.e. along z) variations are adiabatic.

Light propagation then obeys the single scalar Helmholtz equation[42]

$$\begin{aligned} &\frac{\partial^2 H_x}{\partial z'^2} + D_y \frac{\partial^2 H_x}{\partial y'^2} + k_0^2 n_e^2 H_x \\ &\quad = -i\omega \frac{\partial (P_{\mathrm{NL},y} + \tan\delta P_{\mathrm{NL},z})}{\partial z'} + \frac{i\omega n_e^2 \epsilon_0}{\epsilon_{zz}} \frac{\partial P_{\mathrm{NL},y}}{\partial y'}, \quad (10.31) \end{aligned}$$

where the local, extraordinary, refractive index $\epsilon_0 n_e^2 = \epsilon_{yy} - \frac{\epsilon_{yz}\epsilon_{zy}}{\epsilon_{zz}}$ is introduced together with the diffraction coefficient $D_y = \frac{\epsilon_0 n_e^2}{\epsilon_{zz}}$. The expression for n_e shows that the physics of anisotropic media is richer than in the isotropic case: a photonic potential can arise from changes on either of the two eigenvalues $\epsilon_\perp$ and $\epsilon_\parallel$, as well as from space-dependent rotations of the optic axis (when θ is a function of position). The diffraction coefficient

$$D_y = -k_0 n_0 \left. \frac{\partial^2 k_z}{\partial k_y^2} \right|_{k_y=0}$$

accounts for the local concavity of the spatial dispersion $k_z(k_y)$. In isotropic media, the dispersion curve is a circle (in the 2D limit),

yielding $D_y = 1$. In the anisotropic case, the dispersion curves are ellipses due to the direction-dependent refractive index, thus the diffraction coefficient depends on the direction of the optic axis, i.e. on θ.

10.3.1.2. *Hyperbolic dispersion*

To address the main features of bright spatial solitons in the presence of hyperbolic dispersion, it is essential that the role of D_y in the light evolution is examined. To achieve this outcome, consider Eq. (10.31), for paraxial waves, ($\partial^2_{z'} H_x \approx 0$), in the linear regime ($\boldsymbol{P}_{\rm NL} = 0$). For the slowly varying envelope of the magnetic field A,

$$i\frac{\partial A}{\partial z'} = -\frac{D_y}{2k_0 n_0}\frac{\partial^2 A}{\partial y'^2} - \frac{k_0}{2n_0}\left(n_e^2 - n_0^2\right)A, \tag{10.32}$$

where n_0 is the average extraordinary refractive index and A is defined through $H_x = A(y', z')e^{ik_0 n_0 z'}$.

Equation (10.32) is a Schrödinger-like equation, with z' playing the role of an effective time and where the substitution $\hbar \to 1$ must be performed. The effective mass is then given by $k_0 n_0/D_y$, and the effective photonic potential is $-k_0(n_e^2 - n_0^2)/(2n_0)$. Thus, the diffraction coefficient D_y is inversely proportional to the equivalent mass of the light beam.

Striking differences in electromagnetic propagation arise from the signs of the two dielectric eigenvalues, $\epsilon_\perp$ and $\epsilon_\parallel$. When both eigenvalues are positive, the standard anisotropic medium features an elliptic dispersion. A topological transition in spatial dispersion $k_z(k_y)$ occurs when $\epsilon_\perp \epsilon_\parallel < 0$: the dispersion curves become hyperbolae[44] and the corresponding metamaterials are known as hyperbolic, or indefinite. The different signs of the eigenvalues imply dielectric, or metallic, responses according to the polarization of the impinging field. A hyperbolic dispersion can be achieved in metamaterials consisting of planar multilayers of dielectrics and metals, or in the presence of metallic nanowires embedded in a dielectric.[45]

Another property is how D_y depends on the geometric features of the dispersion curves. For elliptic dispersion, $D_y > 0$, and light beams propagate as massive particles subject to a potential given by the

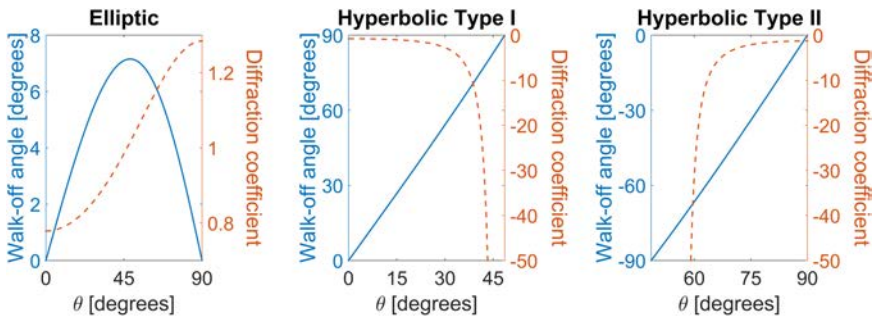


Fig. 10.9. Typical behavior of beam walk-off δ (left axis, blue solid lines) and diffraction coefficient D_y (right axis, dashed red lines) versus the angle θ between the wavevector and the optic axis, for elliptic and hyperbolic materials. The permittivity eigenvalues are $|\epsilon_\perp|/\epsilon_0 = 2.25$ and $|\epsilon_\parallel|/\epsilon_0 = 2.89$. In the hyperbolic case, the plotted range of θ corresponds to the existence of non-evanescent plane waves.

sign-inverted change in refractive index. For hyperbolic dispersion, the diffraction coefficient is always negative due to the opposite curvature of the dispersion curves $k_z(k_y)$ with respect to the elliptic case. In fact, using the relationship $\epsilon_0 n_e^2 = \epsilon_\perp \epsilon_\parallel / \epsilon_{zz}$, the diffraction coefficient can be rewritten as

$$D_y = \frac{\epsilon_\perp \epsilon_\parallel}{\epsilon_{zz}^2}. \tag{10.33}$$

The dependence of the diffraction coefficient D_y and the walk-off angle δ on the optic axis position, θ, is shown in Fig. 10.9. For elliptic dispersion, the walk-off angle reaches a maximum value around $\theta = 45°$, and the sign is that of the anisotropy ϵ_a. The diffraction coefficient is always positive. In the hyperbolic case, two qualitatively different behaviors arise, according to whether $\epsilon_\parallel < 0$ (Type I) or $\epsilon_\parallel > 0$ (Type II). For Type I dispersion, the walk-off angle is always positive, because anisotropy ϵ_a is negative. Furthermore, δ increases monotonically with θ and reaches 90° at the edge of the existence region. The diffraction coefficient, D_y, starts from a value close to -1 for $\theta = 0$, and then decreases, continuously, with θ, finally reaching a singularity at the border of the existence range. Type II material has a similar behavior, but now the walk-off angle is always negative.

The similar behavior for Type I and Type II stems from the fact that they are related by a 90° rotation in the 2D case being treated here.

A negative diffraction coefficient heavily affects light propagation: beams now behave like particles with negative masses. This means that light gets attracted towards regions where the refractive index is lower, in sharp contrast with the standard Fermat's principle. Thus, graded-index (GRIN) waveguides in hyperbolic media require a core with lower refractive index.

Starting from the results in the linear case, it is straightforward to elucidate the main properties of bright spatial solitons in hyperbolic media. To address light self-localization in hyperbolic materials, is assume that a local Kerr-like nonlinear response, leading to a refractive index variation, dependent on the beam intensity, i.e.

$$n_e^2 - n_0^2 = n_2|A|^2. \tag{10.34}$$

The detailed analysis of other nonlinearities can be performed starting from Eq. (10.31). Substituting Eq. (10.34) into Eq. (10.32) results in a standard nonlinear Schrödinger equation. A simple analogy with temporal solitons in fibers can be carried out.[46,47] Due to the negative diffraction (dispersion in time), the formation of bright spatial solitons requires $n_2 < 0$, whereas for $n_2 > 0$ dark solitons can exist. This is in agreement with the theoretical results reported in the literature and based on coupled mode theory,[48,49] or in the analogy with gravitational systems.[50] Figure 10.10 shows the evolution of a Gaussian beam in the presence of a focusing/defocusing nonlinearity in a Type I material for $\theta = 0$, computed using a beam propagation code for $\theta = 0$, i.e. in the absence of walk-off. When $n_2 < 0$, the beam undergoes self-focusing and a hyperbolic secant bright soliton emerges after radiation sheds away (Fig. 10.10(a)). For $n_2 > 0$, a Gaussian beam undergoes enhanced diffraction with respect to the linear case (Fig. 10.10(b)). Figure 10.10(c) finally shows the beam cross-sections for a given propagation distance versus the effective excitation n_2P: the transition from a self-trapped state to a beam wider than in the linear case is clearly appreciable.

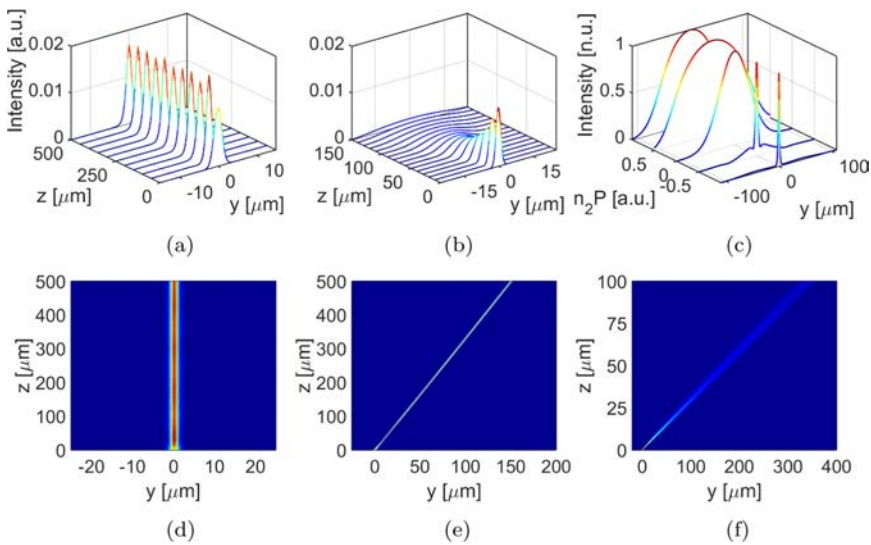


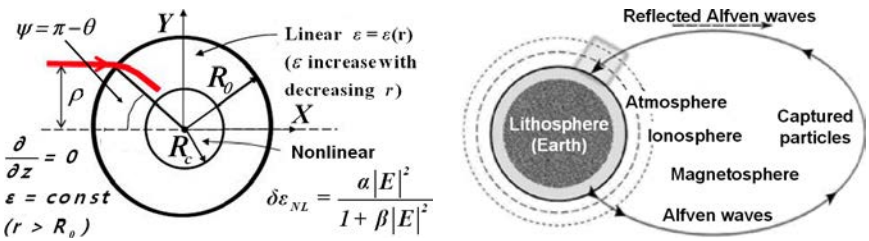
Fig. 10.10. Light self-trapping in hyperbolic metamaterials. Evolution of a Gaussian beam of waist 2 μm in the presence of local Kerr (a) focusing and (b) defocusing nonlinearity for the same input power. (c) Normalized beam cross-section versus y at $z = 500$ μm versus the normalized excitation n_2P. Intensity profile on the plane yz for a fixed input beam when for (d) $\theta = 0°$, (e) $\theta = 10°$, and (f) $\theta = 40°$. Input beam is launched normal to the sample and the observed deflection is exclusively due to the linear walk-off. Here, $\lambda = 1\,\mu$m, $\epsilon_\perp = 2.26$ and $\epsilon_\parallel = -2.87$.

Figures 10.10(d)–10.10(f) show the soliton dependence on the optic axis angle, θ, for fixed n_2P. First, the solitonic beam follows the linear walk-off plotted in Fig. 10.9, thus it is possible to steer the self-confined beam by rotating the optic axis.[51] Second, despite the fact that n_2P is fixed, the soliton width changes with θ due to the variation of the diffraction coefficient, because the amount of self-focusing necessary to form a soliton depends on the diffraction spreading.

10.4. Nonlinear Plasmonic Aspects of Metamaterials

10.4.1. *Metamaterial field concentrator*

The idea of a nonlinear "transformational optics-inspired" electromagnetic cylindrical energy concentrator can be appreciated from



(a) (b)

Fig. 10.11. (a) The geometry of isotropic graded-index nonlinear field superconcentrator. One incident beam is shown. Angle coordinate in the cylindrical system is $\theta = \pi - \Psi_0$. The mode (H_r, H_θ, E_z) is considered for isotropic concentrator. (b) The system "Lithosphere–Atmosphere–Ionosphere–Magnetosphere", where the method of modelling, similar to that used for the field concentrators, is applicable. The conductivity of the lithosphere is supposed, for the simplification, to be the homogeneous one and the same as everywhere in the Earth.

Refs. 37, 52–57. Such a concentrator includes both an external linear region with a dielectric constant increasing towards the center and an internal region with nonlinearity characterized by constant coefficients. Figure 10.11(a) shows the cross-section of such a cylindrical concentrator. The coordinate system is defined as (r, θ, z) and the cylinder axis coincides with the z-axis. The linear external region $(r > R_0)$ is occupied by a homogenous medium, with a relative permittivity of ε_0, and the inner core $(r \leq R_c)$ is nonlinear, homogeneous and non-magnetic.

The nonlinearity of the core capture region is expressed through a change to its linear dielectric behavior through the addition of a saturable nonlinear permittivity $\delta\varepsilon_{\mathrm{NL}}$. The outer cylindrical shell, $(R_c \leq r \leq R_0)$ is both inhomogeneous and linear. The inhomogeneous permittivity increases with decreasing the radius, thus ensuring light capture through a graded-index technique.[37] In a semiconductor, such a profile of the permittivity is due to non-uniform doping and, thus, a non-uniform steady electron concentration.

The cylindrical shell structure can capture electromagnetic energy approaching it externally, through a form of graded-index in the outer cylindrical shell. Ideally, an isotropic form of permittivity

can be used that will capture incoming electromagnetic energy.[37] This energy is deposited into the core of the concentrator. The creation of such a concentrator, which will capture energy, will not engage in reflections that will diminish its effectiveness. This scenario can be created using a straightforward form of radially dependent permittivity distribution,[37] to which now will be added the possibility that the core can become strongly nonlinear.

The form of the relative permittivity distribution to be adopted here, including absorption in the outer cylindrical shell and a nonlinear core, can be written very simply as

$$\varepsilon(r) = \begin{cases} \varepsilon_0, r > R_0, \\ \varepsilon_0(R_0^2/r^2) + \dfrac{i\gamma_{\text{loss}}}{(R_0^2/R_c^2) - 1}((R_0^2/r^2) - 1), \\ R_c < r < R_0 \\ \varepsilon_c + \delta\varepsilon_{\text{NL}} + i(\gamma_{\text{loss}} - \gamma_{\text{gain}}), r < R_c. \end{cases} \tag{10.35}$$

In (10.35), γ_{loss} and γ_{gain} are (linear) absorption and gain parameters, respectively. An absorption parameter γ_{loss} has been included for the core and also the external cylindrical shell. Note that in the latter, the damping changes with radius r and goes to zero at $r = R_0$ and goes to γ_{loss} when r approaches R_c. The elementary form $\varepsilon_c = \varepsilon_0(R_0^2/R_c^2)$ ensures that, as $r \to R_c$, there is complete continuity across the interface in the absence of amplification ($\gamma_{\text{gain}} = 0$).

The nonlinear addition to the relative permittivity of the core is set to the saturable Kerr form

$$\delta\varepsilon_{\text{NL}} = \alpha|E_z|^2/(1+\beta|E_z|^2+\gamma|E_z|^4)+i(\alpha_{s2}|E_z|^2+\alpha_{s4}|E_z|^4), \tag{10.36}$$

in which α is complex and β, γ, α_{s2} and S_4 real constants. α describes complex saturable nonlinearity, coefficients β, γ determine nonlinear saturation (including saturation of fourth order with a coefficient γ) and coefficients α_{s2} and α_{s4} determine additional nonsaturable nonlinear losses, again, including the fourth-order nonlinear term ($\sim \alpha_{s4}$). The contact with a previously used simple form[37] of the permittivity is achieved by setting $\gamma = 0$, and setting $\delta\varepsilon_{\text{NL}} = 0$, in order to stay in the linear domain. It is easy to see that, in a linear

environment, each dielectric permittivity has been set to merge into the other as the appropriate radius is reached.

Again, the transverse electric mode (H_r, H_θ, E_Z) is considered, and the corresponding full-wave equation is

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial E_z}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 E_z}{\partial \theta^2} + k_0^2 \varepsilon(r) E_z = 0. \tag{10.37}$$

In order to simulate this type of structure, the complex geometrical optics, within the inhomogeneous region $R_c \leq r \leq R_0$, is developed to get the precise matching on the boundary $r = R_{\mathrm{c}}$. Inclusion of the saturation, characterized by β in nonlinear electrodynamic region, is very important, because any superfocusing will lead to such strong field concentrations that the core will move rapidly out of the weakly nonlinear regime and beyond the scope of the recently reported[57] nonlinear transformation optics. Generally, several beams are incident upon the concentrator simultaneously.

Within the nonlinear core, the nonlinear equation has been solved directly by means of the expansion of E_z and the permittivity by the Fourier series on the angle θ. The beams incident to the core figure in the boundary conditions at $r = R_c$.

A new nonlinear effect called "switching nonlinear focusing" is found. Accurately speaking, the tendency to "hot spot" formation (small region, of order of wavelength with very strong field concentration) is found, and "switching" to such a new regime is achieved, when amplitude(s) of input beam(s) exceeds some threshold value. For moderate values of "impact distance" $\rho(\rho \leq 0.75\ R_0)$ and incident beams propagating parallel to the axes X (Fig. 10.11), a "hot spot" is placed near the coordinate

$$(x \approx -R_c, y \approx 0). \tag{10.38}$$

It is interesting that the point, where strong nonlinear focusing is possible, corresponds to the position of focusing of *linear* beams (in the absence of any core nonlinearity) in the CGO approximation, which was demonstrated directly by means of numerical modelling for the chosen distribution of linear part of electric permittivity (see

Eq. (10.35)). Note first that a tendency to such a regime is rather universal. Namely, it will be demonstrated now that formation of hot spots may start as a result of any of the following processes: increase of input amplitude (A_{inp}), increase of linear amplification (γ_{gain}) (or decrease of linear losses γ_{loss}) in the central nonlinear region, decrease in nonlinear losses ($\text{Im}(\alpha), \alpha_{s2}$ or α_{s4}) or decrease in saturation of nonlinearity (β or γ) (see Eq. (10.36)). The "start of hot spot formation" or "tendency to hot spot formation" reveals itself in increase of field intensity in the vicinity of the point of "possible hot spot" with coordinates determined by Eq. (10.37).

Figures 10.12(a)–10.12(c) illustrate a tendency to "hot spot" formation when an input amplitude, defined here as a normalized quantity A_{inp} tends to the threshold value, which is a bit larger than 0.4445 as shown in Fig. 10.12(c). In the absence of nonlinearity ($\delta\varepsilon_{\text{NL}} = 0$), some diffraction and interference maxima and minima are produced as shown in Fig. 10.12(a). In the presence of focusing nonlinearity, however, with $\delta\varepsilon_{\text{NL}} \neq 0$, the input amplitude plays a critical role in what the final outcome will be. Below a threshold value, of the order of 0.4445, nonlinear focusing does take place within the nonlinear region, but with a larger amplitude, and smaller peak area (Figs. 10.12(b) and 10.12(c)) than could be achieved for an entirely linear system. When the input amplitude exceeds a threshold value, the peak position "jumps" from some point inside the nonlinear cylinder (Figs. 10.12(b) and 10.12(c)) to a point near to the interface between the linear and nonlinear cylinders, and Fig. 10.12(c) shows "the start of preparation" of such a "jump". Figure 10.12(d) illustrates (only) a tendency to the "hot spot" formation, when an incident beam amplitude slightly exceeds the threshold value. The "hot spot" can be considered as a self-consistent strongly nonlinear surface resonator on the boundary between the linear and nonlinear region (in the vicinity of the point) $x = -R_c, y = 0$.

Thus, in this nonlinear system the nonlinear focusing takes place at the input amplitudes below the threshold, or, in other words, the formation of spatial solitons occurs. Also, the hot spot formation may occur above the threshold value in the near-surface region of the internal cylinder. After exceeding by an input amplitude a

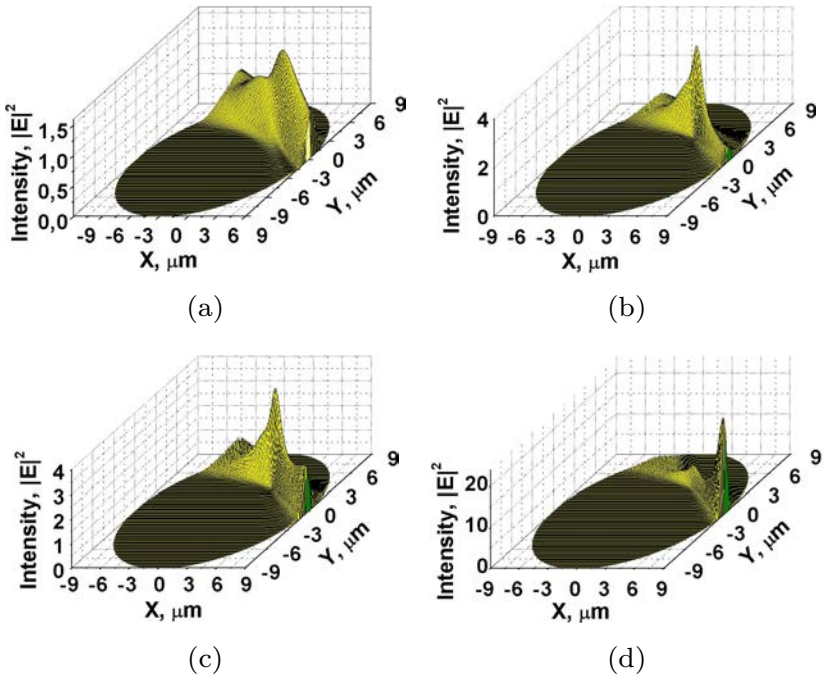


Fig. 10.12. Tendency to hot spot formation, while input amplitude of incident beam(s) to a tends threshold value. Number of incident beams $N_b = 2$; (a) field distribution in the core for purely linear system ($a = 0, \alpha_{s2} = 0, \alpha_{s4} = 0$) for normalized input amplitude (of each of the incident beams) $A_{\text{inp}} = 0.444$; (b) field distribution in the nonlinear core for the same parameters, as in the case (a), while incident amplitude is below the threshold value; (c) incident amplitude is very close to but still below a threshold value, $A_{\text{inp}} = 0.4445$; (d) tendency to the formation of "hot spot" in the vicinity of the point ($x = -R_c, y = 0$) when incident normalized amplitude is slightly above a threshold value, $A_{\text{inp}} = 0.452$. Numerical convergence in this range of amplitude is problematic due to non-regular dependence of "output" on "input" amplitude; here "input" and "output" amplitudes mean an amplitude of an incident beam(s) and an amplitude in the "hot spot".

threshold value (Fig. 10.11(d)) the nonlinearity is strong and cannot be considered as moderate. Moreover, it should be considered in a non-local manner. Therefore, in this range of amplitudes, we can say only about an illustration of a tendency to "hot spot" formation (see also a caption to Fig. 10.11(d). Let us emphasize that the numerical convergence in the "above-threshold" region of parameters, to which Fig. 10.11(d) does correspond, is problematic, because in such a

region a non-regular dependence of "output" on "input" amplitude is observed. To clarify the question of "above-threshold behavior", now a new, non-phenomenological model of an active medium accounting for non-locality and non-small nonlinearity is under development.

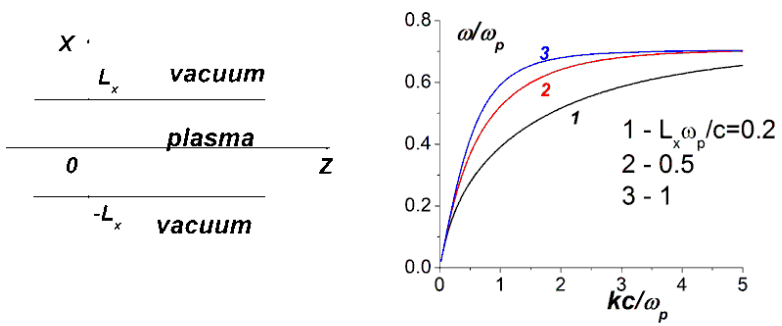
10.4.2. *Nonlinear plasma waves in near-zero metamaterial structures*

10.4.2.1. *Basic equations*

Consider the excitation of electron plasma waves in a plasma slab, see Fig. 10.13(a). The slab is bounded by thin smooth dielectrics at $z = \pm L_z$ with a low dielectric permittivity. Outside there is a vacuum, so the dielectric permittivity in the remaining space is $\varepsilon = 1$. The thickness of the plasma slab is $2L_x$. The electron plasma waves are investigated with the components $E_z, H_y \equiv H, E_x, v_x, v_z$, which are symmetric with respect to E_x and v_x components. The ions are considered as immobile.

The hydrodynamic equations for the volume concentration n and the velocity $\boldsymbol{v}$ of the electron gas are[58,59]:

$$\frac{\partial \vec{v}}{\partial t} + (\vec{v}\cdot\vec{\nabla})\vec{v} = \frac{e}{m}\left(1-\frac{v^2}{c^2}\right)^{1/2}\left(\vec{E}+\frac{1}{c}\vec{v}\times\vec{H}-\frac{\vec{v}}{c^2}(\vec{v}\cdot\vec{E})\right) - \nu\vec{v} - \left(1-\frac{v^2}{c^2}\right)^{1/2}\frac{v_T^2}{n}\vec{\nabla}n; \quad \frac{\partial n}{\partial t} + \vec{\nabla}\cdot(n\vec{v}) = 0. \tag{10.39}$$



(a) (b)

Fig. 10.13. (a) is the geometry of problem. (b) is the wave dispersion for linear electron plasma waves in the plasma layer.

The moving charges are assumed as positive, $e > 0$. The weakly relativistic case is considered: $v/c \leq 0.2$, where c is the velocity of light in vacuum. Here, $v_T = (k_B T/m)^{1/2}$ is the electron thermal velocity. In metallic plasmas, it is replaced by the Fermi velocity $c_F \sim 10^8$ cm/s.

Maxwell equations for this case are:

$$\vec{\nabla} \times \vec{H} = \frac{1}{c}\frac{\partial \vec{E}}{\partial t} + \frac{4\pi}{c} en\vec{v};$$
$$\vec{\nabla} \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{H}}{\partial t}; \quad n\vec{v} = 0 \quad \text{for } |z| > L_z. \tag{10.40}$$

At the boundary $z = L_z$, the conditions are the continuity of EM field components, also the normal component of electron velocity is $v_z = 0$.

The dispersion equation for the symmetric $v_x(-z) = v_x(z)$ linear electron plasma waves in the plasma layer is (the cold plasma, $T = 0$):

$$\tan h(rL_x) + \frac{r}{p\varepsilon(\omega)} = 0, \quad \text{where } r = \left(k^2 - \frac{\omega^2}{c^2}\varepsilon(\omega)\right)^{1/2},$$
$$p = \left(k^2 - \frac{\omega^2}{c^2}\right)^{1/2}, \quad \varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2}. \tag{10.41}$$

For a linear, cold plasma wave when $T = 0$, the volumetric oscillations of the electron concentration are absent: $n = n_0$, whereas the oscillations of the surface charge are present. The dispersion curves for different parameters are presented in Fig. 10.13(b).

In the following development, the simulations of the resonant second-harmonic in the bounded electron plasma are presented and the surface nonlinearity is dominant and is assumed to be moderate.

10.4.2.2. *Giant generation of second-harmonics in permittivity-near-zero plasmonic metamaterial*

The nonlinear phenomenon that provides localized generation of the second-harmonic with a high amplitude is very useful for sensor

applications. The general method for the demonstrating the existence of nonlinear waves in the layered structures using a plasma, and accounting for both volume and surface nonlinearities, is presented. Once this is achieved, estimations based around gaseous and semiconductor plasmas and metals are presented, with a discussion of possible plasmonic metamaterial structures using arrays of nanoparticles.

Consider the nonlinear motion in an electron plasma in the simplest layered structure supporting surface plasmons, namely, the structure plasma-linear dielectric, see Fig. 10.13(a). Suppose that the plasma occupies the half-space $x \leq 0$ and that a (linear) dielectric, with permittivity constant ε occupies half-space $x > 0$, while a wave packet propagates along the x-axis. If $\varepsilon = 1$ (dielectric is vacuum), the linear dispersion equation for surface plasmons is straightforward to obtain when the width of the plasma layer is large.

If the field components vary as $\sim \exp[i(\omega t - kz)], \omega$ and k are frequency and wavenumber, respectively. In the nonlinear problem they would play the roles of carrier frequency and wavenumber of a narrow wave packet. ω_p and c are plasma frequency and speed of the light, respectively. The evolution of the coupled main and second-harmonics of surface plasmons, uses the methods reported earlier.[31–35] A weak nonlinearity and a slowly varying amplitude will now be assumed together with the introduction of the variation of the plasma concentration, velocity and electric and magnetic fields in the example form $\tilde{n}_2^*, \vec{v}_2^*, \vec{E}_2^*, \vec{H}_2^*$. If the suffix 1 is used, such values would correspond to surface nonlinear plasmons and are proportional to the slowly varying nonlinear amplitude. The index "2" describes corresponding linear waves in the same system, but without any nonlinearity. The upper index "*" simply means complex conjugation. Corresponding equations of motion in nonlinear plasma have the form

$$\frac{\partial \tilde{n}_1}{\partial t} + \mathrm{div}(n\vec{v}_1) = -\mathrm{div}(\tilde{n}\tilde{\vec{n}}), \quad \frac{\partial \vec{v}_1}{\partial t} + \frac{e}{m}\vec{E}_1 = -(\vec{v}_1\vec{\nabla})\vec{v}_1 - \frac{e}{m}[\vec{v}_{1x}\vec{H}_1]$$

$$\mathrm{curl}l\vec{H}_1 = \frac{1}{c}\frac{\partial E_1}{\partial t} + \frac{4\pi}{c}(-en_0\vec{v}_1 - e\tilde{n}\vec{v}),$$

$$\mathrm{curl}l\vec{E}_1 = -\frac{1}{c}\frac{\partial \vec{H}_1}{\partial t}. \tag{10.42}$$

Similar equations (but without nonlinear terms) can be written for the linear components with indices "2". In Eqs. (10.42), we put electron concentration n to be equal to $n = n_0 + \tilde{n}$, where n_0 and $\tilde{n}$ are stationary and variable parts of the electron concentration, and e and m are electron charge and mass, respectively. The corresponding system of complex conjugated equations for linear plasma, and the procedure similar to that used for the derivation of energy conservation law.[60] We account for boundary conditions at the interface ($z = 0$) "plasma-dielectric", the presence of free carriers, and, respectively, surface charge and surface current j_{surf}, namely

$$H_{1y}|_{x=+0} - H_{1y}|_{x=-0} = -\frac{4\pi}{c} j_{\mathrm{zsurf}}, j_{\mathrm{zsurf}} = -e\delta\vec{v}_z|_{x=+0} \tag{10.43a}$$

Here, δ is the surface charge concentration. Using (10.43a), we get

$$[\vec{E}_1^* \times \vec{H}_2]_{x:x=+0} - [\vec{E}_2^* \times \vec{H}_1]_{x:x=-0} = \frac{4\pi}{c} E_{2z}^* e\delta v_z|_{x=+0}. \tag{10.43b}$$

In Eq. (10.43b), a notation like $[\ldots]_{x:x\pm0}$ means x-component of the vector $[\ldots]$, taken at $x = \pm 0$.

Finally, the integral nonlinear relation is obtained in the form

$$\begin{aligned}
&\frac{\partial}{\partial t}\int_{-\infty}^{x}\left\{\frac{1}{4\pi}(\vec{E}_1\vec{E}_2^* + \vec{H}_1\vec{E}_2^*) + mn_0\vec{v}_1\vec{v}_2^*\right\}dx \\
&\quad + \frac{c}{4\pi}\frac{\partial}{\partial z}\int_{-\infty}^{\infty}\{[\vec{E}_1 \times \vec{H}_2^*]_z + [\vec{E}_2^* \times \vec{H}_1]_z\}dx - eE_{2z}^*\sigma v_x|_{x=+0} \\
&= -m\int_0^\infty n_0\vec{v}_2^*(\vec{v}\vec{\nabla})\vec{v}dx - \frac{e}{c}\int_0^\infty n_0\vec{v}_2^*[\vec{v} \times \vec{H}]dx \\
&\quad + e\int_0^\infty E_2^*(\tilde{n}\vec{v})dx - \frac{e}{c}\int_0^\infty n_0\vec{v}_2^*[\vec{v} \times \vec{H}]dx + e\int_0^\infty E_2^*(\tilde{n}\vec{v})dx.
\end{aligned} \tag{10.44}$$

Note that plasma concentration n_0 in the first term in Eq. (10.44) is non-zero only in the region $x \leq 0$, and, in the expressions like $E_2^*(\tilde{n}\vec{v})$, etc., in $(\tilde{n}\vec{v})$, the terms $\sim \exp(i\omega t)$ are revealed. The terms in the integral relation (3.4) have a clear physical sense. Accounting for the fact that nonlinear terms with index "1" are proportional to slowly varying amplitude A_i, it is possible to show that the first and second terms in the left-hand part of Eq. (10.44) are proportional to the time and spatial derivatives of the slowly varying amplitude, respectively, while the coefficients of proportionality are integral densities of linear wave energy and energy flow, respectively. Moreover, more detailed analysis shows that the total linear operator, describing the evolution of the slowly varying amplitude in the parabolic approximation,[61] may be revealed from the left-hand part of the Eq. (10.44), including linear dispersion and diffraction. The third term in the left-hand part of Eq. (10.44) describes the surface nonlinearity, while the right-hand part of this equation describes the volume nonlinearity. The first, second and third terms in the right-hand part of Eq. (10.44), correspond to the substantial, Lorentz and concentration volume nonlinearities, respectively. Therefore nonlinear evolution equation for slowly varying amplitude can be obtained from Eq. (10.44). To do this, the presence of zero and second-harmonics in the nonlinearities contained by Eq. (10.44) should be accounted for, and these higher harmonics should be expressed in the terms of the slowly varying amplitude A_1 of the main harmonic. Then, the terms $\sim\exp(i\omega t)$ are revealed from the all combinations of values, which include nonlinear fields in Eq. (10.44), while (complex conjugated) linear fields (with index "2") are proportional to $\sim\exp(-i\omega t)$. Finally, the "fast phase" is totally removed from Eq. (10.44) and the evolution equation for slowly varying amplitude of the main harmonic is obtained. To find zero and second-harmonics, they are considered as "effective forced oscillators" with "external force" determined in the terms of the amplitude A_1 of main harmonic. Proper boundary conditions and surface nonlinearities (due to the motion of nonlinear surface charge) are accounted for. All zero harmonic components are found in this way. Second-harmonic components $\vec{H}_2, \vec{E}_2 \sim \exp[2i(\omega t - kz)]$

are found from the set of Maxwell and plasma equations (similar to the system of equations (10.42)). For example, the second-harmonic of x-component of electric field in plasma region ($z \leq 0$) is obtained in the form

$$E_z^{(2)} = E_0 e^{q(2\omega)x} - iek\omega_p^2 \varepsilon(\omega)[8me^2q^2\omega^2\varepsilon(2\omega)]^{-1} A_1^2 e^{2qz}, \quad z \geq 0, \tag{10.45a}$$

where $q(2\omega) = (k_2^2 - k_{02}^2\varepsilon(\omega_2))^{1/2} k_2 = 2k, k_{02} = 2\omega/c, \omega_2 = 2\omega$

$$E_0 = \frac{iek\omega_p^2 A_1^2}{2m\omega^4 q} \frac{1 + \omega^2\varepsilon(\omega)[4c^2pq\varepsilon(2\omega)]^{-1}}{2\varepsilon(2\omega)q(2\omega)^{-1} + 1}. \tag{10.45b}$$

As seen from Eqs. (10.45a) and (10.45b), under the condition of "temporal resonance", corresponding to the regime of near-zero-epsilon metamaterial, namely

$$\varepsilon(2\omega) \approx 0, \quad \omega \approx \omega_p/2, \tag{10.46}$$

the amplitude of the second-harmonic can become very large. The term "temporal resonance" is connected with the fact that, at the same frequency, "spatial resonance" is absent because $k(2\omega) \neq 2k$. In the vicinity of temporal resonance frequency (Eq. (10.46)), the amplitude of the second-harmonic A_2, should be described by the separate evolution equation. In the near-resonant region (see relation (3.46)), equations for the second-harmonic (see Eqs. (10.45a) and (10.45b) reduce to the form

$$E_z^{(2)} = (1/2)A_2 e^{'2(\omega t - kz)} f_2(x), \quad \varepsilon(2\omega)A_2 f_2(x) = RA_1^2 f_2(x)$$
$$R = iek\omega_p^2\varepsilon(\omega)[8mc^2q^2\omega^2]^{-1}, \quad f_2(x) = e^{q(2\omega)x} - e^{2qx}.$$

Then in the vicinity of the resonant frequency, $\omega_2 = 2\omega \approx \omega_p$, we could write $\varepsilon(2\omega) \equiv \varepsilon(\omega_2) = \varepsilon(\omega_2 - \omega_{20} + \omega_{20} - \omega_{pc} + \omega_{pc}) = \varepsilon(\omega_{pc}) + (d\varepsilon/d\omega)(\Delta\omega_0 + \Delta\omega), \varepsilon(\omega_{pc}) = 0$, where the values $\Delta\omega_0 = \omega_{20} - \omega_{pc}, \Delta\omega = \omega_2 - \omega_{20}$ are small. Using proper Fourier transform and making the replacement $i\Delta\omega \to \partial/\partial t$, it is possible to obtain, for the amplitude of the second harmonic, the equation

$$\partial A_2/\partial t + i\Delta\omega A_2 = (\partial\varepsilon/\partial\omega)^{(-1)} RA_1^2. \tag{10.47}$$

Using the relations (10.44) and (10.47), it is possible to obtain, in the near-resonant region, the following set of equations for the main and second (resonant) harmonics

$$\frac{\partial A_1}{\partial t} + v_g \frac{\partial A_1}{\partial z} + \frac{i\omega_{kx}}{2}\frac{\partial^2 A_1}{\partial z^2} + ig_2 \frac{\partial^2 A_1}{\partial y^2} + \alpha_1 A_1^* A_2 + i\alpha_0 |A_1|^2 A_1 + \gamma_1 A_1 = 0, \quad (10.48a)$$

$$\frac{\partial A_2}{\partial t} + (i\Delta\omega_0 + \gamma_2) A_2 = \alpha_2 A_1^2. \quad (10.48b)$$

Here, $v_g, \omega_{kk}, g_2, \Delta\omega_0$ and $\gamma_{1,2}$ are group velocity, coefficients of (linear) dispersion and diffraction, deviation of the frequency of the main harmonic from the resonant frequency and losses of the main and second harmonics, respectively. Note that both quadratic (with coefficients $\alpha_{1,2}$) and cubic (with coefficient α_0) nonlinearities are included into the system (10.48a) and (10.48b). The detailed forms of the nonlinear coefficients are not presented here, because they are rather cumbersome. Nonlinear coefficients in the equations for the main harmonic include contributions from "volume" and "surface" nonlinearities: $\alpha_0 = \alpha_{0\text{surf}} + \alpha_{0\text{volume}}$, $\alpha_1 = \alpha_{0\text{surf}} + \alpha_{0\text{volume}}$, while the main contribution into nonlinear coefficient α_2 in the equation for resonant second-harmonic comes from the volume nonlinearity. We would like to emphasize that the contribution from surface and volume nonlinearities have different signs

$$\alpha_{0\text{surf}} \cdot \alpha_{0\text{volume}} < 0, \quad \alpha_{1\text{surf}} \cdot \alpha_{1\text{volume}} < 0 \quad (10.49a)$$

and the main contributions, in fact, come from the surface nonlinearities, because

$$|\alpha_{0\text{surf}}|/|\alpha_{0\text{volume}}| \sim 3, \quad |\alpha_{1\text{surf}}|/|\alpha_{1\text{volume}}| \sim 7. \quad (10.49b)$$

The inclusion of the surface nonlinearity may change the sign of the "effective nonlinear coefficient", and a qualitative change of the behavior of a nonlinear system due to the surface nonlinearity can be expected. In particular, the estimations show that for $\omega/\omega_{pe} > 0.65$, the time derivative in Eq. (10.48a) can be neglected, and the system

of equations (10.48a) and (10.48b) reduces to standard nonlinear Schrödinger equation (NSE) with the effective nonlinear coefficient

$$\alpha_{eff} = \alpha - \alpha_1\alpha_2/\Delta\omega. \tag{10.50}$$

In this equation, $\omega_{\rm kk} < 0, \alpha_{eff} > 0$ and "formal" Lighthill criterion[61] (in other words, the criterion obtained using Eq. (10.50) for effective nonlinear coefficient) are not fulfilled,

$$\alpha_{eff}\,\omega_{\rm kk} < 0, \tag{10.51}$$

therefore solitons are impossible. In distinction to this case, if surface nonlinearity is absent, effective nonlinear coefficient α_{eff} changes sign, and solitons would be possible. Numerical calculations based on the system (10.48a) and (10.48b),[34] are illustrated in Fig. 10.14. The giant second-harmonic generation creates an amplitude that is greater than the first harmonic; compare Fig. 10.14(b) with Fig. 10.14(a). It is shown, in accordance with relation (10.49b), that

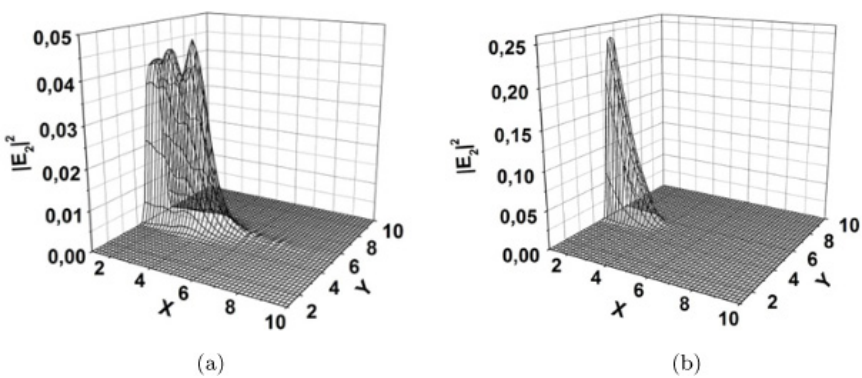


Fig. 10.14. The dependence of the normalized amplitude of the first (a) and second (b) harmonics of normalized coordinates for the normalized time value $T = 4.5$, which took place from the "switching on" the main harmonic (pump pulse) at input of the system. Both cubic and quadratic nonlinearities are present, and the sign of full nonlinear coefficient corresponds to the surface of nonlinearity. The pumping field (main harmonic) on the input of the system is a step function, which can be considered as approximately stationary after starting pumping, $E_1(x = 0) \approx \text{const} = 0.2$. Normalized (to the frequency to fundamental harmonic) frequency mismatch and the decrement of decay for the first and second harmonics are equal to each other $(-\Delta\omega) = \gamma_{1,2} = 0.01$.

the contribution of the surface nonlinearity in the nonlinear coefficients, which determine the generation of the giant second-harmonic resonance, is about three times greater than the contribution of the volume nonlinearity. During the calculations, the control of the parameters characterizing the smallness of the relative values of zero and second-higher harmonics of electron density, respectively, is provided.

Evaluations can now be made of typical parameters corresponding to possible experimental observations of nonlinear effects in different layered plasmonic structures.

(a) Gaseous plasma–dielectric structure. The analysis of the nonlinear coefficients in the Eqs. (10.48a) and (10.48b) shows that, by the order of values,

$$|\alpha_0| \sim \frac{\omega_{pe}}{8E_e^2}, \quad |\alpha_1| \sim \frac{\omega_{pe}}{60E_e^2}, \quad |\alpha_2| \sim \frac{\omega_{pe}}{5E_e^2}, \quad \text{where } E_c = \frac{mc\omega_{pe}}{e}.$$

Typical parameters for observation of nonlinear wave phenomena including giant generation of the second-harmonic in the structure "gaseous plasma-dielectric" are the following. For the plasma with electrons concentration $n_0 \sim 10^{11}\,\mathrm{cm}^{-3}$ and $\omega_{pe} \sim 1.6.10^{10}\,\mathrm{s}^{-1}$, the amplitude of the input pulse should be $A_1(x=0) \sim 0.1E_e \sim 100\,\mathrm{Gs}$, and the intensity of pumping wave $\sim (c/8\pi)A_1^2 \sim 1\,MW/\mathrm{cm}^2$ at the frequency $(\omega/2\pi) \sim 1.5\,\mathrm{GHz}$. The localization depth is, in this case, of order of 3 cm, and the duration of the pumping pulse of the first harmonic should be no less than 100 ns.

(b) Semiconductor–dielectric structure. For narrow-gap semiconductor (n-InSb), the essential nonlinearity is the Kane one.[61] Corresponding dispersion law is an analog of the relativistic law, namely $E^2(p) = m^*(0)^2v_n^2 + p^2v_n^2$, where E, p are energy in the conduction band and quasi-momentum, respectively, $m^*(0)$ is the effective mass at the bottom of the conduction band, $v_n = [E_g/2\mathrm{m}^*(0)]^{1/2} \sim 1.8\cdot10^8$ cm/s (for InSb) is the characteristic velocity, E_g is the forbidden gap ($\sim 0.2\,\mathrm{eV}$ for InSb), and "effective quasi-relativistic mass" is $\mathrm{m}^* = \mathrm{m}^*(0)(1-v/v_n^2)^{-1/2}$. For InSb with $n_0 \sim (10^{15}-10^{17})\,\mathrm{cm}^{-3}$ and

$\omega_{pe} \sim (10^{13} - 10^{14})\,\mathrm{s}^{-1}$, $\omega \sim \omega_{pe}/\varepsilon_L^{1/2}$ ($\varepsilon_L \sim 10$ is the lattice dielectric constant), the estimation for the first harmonic amplitude yields $\sim 0.1(m^* v_n \omega_{pe}/e) \sim 3.10^1$ Gs, the pulse duration should be $\sim$100 ps, and input intensity of the first harmonic should be $\sim$100 kW/cm^2.

(c) Metal–dielectric structure. For metals like Au, with $n_0 \sim 10^{23}\,\mathrm{cm}^{-3}$, $\omega_{pe} \sim 2.10^{16}\,\mathrm{cm}^{-3}$, characteristic nonlinear field is $E_e \approx 1/(\chi^{(3)})^{1/2}$, where nonlinear cubic constant is $\chi^{(3)} \sim 10^{-9}\,\mathrm{Gs}^{-2}$, and the estimation for the first harmonic yields $\sim 0.1E_c \sim 3.10^3$ Gs, the input pulse duration should be $\sim$100 fs, and the input intensity of the first harmonic should be $\sim$10 GW/cm^2. The estimations confirm also a possibility of observations of solitons in the "metal–dielectric" structures.[30] Note also that a possibility of generation of the giant plasmonic second-harmonic in the metal–dielectric structures is connected with a presence of proper quadratic nonlinearity. The best way to achieve this may be using metamaterial structures like metal-dielectric, with a set of nonlinear nanoparticles with proper asymmetry, which allows a presence of quadratic nonlinearity. In this connection, let us note that excitation of the second-harmonic in the optical range in the system, containing the asymmetric set of metal nanoparticles, with intensity of pumping wave of order of 1 GW/cm^2, was considered in Ref. 63. Note that such a value lies in the range of the necessary intensities of the main harmonic, in accordance with the estimations, presented in our chapter earlier.

A possibility of "giant" generation of the second-harmonic of the surface plasmons in the regime of the "temporal" resonance of such a harmonic is now discussed, as well as a possibility of soliton and quasi-soliton pulse propagation. Both quadratic and cubic nonlinearities are accounted for. The general method for the derivation of the evolution equation in the layered structures with a nonlinear plasma, including both volume and surface nonlinearities, is presented. An issue of importance of surface nonlinearity, which may change qualitatively a behavior of a nonlinear system, is addressed. It is shown that under a condition of the temporal resonance of the second-harmonic, modulation-like instability and quasi-soliton propagation

are possible even in the range of parameters out of fulfillment of "formal" Lighthill criteria. A possibility of observations of nonlinear surface plasmons in different layered structures, including gaseous, semiconductor and metal plasma and plasmonic metamaterials, is discussed. The considered phenomena may be useful for sensors, subwavelength imaging and other applications.

10.4.2.3. *Strong surface nonlinearity and collapse*

For numerical simulations, the basic equations have been rewritten in unidimensional form, Eqs. (10.52) and (3.53). The time is normalized to $t_n = \omega_p^{-1}$, where $\omega_p = (4\pi e^2 n_0/m)^{1/2}$ is the plasma frequency, the distances are normalized to $l_n = ct_n$. The electromagnetic field is normalized to $E_n = mc\omega_p/e$, the electron velocity is normalized to c. The electron concentration is normalized to the equilibrium one n_0.

$$\begin{aligned}
&\frac{\partial n}{\partial t} + \frac{\partial}{\partial x}(nv_x) + \frac{\partial}{\partial z}(nv_z) = 0,\\
&\frac{\partial v_x}{\partial t} + v_x\frac{\partial v_x}{\partial x} + v_z\frac{\partial v_x}{\partial z}\\
&\quad = \gamma^{-1}\left(E_x - \frac{1}{c}v_z H - v_x(\vec{v}\cdot\vec{E})\right) - \nu v_x - \gamma^{-1}\frac{c_s^2}{n}\frac{\partial n}{\partial x},\\
&\frac{\partial v_z}{\partial t} + v_x\frac{\partial v_z}{\partial x} + v_z\frac{\partial v_z}{\partial z}\\
&\quad = \gamma^{-1}\left(E_z + \frac{1}{c}v_x H - v_z(\vec{v}\cdot\vec{E})\right) - \nu v_z - \gamma^{-1}\frac{c_s^2}{n}\frac{\partial n}{\partial z},
\end{aligned} \tag{10.52}$$

where $\gamma^{-1} = (1-v^2)^{1/2}$, $c_s^2 = v_T^2/c^2 \ll 1$,

$$\frac{\partial H}{\partial t} = \frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z};\quad \frac{\partial E_x}{\partial t} = -\frac{\partial H}{\partial z} - nv_x;\quad \frac{\partial E_z}{\partial t} = \frac{\partial H}{\partial x} - nv_z,$$
$$|x| > L_x : n\vec{v} = 0. \tag{10.53}$$

Here, c_s^2 is unidimensional electron temperature, in the cold electron plasma its value is $c_s^2 \le 10^{-3} \ll 1$.

It is considered primarily that the initial excitation of the perturbation of the electron concentration is $\tilde{n} \equiv n - 1$ at $|x| \leq L_x$:

$$\tilde{n}(t = 0, z, x) = a_0 \sin\left(\frac{z - z_1}{z_{02}}\right) \exp\left(-\left(\frac{z - z_1}{z_{01}}\right)^6\right); \; v_z(t = 0) = 0. \tag{10.54}$$

The distance z_{01} determines the modulation scale, whereas the distance $z_{02} \ll z_{01}$ determines the spatial oscillation period. The amplitude of the initial excitation is $a_0 \ll 1$.

The initial values of E_x and E_z should satisfy the Maxwell equation:

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_z}{\partial z} = \tilde{n} \equiv n - 1. \tag{10.55}$$

Therefore, the initial values of E_x and E_z are calculated from the electric potential φ:

$$\vec{E}(t = 0) = -\vec{\nabla}\varphi, \quad \Delta\varphi = -\tilde{n}. \tag{10.56}$$

Initially, the electron velocity is absent: $v_x = v_z = 0$. The case when $v_z \neq 0$ is also considered below. The magnetic field is absent initially, $H(t = 0) = 0$. The Poisson equation for the electric potential φ has been solved by the fast Fourier transform with respect to z.

In the numerical methods, the cross-like explicit difference schemes have been used[64]:

$$\frac{n_{j,k}^{p+1} - n_{j,k}^{p-1}}{2\tau} - \frac{h_{x2}(nv_x)_{j,k-1}^{p}}{h_{x1}(h_{x1} + h_{x2})} + \frac{(h_{x2} - h_{x1})(nv_x)_{j,k}^{p}}{h_{x1}h_{x2}} + \frac{h_{x1}(nv_x)_{j,k+1}^{p}}{h_{x2}(h_{x1} + h_{x2})} + \frac{(nv_z)_{j+1,k}^{p} - (nv_z)_{j-1,k}^{p}}{2h_z} = 0. \tag{10.57}$$

Here, j, k are numbers of nodes for z and x; p is the node for t; τ, h_z are steps for t, and z, h_{x1} and h_{x2} are the steps for x near the node with the number k. Generally, the variable steps for x are used: $h_{x1} \neq h_{x2}$. The step h_z for z is constant. Near the boundary $x = L_x$, the steps for x should be much smaller than in the center $x = 0$ because the details of the dynamics of the electron concentration are

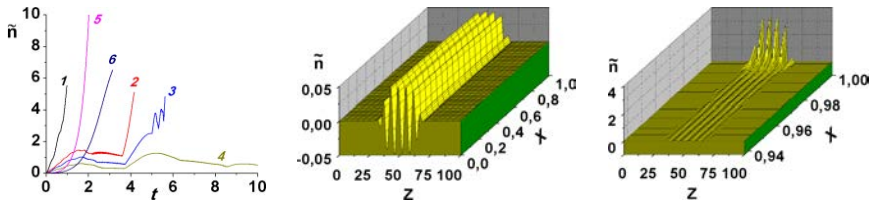
important there, as shown below. The temporal step τ is chosen from the condition of stability.

At the boundary $x = L_x$, the equation for the concentration n has resulted from the balance of electrons near the boundary $x = L_x$ and from the condition $v_x = 0$ at this boundary.

The initial excitation of electron plasma waves in the plasma layer has been considered at low amplitudes of the perturbations of the concentration $a_0 \ll 0.1$, when the volume nonlinearity is negligibly small. Initially, the surface charge near the boundaries is absent and is excited during the evolution of the plasma wave.

When the amplitudes are very small, $a_0 \leq 0.01$, the initial excitation results in oscillations of the electron concentration near the plasma boundary. But at higher initial amplitudes, $a_0 \ll 0.01$, the dynamics change drastically. The sharp peaks of the electron concentration near the surfaces are formed, which are localized both in longitudinal (z) and transverse (x) directions. The thickness of the localization for the transverse coordinate x is about 0.01. The maximum values of $\tilde{n} > 4$ are 50–100 times higher than within the volume. The nonlinearity ceases to be moderate near the surface. Moreover, under the values of $a_0 > 0.01$, the calculations overflow. It is possible to interpret this phenomenon of the sharp growing and the narrowing of the peaks of the electron concentration near the surface as a manifestation of the wave collapse.[65] The dynamics of the wave collapse depend on the electron temperature. Namely, at unidimensional electron temperatures, $c_s^2 > 0.005$, the dynamics of surface charge in the electron plasma waves do not lead to collapse phenomena. Therefore, the electron heating can be a mechanism of stabilization of the wave collapse.

In Fig. 10.15(a), the dependencies of the maximum values of perturbations of the electron concentration $\tilde{n}$ on the time t in the plasma layer for different values of the initial amplitudes a_0 are shown. The sizes of the plasma slab are: $L_z = 100, 2L_x = 2$. The parameters on the initial excitation are $z_{01} = 15$ and $z_{02} = 1$. The unidimensional temperature is $c_s^2 = 10^{-4}$. The wave dynamics is tolerant to changes of the sizes of the plasma layer $2L_x, L_z$ and to the values of z_{01}, z_{02}. For



(a) (b) (c)

Fig. 10.15. (a) is dependencies of maximum values of perturbations of electron concentration $\tilde{n}$ on time t. Curve 1 is for the initial amplitude of the perturbation of the volume concentration $a_0 = 0.05$, curve 2 is for $a_0 = 0.02$, curve 3 is for $a_0 = 0.015$ and curve 4 is for $a_0 = 0.01$. Curves 5 and 6 are for $a_0 = 0.05$ and for $a_0 = 0.02$ for the non-zero initial distribution of the longitudinal electron velocity v_z with the amplitude a_0. Figures 10.15(b) and 10.15(c) are the dynamics of the wave collapse at a relatively high initial amplitude $a_0 = 0.05$ above the threshold, see curve 1 in Fig. 10.2; Fig 10.15(b) is the perturbation of the electron concentration $\tilde{n}$ for $t = 0$; (c) is $\tilde{n}$ for $t = 0.8$.

the initial amplitudes that correspond to the curves 1, 2, 3, 5 and 6, the wave collapse takes place. In the numerical simulations, this manifests in the overflowing after some time of calculations, when the perturbations of the electron concentration near the plasma boundary reaches high values, $\tilde{n} > 5$. Note that the results of simulations do not change when the spatial or temporal steps have been varied.

At low initial amplitudes, the oscillatory behavior of the surface charge takes place. The values of the concentration near the boundary exceed the values within the volume 10 times only. At higher amplitudes, the surface nonlinearity cannot be considered as moderate, and wave collapse occurs.

The distributions of the electric field component E_z and the component of the electron velocity v_z on the coordinate x are smooth.

Also, the excitation of plasma waves has been considered when the initial non-zero value of the longitudinal component of the electron velocity v_z has been given. Typical results are similar to the excitation by perturbations of the electron concentration.

It is possible to conclude that the wave collapse, i.e. the sharp increase of the plasma concentration near the boundaries, is a general phenomenon. It occurs when the initial excitation of the plasma

waves takes place within the volume and initially the surface charge near the plasma boundaries is absent.

Because at higher electron temperatures the wave collapse disappears, a possible mechanism of limitation of high values of the electron concentration can be electron heating.

10.4.2.4. *Final remarks*

The exciting possibility of nonlinear control over strongly nonlinear phenomena is discussed in terms of a self-consistent nonlinear surface concentrator. It involves jumping the focusing point and a tendency to "hot spot" formation in a metamaterial nonlinear concentrator (nonlinear electromagnetic "black hole") by means of applying an additional "controlling" beam. The tendency to hot spot formation reveals itself in any of the following processes: an increase of input amplitude; an increase of linear amplification in the central nonlinear region; a decrease in nonlinear losses, or decrease in saturation of nonlinearity. Therefore a tendency to a formation of "hot spots" is a rather universal feature in the strongly nonlinear behavior of the "nonlinear resonator" system. At the same time, the system is not sensitive to the "prehistory" of approaching a nonlinear threshold intensity (amplitude). The analysis given here shows that the proposed method of matching Complex Geometrical Optics with full-wave electromagnetic (nonlinear) solution can be used for modelling a field concentrator based on isotropic and strongly nonlinear metamaterials. Nevertheless, this method can be used also for anisotropic (hyperbolic) metamaterials as well as propagation of electromagnetic waves through the layered anisotropic system "Lithosphere–Atmosphere–Ionosphere–Magnetosphere",[66] and interpretation of data of observations in space projects.

Systems under consideration could be used for creation of antenna with strongly nonlinear focusing and for providing strongly nonlinear effects such as higher harmonic generation, etc.

A general method for the derivation of a nonlinear evolution equation for (wave processes in) layered structures (NEELS) with surface and volume nonlinearities is developed and the importance

of surface nonlinearity in the nonlinear processes is addressed. The methods for the nonlinear wave processes in the layered structures[31–35,67] and[36] are, in fact, the parts of the general metamaterial approach to the wave processes, applicable for both an artificial[26] and natural[36] media.

A possibility of "giant" resonant generation of the second-harmonic of the surface plasmons in the regime of the "temporal" resonance of such a harmonic is shown clearly in this chapter. A possibility of observations of nonlinear surface plasmons in different layered structures, including gaseous, semiconductor and metal plasma and plasmonic metamaterials, is discussed. The considered phenomena may be useful for sensors, subwavelength imaging and other applications.

Acknowledgment

One of the authors (Yu. G. Rapoport) would like to thank the STCU (Project 6060, Theoretical and experimental investigations of the resonant phenomena in the near-space plasma) for partial support of this work.

References

1. Remoissenet, M. (1996). *Waves Called Solitons*, 2nd Edn. (Springer, Berlin).
2. Boardman, A. D. and Sukhorukov, A. P. (2000). *Soliton-driven Photonics* (Kluwer Academic Publishers, Dordecht).
3. Trillo, S. and Torruellas, W. (2001). *Spatial Solitons* (Springer, Berlin).
4. Ferrari, L., Wu, C., Lepage, D., Zhang, X. and Liu, Z. (2015). Hyperbolic metamaterials and their applications. *Progress in Quantum Electronics*, **40**, pp. 1–40.
5. Aitchison, J. S., Weiner, A. M., Silberberg, Y., Leaird, D. E., Oliver, M. K., Jackel, J. L. and Smith, P. W. (1990). Experimental observation of spatial solitons in a nonlinear glass waveguide. *Opt. Lett.* **15**, p. 471.
6. Ballav, M. and Chowdhury, A. R. (2006). On a study of diffraction and dispersion-managed solitons in cylindrical media. *Prog. Electromagn. Res.* **63**, pp. 33–50.
7. Ablowitz, M. J. and Musslimani, Z. H. (2001). Discrete diffraction-managed spatial solitons. *Phys. Rev. Lett.* **87**, p. 254102.

8. Boardman, A. D., Marinov, K., Pushkarov, D. I. and Shivarova, A. (2000). Influence of nonlinearly induced diffraction on spatial solitary waves. *Opt. Quant. Electron.* **32**(1), pp. 49–62.
9. Boardman, A. D., Marinov, K., Pushkarov, D. I. and Shivarova, A. (2000). Wave-beam coupling in quadratic nonlinear optical waveguides:Effects of nonlinearly induced diffraction. *Phys. Rev. E* **62**, pp. 2871–2876.
10. Kockaert, P., Tassin, P., Van Der Sande, G., Veretennicoff, I. and Tlidi, M. *et al.* (2006). Negative diffraction pattern dynamics in nonlinear cavities with left-handed materials. *Phys. Rev. A* **74**, p. 033822.
11. Boardman, A. D., King, N. and Velasco, L. (2005). Negative refraction in perspective. *Electromagn.* **25**, pp. 365–389.
12. Infeld, E. and Rowlands, G. (2000). *Nonlinear Waves, Solitons and Chaos* (Cambridge University Press, Cambridge).
13. Scalora, M., Syrchin, M. S., Akozbek, N., Poliakov, E. Y., Aguanno, G. D., Mattiucci, N., Bloemer, M. J. and Zheltikov, A. M. (2005). Generalized nonlinear Schrödinger equation for dispersive susceptibility and permeability: Application to negative index materials. *Phys. Rev. Lett.* **95**, p. 013902.
14. Wen, S. C., Wang, Y. W., Su, W. H., Fu, X. Q. and Fan, D. Y. (2006). Modulation instability in nonlinear negative-index material. *Phys. Rev. E* **73**, p. 036617.
15. Wen, S., Xiang, Y., Dai, X., Tang, Z., Su, W. and Fan, D. (2007). Theoretical models for ultrashort electromagnetic pulse propagation in nonlinear metamaterials. *Phys. Rev. A* **75** p. 033815.
16. Wen, S. C., Xiang, Y. J., Su, W. H., Hu, Y. H., Fu, X. Q. and Fan, D. Y. (2006). Role of the anomalous self-steepening effect in modulation instability in negative-index material. *Opt. Exp.* **14** pp. 1568–1575.
17. Xiang, Y., Wen, S., Dai, X., Tang, Z., Su, W. and Fan, D. (2007). Modulation instability induced by nonlinear dispersion in nonlinear metamaterials. *J. Opt. Soc. Am. B* **24**, pp. 3058–3063.
18. D'Aguanno, G., Mattiucci, N. and Bloemer, M. J. (2008). Ultraslow light pulses in a nonlinear metamaterial. *J. Opt. Soc. Am. B* **25**, pp. 1236–1241.
19. Boardman, A. D., King, N., Mitchell-Thomas, R. C., Malnev, V. N. and Rapoport, Y. G. (2008). Gain control and diffraction-managed solitons in metamaterials. *Metamaterials* **2**, pp. 145–154.
20. Boardman, A. D., Rapoport, Y. G., King, N. and Malnev, V. N. (2007). Creating stable gain in active metamaterials. *J. Opt. Soc. Am. B* **24**, pp. A53–A61.
21. Webb, K. J and Ludwig, A. (2009). Lossless negative dielectric constant optical material from a semiconductor quantum dot mixture. *Phys. Rev. B* **78** pp. 1–2.
22. Kinsler, P. and McCall, M. W. (2008). Causality-based criteria for a negative refractive index must be used with care. *Phys. Rev. Lett.* **101**, p. 167401.
23. Kadomtsev, B. B. and Karpman, V. I. (1971). Nonlinear waves. *Phys. Usp.* **14**, pp. 40–60.

24. Wang, H. C. and She, W. L. (2005). Modulation instability and interaction of non-paraxial Kerr media. *Opt. Commun.* **254**, pp. 145–151.
25. Zayats, A. V. and Maier, S. A., (eds.) (2013). *Active Plasmonics and Tuneable Plasmonic Metamaterials* (John Wiley & Sons, New York).
26. Rapoport, Yu. G., Boardman, A. D., Grimalsky, V. V., Ivchenko, V. M. and Kalinich, N. (2014). Strong nonlinear focusing of light in nonlinearly controlled electromagnetic active metamaterial field concentrators. *J. Opt.* **16**, p. 055202.
27. Paul, T., Rockstuhl, C. and Lederer, F. (2011). Integrating cold plasma equations into the fourier modal method to analyze second harmonic generation at metallic nanostructures. *J. Mod. Opt.* **58**, pp. 438–448.
28. Mochan, W. L. and Maytorena, J. A. (2004). Theory of surface second harmonic generation. *Proc. of EPIOPTICS-8 - 33rd Course of the International School of Solid State Physics*, Erice, Italy, 20–26 July 2004, pp. 1–29.
29. Zheludev, N. I. and Emel'yanov V. I. (2004). *J. Opt. A: Pure Appl. Opt.* **6**, p. 26.
30. Samson, Z. L., Horak, P., MacDonald, K. F. and Zheludev, N. I. (2011). *Opt. Lett.* **36**, pp. 250–252.
31. Grimalsky, V. V., Rapoport, Yu. G. and Slavin, A. N. (1997). Nonlinear diffraction of magnetostatic waves in ferrite films. *J. Phys. IV Colloque* **7**, pp. 393–394.
32. Boardman, A. D., King, N. and Rapoport, Yu. (2007). Metamaterials driven by gain and special configurations, *SPIE Proc. of Metamaterials II*, **6581**(658108), pp. 1–10.
33. Grimalsky, V. V. and Rapoport, Yu. G. (1998). Modulational instability of surface plasma waves in the second–harmonic resonance region. *Plasma Phys. Rep.* **24**, pp. 980–982.
34. Rapoport, Yu. G. and Grimalsky, V. V. (2011). Modulational instability of surface plasma waves in the second–harmonic resonance region, *Proc. International Conference Days on Diffraction 2011, DD 2011, St. Petersburg, Russian Federation*, 30 May-3 June 2011, Art. No 6094387, pp. 168–173.
35. Rapoport, Yu. G. (2014). General method for the derivations of the evolution equations and modeling, *Bullet. Taras Shevchenko Nat. Univ. of Kyiv, Series Phys. & Math.* **1**, pp. 281–288; http://www.library.univ.kiev.ua/ukr/host/10.23.10.100/db/ftp/visnyk/fiz_mat_1_2014.pdf
36. Grimalsky, V. V., Kremenetsky, I. A. and Rapoport, Yu. G. (1999). Excitation of electromagnetic waves in the lithosphere and their penetration into ionosphere and magnetosphere. *JAE* **19**, pp. 101–117.
37. Narimanov, E. E. and Kildishev, A. V. (2009). Optical black hole: Broadband omnidirectional light absorber. *Appl. Phys. Lett.* **95**, p. 041106.
38. Kildishev, A. V. and Shalaev, V. M. (2011). Transformation optics and metamaterials. *Phys. Usp.* **54**, pp. 53–64.
39. Greegor, B., Parazzoli, C. G., Nielsen, J. A., Thompson, M. A., Tamelian, M. H. and Smith, D. R. (2005). Simulation and testing of a graded negative index of refraction lens. *Appl. Phys. Lett.* **87**, pp. 091114–091117.

40. Boardman, A. D., Mitchell-Thomas, R. C., King, N. J. and Rapoport, Y. G. (2010). Bright spatial solitons in controlled negative phase metamaterials. *Opt. Commun.* **283**, pp. 1585–1597.
41. Scalora, M., Caglia, D., D'Aguanno, G. Mattiucci, N., Akozbek, N., Centini, M. and Bloemer, M. J. (2007). Gap solitons in a nonlinear quadratic negative-index cavity. *Phys. Rev. E* **75**, p. 066606.
42. Alberucci, A. and Assanto, G. (2011). Nonparaxial $(1 + 1)$D spatial solitons in uniaxial media. *Opt. Lett.* **36**, pp. 193–195.
43. Alberucci, A. and Assanto, G. (2011). On beam propagation in anisotropic media: One-dimensional analysis, *Opt. Lett.* **36**, pp. 334–336.
44. Smith, D. R. and Schurig, D. (2003). Electromagnetic wave propagation in media with indefinite permittivity and permeability tensors. *Phys. Rev. Lett.* **90**, p. 077405.
45. Poddubny, A., Iorsh, I., Belov, P. and Kivshar, Y. (2013). Hyperbolic metamaterials. *Nat. Photon.* **7**, pp. 958–967.
46. Hasegawa, A. and Tappert, F. (1973). Transmission of stationary nonlinear optical pulses in dispersive dielectric fibers. I. anomalous dispersion. *Appl. Phys. Lett.* **23**, pp. 142–144.
47. Hasegawa, A. and Tappert, F. (1973). Transmission of stationary nonlinear optical pulses in dispersive dielectric fibers. II. normal dispersion, *Appl. Phys. Lett.* **23**, pp. 171–172.
48. Kou, Y., Ye, F. and Chen, X. (2011). Multipole plasmonic lattice solitons. *Phys. Rev. A* **84**, p. 033855.
49. Silveirinha, M. G. (2013). Theory of spatial optical solitons in metallic nanowire materials. *Phys. Rev. B* **87**, p. 235115.
50. Smolyaninov, I. I. (2013). Analog of gravitational force in hyperbolic metamaterials. *Phys. Rev. A* **88**, p. 033843.
51. Peccianti, M., Conti, C., Assanto, G., De Luca, A. and Umeton, C. (2004). Routing of anisotropic spatial solitons and modulational instability in liquid crystals. *Nature* **432**, pp. 733–737.
52. Lu, W., Jin, J., Lin, Z. and Chen, H. (2010). A simple design of an artificial electromagnetic black hole. *J. Appl. Phys.* **108**, p. 064517.
53. Cheng, Q., Cui, T. C., Jiang, W. X. and Cai, B. G. (2010). An omnidirectional electromagnetic absorber made of metamaterials. *New J. Phys.* **12**, p. 063006.
54. Zhou, J., Cai, X., Chang, Z. and Hu, G. (2011). Experimental study on a broadband omnidirectional electromagnetic absorber. *J. Opt.* **13**, p. 085103.
55. Yang, Y., Leng, L. Y., Wang, N., Ma, Y. and Ong, C. K. (2012). Electromagnetic field attractor made of gradient index metamaterials. *J. Opt. Soc. Amer. A* **29**, pp. 473–475.
56. Leonhardt, U. and Philbin, T. G. (2009). Transformation optics and the geometry of light. *Prog. Opt.* **53**, pp. 69–152.
57. Bergman, L., Alitalo, P. and Tretyakov, S. A. (2011). *Phys. Rev. B* **84** p. 205103.
58. Alexandrov, A. F., Bogdankevich, L. S. and Rukhadze, A. A. (1984). *Principles of Plasma Electrodynamics* (Springer, New York).

59. Aliev, Yu. M., Schlüter, H. and Shivarova, A. (2000). *Guided-Wave-Produced Plasmas* (Springer, Berlin).
60. Landau, L. D. and Lifshitz, E. M. (1984). *Electrodynamics of Continuous Media* (Pergamon Press, Oxford).
61. Kadomtsev, B. B. (1988). *Collective Phenomena in Plasma* (Nauka, Moscow) (in Russian).
62. Chu, J. and Sher, A. (2008). *Physics and Properties of Narrow Gap Semiconductors* (Springer, Berlin).
63. Zheludev, N. I. and Emel'yanov. V. I. (2004). Phase matched second harmonic generation from nanostructured metallic surfaces. *J. Opt. A* **6**, pp. 6–28.
64. Roache, P. J. (1998). *Fundamentals of Computational Fluid Dynamics* (Hermosa Publ., Socorro).
65. Berge, L. (1998). Wave collapse in physics: Principles and applications to light and plasma waves. *Phys. Rep.* **303**, pp. 259–370.
66. Rapoport, Yu., Boardman, A., Grimalsky, V., Selivanov, Yu. and Kalinich, N. (2012). Metamaterials for space physics and the new method for modeling isotropic and hyperbolic nonlinear concentrators. *Proc. International Conference on Mathematical Methods in Electromagnetic Theory.* MMET 2012, Kharkiv, Ukraine, 28–30 August 2012, Art. No 6331154, pp. 76–79.
67. Pitilakis, A., Chatzidimitriou, D., Kriezis, E. E. (2016). Theoretical and numerical modeling of linear and nonlinear propagation in graphene waveguides. *Opt. Quant. Electron.* **48**, pp. 1–22.

CHAPTER 11

Metamaterial Catheter Receivers for Internal Magnetic Resonance Imaging

RICHARD R. A. SYMS,* IAN R. YOUNG and LASZLO SOLYMAR

Imperial College London, UK

Metamaterial catheter receivers for high-resolution internal magnetic resonance imaging (MRI) of the biliary system are described. The clinical goal, early detection and staging of bile duct cancer, is first introduced. Current approaches to diagnosis are then described, with an emphasis on MRI, and the advantages of internal imaging with an endoscopically-delivered receiver are highlighted. Magneto-inductive (MI) waveguides — linear arrays of magnetically-coupled L-C resonators — are proposed as a solution. MI waves and waveguides are reviewed, thin-film MI cables are introduced as a low-loss and stable format that can easily be mounted on a flexible catheter and the layouts needed for an inherently safe receiver are described. Experiments designed to demonstrate endoscopic compatibility, confirm safe operation and demonstrate high-resolution imaging are then presented. Frequency scaling is discussed, and experimental receivers are shown to out-perform surface coil arrays *in vitro*, over a limited field of view.

Corresponding author: *r.syms@imperial.ac.uk

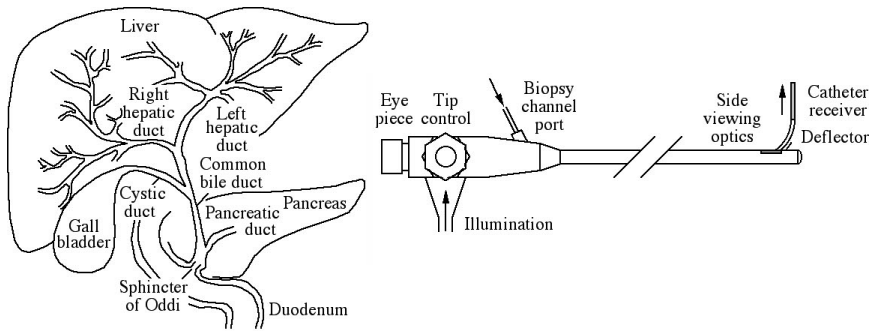
11.1. Introduction

The observation by Pendry[1] that periodic arrangements of structured conductors could be used to form artificial media with novel electromagnetic properties and hence realize the predictions of Veselago[2] concerning media with negative permittivity and permeability has led to the new field of metamaterials. In the following decade and a half, interest has been immense. Many different unit cells have been proposed, a wide variety of lattices have been explored and the operating range has been raised from radio frequency to the optical domain.[3–5] In addition to identifying opportunities, this activity has also (naturally enough) stumbled upon difficulties. Metallic metamaterials are inherently narrow band and lossy, and it has proved difficult to manufacture large arrays at low cost, especially at high frequency when expensive patterning is needed to fabricate the fine details involved. As a result, applications are likely to be limited to those for which the novel properties of metamaterials can overcome these drawbacks.

The aim of this chapter is to demonstrate an application that appears to fit the bill, in that a metamaterial approach seems capable of out-performing conventional techniques. The application is in medical imaging, specifically in internal magnetic resonance imaging (MRI) of cholangiocarcinoma (CCA). The disease problem is serious and without a current solution, encouraging the idea that metamaterials could help to solve a global problem. To explain how this may be, we shall first give some background on CCA and on current imaging approaches, explaining their limitations. We will then introduce the particular metamaterial involved — the magneto-inductive (MI) waveguide — and explain how it may be used in internal MRI. We will then give examples of imaging performance, and conclude by outlining the path to clinical trials.

11.2. Clinical and Imaging Background

CCA[6] is a cancer of the bile ducts that lie between the liver and the duodenum (Fig. 11.1(a)). CCA is rare in the West (US annual



(a) (b)

Fig. 11.1. (a) Anatomy of the biliary ductal system and (b) flexible duodenoscope.

incidence 1 in 100,000) although incidence is rising.[7] Incidence is much higher in the Far East, especially in Thailand, due to infestation of the local population with parasitic liver flukes (*O. viverrini* and *C. sinensis*) following ingestion of cysts in uncooked or pickled river fish.[8] An estimated 9.4% of the Thai population (6 million) is infected, and the annual CCA incidence is >100 per 100,000 in Khon Kaen province, the region with the highest worldwide incidence. As the flukes develop, they induce chronic irritation and hyperplasia of the duct lining or epithelium, leading to malignant transformation and adenocarcinoma. Unfortunately CCA is resistant to chemotherapy and radiotherapy, and surgical resection or liver transplantation offer the only real hope of a cure.[9] Several factors combine to yield a dismal 5-year survival. CCA generally recurs unless the lesions are detected early, when they are less than 1 mm in size. However, the initial symptoms are mild, so CCA patients present late. Furthermore, existing imaging methods lack the resolution and contrast for diagnosis and accurate staging, contributing to poor prognosis even when surgery is feasible.

The biliary and pancreatic ducts are normally investigated using endoscopic retrograde cholangio-pancreatography (ERCP).[10] A flexible duodenoscope (Fig. 11.1(b)) is passed down the throat to the duodenum, using a side viewing optical system to locate the sphincter of

Oddi (the common opening to the ducts). The instrument contains a biopsy channel with a steerable side port, allowing the ductal system to be cannulated with catheter tools. Usually a guidewire system is used; a wire is first inserted into the duct, and a hollow catheter is then inserted over the wire. For X-ray imaging, the wire is then withdrawn and a contrast agent containing iodine is injected through the catheter. This approach allows imaging of the entire ductal system using X-ray fluoroscopy, and observation of constrictions can then be used to identify the location of lesions. However, owing to the lack of soft tissue contrast, such images only allow visualization of the duct boundaries and cannot differentiate between malignant and non-malignant strictures.

Endoscopic ultrasound (EUS) using a duodenoscope with a mechanically or electrically scanned transducer at its tip allows acoustic imaging to be carried beyond the duodenum.[11] Although EUS does provide soft tissue contrast, there is a trade-off between range and resolution caused by increasing acoustic attenuation at high frequency. Furthermore, EUS suffers from near-field artifacts arising from reflections in the transducer housing that tend to obscure important details. Strongly reflecting anatomical features also cause reverberation artifacts or acoustic shadowing. Intraductal ultrasound (IDUS) using a catheter with a much smaller, electronically scanned, transducer extends ultrasonic imaging to the duct itself, but with an even shorter range.[12]

Improvements in contrast are offered by MRI, which operates by controlled manipulation of nuclear magnetic dipoles.[13] The most abundant nuclear species in the body is the proton, so ^{1}H MRI is commonly used for medical imaging; however, other nuclei such as ^{13}C, ^{14}N, ^{17}O and ^{31}P are also used. The patient lies in a superconducting magnet, whose effect is to align the magnetization vector $\mathbf{M}$ of dipole ensembles almost parallel to the static magnetic field B_0 (normally, the z-direction). However, the alignment is not perfect at finite temperature; instead, $\mathbf{M}$ precesses around the z-axis at the Larmor frequency ω_L, an angular frequency related to the magnetic field and the gyromagnetic ratio γ of the nuclei. The

magnetization vector then has two components, a longitudinal component M_z parallel to the z-axis and a transverse component M_{xy} rotating around it; however, in equilibrium, M_{xy} is much smaller than M_z. For protons, $\gamma/2\pi = 42.576\,\mathrm{MHz/T}$, so in a 1.5 T magnetic field (the strength of many clinical magnets) the Larmor frequency is $f_L = \omega_L/2\pi = 63.86\,\mathrm{MHz}$, a radio frequency.

MRI is a two-step process. The first step is excitation, and it is carried out by applying a short pulse of a time-varying field B_1 that also rotates around the z-axis at the Larmor frequency. The B_1 field exerts a torque that alters the dipole alignment so that $\mathbf{M}$ rotates with respect to the z-axis. The "flip angle" depends on the duration of the excitation pulse; a 90° pulse rotates $\mathbf{M}$ until it is perpendicular to the z-axis. The second step is detection. The excitation is removed, and a resonant circuit tuned to ω_L is used to detect the signals induced by the precessing dipoles as they return to equilibrium. They relax exponentially, but with different time constants T1 and T2 for M_z and M_{xy}. Importantly, these vary from tissue to tissue, so the differences can be converted into amplitude variations that provide contrast. Different excitation sequences can also be used to render the contrast more sensitive to variations in T1 or T2, and contrast agents can be used to enhance the difference still further. Because the excitation is only effective if its frequency matches ω_L, an axial gradient in B_0 (which spatially varies ω_L) can localize the excitation to a slice of tissue. Transverse gradients then encode the origin of the detected signal in frequency and phase, allowing a 2D image to be recovered by fast Fourier transformation.

By analogy with ERCP, the technique used for ductal imaging is known as magnetic resonance cholangio-pancreatography (MRCP).[14] Excitation is carried out using the system body coil, a large RF coil in the magnet tunnel, while detection is carried out using a surface coil array on the patient's abdomen (Fig. 11.2(a)). Unfortunately, the resolution achievable is directly linked to the signal-to-noise-ratio (SNR) of the detected signal. The largest source of electrical noise is thermal noise arising from the patient, whose soft tissue is a lossy dielectric at radio frequency.[15] For surface coils, which have a large

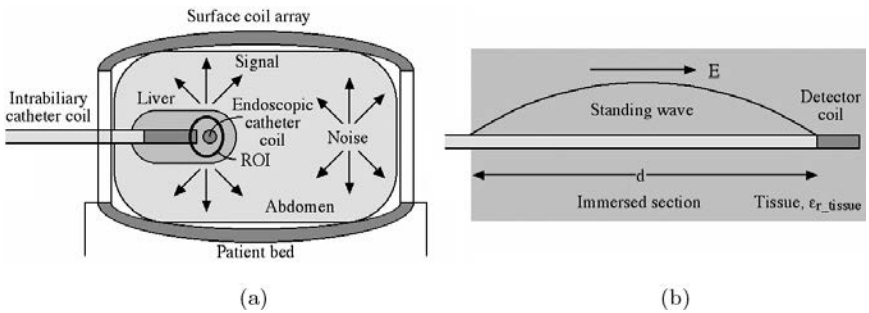


Fig. 11.2. (a) Alternative locations for signal detectors and (b) standing waves on conductors.

field of view (FOV), this noise is extremely significant and cannot be increased significantly by averaging due to patient motion. As a result, it is difficult to achieve the sub-millimeter resolution required for imaging cancers such as CCA.

Since they have a much smaller FOV for noise, small internal coils can have much higher SNR provided they can be located close enough to the region of interest (ROI).[16] A variety of catheter-based coils have been developed for intravascular imaging,[17,18] and MRCP has been adapted to use these coils, inserted through percutaneous drainage tubes into the biliary region in a procedure known as intrabiliary MRI[19] (Fig. 11.2(a)). The image quality is much higher than with external coils, but the procedure is inherently invasive and generally involves surgery.

Internal coils also suffer from other drawbacks. The B_1 field of the transmitter must be prevented from coupling directly to the detection coil, or it will induce large unwanted currents. Less obviously, there is generally a large RF electric field E near capacitors in the body coil. This field can excite surface waves on long conductors, and if the frequency is such that standing waves are excited (Fig. 11.2(b)), the resulting current can be high. First-order resonances will be excited if the conductor length d is half a wavelength, so that $d = c/(2f_L\sqrt{\varepsilon_{r\_\text{tissue}}})$, where c is the velocity of light and $\varepsilon_{r\_\text{tissue}}$ is the relative permittivity of the surrounding tissue. At 63.86 MHz, $\varepsilon_{r\_\text{tissue}} \approx 77$, so the critical length can be as short as

0.27 m. There can then be significant tissue heating near the conductor ends, and problems have been identified with many metallic medical components designed for internal use such as guide-wires and leads.[20,21] Several solutions have been proposed, including the insertion of chokes or tank filters at intervals to block the flow of currents,[22] and transformer segmentation.[23]

The arguments above suggest a need for alternative approaches that might allow high-resolution MRI of CCA. To increase SNR, internal coils should be used; however, to access the biliary ducts without surgery, they should be endoscopically delivered. Consequently, they should be catheter-mounted and capable of passing through the biopsy channel of a duodenoscope, which has a length of ≈1.5 m and an internal diameter of ≈3 mm. The catheter should be flexible, so it can be bent through around 90° for insertion into the sphincter of Oddi. To allow efficient cannulation, the catheter should be capable of passing over a guidewire. Finally, for intrinsic safety, the system should contain circuitry to minimize direct coupling between the receiver and the scanner. These requirements are challenging, but can be met using a metamaterial construction based on MI waveguides, as we now show.

11.3. Magneto-Inductive Systems

Magneto-Inductive (MI) media are periodic arrays of magnetically coupled L-C resonators.[24] The arrays can exist in one, two or three dimensions,[25] but the simplest arrangement involves a linear waveguide with period a. There are two conventional configurations: axial and planar (Fig. 11.3(a)).

In each case, the resonant elements are planar loops. However, in the axial geometry, the elements lie in stacked planes, with their magnetic fields parallel to the axis of the guide. In the planar geometry, the elements lie in the same plane with their fields perpendicular to the axis. An equivalent circuit suitable for both variants is shown in Fig. 11.3(b). Here, magnetic coupling is represented by mutual inductance M between neighboring elements, and losses by resistors R. In the absence of voltage sources, Kirchhoff's voltage law implies

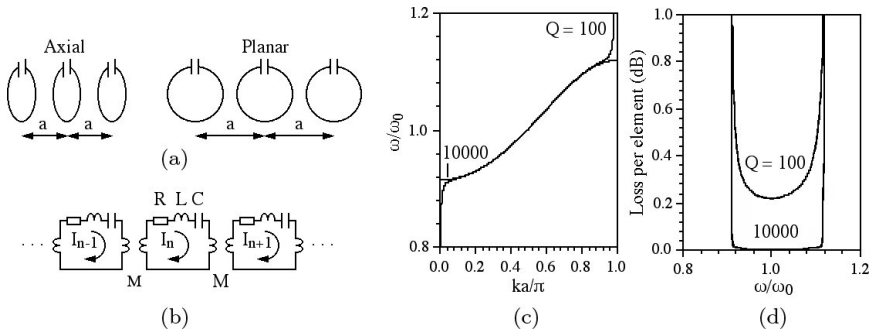


Fig. 11.3. (a) MI waveguides in axial and planar configurations and (b) equivalent circuit, (c) and (d) example dispersion and loss variations.

that the current I_n in the nth element at angular frequency ω can be related to the currents I_{n-1} and I_{n+1} in its neighbors by the recurrence relation:

$$(R + j\omega L + 1/j\omega C)I_n + j\omega M(I_{n-1} + I_{n+1}) = 0. \tag{11.1}$$

Solution of an infinite set of equations of this type can be achieved by assuming currents in the form of traveling waves, as $I_n = I_0 \exp(-\text{jnka})$ where I_0 is the wave amplitude and k is the propagation constant. Substitution into Equation (11.1) yields the dispersion relation:

$$1 - {\omega_0}^2/\omega^2 - j/Q + \kappa\cos(ka) = 0. \tag{11.2}$$

Here, $\omega_0 = 1/\sqrt{(LC)}$ is the angular resonant frequency of the elements, $Q = (\omega/\omega_0)Q_0$ where $Q_0 = \omega_0 L/R$ is the quality factor and $\kappa = 2M/L$ is the coupling coefficient. Assuming that $k = k' - jk''$, and that losses are small, Eq. (11.2) may be approximated as

$$\begin{aligned} 1 - {\omega_0}^2/\omega^2 + \kappa\cos(k'a) &= 0 \\ k''a &= 1/\{\kappa Q\sin(k'a)\}. \end{aligned} \tag{11.3}$$

The upper equation is the dispersion relation for lossless waves. In the axial configuration, M and κ are positive, the guide supports forward waves (whose phase and group velocities have the same sign) and propagation is obtained only over the frequency band

$1/\sqrt{(1+|\kappa|)} \leq \omega/\omega_0 \leq 1/\sqrt{(1-|\kappa|)}$. The lower equation shows that losses depend on k' and hence on frequency, and are smallest when $k'a = \pi/2$ and $\omega = \omega_0$. In the planar configuration, κ is negative and backward waves are supported. Figure 11.3(c) shows dispersion characteristics for the axial case, assuming a coupling coefficient $\kappa = 0.2$ for different Q-factors. For $Q = 10,000$ (achievable using superconducting elements), the results are indistinguishable from the lossless case. For $Q = 100$ (more realistic for conventional elements), there are clearly departures from ideal behavior. Differences now occur at the band edges, and the effect of loss is to allow propagation outside the ideal band. Figure 11.3(d) shows the frequency variation of loss; losses rise rapidly at the band edges.

A further important characteristic of any waveguide is its characteristic impedance. By inserting an impedance Z_{0m} into the final element of a finite guide, and choosing its value to eliminate reflection, it can be shown[26] that the characteristic impedance of a MI waveguide is

$$Z_{0m} = j\omega M \exp(-jka). \tag{11.4}$$

Generally, Z_{0m} is complex and varies with frequency. However, for lossless systems at resonance, it simplifies to the real value $Z_{0M} = \omega_0 M$.

This argument also implies that a resonant element can act as a transducer connecting the MI waveguide to a terminating impedance Z_0, which might represent a conventional system with real impedance. Overall behavior can then be modelled by writing all the circuit equations, including those for transducers, in the form $\mathbf{V} = \mathbf{Z}\,\mathbf{I}$. Here, $\mathbf{V}$ and $\mathbf{I}$ are column vectors describing the voltages and currents, and $\mathbf{Z}$ is a square matrix of self and mutual coupling impedances. The equations may be solved to find the unknown currents in terms of known voltages, as $\mathbf{I} = \mathbf{Z}^{-1}\mathbf{V}$. The reflection coefficient may then be extracted by resolving the input current into incident and reflected waves, which in turn allows the transmission coefficient to be found.

The characteristics above have been verified experimentally using wire-wound elements,[27] and MI waves have been observed in other

magnetically coupled media at gradually rising frequency, including "Swiss rolls",[28] solenoids,[29] split-ring resonators[30] and negative index photonic media.[31] Consequently, MI waves are now well established. However, performance must be high for practical systems. Dispersion is one problem; however, most applications proposed have been narrow band. Loss is a more significant obstacle. Equation 11.3 imply that low loss requires a high Q-factor and a high coupling coefficient. The former is limited by the conductivity of the material used (normally copper) and the skin effect. The latter is controlled by the arrangement of the elements. Planar elements can have a large extent in the direction of propagation, but have a relatively small coupling coefficient since $|\kappa| < 1$ in this geometry. Axial elements have a larger coupling coefficient that can approach $\kappa = 2$, but only if the period is so small that the distance propagated per element is extremely short. The net result is that losses per meter are high by comparison with other guides, since the power dissipated in each element (which depends on R) is high compared with the power flowing through it (which is determined by Z_{0M}).

Efforts have been made to reduce losses, and for a while the best performance quoted was the 0.12 dB/element obtained at 150 MHz using printed circuit board elements arranged in an axial configuration but based on double-sided spirals to maximize κ.[32] Figure 11.4(a) shows the configuration, which has a resonant transducer coupled to the input and output of the guide and designed

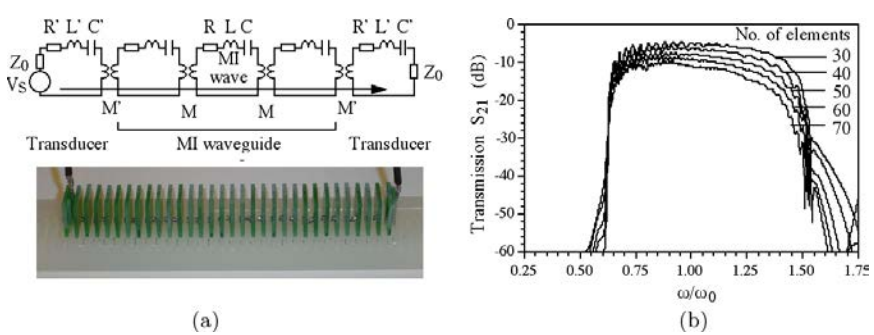


Fig. 11.4. (a) MI link and (b) measured performance.

to match the impedance Z_{0M} to a signal source V_S with output impedance $Z_0 = 50\Omega$ and to a similar load.

Matching requires the inductance L' and capacitance C' of the transducers to satisfy $1/\sqrt{(L'C')} = \omega_0$, and the mutual inductance M' at the terminations to satisfy

$$\omega_0 M' = \sqrt{(Z_0 Z_{0M})}. \tag{11.5}$$

Exact matching is only obtained at resonance, but performance can be surprisingly good. For example, Fig. 11.4(b) shows the measured response with different numbers of elements (in each case, with $Q = 110$ and $\kappa = 0.76$), which shows low-loss transmission over the MI band. However, it is important to note that with the period used ($a = 2.5\,\mathrm{mm}$) the propagation distances are very short (0.175 m with 70 elements).

A further problem is unwanted coupling. The fields are generated by each element spread, leading to magnetic coupling between second and higher neighbors. In an equivalent circuit, this can be represented by (for example) a mutual inductance M_2 between second-neighbors in addition to the desired term M_1 between nearest neighbors (Fig. 11.5(a)), and leads to additional evanescent waves at the same frequency as the MI wave.[33] The effects can be reduced by careful design (in the axial configuration, by using elements that are thick, but closely spaced).

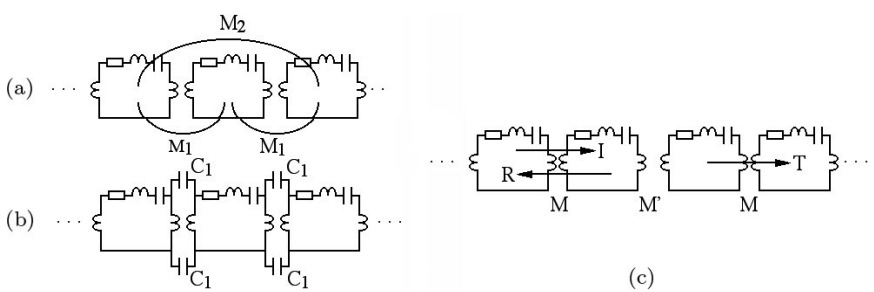


Fig. 11.5. Unwanted effects represented using equivalent circuits: (a) non-nearest neighbor magnetic coupling, (b) electric coupling and (c) reflections at a discontinuity.

Retardation, which leads to phase shifts in the coupled currents, also allows additional modes of propagation.[34] However, this effect occurs at high frequency. The elements may also be coupled electrically[35]; this effect again occurs at high frequency in small, widely spaced elements, but also at low frequency in large, closely spaced elements. This coupling may be represented using capacitors C_1 between nearest neighbors (Fig. 11.5(b)) and gives rise to extra bands above the MI band.

Discontinuities — variations in the resonant frequency ω_0 from element-to-element, or in the coupling coefficient κ between elements — cause further difficulties. An incident wave I will give rise to a reflected wave R in addition to the desired transmitted wave T (Fig. 11.5(c)) with a reflection coefficient that varies slowly across the MI band.[36] However, multiple discontinuities — for example, two consecutive variations in κ — cause multiple reflections, and coherent summation results in Fabry–Perot interference. For rigid elements, it is simple to minimize variations in ω_0, and straight guides may be constructed with a well-controlled spacing as previously shown in Fig. 11.4(a). For the planar geometry, curved guides can easily be defined. However, the same cannot be done in the axial geometry; in this case, bends inevitably degrade performance, introducing rapid variations in transmission and reflection.[37]

These difficulties have spurred the development of a flexible variant, MI cable,[38] which has the inductance and capacitance of resonant elements subdivided into two parts with values $L/2$ and $2C$ (Fig. 11.6(a)). This geometry allows capacitor plates to be formed on

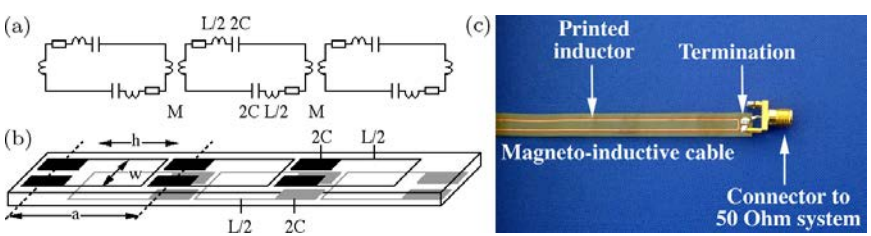
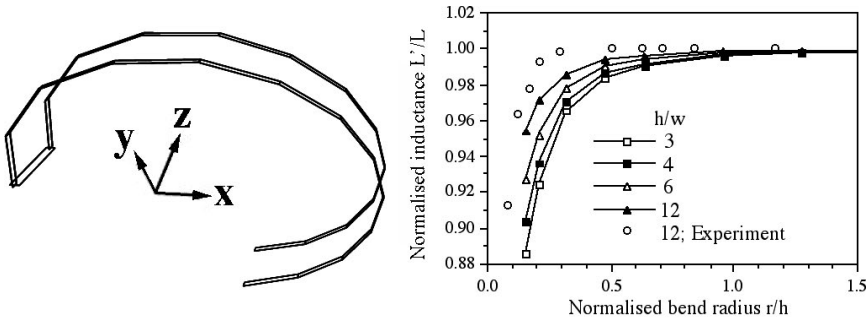


Fig. 11.6. (a) Equivalent circuit and (b) construction of MI cable; (c) completed cable.

either side of a thin dielectric, and linked by inductors that can be overlaid with those of neighboring elements (Fig. 11.6(b)). The entire structure can be formed using patterning and etching of copper-clad polyimide, in much longer lengths. Cables up to 2 m long with periods of $a = 10\,\text{cm}$ have been made with $35\,\mu\text{m}$ thick Cu on $25\,\mu\text{m}$ thick Kapton, using varying inductor widths w and lengths h to control ω_0 and Z_{0M}. As before, a resonant transducer can be used for connection to a 50Ω system (Fig. 11.6(c)). If Z_{0M} corresponds to this value, matching can be achieved at ω_0 using an element with inductance and capacitance L and C. However, improved matching can be achieved with values $L/2$ and $2C$, which can be provided by a halved cable element. This arrangement gives exact matching at two frequencies, ω_0 and $\omega_0/\sqrt{(1-\kappa^2)}$,[39] and more sophisticated transducers have since been developed.[40]

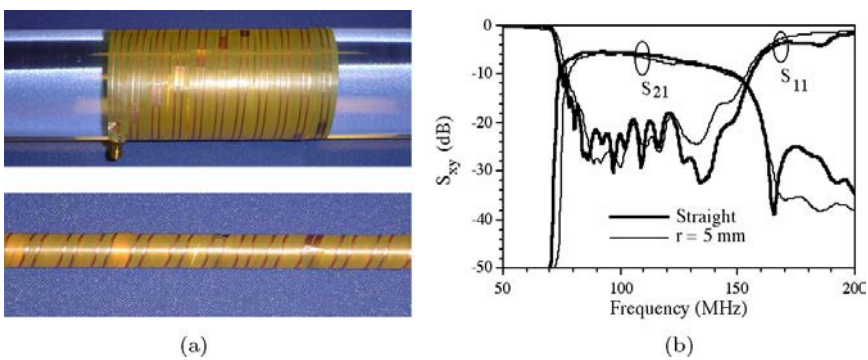
Neighboring elements must be offset to reduce electric coupling via the capacitance between the tracks.[41] However, with this proviso, the overall arrangement has several advantages. High coupling with a positive coefficient (typically, $\kappa > 0.65$) can be achieved, with minimal second neighbor coupling. Losses per element are low, but the geometry is planar, and the period is approximately half the element length. Consequently, losses per meter can also be low; 2.3 dB per meter was demonstrated using elements with $h = 4.7\,\text{mm}$ and $a = 10\,\text{cm}$ operating at 130 MHz.[38] Even lower figures have been measured at Imperial with much larger elements operating at lower frequency (0.5 dB/meter with $h = 40\,\text{mm}$, $a = 600\,\text{mm}$, $f_0 = 20\,\text{MHz}$). Accompanying low loss is low noise. Resistance in the elements results in Johnson noise, which also propagates as MI waves[42] and has been measured as standing waves in isolated arrays.[43] The noise factor of MI cable has recently been measured and was found to be excellent.[44]

The element shape confers further benefits. Simulation shows that a long, thin inductor can be bent through an arc of a circle of radius r (Fig. 11.7(a)) without a significant change in its inductance L', provided r does not fall much below the track separation h (Fig. 11.7(b)). Because the elements are held together by the substrate, the mutual inductance M is also largely unaffected. MI cable



(a) (b)

Fig. 11.7. (a) CAD model of a curved inductor, and (b) variation of inductance with radius.



(a) (b)

Fig. 11.8. (a) Spiralled MI cable; (b) frequency dependence of scattering parameters for straight and spiralled cable.

can therefore be bent through curved paths (Fig. 11.8(a)) without changing ω_0 or κ, and hence without reflections.[37]

As a result, the transmission characteristics of waveguides spiralled with a bend radius as small as a 5 mm are similar to those of a straight guide (Fig. 11.8(b)). The only significant drawback is that the unshielded nature of MI cables allows cross-talk between adjacent guides[45]; however, this characteristic is advantageous for signal detection.

Additional components are needed in a complete system. Many passive MI components, including mirrors, Bragg gratings, Fabry–Perot interferometers, Y-junction splitters and directional couplers have been investigated theoretically.[26,36] Generally, these have involved discrete elements. However, because of the poor performance in this format, few experiments have been performed; the exceptions are detailed investigations of directional couplers.[46] Several cable-based components have been demonstrated, including splitters,[47] phase shifters and Mach–Zehnder interferometers.[48] Active systems have also been proposed to overcome losses; two-frequency parametric amplification has been analyzed theoretically,[49] and three-frequency amplification[50] and negative impedance converters[51] have been demonstrated experimentally. However, such systems are relatively complex.

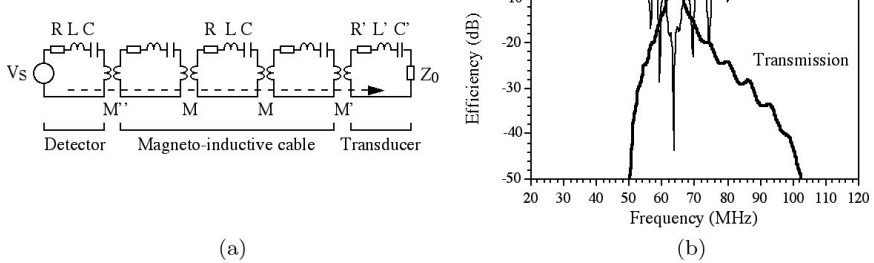
MI waves have been considered for use in low-frequency communications[52] (particularly in underground,[53] underwater[54] or challenged environments[55]). Systems for contact-less power transfer[56] and combined communications and power[57] have also been described. Distributed sensing has also been considered, exploiting impedance discontinuities caused by movement to monitor structural health,[58] or changes in loading to measure conductivity.[59] However, MRI appears to offer the greatest potential. For example, arrays of "Swiss rolls"[60] and wires[61] have been proposed for near-field image transfer. Applications involving MI waves have included RF flux concentrators based on coupled loops of decreasing size,[62] planar imaging lenses based on sheets of coupled loops[63] and bird-cage coils based on resonant rings.[64] In each case, advantages have been claimed. However, some systems have required fundamental changes to MRI (for example, parallel imaging in place of frequency-coded serial imaging), while others have been too bulky, or yielded only modest improvements in SNR. As a result, the advantages have not been overwhelming. However, the realization that magnetic segmentation of a signal-carrying path can provide intrinsic safety[65] in an internal imaging system with an inherent gain in SNR may now be

altering the balance. Such systems can be based on MI cable, as we now show.

11.4. Magneto-Inductive Catheter Receivers

A length of MI cable can be converted into an RF receiver for internal MRI as follows.[66] The resonant elements must first be tuned to the Larmor frequency. A voltage V_S may then be induced by precessing dipoles at in a detecting element at one end of a waveguide (Fig. 11.9(a)). The element shape — a long, thin rectangle — results in a detection sensitivity that falls off as $1/r^2$, where r is the radial distance from the axis. Consequently, the sensitivity is non-uniform and the FOV restricted to a multiple of the track spacing h. The induced signal may then be transmitted as an MI wave to the scanner input at the other. Resonant elements located part way along the cable may of course act as detectors as well. In this case, the induced voltage will generate two MI waves, one of which travels to the scanner input, and the other to the detector, where it is absorbed.

For efficient transmission, the output transducer should satisfy the previous matching condition $\omega_0 M' = \sqrt{(Z_0 Z_{0M})}$. The mutual inductance M'' at the detector (which for light loads has an output impedance at resonance equal to the loop resistance R) should satisfy a modified condition $\omega_0 M'' = \sqrt{(R Z_{0M})}$. Although Z_{0M} may be comparable to Z_0, R is generally much smaller. As a result, matching



(a) (b)

Fig. 11.9. (a) Equivalent circuit model of catheter receiver with signal injected at tip; (b) simulation of receiver performance.

is narrow band, and the frequency variation of sensitivity is a composite of a resonant response near ω_0 and a band response elsewhere (Fig. 11.9(b)). Here, parameters comparable to those of experimental devices ($Q = 40$, $Z_{0M} = 40\Omega$ and $\kappa = 0.63$) have been assumed.

The entire circuit can be realized in thin-film form, by careful design of the resonant elements. Protection against coupling to external E fields can be obtained by making the element length $2a$ less than the resonant length d of an element in tissue. Protection against uniform B_1 fields can also be provided, by twisting the shape of each element into a figure-of-eight so the signals induced in each half of the loop cancel (Fig. 11.10(a)). Long, thin capacitors can be used to increase flexibility. Impedance matching to the detector element can be achieved by adjusting the overlay b between the waveguide and the final detector (Fig. 11.10(b)). Thin-film circuits can be produced in batches, as a closely spaced array. Individual circuits can then be separated with a scalpel and mounted on a hollow catheter scaffold with heat-shrink tubing so the long coil conductors lie on a diameter (Fig. 11.10(c)). The internal lumen acts as a passage for a guidewire during clinical use and allows construction to be carried out on a wire rail. If the resonant element connected to the MRI system is constructed from a separate thin-film circuit, the output transducer can be demountable (Fig. 11.10(d)).

Prototype receivers have been constructed in batches of 24. Each receiver was approximately 1.6 m long, and formed from 15 resonant

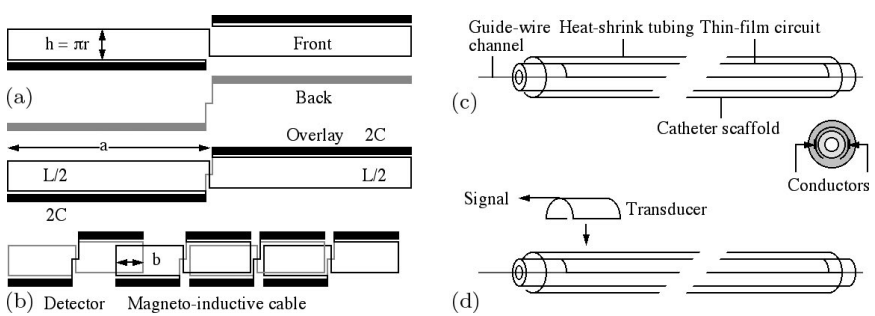


Fig. 11.10. Layout of (a) resonant elements and (b) complete receiver; (c) catheter construction; (d) transducer operation.

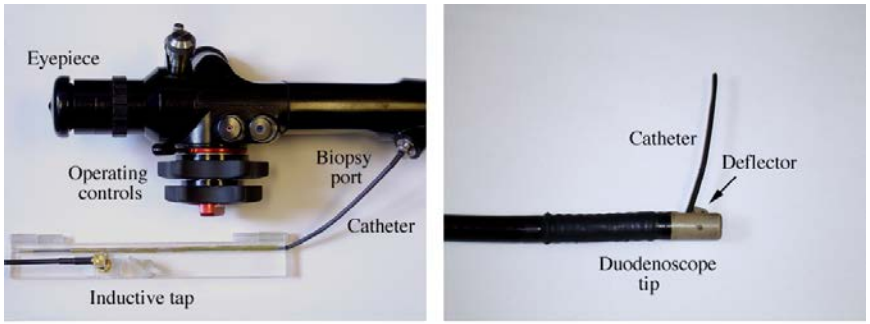


Fig. 11.11. Use with a duodenoscope: (a) insertion into biopsy channel, (b) exit at tip.

elements with a period of $a = 100\,\text{mm}$. The element length was then 0.2 m, significantly less than the critical value of 0.27 m described earlier. The scaffold was a 2.25 mm diameter PTFE tube, and the heat-shrink was a thin-walled polyolefin. The transducer was a two-turn thin-film inductor, which was epoxied onto the inside of a split Perspex clamp attached to the outside of the catheter. The result was a low-cost, disposable receiver that can be inserted into the biopsy channel of a non-magnetic duodenoscope (Fig. 11.11(a)) and emerge from the side-port at the distal tip (Fig. 11.11(b)).[67]

Experimental measurement of individual and paired elements suggested a Q-factor of 30–40, a characteristic impedance at resonance of 40 Ω, a coupling coefficient of 0.63 and a propagation loss of 4.5 dB/m.

In practice, overlay of the elements implies that signal detection must be more complicated than suggested by Fig. 11.9(a). Depending on its position, an external dipole (which can be represented as a resonant loop with a voltage source in an equivalent circuit model) can couple to more than one element (Fig. 11.12(a)). As a result, several MI waves will normally be generated, and the signal reaching the receiver is their coherent sum. These effects may easily be modelled[66]; simulation of the transmission between a source at different positions along the receiver and the load shows that the

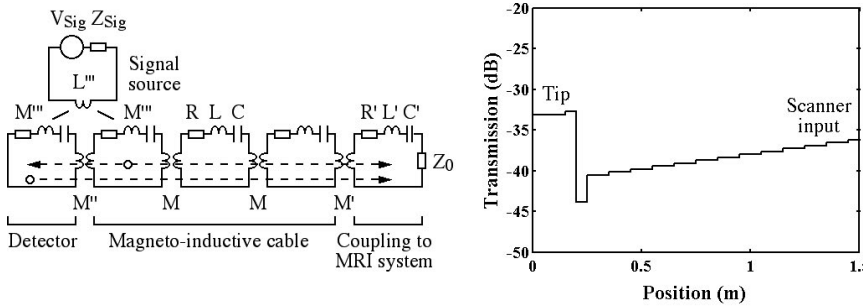


Fig. 11.12. (a) Equivalent circuit model of external signal detection; (b) simulated axial variation in sensitivity.

sensitivity must be high at the resonant tip, falls at the junction to the cable, and then gradually rises towards the load (Fig. 11.12(b)). The receiver will therefore detect MRI signals along its whole length, but the brightest images will be at the tip.

Experiments to demonstrate ^{1}H MRI and verify intrinsic safety with prototype catheter receivers were performed using a 1.5 T GE Signa Excite clinical scanner at St Mary's Hospital, London. Imaging was carried out *in vitro* using phantoms filled with a solution doped with NaCl and NiC_2 to obtain relaxation time constants T1 and T2 similar to those of tissue. A variety of gradient-recalled echo (GRE) and 2D spin echo (SE) sequences were used for imaging, while more specialized sequences were used to investigate the possibility of RF-induced heating.

Decoupling from external B_1 fields was investigated by completely immersing individual catheter-mounted resonant elements in signal source (Fig. 11.13(a)). This arrangement was then excited and imaged using the system body coil. The resulting sagittal slice image (Fig. 11.13(b)) shows little perturbation to the magnetization pattern, implying that the figure-of-eight layout has indeed largely eliminated coupling to the B_1 field. Similar experiments were carried out using unmounted elements, held flat against a cuboid phantom filled with the same signal source, or bent through increasing angles θ (Fig. 11.13(c)). The resulting axial MR images (Fig. 11.13(d)) show

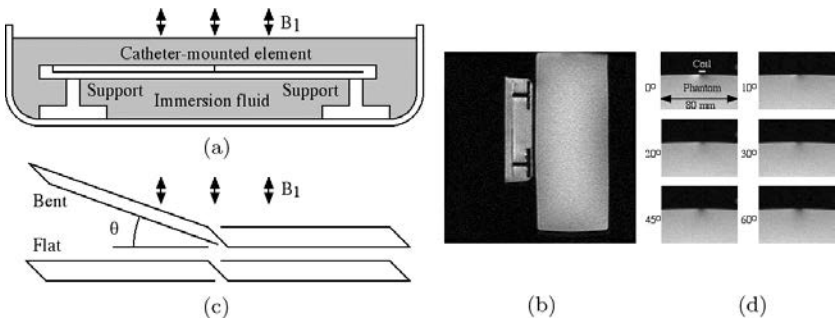


Fig. 11.13. Arrangements for investigating B_1 field coupling to MI cable with (a) mounted and (b) unmounted elements; (c) and (d) corresponding body coil images.

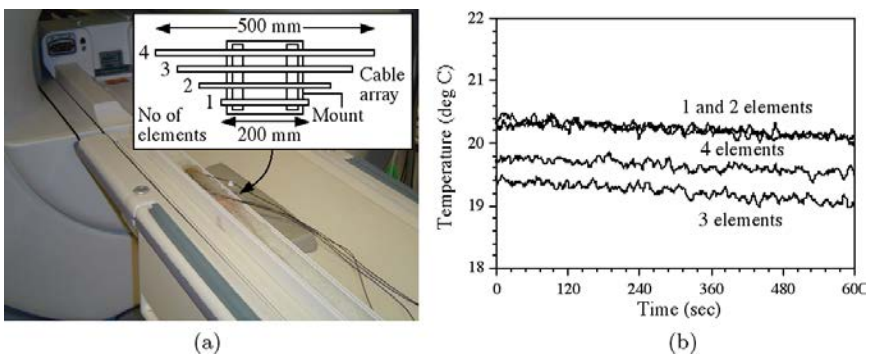
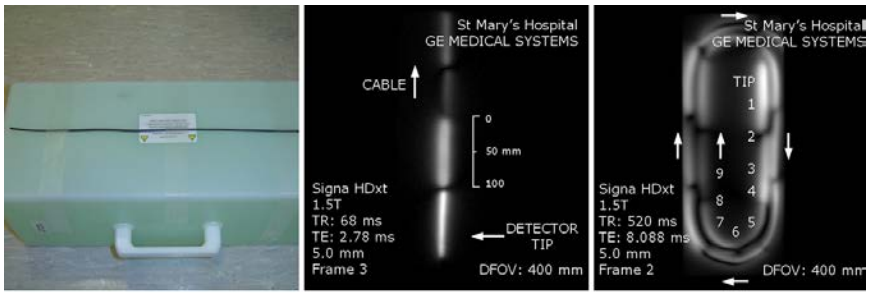


Fig. 11.14. (a) Arrangements for investigating E field coupling; (b) variation of temperature with time turing RF heating experiments.

that there is little change in the magnetization beneath the element with θ, implying that the figure-of-eight arrangement is still effective even when it is abruptly bent.

Decoupling from external E fields was investigated by immersing arrays of catheter-mounted cables containing different numbers of resonant elements in a doped polyacrylic gel whose dielectric properties simulated those of the tissue (Fig. 11.14(a)). Temperature measurements made using fiber-optic thermometers at the midpoints and ends of the cables confirmed the absence of any temperature rise, even when the coil array was placed close to the body coil and when



(a) (b) (c)

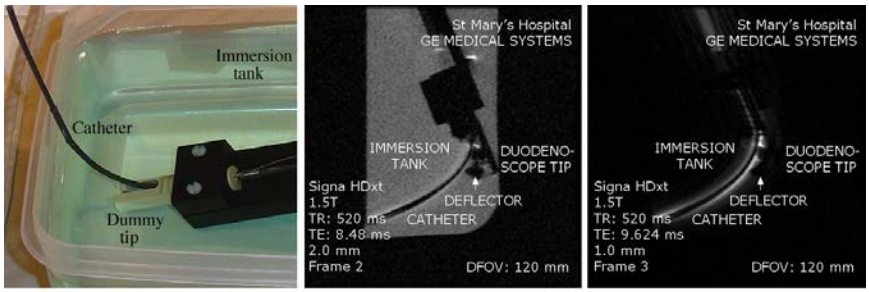
Fig. 11.15. (a) Arrangement for imaging, (b) and (c) coronal images of straight and spiral tracks.

excitation sequences involving high RF power were used for long periods (Fig. 11.14(b)). These results have been confirmed by numerical simulation of thin-film circuits.[68]

Imaging experiments were performed with a complete receiver on a cuboid phantom (Fig. 11.15(a)). The resulting coronal slice image (Fig. 11.15(b)) shows that signals are only detected from regions in the immediate vicinity, in discrete segments corresponding to the loops of the figure-of-eight elements. The image is clearly brightest at the resonant tip. However, the receiver does form an image along its entire length, and functions correctly even when its track is bent into a spiral (Fig. 11.15(c)).

Imaging was also carried out with a dummy component representing the tip of a non-magnetic duodenoscope,[69] using a tank of immersion fluid as a signal source (Fig. 11.16(a)). Sagittal images obtained using the body coil with the receiver in place (Fig. 11.16(b)) show little perturbation to the magnetization. Similarly, local images were successfully acquired using the catheter receiver even when its track was bent through 90° to simulate cannulation during ERCP (Fig. 11.16(c)).

Catheter receivers have also been modified to operate in higher magnetic fields.[70] Scaling from 1.5 T to 3 T implies that the Larmor frequency ω_L must double, and also (provided the dielectric constant of tissue is slowly varying) that the critical length d of an



(a) (b) (c)

Fig. 11.16. (a) Catheter receiver in dummy duodenoscope; (b) and (c) sagittal images obtained with the body coil and catheter receiver.

immersed element must halve. Conveniently, halving the period of an MI waveguide to satisfy this new requirement will double the resonant frequency of the elements, since the process reduces the capacitance C and the inductance L by a factor of two. Using this simple scaling, similar frequency variations in detection sensitivity were obtained from catheter receivers operating at 1.5 T and 3 T.

In vitro ^{1}H MRI with prototype 3 T catheter receivers was performed using a GE Discovery MR750 scanner, again at St Mary's Hospital. Comparative measurements of the SNR achievable with an eight-element chest coil array and a catheter receiver when both are loaded with cuboid phantoms designed to mimic the human torso (Fig. 11.17(a)) highlight the advantage of a restricted FOV. The spatial variation of SNR obtained in the array is almost independent of position, and it provides approximately constant brightness in axial images (Fig. 11.17(b)). However, the SNR is relatively low (here, around 100). In contrast, the SNR achieved using the catheter receiver with the same sequence varies as $1/r^2$, where r is the distance from the catheter axis. Correction for this variation can be provided using software. Importantly, the SNR is highest near the catheter, and hence should allow high-resolution imaging of duct walls.

The SNR peaks at around 500. As a result, the catheter receiver can outperform the array coil for radii less than 20 mm — which represents the useful FOV — but is clearly less effective outside this

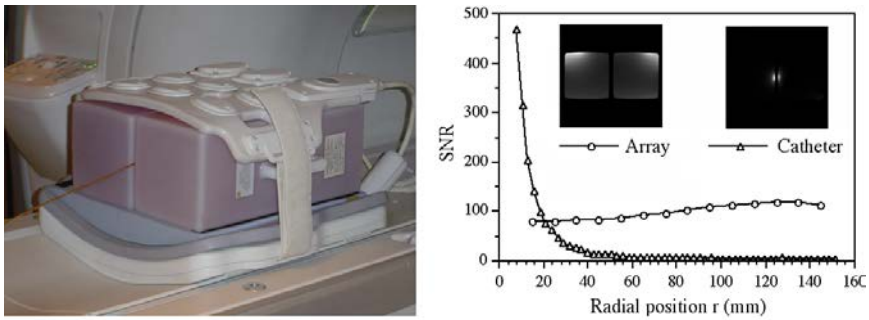


Fig. 11.17. (a) Arrangement for comparison of catheter receiver with 3 T surface coil array; (b) axial images and spatial variations in SNR.

range. The peak SNR using the catheter receiver is around five times larger. This increase could in principle be recovered using the array by averaging the results of multiple signal acquisitions. However, the statistics of signal averaging imply that the number needed is around $5^2 = 25$, prohibitively large. Not only would the duration of the MRI examination increase drastically, but the advantage in SNR gained by averaging would be nullified by the effects of patient motion; organs such as the liver are significantly displaced and distorted by the expansion of the lungs in breathing. Thus, provided it can be safely delivered into the body, the metamaterial design appears to offer a performance that conventional approaches struggle to match. Considerable further work is required before this can be verified *in vivo*, but a program of early stage trials at Khon Kaen University Hospital is now in place.

11.5. Conclusions

Catheter-based receivers for internal MRI are now showing considerable promise. The receivers are based on a particular type of metamaterial, the MI waveguide, which consists of a linear array of magnetically coupled L-C resonators. MI waveguides can be realized as a thin-film cable by double-sided patterning of copper-clad Kapton. MI cable is highly flexible and can easily be mounted on a

catheter, but has low propagation loss and is relatively immune to bending. Segmentation into resonant elements with a short, figure-of-eight shape prevents direct coupling to electric and magnetic RF fields during the excitation phase of MRI, and hence provides intrinsic patient safety. The receiver has a sensitive tip, but can detect an image along its entire length in sections that match the cable divisions. Its limited FOV reduces the mount of body noise detected. As a result, the receiver can provide a gain in SNR compared to an external detector, but only over a limited volume. If this matches the target tissue, improved performance can be obtained. The proposed application is imaging of the biliary ductal system, where very high resolution is required for diagnosis and staging of CCA. The receivers have been shown to be compatible with the side-opening duodenoscopes used for introduction of catheters into the bile ducts. Early stage trials are now taking place in Khon Kaen University Hospital, Thailand, where CCA is epidemic due to liver fluke infestation.

Acknowledgments

The authors gratefully acknowledge the contributions of their colleagues Dr Munir Ahmad, Dr Timmy Floume, Dr Marc Rea, Dr Khoonsake Segkhoonthod, Dr Chris Wadsworth, Prof Katya Shamonina, Dr Oleksiy Sydoruk, Prof. Simon Taylor-Robinson and Dr Mike Wiltshire.

References

1. Pendry, J. B., Holden, A. J., Robbins, D. J. and Stewart, W. J. (1999). *IEEE Trans. Micr. Theor. Tech.* **47**, p. 2075.
2. Veselago, V. G. (1968). *Sov. Phys. Usp.* **10**, p. 509.
3. Caloz, C. and Itoh, T. (2005). *Electromagnetic Metamaterials: Transmission Line Theory and Microwave Applications* (John Wiley and Sons, New York).
4. Engheta, N. and Ziolkowski, R. W. (2006). *Metamaterials: Physics and Engineering Explorations* (John Wiley and Sons, New York).
5. Solymar, L. and Shamonina, E. (2009). *Waves in Metamaterials* (Oxford University Press, Oxford).
6. Nakeeb, A., Pitt, H., Sohn, T. *et al.* (1996). *Ann. Surg.* **224**, p. 463.
7. Khan, S., Taylor-Robinson, S., Toledano, M. *et al.* (2002). *J. Hepatol.* **37**, p. 806.

8. Sriamporn, S., Pisani, P., Pipitgool V. *et al.* (2004). *Trop. Med. Int. Health* **9**, p. 588.
9. Hashmonai, M., Lev, L., Schramek, A. *et al.* (1980). *J. Surg. Oncol.* **13**, p. 231.
10. Baillie, J., Paulson, E. K. and Vitellas, K. M. (2003). *Gastroenterology* **124**, p. 1686.
11. Gaschen, L., Kircher, P. and Lang, J. (2003). *Ultrasound* **44**, p. 665.
12. Inui, K. and Miyoshi, H. (2005). *Gastrointest. Endosc.* **15**, p. 143.
13. McRobbie, D. W., Moore, E. A., Graves, M. J. and Prince, M. R. (2006). *MRI from Proton to Picture* (Cambridge University Press, Cambridge).
14. Manfredi, R., Barbaro, B., Masseli, G. *et al.* (2004). *Semin Liver Dis.* **24**, p. 155.
15. Hoult, D. I. and Lauterbur, P. C. (1979). *J. Magn. Reson.* **34**, p. 425.
16. Celik, H., Eryaman, Y., Altinas, A. *et al.* (2004). *Mag. Res. Med.* **52**, p. 640.
17. Kantor, H. L., Briggs, R. W. and Balaban, R. S. (1984). *Circ. Res.* **55**, p. 261.
18. Quick, H. H., Ladd, M. E., Zimmermann-Paul, G. G. *et al.* (1999). *Mag. Res. Med.* **41**, p. 751.
19. Weiss, C. R., Georgiades, C., Hofmann, L. V. *et al.* (2006). *J. Vasc. Interv. Radiol.* **17**, p. 845.
20. Konings, M. K., Bartels, L. W., Smits, H. F. M. *et al.* (2000). *J. Magn. Reson. Imag.* **12**, p. 79.
21. Park, S. M., Kamondetdacha, R., Amjad, A. *et al.* (2005). *IEEE Trans. Magn.* **41**, p. 4197.
22. Ladd, M. E. and Quick, H. H. (2000). *Magn. Reson. Med.* **43**, p. 615.
23. Weiss, S., Vernickel, P., Schaeffter, T. *et al.* (2005). *Magn. Reson. Med.* **54**, p. 182.
24. Shamonina, E., Kalinin, V. A., Ringhofer, K. H. and Solymar, L. (2002). *Electron. Lett.* **38**, p. 371.
25. Shamonina, E., Kalinin, V. A., Ringhofer, K. H. and Solymar, L. (2002). *J. Appl. Phys.* **92**, p. 6252.
26. Shamonina, E. and Solymar, L. (2004). *J. Phys. D. Appl. Phys.* **37**, p. 362.
27. Wiltshire, M. C. K., Shamonina, E., Young, I. R. and Solymar, L. (2003). *Elect. Lett.* **39**, p. 215.
28. Wiltshire, M. C. K., Shamonina, E., Young, I. R. and Solymar, L. (2004). *J. Appl. Phys.* **95**, p. 4488.
29. Maslovski, S., Ikonen, P., Kolmakov, I. and Tretyakov, S. (2005). *PIER* **54**, p. 61.
30. Shadrivov, I. V., Reznik, A. N. and Kivshar, Y. S. (2007). *Physica B* **394**, p. 180.
31. Dolling, G., Wegener, M., Schädle, A. *et al.* (2006). *Appl. Phys. Lett.* **89**, p. 231118.
32. Syms, R. R. A., Young, I. R. and Solymar, L. (2006). *J. Phys. D. Appl. Phys.* **39**, p. 3945.

33. Syms, R. R. A., Sydoruk, O., Shamonina, E. and Solymar, L. (2007). *Metamaterials* **1**, p. 44.
34. Radkovskaya, A., Shamonin, M., Stevens, C. J. *et al.* (2006). *J. Magnetism Magn. Mats.* **300**, p. 29.
35. Hesmer, F., Tatartschuk, E., Zhuromskyy O. *et al.* (2007). *Phys. Stat. Sol. (b)* **244**, p. 1170.
36. Syms, R. R. A., Shamonina, E. and Solymar, L. (2006). *IEE Proc. Micr. Ant. Propag.* **153**, p. 111.
37. Syms, R. R. A. and Solymar, L. (2010). *Metamaterials* **4**, p. 161.
38. Syms, R. R. A., Young, I. R., Solymar, L. and Floume, T. (2010). *J. Phys. D. Appl. Phys.* **43**, p. 055102.
39. Syms, R. R. A., Solymar, L. and Young, I. R. (2010). *J. Phys. D. Appl. Phys.* **43**, p. 285003.
40. Sydoruk, O. (2011). *Metamaterials* **5**, p. 42.
41. Syms, R. R. A. and Floume, T. (2013). *Proc. 7th Int. Cong. on Advanced Electromagnetic Metamaterials in Microwaves and Optics*, September 16–21, Bordeaux, p. 148.
42. Syms, R. R. A. and Solymar, L. (2011). *J. Appl. Phys.* **109**, p. 124909.
43. Wiltshire, M. C. K. and Syms, R. R. A. (2014). *J. Appl. Phys.* **115**, p. 084905.
44. Wiltshire, M. C. K. and Syms, R. R. A. (2014). *J. Appl. Phys.* **116**, p. 034503.
45. Syms, R. R. A. and Solymar, L. (2011). *J. Appl. Phys.* **109**, p. 044902.
46. Sydoruk, O., Radkovskaya, A., Zhuromskyy, O. *et al.* (2006). *Phys. Rev. B* **73**, p. 224406.
47. Sydoruk, O. (2011). *Elect. Lett.* **47**, p. 549.
48. Syms, R. R. A. and Solymar, L. (2011). *Metamaterials* **5**, p. 155.
49. Sydoruk, O., Shamonina, E. and Solymar, L. (2007). *J. Phys. D. Appl. Phys.* **40**, p. 6879.
50. Syms, R. R. A., Young, I. R. and Solymar, L. (2008). *Metamaterials* **2**, p. 122.
51. Rajab, K. Z., Fan, Y. F. and Hao, Y. (2012). *J. Opt.* **14**, p. 114004.
52. Stevens, C. J., Chan, C. W. T., Stamatis, K. and Edwards, D. J. (2010). *IEEE Trans. Micr. Theory Tech.* **58**, p. 1248.
53. Sun, Z. and Akyildiz, I. F. (2010). *IEEE Trans. Antennas Propag.* **58**, p. 2426.
54. Gulbahar, B. and Akan, O. B. (2012). *IEEE Trans. Wireless Comm.* **11**, p. 3326.
55. Sun, Z., Aykildiz, I. F. and Kisseleff, S. (2013). *IEEE Trans. Comm.* **61**, p. 3943.
56. Zhong, W., Lee, C. K. and Hui, S. Y. R. (2013). *IEEE Trans. Indust. Electr.* **60**, p. 261.
57. Agbinya, J. I. (2013). *PIER C* **37**, p. 15.
58. Chen, Y., Munukutla, S., Pasupathy, P. *et al.* (2010). *SPIE Proc.* **7647**, p. 764749.
59. Floume, T. (2011). *Metamaterials* **5**, p. 206.
60. Wiltshire, M. C. K., Hajnal, J. V., Pendry, J. B. *et al.* (2003). *Opt. Expr.* **11**, p. 709.

61. Radu, X., Garray, D. and Craeye, C. (2009). *Metamaterials* **3**, p. 90.
62. Wiltshire, M. C. K., Shamonina, E., Solymar, L. and Young, I. R. (2004). *Proc. 12th Ann. Meet. of ISMRM, Kyoto*, p. 1582.
63. Freire, M. J., Marques, R. and Jelinek, L. (2008). *Appl. Phys. Lett.* **93**, p. 231108.
64. Syms, R. R. A., Floume, T., Young, I. R. *et al.* (2010). *Metamaterials* **4**, p. 1.
65. Syms, R. R. A., Solymar, L. and Young, I. R. (2010). *IEEE J. Sel. Top. Quant. Elect.* **16**, p. 433.
66. Syms, R. R. A., Young, I. R., Ahmad, M. M. and Rea, M. (2012). *J. Appl. Phys.* **112**, p. 114911.
67. Syms, R. R. A., Young, I. R., Ahmad, M. M. *et al.* (2013). *IEEE Trans. Biomed. Engng.* **60**, p. 2421.
68. Segkhoonthod, K., Syms, R. R. A. and Young, I. R. (2014). *IEEE Sensors J.* **14**, p. 1505.
69. Syms, R. R. A., Young, I. R., Ahmad, M. M. *et al.* (2013). *Proc. 7th Int. Cong. on Advanced Electromagnetic Metamaterials in Microwaves and Optics, Sept. 16–21, Bordeaux*, p. 25.
70. Syms, R. R. A., Syms, I. R., Young, I. R. and Rea, M. (2013). *Proc. 17th Transducers Conf., Barcelona, June 16–20*, M3P p. 147.

CHAPTER 12

Microwave Sensors Based on Symmetry Properties and Metamaterial Concepts

JORDI NAQUI*, ALI KARAMI HORESTANI†, CHRISTOPHE FUMEAUX†,‡ and FERRAN MARTÍN*,§

*Universitat Autònoma de Barcelona, Spain
†The University of Adelaide, Australia

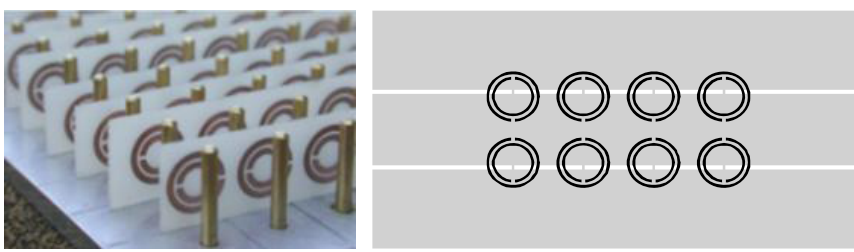
Microwave sensors based on the symmetry properties of transmission lines loaded with metamaterial resonators are reviewed in this chapter. In most microwave sensors based on resonant elements, the physical variable to be measured modifies the resonance frequency, phase or quality factor of the sensing resonant structure. In this chapter, a novel sensing principle, based on the disruption of symmetry, is studied. The proposed sensors are implemented by loading a transmission line either with symmetric resonators (typically, although not exclusively, resonant elements useful for the implementation of metamaterials) or with symmetric configurations of resonator pairs. In the unperturbed state, the whole structure (line and resonator/s) is symmetric, and it is designed to exhibit

Corresponding authors: ‡christophe.fumeaux@adelaide.edu.au
§ferran.martin@uab.cat

either an all-pass behavior (Type I sensors) or a single transmission zero (Type II sensors). Conversely, when symmetry is broken by the physical effect to be sensed (e.g. a linear or angular displacement, dielectric loading, etc.), a notch appears in Type I sensors, whereas a split-off (i.e. two transmission zeros) emerges in Type II sensors. Hence, Type I and Type II sensors can be designated as resonance-based and frequency-splitting sensors, respectively. These sensors are of special interest as comparators and are robust against changes in environmental conditions. Potential applications include contactless linear and angular displacement and velocity sensors, alignment sensors, permittivity sensors, etc.

12.1. Introduction

Transmission line metamaterials based on split-ring resonators (SRRs) were first proposed in Ref. 1, where it was demonstrated that a coplanar waveguide (CPW) transmission line loaded with pairs of SRRs (etched on the back substrate side) and shunt inductive strips mimic the first left-handed (LH) bulk metamaterial, reported in 2000[2] (see Fig. 12.1). In the artificial medium of Fig. 12.1(a), left-handedness is manifested as long as the incident radiation is polarized with the electric field parallel to the metallic posts and the magnetic field axial to the SRRs. Under these conditions, the metallic posts exhibit a negative effective permittivity up to the so-called plasma frequency, which depends on the distance between posts and their



(a) (b)

Fig. 12.1. (a) Photograph (courtesy by D.R. Smith) of the first artificial LH medium, based on a combination of metallic posts and SRRs. (b) Layout of the first LH (actually CRLH) line based on a CPW loaded with shunt inductive strips and SRRs. In (b), the SRRs (in black color) are etched on the back substrate side.

radius,[3] and can be tailored and set to frequencies in the GHz range. The array of SRRs is responsible for the negative effective permeability, which takes place in a narrow band above the SRR fundamental resonance frequency.[2,4] If that frequency is set below the electric plasma frequency of the posts, a narrow band with simultaneous negative permittivity and permeability arises, and wave propagation is LH (or backward) in that band.[5] Despite the fact that the fields generated in a CPW transmission line are not uniform, the orientation and position of the shunt strips and pairs of SRRs make the structure of Fig. 12.1(b) to behave as a 1D LH medium.[a] Indeed, the structures of Figs. 12.1(a) and 12.1(b) can be described by identical circuit models,[1,6] i.e. a transmission line circuit model magnetically coupled to resonators and with shunt connected inductors.[b] From the analysis of the dispersion relation of such circuit model, it can be easily deduced that the structures of Fig. 12.1 exhibit LH behavior in a certain (narrow) band, and right-handed (RH)[c] behavior at higher frequencies.[9]

The band pass functionality of SRR/strip-loaded CPW transmission lines switches to a stop band (in the vicinity of the SRR resonance frequency) if the strips are removed. This stopband behavior can be interpreted as due to the negative effective permeability (related to the presence of the SRRs) above the SRR fundamental

[a]Note that the shunt strips are parallel-oriented to the electric field generated by the line, and a significant component of the magnetic field generated by the line on the back substrate side is axial to the SRRs.

[b]The former circuit model of a CPW loaded with SRRs and shunt strips was reported in Ref. 1, and revised in Ref. 7. The effect of coupling between resonators of adjacent elemental cells was studied in Ref. 8, whereas the interaction between the resonators belonging to the same unit cell was treated in Ref. 59.

[c]The term RH is used here as synonymous of "forward", to indicate that above the LH band there is a region where wave propagation if forward or RH (i.e. the propagation vector, the electric field intensity vector and the magnetic field intensity vector form a RH triplet, in contrast to the LH triplet that form such vectors when propagation is backward). Transmission lines exhibiting LH and RH wave propagation have been called composite right/left-handed (CRLH) lines.[9] Indeed, purely LH lines do not actually exist since at sufficiently high frequencies, the host line elements dominate over the loading elements (shunt strips and SRRs).

resonance and to the high positive permittivity (causing a strong mismatch) below that frequency. SRR-loaded lines have been applied to the design of notch and stopband filters,[10] where the resonance phenomenon of the loaded lines is exploited.

Although the genuine applications of metamaterial transmission lines concern the design of microwave components on the basis of impedance and dispersion engineering,[4,11–13d] transmission lines loaded *only* with electrically small resonators, such as SRRs or other resonators formerly used for the synthesis of metamaterials, are of interest in certain applications, including stopband and notch filters (as mentioned above) multiband printed dipole and monopole antennas,[14–16] multiband matching networks,[17] and microwave sensors (see for instance Ref. 18), among others. As long as the loading resonators of these lines are used for the synthesis of metamaterials, and the working principle of these lines is the controllability of the resonance phenomenon (rather than the controllability of the phase constant and characteristic impedance), such resonator-loaded lines can be designated as transmission lines with metamaterial loading.[13]

The majority of the microwave sensors studied in this chapter are based on transmission lines with metamaterial loading. Most of the considered sensors are implemented by using SRRs (or variations of the typical topology of this particle). The reasons that have led us to consider a specific resonator topology will be given in detail later. Microwave sensors based on the variation of the resonance frequency, phase or quality factor of resonator-loaded lines have been reported in the last few years.[18–26] These electrical parameters may be influenced by environmental changes (temperature, moisture, etc.),[27] and hence proper calibration is required, in general, before sensing (exceptions may be sensors based on differential measurements). Otherwise, measurements are likely to be affected by random errors as a result of changing ambient factors.

[d]To this end, CRLH lines, such as CPW transmission lines loaded with shunt strips and SRRs or many other lines loaded with different combinations of reactive elements, are required.

Alternatively, a novel sensing principle, where the symmetry properties of transmission lines loaded with electromagnetic resonators are exploited, has been recently pointed out.[28,29] The idea in this approach is to symmetrically load the transmission line with a single resonant element, which must be chosen and configured to achieve all-pass behavior (total transmission) in the unperturbed symmetric state, while creating a notch (transmission zero) when symmetry is disrupted. In these sensors, the notch magnitude varies with the level of asymmetry, and therefore these sensors can be identified as notch-depth sensors or resonance-based sensors.[e] The sensors based on this design principle are robust against variations of environmental conditions and especially suitable for alignment purposes, since alignment is a geometrical property invariant to these conditions. Prototypes of linear and angular displacement sensors[28–39] as well as angular velocity sensors[37,38] have been demonstrated. The purpose of this chapter is to describe the operation principle of these sensors and to report some examples. Nevertheless, some immunity to variable ambient conditions can also be attained using pairs of identical resonators. In this approach, the loaded line exhibits a single transmission zero in the unperturbed (symmetric) state, and a split-off (i.e. two notches) emerges when symmetry is broken. On the basis of this frequency-splitting principle, displacement sensors[40] and sensors for dielectric measurements[41] have been proposed.

The principle of operation of sensors based on symmetry rupture, namely resonance-based sensors and frequency-splitting sensors, is explained in detail in Section 6.2. In Section 6.3, some prototype sensors are reported, including linear and angular displacement sensors, rotation speed sensors and permittivity sensors. Advantages and disadvantages of these sensors, compared to other microwave sensors

[e]This nomenclature obeys to the fact that the resonance phenomenon appears when symmetry is broken. As will be shown, in the symmetric state, the resonator is not effectively coupled to the line and resonance is prevented. If symmetry is disrupted, the coupling level (and hence the notch magnitude) varies with the level of asymmetry.

based on different approaches, are also discussed. Finally, the main conclusions are highlighted in Section 6.4.

12.2. Microwave Sensors Based on Symmetry Properties of Resonator-Loaded Lines: Principle and Potential Configurations

In this section, the working principle of resonance-based and frequency-splitting sensors is presented, and different combinations of transmission lines and resonators, useful for the implementation of such sensors, are pointed out.

12.2.1. *Resonance-based sensors*

As will be shown, the principle of operation of resonance-based sensors relies on controlling a resonance by means of the coupling between a transmission line and a resonator. It is for this reason that these sensors may be said to be based on a coupling-modulated resonance[29] or on a coupling-controlled resonance. However, for simplicity, these sensors are referred to as resonance-based sensors throughout this chapter.

The most usual transmission lines for the implementation of microwave sensors based on symmetry properties are microstrip lines and CPWs. The 3D view of these lines and the field distributions for the fundamental modes are depicted in Fig. 12.2. Both lines exhibit a magnetic wall (H-wall) at the symmetry plane. Let us now assume that the line is loaded with a single symmetric and planar resonator, parallely oriented to the plane of the line (e.g. etched in the back substrate side of the CPW, or in the ground plane of the microstrip line, provided a slot resonator is considered). Typically, when a transmission line is loaded with a resonator coupled to it, the transmission coefficient exhibits a stopband (a transmission zero or notch) at the fundamental resonance frequency. Suppression of the signal transmission occurs because the injected power is reflected back to the source. However, it is possible to prevent resonance and the appearance of the associated notch in the transmission coefficient.

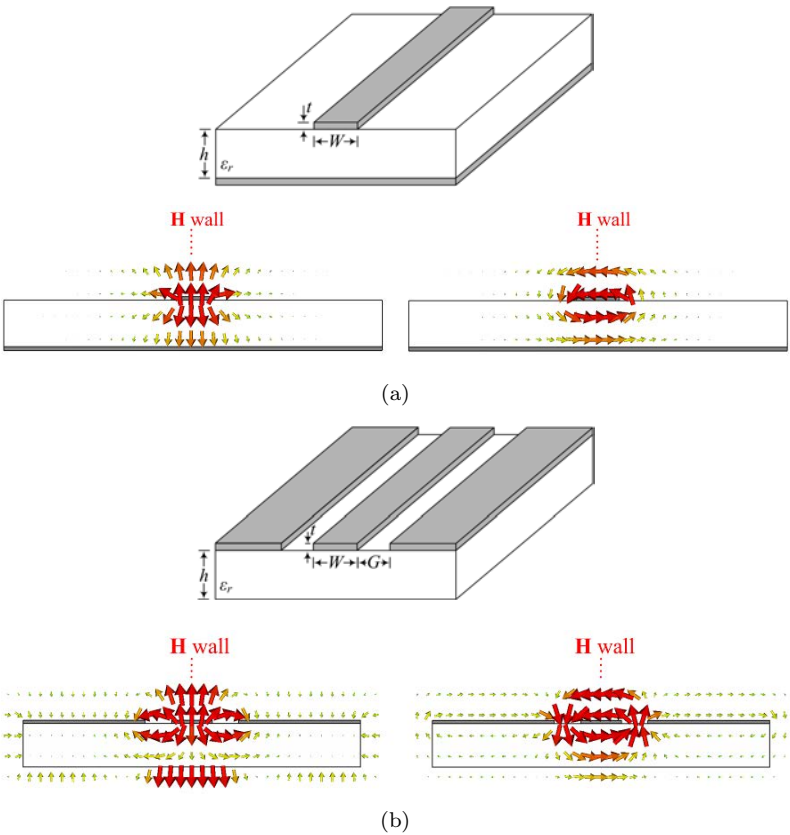


Fig. 12.2. Three-dimensional views and electric (left) and magnetic (right) field distributions of a microstrip line (a) and CPW (b) for the fundamental (even) modes.

The absence or presence of resonance is determined by the alignment of the symmetry planes of the line and the resonator as follows. Assume that the symmetry plane of the resonator is perfectly aligned with the symmetry plane of the transmission line, so that the structure is perfectly symmetric with regard to its longitudinal midplane. If the two symmetry planes are of the same electromagnetic nature (either electric or magnetic walls), the resonator is excited. The structure is resonant and signal transmission is suppressed in the vicinity of the resonance frequency, resulting in a stopband

response ($|S_{21}| \approx 0$). The resonator may be electrically, magnetically or magnetoelectrically coupled to the line. Conversely, if the two symmetry planes are of different electromagnetic nature (one being a magnetic wall and the other one an electric wall) the resonator cannot be excited. Because of perfect symmetry, the net electric and magnetic field components "illuminating" the resonator exactly cancel (the electric and magnetic fluxes over the resonator area are zero). Therefore, the resonance condition is inhibited because it cannot be established by either electric or magnetic coupling. Accordingly, the whole structure behaves as a transmission line exhibiting an all-pass response ($|S_{21}| \approx 1$).

To clarify the previous assertions related to the resonance prevention, consider as an illustrative example a CPW transmission line (see Fig. 12.3). The fundamental CPW mode of interest is an even mode (see Fig. 12.2) and hence the transmission line exhibits a magnetic wall at its symmetry plane. If we align with such a line an SRR

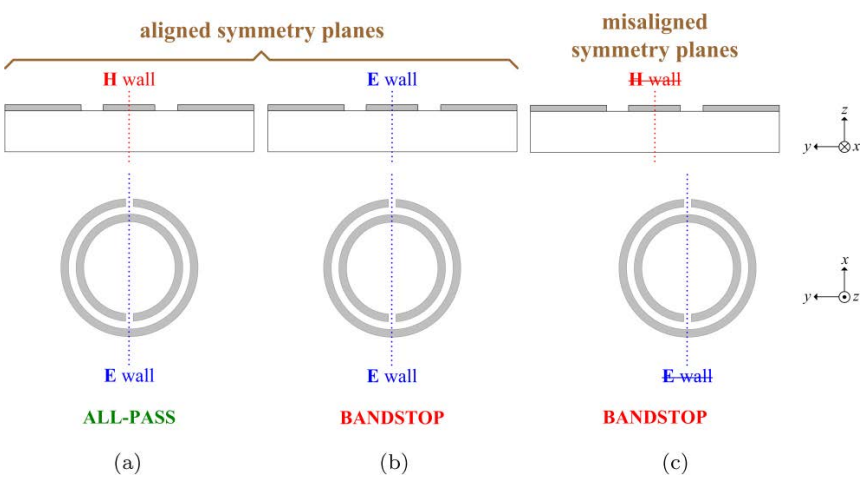


Fig. 12.3. CPW transmission line loaded with a single SRR. (a) Line and resonator aligned and excitation with the fundamental (even) mode; (b) line and resonator aligned and excitation with the slot (odd) mode; (c) line and resonator misaligned and excitation with the fundamental (even) mode. It is apparent that, for misaligned loadings, the magnetic wall at the symmetry plane of the line is altered.

(it exhibits an electric wall at its symmetry plane at the fundamental resonance), the resonator will not be excited (Fig. 12.3(a)). However, for the slot (odd) mode of the CPW, the resonator will be excited, since the symmetry plane of the line is an electric wall for that mode (Fig. 12.3(b)).

In the situations considered above, the symmetry planes of the line and the resonator are aligned. Let us now consider an all-pass structure in the unperturbed symmetric state (for instance the one of Fig. 12.3(a)), where the resonator is movable (e.g. by etching it in another substrate). If the resonator is rotated and/or laterally displaced, the perfect cancellation of fields over the resonator no longer holds, and a notch in the transmission coefficient emerges (see Fig. 12.3(c)). Due to symmetry disruption, the resonator is effectively coupled to the line, the coupling level (and hence the notch magnitude) being intimately related to the level of asymmetry. From the previous words, it follows that transmission lines loaded with movable symmetric resonators can be used for sensing linear and angular displacements and velocities. Obviously, the line and resonator geometries need to be optimized in order to achieve optimum sensor performance (linearity, sensitivity, dynamic range, etc.). This depends on the specific type of measurement and is left for the next section, where various examples are reported. Since symmetry is not affected by changes in environmental conditions, these sensors based on symmetry properties are robust against variable ambient conditions (temperature, moisture, etc.). It should be noted that these sensors are specially suited to detect a misalignment between the resonator and the line, and are therefore useful as alignment sensors.

So far, the considered symmetry disruption mechanism is a linear or angular displacement, with the result of a misalignment between the line and the resonator. However, the lack of symmetry can be caused by other mechanisms which do not involve the relative orientation between the line and the resonator. For instance: (a) the resonator symmetry may be topologically destroyed; (b) other elements can be included (e.g. dielectric loads); or (c) the surrounding medium can be subjected to inhomogeneities. In those cases where

the symmetry rupture is not based on geometrical misalignments, the propagation characteristics of the structure should be defined in a general sense in terms of symmetry/asymmetry rather than on alignment/misalignment. Provided that a transmission line is loaded with a resonator, and that the line may also be loaded with other inclusions (see Fig. 12.4), if the whole structure is symmetric (for which the line and the resonator are necessarily symmetric) and the symmetry planes of the line and the resonator are of distinct nature, resonance is inhibited ($|S_{21}| \approx 1$). Obviously, resonance is established if the symmetry planes are of the same electromagnetic nature ($|S_{21}| \approx 0$). However, if the whole structure is asymmetric, resonance is allowed, the stopband characteristics being dependent on the asymmetry level ($0 < |S_{21}| < 1$). Therefore, the stopband characteristics may be modulated by symmetry properties as a whole. The previous statements are valid regardless of the symmetry disruption mechanism.

By means of the sensing principle based on the symmetry properties of resonator-loaded lines, it is possible to implement spatial sensors, sensors for dielectric characterization, comparators, etc. There are many potential combinations of resonators and lines that can be useful for the implementation of these sensors, besides the one

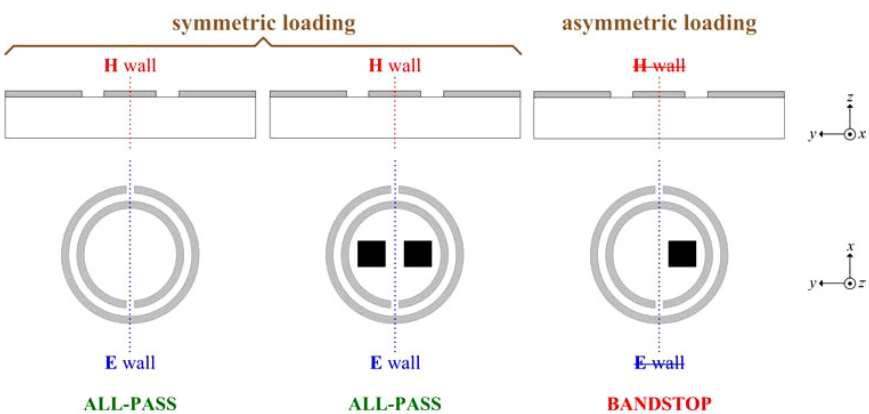
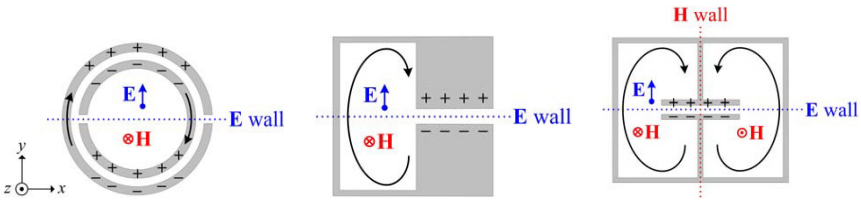


Fig. 12.4. Illustration of symmetry rupture by square-shaped inclusions in the resonance approach based on a single resonator, where a CPW is loaded with an SRR.



(a) (b) (c)

Fig. 12.5. Metallic symmetric resonators. (a) SRR; (b) FSIR; (c) ELC. The driving mechanisms (electric and/or magnetic field) and their direction are indicated.

considered to describe the sensing principle (an SRR-loaded CPW). The main line types are those depicted in Fig. 12.2. Concerning the resonant elements, multiple possibilities arise. Let us briefly review some of these resonators, dividing them into metallic and slot resonators.

Figure 12.5 depicts typical topologies of metallic resonant elements, i.e. the SRR, the folded stepped impedance resonator (FSIR), and the electric-field-coupled LC resonator or simply electric LC (ELC) resonator. The distribution of charges and currents, as well as the electromagnetic nature of the symmetry plane (E-wall or H-wall), at the fundamental resonance are also indicated. The SRR, proposed by Pendry *et al.* in 1999,[42] is a well-known particle for the implementation of negative effective permeability media. As mentioned before, it exhibits an electric wall at the symmetry plane at the fundamental resonance, and its electrical size is small by virtue of the coupling between the two metallic open loops of the resonator. The size of the particle can be made even smaller by etching the two metallic loops in opposite sides of a narrow substrate (the resulting particle has been denominated broadside-coupled SRR, BC-SRR[43]), and electrical sizes (side length for a square geometry) as small as $\lambda/16$, λ being the wavelength in the considered substrate material at the fundamental resonance, have been demonstrated.[44] One further advantage of this particle is that by etching the two loops of the resonator on different substrates (with relative motion), it is possible to implement displacement sensors based on them, as will be shown

later. The FSIR also exhibits an electric wall at its symmetry plane at the fundamental resonance. Loaded to a CPW or to a microstrip line, the dominant coupling mechanism is electric (at least in the configurations of Refs. 30, 33) and, by virtue of the broadside capacitance between the line and the resonator, the physical size of the particle, for a given resonance frequency (a requirement in many applications), can be very small.[30,33] A very interesting particle for sensing purposes is the ELC (Fig. 12.5(c)). This is a bisymmetric resonator simultaneously exhibiting an electric wall and a magnetic wall (orthogonally oriented) at the fundamental resonance. This particle was proposed in order to achieve resonant-type negative effective permittivity media.[45] Note that the particle cannot be excited by a uniform time-varying magnetic field orthogonal to the plane of the particle, but it can be excited by an electric field applied in the direction orthogonal to the electric wall. According to the current distribution at the fundamental resonance, the particle can also be excited by counter (i.e. non-uniform) magnetic fields applied to the loops, as those present in the slots of a CPW for the fundamental CPW mode.

The slot resonators depicted in Fig. 12.6 are the dual (or complementary) counterparts of those represented in Fig. 12.5. The slot version of the SRR has been designated as complementary SRR (CSRR).[46] The slotted ELC has been called complementary ELC,[47] and magnetic LC (MLC),[48] but we prefer the latter designation to emphasize the fact that the particle can be excited by a uniform magnetic field (rather than electric field) applied in the plane of the

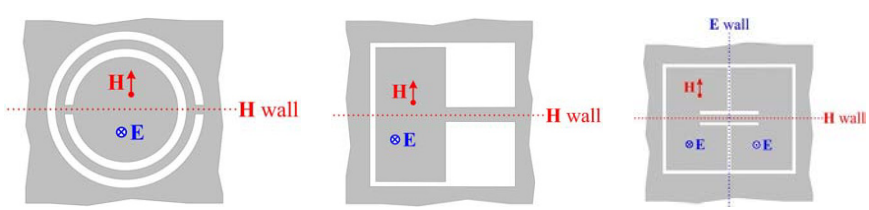


Fig. 12.6. Slotted symmetric resonators. (a) CSRR; (b) complementary FSIR; (c) MLC. The driving mechanisms (electric and/or magnetic field) and their direction are indicated.

particle. The roles of the magnetic and electric walls, as well as the driving mechanisms (electric and magnetic fields), are interchanged in the metallic and slot resonators, as can be appreciated by comparing Figs. 12.5 and 12.6.

By adequately combining the previous resonant elements with the transmission lines depicted in Fig. 12.2 (or with other transmission lines not shown), it is possible to implement microwave sensors based on symmetry disruption. A possible combination is the CPW transmission line with the SRR, as considered in Figs. 12.3 and 12.4. Another prominent resonant element for the implementation of sensors in CPW technology, especially for 90° angular displacement measurements, is the ELC resonator (see Fig. 12.7). The reason is that the particle is bisymmetric. Thus, when the electric wall is aligned with the line, the structure is transparent, but particle rotation induces magnetic coupling between the line and the resonator. The tightest coupling corresponds to the case when the magnetic wall of the particle is aligned with the line, causing a deep notch in the transmission coefficient. According to these words, this configuration is expected to exhibit high linearity of the notch depth with the rotation angle (this aspect will be treated later).

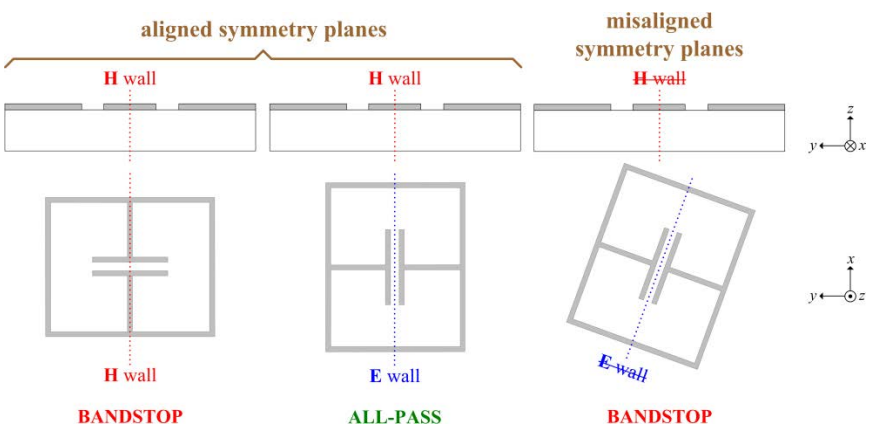


Fig. 12.7. Illustration of the modulation in the suppression by alignment/misalignment of the symmetry planes in a CPW loaded with an ELC resonator. The stopband characteristics may be modulated by rotating the resonator.

In principle, any combination of transmission line and symmetric resonator exhibiting different electromagnetic walls at the symmetry plane for the unperturbed symmetric state can be useful for sensing purposes on the basis of symmetry disruption. By etching FSIRs (or even unfolded SIRs) in a movable substrate, these resonators are also useful for the implementation of displacement sensors in both CPW[30] and microstrip[33] technology. Note also, for instance, that differential microstrip lines (excited by differential signals) can be combined with CSRRs, or with any other resonant element exhibiting a magnetic wall in its symmetry plane (e.g. ELC and MLC) at the fundamental resonance.[49]

12.2.2. *Frequency-splitting sensors*

Thus far, we have analyzed the symmetry properties of transmission lines loaded with a single resonator, and it has been demonstrated that these structures are potentially useful for sensing purposes. Symmetry properties in transmission lines loaded with any other number of resonators (arranged transversally to the line) can also be derived. Let us consider the particular case with pairs of resonators. The generalization of the symmetry rupture for pairs of resonators is subjected to using identical resonators (symmetric resonators are not necessary) in order to exhibit a plane of symmetry (at the midway between the resonators). Thereby, the loaded line can be symmetric by aligning the plane of symmetry of the paired resonators with the plane of symmetry of the line. However, there is a fundamental difference as compared to using a single resonator: if symmetric resonators are used, the two resonators cannot be simultaneously aligned with the line (we can align only one resonator). Therefore, only one resonance frequency can be inhibited at most in transmission lines loaded with pairs of resonators (and in the general case with more than a single resonator). The relevant conclusion is that resonance always occurs. In terms of absolute symmetry, provided that a transmission line is loaded with a pair of resonators, and that the line may also be loaded with other inclusions, if the whole structure is symmetric (where the line must be symmetric and

the pair of resonators must be identical and symmetrically oriented to the line axis), there is only one transmission zero because the resonance frequencies of the two resonators degenerate to the same value ($|S_{21}| \approx 0@f_0$). However, if the whole structure is asymmetric, in general there are two transmission zeros because two different resonances arise ($|S_{21}| \approx 0@f_1; f_2$), i.e., frequency splitting. It should be highlighted that resonance frequency-splitting is used in this chapter in a broad sense involving splitting as well as shifting. For instance, if a transmission line is loaded with a pair of resonators with different resonance frequencies, it is apparent that there is actually no split from a single resonance frequency.

Consequently, in this approach based on pairs of resonators, a high controllability in regard to the number and frequency of resonances as a function of symmetry may be achieved (the bandwidth and depth of the notches, however, may also depend on symmetry). Figure 12.8 illustrates the above-mentioned assertions where symmetry is disrupted by inclusions. As can be seen, an interesting aspect of this approach is that there is no restriction in the orientation of the individual resonators in regard to the line. It would be interesting to cancel mixed-coupling effects in certain topologies.[50] For instance,

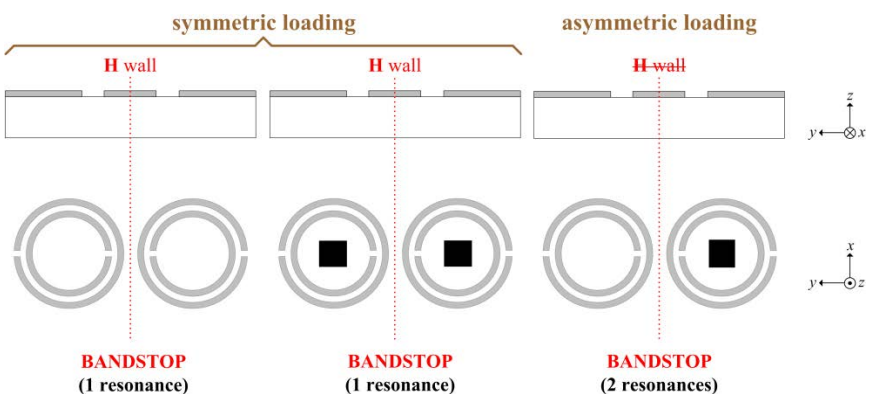


Fig. 12.8. Illustration of symmetry rupture by square-shaped inclusions in the frequency-splitting approach based on pairs of resonators, where a CPW is loaded with a pair of SRRs. Note that the symmetry planes of the SRRs are not required to be parallel to the longitudinal x-direction, preventing from mixed coupling.

if symmetric SRRs are considered, their symmetry planes do not need to be oriented longitudinally and the SRRs are coupled only magnetically to the line.

Symmetry rupture by adding inclusions to the structure may be useful for the implementation of differential sensors and comparators. Potential applications are dielectric characterization, defect detection in samples (by comparison to a reference), biosensors, etc. It is also possible to use this sensing principle based on frequency splitting for the implementation of spatial sensors, as will be shown later.

12.3. Microwave Sensors Based on Symmetry Properties of Resonator-Loaded Lines: Examples

In this section, several examples of sensors (or proof-of-concept demonstrators) based on the principles pointed out in the preceding section are reported. Specifically, we will present linear displacement sensors, angular displacement and velocity sensors, and sensors for dielectric measurements.

12.3.1. *Linear displacement sensors*

Resonance-based[28–31,33,34,36] and frequency-splitting[40] sensors for linear displacement measurement have been proposed. In Ref. 31, it is demonstrated that a right-angle bended CPW transmission line loaded with four single-loop SRRs can be used as a 2D displacement sensor, and more precisely, as a 2D position sensor, able to distinguish between displacements in the positive or negative directions of the 2D space (Fig. 12.9). For the implementation of the sensor, four SRRs of different dimensions (and hence fundamental resonance) must be etched on a movable substrate (with relative motion with regard to the CPW transmission line). The resonators $\mathrm{SRR}_{\Delta x}$ and $\mathrm{SRR}_{\Delta y}$ are called displacement sensing resonators, and when they are laterally displaced (with regard to the corresponding line section axis), a notch in the transmission coefficient appears. To distinguish between the $\pm x$ and $\pm y$ directions, the resonators designated as $\mathrm{SRR}_{\pm x}$ and $\mathrm{SRR}_{\pm y}$ play the key role, namely, the corresponding

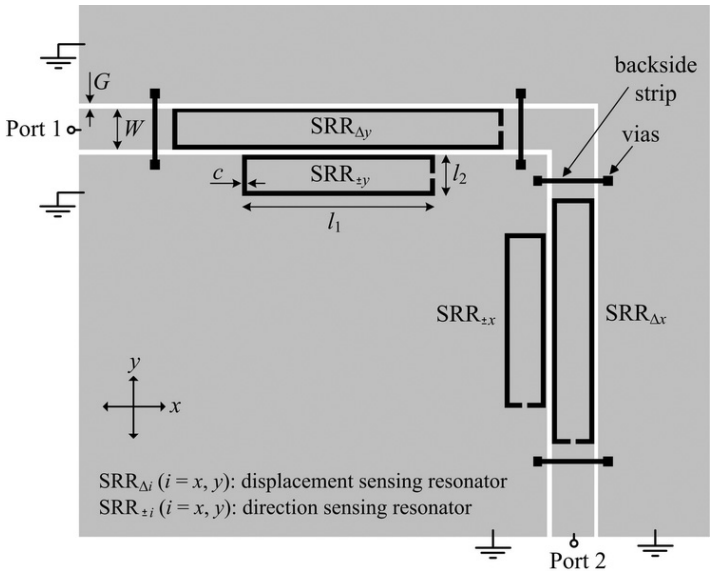
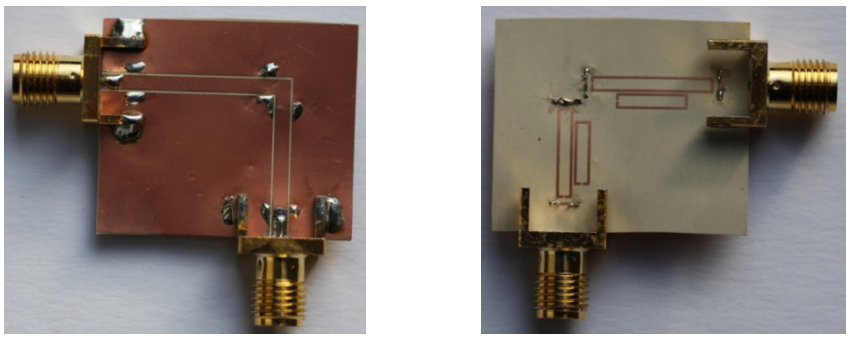


Fig. 12.9. Topology of the 2D displacement sensor based on the resonance approach. The SRRs must be etched on a movable substrate. The CPW dimensions are $W = 1.67$ mm and $G = 0.2$ mm, respectively. The dimensions of the SRRs are: $l_1(\mathrm{SRR}_{\Delta x}) = 9.95$ mm, $l_1(\mathrm{SRR}_{\pm x}) = 7.05$ mm, $l_1(\mathrm{SRR}_{\Delta y}) = 13.4$ mm, $l_1(\mathrm{SRR}_{\pm y}) = 7.8$ mm, $l_2 = 1.67$ mm and $c = 0.2$ mm. The considered substrate is *Rogers RO3010* with relative permittivity $\varepsilon_r = 10.2$, thickness $h = 127$ μm and loss tangent $\tan\delta = 0.0023$.

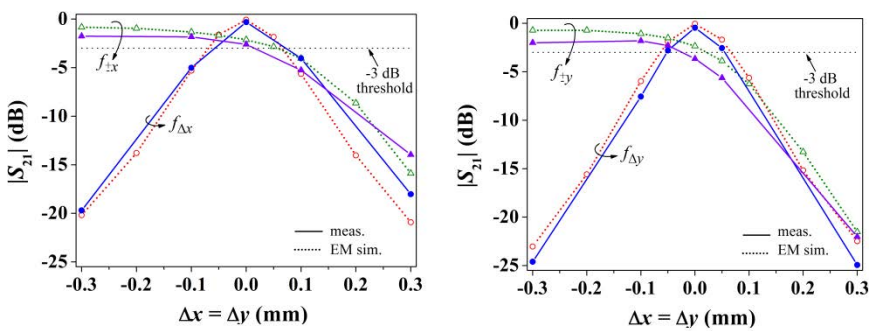
resonance frequency (notch) is "activated" or not, depending on the sign of the displacement. Activation occurs only for positive displacement in x- or y-directions, since for negative shifts the resonators $\mathrm{SRR}_{\pm x}$ and $\mathrm{SRR}_{\pm y}$ are out of the influence of the CPW and are not coupled to it. Indeed, resonators $\mathrm{SRR}_{\pm x}$ and $\mathrm{SRR}_{\pm y}$, whose sensing principle is not based on symmetry arguments, act as flags that are activated for positive displacements.

To demonstrate the potential of this approach, a proof-of-concept was reported in Ref. 31, where several samples with SRRs etched on the back substrate side of the CPW but at different locations (corresponding to different displacements) were fabricated. The device corresponding to the unperturbed state is depicted in Fig. 12.10, and the notch magnitudes of the transmission coefficient at the frequencies of



(a) (b)

Fig. 12.10. Photograph of the proposed proof-of-concept device for the aligned position; (a) top and (b) bottom face.



(a) (b)

Fig. 12.11. Notch magnitude of the transmission coefficient S_{21} at the indicated frequencies for $x = y$-oriented displacement; results for (a) x- and (b) y-axis position sensing. Reprinted with permission from Ref. 31.

the different resonators for $x = y$-oriented displacement is depicted in Fig. 12.11. It can be appreciated that for positive displacements, the $\text{SRR}_{\pm x}$ and $\text{SRR}_{\pm y}$ resonators are activated as is manifested by a clear increase in the notch at $f_{\pm x}$ and $f_{\pm y}$, whereas the threshold level (−3 dB) is not exceeded for negative displacements (indicating that the shift is in the negative direction).

In Ref. 34, the proposed sensor (1D) uses also single-loop SRRs coupled to a CPW transmission line, but the shape is optimized to

improve the linearity and dynamic range. In addition, the sensor is able to operate at a fixed frequency since the notch frequency is constant with the displacement, bypassing the need for a frequency sweeping source and measurement system.

A different approach is used in Ref. 36, where a 2D displacement and alignment sensor is implemented on the basis of the reflection coefficients of orthogonally oriented open microstrip lines loaded with SRRs. An original aspect of these sensors is that they work in reflection, rather than transmission, but the symmetry-related operating principle is the same. The layout of the device is depicted in Fig. 12.12. For the unperturbed state, the resonators are not excited and the injected signals to both ports are reflected back to the source. Hence, the reflection coefficients are $|S_{11}| = |S_{22}| = 1$. However, if the alignment is broken by a displacement in the x- and y-directions, the SRRs will be excited and notches will appear in the reflection coefficients. Further displacement in the x- and y-directions gives rise

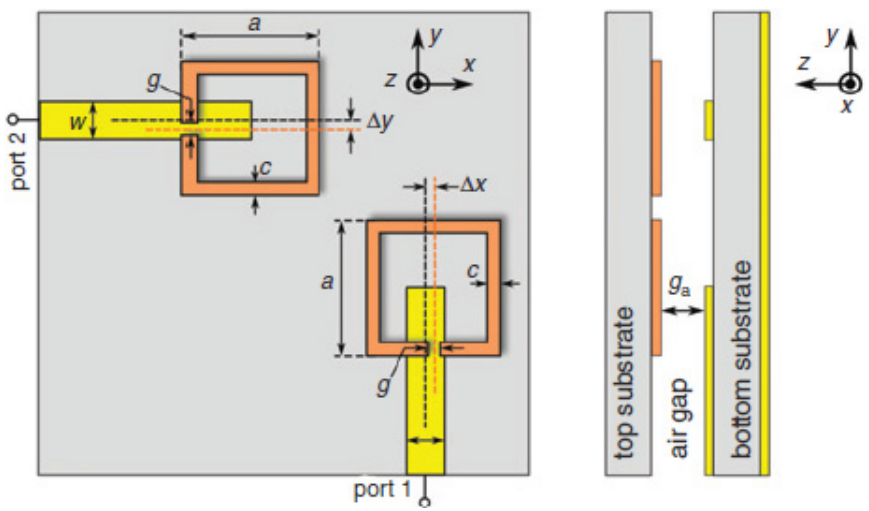


Fig. 12.12. Top (a) and side (b) views of the 2D displacement sensor based on the resonance in the reflection coefficients. The structures are patterned on *Rogers RO4003* substrates with 0.81 mm thickness. The geometrical dimensions of the line and SRR are $w = 1.84$ mm, $a = 7$ mm, $g = 0.5$ mm and $c = 0.5$ mm. There is an air gap with $g_a = 0.76$ mm between the two substrates.

to a stronger coupling between the transmission lines and the SRRs, in turn resulting in deeper notches in the reflection coefficients at the resonance frequency of the SRRs. These notches result from radiation effects,[39] and this is another relevant feature of this sensing strategy based on (one-port) open-ended transmission lines.

Note that displacement in the x-direction has no effect on the depth of notch in $|S_{22}|$ (nor displacement in the y-direction onto $|S_{11}|$). Thus, misalignment in the x and y directions can be independently sensed from the depth of notches in $|S_{11}|$ and $|S_{22}|$, respectively. One advantage of this method is that both SRRs can be designed to operate at the same resonance frequency. This is in contrast to the previous 2D displacement sensor based on the transmission characteristics of a CPW, where SRRs needed to have distinct resonance frequencies. Note that the direction of motion cannot be determined with the sensor depicted in Fig. 12.12. Nevertheless, additional resonant elements, tuned at different frequencies, can be added to the movable substrate in order to distinguish the directions of motion, similar to the strategy of the sensor of Fig. 12.10.

The photograph of the fabricated device is shown in Fig. 12.13. Figure 12.14 depicts the reflection coefficient measured from any of

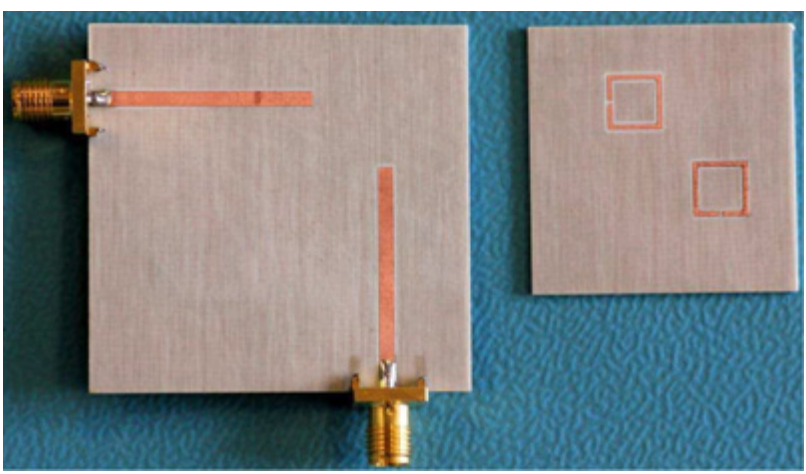


Fig. 12.13. Photograph of the fabricated sensor implemented through open-ended transmission lines.

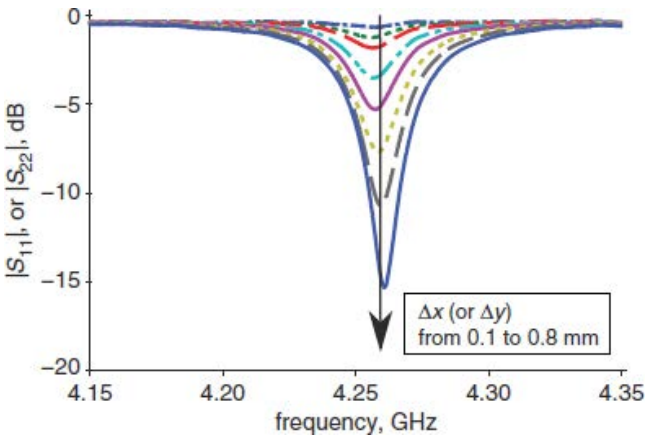


Fig. 12.14. Measured $|S_{11}|$ (or $|S_{22}|$) for different values of displacement Δ_x (or Δ_y) from 0.1 to 0.8 mm in steps of 0.1 mm. Reprinted with permission from Ref. 36.

the ports for different lateral displacements (seen from the corresponding port). One important feature of the proposed sensor is that displacement affects only the depth of the notch, and leaves the resonance frequency nearly intact. As mentioned earlier in conjunction with Ref. 34, a fixed resonance frequency can be an important feature that enables the proposed sensor to operate at a fixed frequency. To validate this important feature, Fig. 12.15 depicts the simulated and measured $|S_{11}|$ and $|S_{22}|$ at a fixed frequency of 4.253 GHz against Δx, while Δy is 0 mm. The figure clearly shows that the movement in the x-direction can be sensed from $|S_{11}|$, whereas $|S_{22}|$ remains unaffected and shows the alignment in the y-direction.

Frequency-splitting 2D alignment and displacement sensors based on broadside-coupled split ring resonators (BC-SRRs[43]) were reported for the first time in.[40] The side and top views of the proposed sensor are depicted in Fig. 12.16. Note that a right-angle bended microstrip line is considered for sensing, where the two resonant elements are designed to resonate at different frequencies. The typical topology of the BC-SRR was modified so that the resonant frequency is altered by the displacement in one direction only. To this end, the upper rings are replaced with straight strips. In this

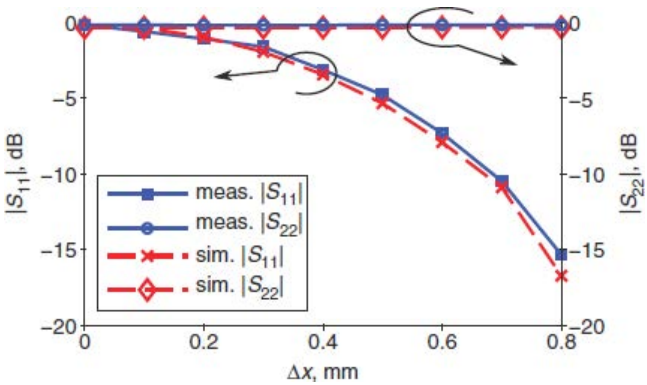


Fig. 12.15. Comparison between measured and simulated $|S_{11}|$ and $|S_{22}|$ at a fixed frequency $f = 4.253$ GHz against Δx, for $\Delta y = 0$ mm. Owing to the symmetry of the proposed sensor, identical comparison is valid between $|S_{11}|$ and $|S_{22}|$ against Δy. Reprinted with permission from Ref. 36.

configuration, provided the strips are long enough, the pair of modified resonators that are coupled to the horizontal section of the microstrip line is only sensitive to a displacement in y direction, while the pair of the modified BC-SRRs coupled to the vertical section of the line is only responsive to a displacement in x direction. An air gap, to avoid abrasion of metallic layers and substrates, is added between the two substrates, as indicated in Fig. 12.16.

Figures 12.17(a) and 12.17(b) depict the simulated response of the proposed 2D sensor to displacements in x and y directions, respectively. Figure 12.17(a) depicts the simulated transmission coefficients of the proposed 2D sensor for different values of displacement in x direction from 0 mm to 3 mm in steps of 1 mm when no displacement in y direction is applied, i.e. $\Delta y = 0$ mm. At the initial position, when both pairs of the modified BC-SRRs are aligned with the symmetry plane of the corresponding sections of the microstrip line, only two notches at f_{x0} and f_{y0} appear in the transmission spectrum of the line. The resonances at frequencies f_{x0} and f_{y0} are associated with Pair 1 and Pair 2, respectively. However, as shown in the figure, a displacement in x direction results in the splitting of the resonance of Pair 1 in two notches at f_{x0} and f_{x1}. The difference

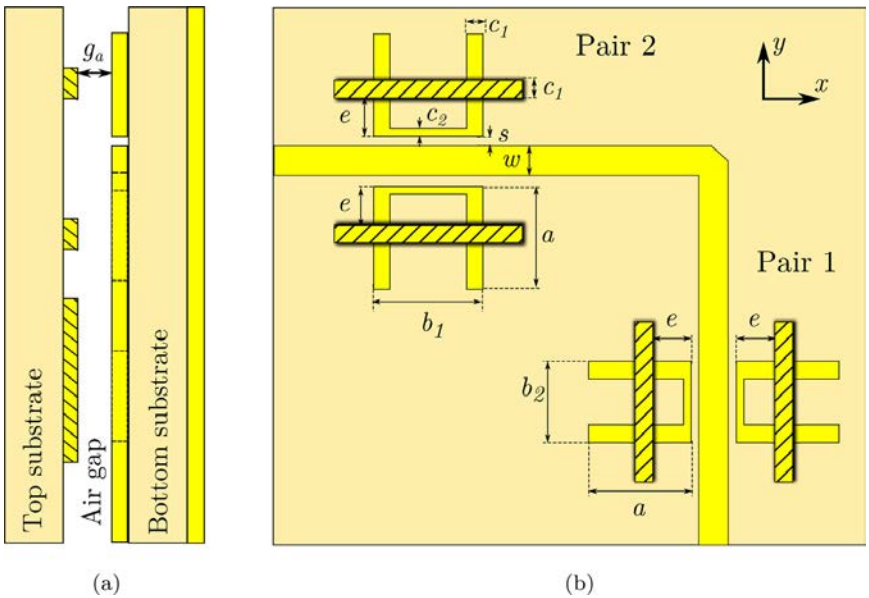


Fig. 12.16. (a) Side view and (b) top view of the proposed 2D displacement sensor. *Rogers RO4003* substrate with relative permittivity $\varepsilon_r = 3.38$ and thickness $h = 0.81$ mm is used for the bottom and top substrates, which are separated with an air gap $g_a = 0.25$ mm. The microstrip line has a width $w = 1.85$ mm, which corresponds to a 50-Ω line. The dimensions of the resonators are as follows: $a = 12.2$ mm, $b_1 = 15.5$ mm, $b_2 = 10.5$ mm, $c_1 = 2.5$ mm, $c_2 = 1.2$ mm, $s = 0.2$ mm and at initial position $e = 4.8$ mm.

between the two frequencies, i.e. $\Delta f_x = f_{x1} - f_{x0}$, is increased with an increase in displacement in x direction, thus it can be used for sensing the value of Δx. Similarly, Fig. 12.17(b) shows the simulated transmission coefficients of the proposed sensor for variation of Δy from 0 mm to 3 mm in steps of 1 mm while the Pair 1 is aligned with the symmetry plane of the vertical section of the microstrip line, i.e. $\Delta x = 0$ mm. The figure clearly shows that $\Delta f_y = f_{y1} - f_{y0}$ increases with Δy, allowing for sensing a displacement in y direction.

In order to validate the concept and the simulation results, the designed 2D sensor was fabricated and measured. The photographs of the top and bottom substrates of the fabricated prototype are

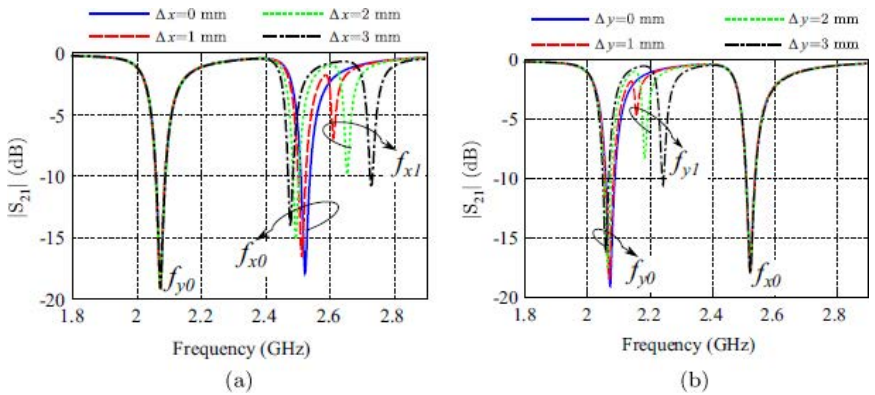


Fig. 12.17. Simulated transmission coefficients of the proposed sensor (a) for different values of displacement in x direction from 0 mm to 3 mm in steps of 1 mm, while $\Delta y = 0$ mm, and (b) for different values of displacement in y direction from 0 mm to 3 mm in steps of 1 mm, while $\Delta x = 0$ mm. Reprinted with permission from Ref. 40.

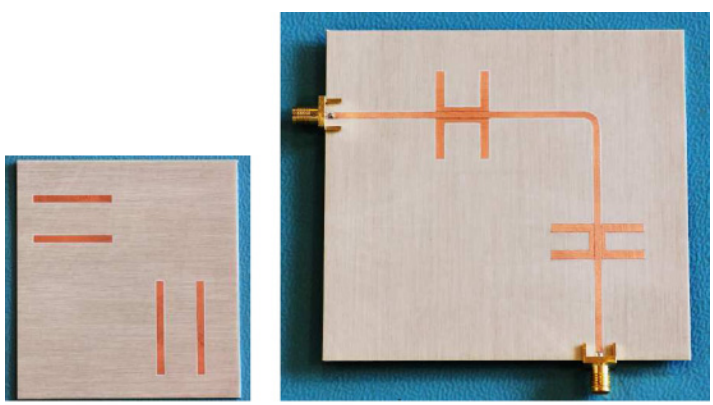


Fig. 12.18. Photographs of the fabricated prototype: (a) top and (b) bottom substrates.

depicted in Fig. 12.18. Figure 12.19 shows a photograph of the measurement setup, which is composed of a pair of micrometer actuators for adjusting the air space between the two substrates, as well as two pairs of micrometer actuators for accurate displacement in x and y

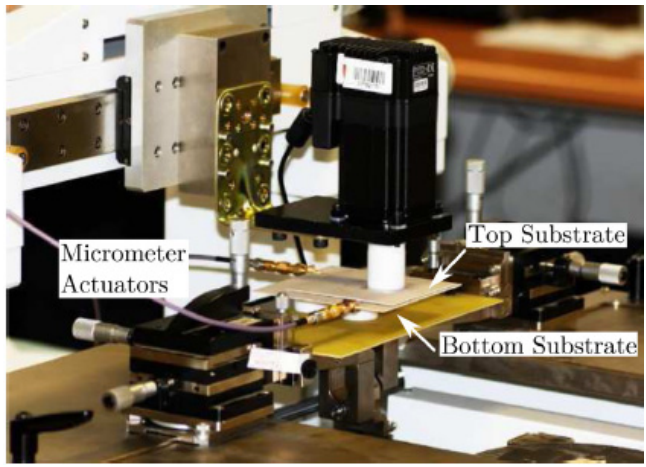


Fig. 12.19. Measurement setup, composed of three sets of micrometer actuators for accurate adjustment of the air gap between the substrates and displacement in x and y directions.

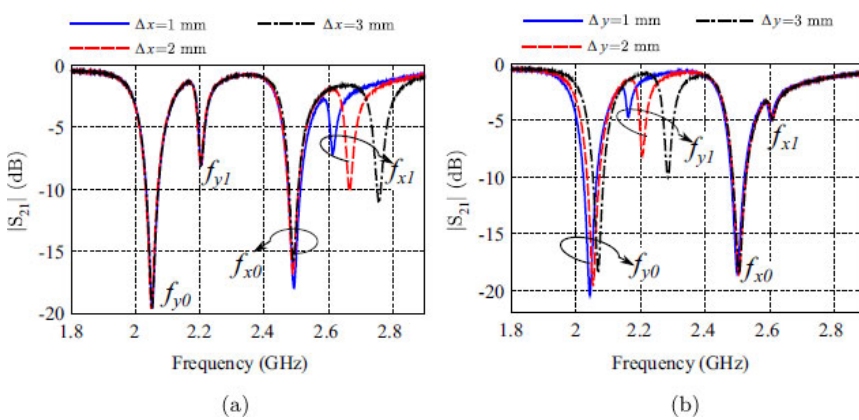


Fig. 12.20. Measured transmission coefficients (a) for three different values of displacement in x direction, while $\Delta y = 2$ mm, and (b) for three different values of displacement in y direction, while $\Delta x = 0.5$ mm. Reprinted with permission from Ref. 40.

directions. The gap uniformity is guaranteed because the considered substrates are of sufficient rigidity.

Figure 12.20(a) depicts the measured transmission coefficients of the sensor for different values of $\Delta x = 1$ mm, 2 mm and 3 mm, while

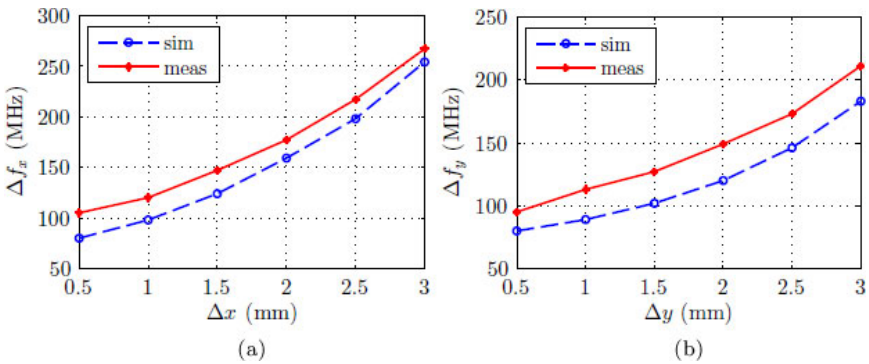


Fig. 12.21. Comparison between measured and simulated spurious resonance frequency versus displacement: (a) Δf_x versus Δx and (b) Δf_y versus Δy. Reprinted with permission from Ref. 40.

the sensor has a fixed displacement of 2 mm in y direction. The figure shows that while frequency difference Δf_x is increased from 105 MHz to 267 MHz, Δf_y does not change. Similarly, Fig. 12.20(b), shows the measured results for a fixed displacement of 0.5 mm in x direction, resulting in fixed notches at $f_{x0} = 2.5$ GHz and $f_{x1} = 2.6$ GHz, while Δy is changed from 1 mm to 3 mm in steps of 1 mm, which can be sensed from the increase in Δf_y from 95 MHz to 210 MHz. The experiment shows that each pair of the modified BC-SRRs are exclusively responsive to a displacement either in x or y direction. Thus, the proposed structure can be efficiently used as a 2D alignment and displacement sensor.

In Figs. 12.21(a) and 12.21(b) the measured Δf_x and Δf_y versus displacement in x and y directions, respectively, are compared with those of the simulation results. The figure shows the satisfactory linearity of the proposed sensor. Note that the air gap distance can modify the resonance of the notch, but if there is alignment, only one notch is expected. Thus, for alignment purposes, the sensor benefits from robustness against environmental changes. Nevertheless, tolerances in the gap distance may also influence the results. This is actually (along with the fabrication and measurement setup tolerances) the reason for the discrepancy between measurements and simulations in the results of Fig. 12.21. However, as clearly shown

in Fig. 12.21, the trend of the measured frequency difference is very well matched with the simulated results. This can be quantitatively verified through Pearson's correlation coefficient between simulated and measured data.[51] The correlation coefficient is better than 0.99 for displacement in both x and y directions, which quantitatively verifies the good agreement between the simulated and measured data. Hence, calibration of the sensor can be used in order to remove the effect of tolerances in fabrication and in the gap distance.

Since the purpose of this chapter is to provide initial proof-of-principles, a vector network analyzer (VNA) is used for the measurement of the reflection/transmission coefficient of all the sensors presented. However, in a practical application, in frequency-splitting sensors, a voltage-controlled oscillator (VCO) can be used to excite the sensor. By sweeping the frequency in the region of interest, by means of an envelope detector cascaded at the output port, the frequencies where the notch (for alignment) or notches (for misalignment) appear can be detected. Similar feeding and readout circuits can be used in resonance-based sensors, where the VCO can be replaced with an oscillator of fixed frequency in the case of a frequency-independent notch.

Furthermore, through an in-between air gap, the displacement and alignment sensing in the proposed sensors can be implemented by displacing the top substrate without any friction over the bottom substrate. Due to a contactless measurement, the sensors benefit from minimum aging effects and a high repeatability.

12.3.2. *Angular displacement and velocity sensors*

Resonance-based angular displacement and velocity sensors have been reported in Refs. 32, 37 and Ref. 38 in CPW and microstrip technology, respectively. In all the cases, a circularly-shaped ELC resonator and line are used since by these means linearity is optimized. For the CPW-based sensor, the line is non-uniform along the propagation direction. However, matching is preserved provided the central strip and slot widths have the appropriate dimensions. For the microstrip-based sensors, the transmission line is actually composed

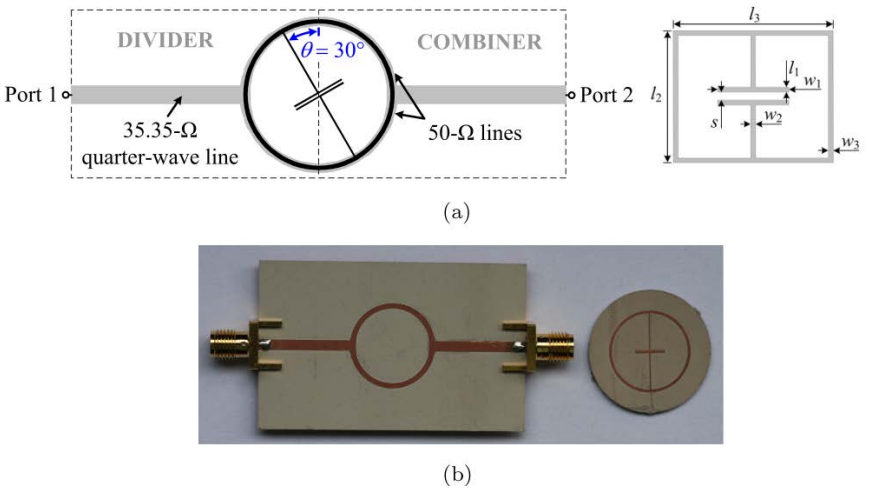
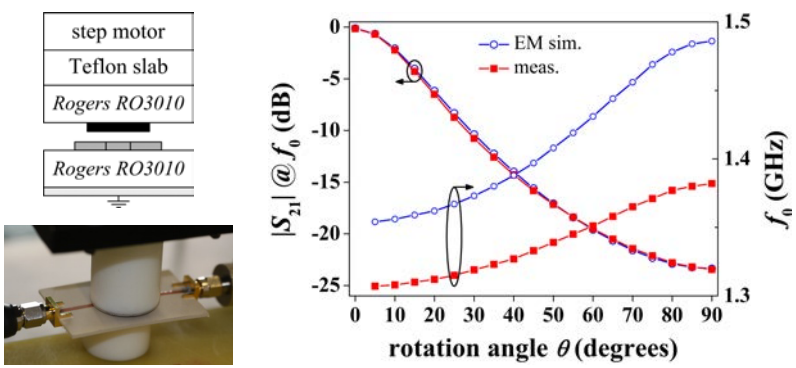


Fig. 12.22. Angular displacement sensor. (a) Layout, and (b) photograph. The substrates are *Rogers RO3010* with relative permittivity $\varepsilon_r = 11.2$, thickness $h = 1.27$ mm (microstrip line) and $h = 0.635$ mm (resonator), and loss tangent $\tan\delta = 0.0023$. The line widths W are 2.06 mm and 1.04 mm. ELC mean radius $r_0 = 8.05$ mm, $w_1 = 6$ mm, $w_2 = l_1 = s = 0.2$ mm and $w_3 = 0.5$ mm.

of a circularly-shaped divider/combiner, as depicted in Fig. 12.22. This is the sensor prototype that has been chosen to be reported in this chapter, hence let us analyze it in detail. Let us consider that a circular ELC resonator (attached to the rotating object) is symmetrically placed above the pair of lines. Under these conditions, the symmetry plane (i.e. the midplane) behaves as a magnetic wall. As proven in,[49] and also according to the discussion of Subsection 6.2.1, the structure is transparent to signal transmission if the ELC electric wall is aligned with the symmetry plane of the line. Contrarily, a transmission zero arises if the alignment is with the ELC magnetic wall. In the former situation, the fundamental ELC resonance cannot be excited. However, the resonator is strongly coupled (magnetically) to the lines if the magnetic walls are aligned. Such behavior, explained through equivalent circuit models in,[49] suggests that rotations with a 90° dynamic range can be measured from the common-mode transmission coefficient.

By feeding the structure with a single-ended signal, the divider acts as a single-ended to common-mode signal transition. Conversely, the combiner converts the common-mode signal into a single-ended signal. The circularly shaped lines exhibit 50-Ω even-mode characteristic impedance. The combiner/divider is implemented with 35.35-Ω impedance inverters (quarter-wavelength lines at 1.45 GHz) to achieve matching to 50-Ω reference ports. The circular 50-Ω line pair (which is responsible for resonator coupling) is arranged face-to-face with the ELC external ring in order to enhance electromagnetic coupling. Furthermore, the line pair as well as the resonator was circularly-shaped to enhance linearity, as mentioned above.

The displacement measurement was performed using an experimental setup similar to that reported in Fig. 12.19, and shown in Fig. 12.23(a). The ELC substrate was attached to a step motor through a Teflon slab with 3.5 cm thickness and relative permittivity $\varepsilon_r = 2.08$. The air gap between the ELC resonator and the microstrip lines was set to 0.254 mm. As shown in Fig. 12.23(b), the rotation angle can be sensed from the notch magnitude (and frequency as well, in this configuration), with a reasonably linear dependence.



(a) (b)

Fig. 12.23. (a) Experimental setup, and (b) notch magnitude and frequency versus the angular displacement. The frequency shift is attributed to substrate and fabrication tolerances (f_0 is very sensitive to the resonator parameters). Reprinted with permission from Ref. 38.

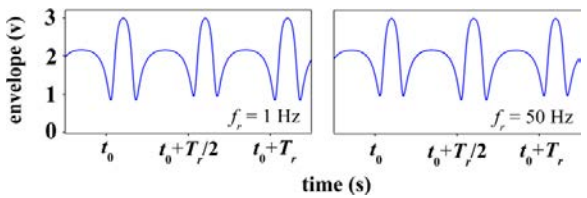


Fig. 12.24. Angular velocity measurements for an arbitrary initial time t_0. Reprinted with permission from Ref. 38.

Let us now consider that the ELC resonator is continuously rotating. To obtain the angular velocity, ω_r, a harmonic signal tuned at a fixed frequency (e.g. the notch frequency corresponding to $\theta = 10°$, $f_0 = 1.308$ GHz) is injected to the input port, whereas an isolator plus an envelope detector are cascaded to the output port. The envelope signals at the output (which are dependent on the transmission coefficient) are captured by an oscilloscope (Fig. 12.24), and the velocities are derived from the time difference between consecutive transmission peaks $T_m = T_r/2 = 1/2f_r$. By configuring the step motor with $f_r = 1$ Hz and 50 Hz ($\omega_r = 2\pi f_r$), the measured velocities are $f_r = 0.998$ Hz and 50.251 Hz, respectively, validating the approach. Rotation speeds as high as required can be measured by means of this approach since the carrier signal frequency (within the microwave region of the electromagnetic spectrum) is much higher than the equivalent linear velocity of actual rotating systems. As long as the rotation speed is constant over transmission peaks, precision can be improved by merely averaging the time between peaks.

12.3.3. *Permittivity sensors*

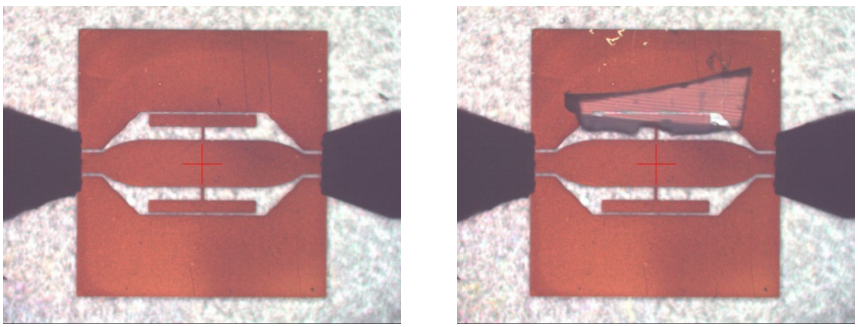
In this subsection, it is shown that frequency-splitting sensors, based on symmetry properties, for dielectric measurements can be implemented. Microwave sensors for dielectric characterization of biological substances (such as organic tissue), for biological cell analysis, or for the determination of component concentration/composition in aqueous (or other liquid) solutions by means of microfluidics have been proposed.[19–22,52–58] In most of these approaches, sensing is based on frequency shift. Let us now explore the principle for the

implementation of microwave sensors for dielectric measurements on the basis of frequency splitting. The main idea was detailed in Section 6.2.2, where a perturbation of the symmetry in a transmission line loaded with a pair of resonators produces a frequency splitting (see Fig. 12.8). In particular, symmetry can be perturbed by asymmetric dielectric loading. Hence, the structures can be utilized to work as differential sensors, and they are of special interest as comparators, where defects or abnormalities of samples as compared to a reference can be easily detected.

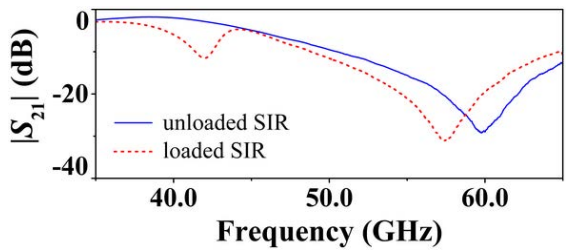
The proof-of-concept demonstrator reported here is based on a CPW transmission line loaded with a pair of step impedance resonators (SIRs). The analysis of these structures, where magnetic coupling between the SIRs is necessary for an accurate modelling, was carried out in Ref. 41. Therein, it was demonstrated that asymmetric SIRs (or, equivalently, identical SIRs with asymmetric loading) provide two resonances (i.e. frequency splitting). However, due to magnetic coupling between the resonant elements, frequency splitting is higher than the frequency span corresponding to the frequencies of the isolated resonators. This aspect is not beneficial in terms of sensitivity for differential measurements, but is not dramatic for the use of the structure as a comparator. Figure 12.25 depicts the fabricated proof-of-concept demonstrator, implemented specifically in conductor-backed CPW (CB-CPW). By loading the structure with a piece of glass ($\varepsilon_r = 4.6$) on top of one of the SIRs, two transmission zeros appear, as compared to the structure (symmetric) without the piece of glass. As an alternative to SIRs, a similar frequency-splitting sensor can be realized by loading a CPW transmission line with a pair of SRRs. An exhaustive analysis of this structure (including magnetic coupling between the SRRs) was carried out in Ref. 59, and similar conclusions (as compared to the SIR-loaded line) relative to the effects of magnetic coupling between resonators were drawn.

12.4. Conclusions

In conclusion, it has been shown in this chapter that microwave sensors inspired by metamaterials and designed on the basis of the



(a)



(b)

Fig. 12.25. (a) Photograph of the proof-of-concept demonstrator without (left) and with (right) a piece of glass on the top; (b) measured transmission coefficient. The dimensions are: $W = 350$ μm, $G = 210$ μm, the lengths of the SIR are $l_1 = 90$ μm (low impedance section) and $l_2 = 100$ μm (high impedance section), and the widths of the SIR are $w_1 = 800$ μm and $w_2 = 20$ μm. The substrate parameters are $\varepsilon_r = 4.6$, $h = 200$ μm and $\tan\delta = 0.0021$. Reprinted with permission from Ref. 41.

symmetry-related electromagnetic properties of transmission lines loaded with electrically small resonators can be implemented. The sensing principle, based on symmetry disruption, has been explained comprehensively. It has been shown that the presented sensors can be classified into two types: resonance-based sensors and frequency-splitting sensors. Several prototype devices to demonstrate the potential of the two approaches have been presented, including linear displacement sensors, angular displacement and velocity sensors and permittivity sensors.

Acknowledgments

The authors would like to thank Gerard Sisó and Javier Hellín the fabrication of several prototypes. This work has been supported by MINECO-Spain (projects TEC2010-17512 METATRANSFER, CONSOLIDER EMET CSD2008-00066, TEC2013-40600-R COM-SEN-RFID and TEC2013-49221-EXP) and by Generalitat de Catalunya (project 2014SGR-157). Ferran Martín has been awarded an ICREA Academia Award. Christophe Fumeaux acknowledges the Australian Research Council Future Fellowship funding scheme under FT100100585.

References

1. Martín, F., Falcone, F., Bonache, J., Marqués, R. and Sorolla, M. (2003). Split ring resonator based left handed coplanar waveguide, *Appl. Phys. Lett.* **83**(22), pp. 4652–4654.
2. Smith, D. R., Padilla, W. J., Vier, D. C., Nemat-Nasser, S. C. and Schultz, S. (2000). Composite medium with simultaneously negative permeability and permittivity, *Phys. Rev. Lett.* **84**(18), pp. 4184–4187.
3. Pendry, J. B., Holden, A. J., Stewart, W. J. and Youngs, I. (1996). Extremely low frequency plasmons in metallic mesostructures, *Phys. Rev. Lett.* **76**, pp. 4773–4776.
4. Marqués, R., Martín, F. and Sorolla, M. (2008). *Metamaterials with Negative Parameters: Theory, Design and Microwave Applications* (John Wiley, New York).
5. Veselago, V. G. (1968). The electrodynamics of substances with simultaneously negative values of ε and μ, *Sov. Phys. Usp.* **10**(4), pp. 509–514.
6. Eleftheriades, G. V., Siddiqui, O. and Iyer, A. (2003). Transmission line models for negative refractive index media and associated implementations without excess resonators, *IEEE Microw. Wireless Compon. Lett.* **13**, pp. 51–53.
7. Aznar, F., Bonache, J. and Martín, F. (2008). Improved circuit model for left handed lines loaded with split ring resonators, *Appl. Phys. Lett.* **92**, p. 043512.
8. Naqui, J., Fernández-Prieto, A., Mesa, F., Medina, F. and Martín, F. (2014). Effects of inter-resonator coupling in split ring resonator (SRR) loaded metamaterial transmission lines, *J. Appl. Phys.* **115**, p. 194903.
9. Caloz, C. and Itoh, T. (2003). Novel microwave devices and structures based on the transmission line approach of metamaterials, *IEEE-MTT Int. Microwave Symp. Dig.* vol. **1**, pp. 195–198.
10. Martín, F., Falcone, F., Bonache, J., Marqués, R. and Sorolla, M. (2003). Miniaturized CPW stop band filters based on multiple tuned split ring resonators, *IEEE Microw. Wireless Compon. Lett.* **13**, pp. 511–513.

11. Caloz, C. and Itoh, T. (2006). *Electromagnetic Metamaterials: Transmission Line Theory and Microwave Applications* (John Wiley, New York).
12. Durán-Sindreu, M., Vélez, A., Sisó, G., Selga, J., Vélez, P., Bonache, J. and Martín, F. (2011). Recent advances in metamaterial transmission lines based on split rings, *Proc. IEEE* **99**, pp. 1701–1710.
13. Martín, F. (2015). *Artificial Transmission Lines for RF and Microwave Applications* (John Wiley, New York).
14. Herraiz-Martínez, F. J., García-Muñoz, L. E., González-Ovejero, D., González-Posadas, V. and Segovia-Vargas, D. (2009). Dual-frequency printed dipole loaded with split ring resonators, *IEEE Ant. Wireless Propag. Lett.* **8**, pp. 137–140.
15. Herraiz-Martínez, F. J., Zamora, G., Paredes, F., Martín, F. and Bonache, J. (2011). Multiband printed monopole antennas loaded with open complementary split ring resonators for PANs and WLANs, *IEEE Ant. Wireless Propag. Lett.* **10**, pp. 1528–1531.
16. Herraiz-Martínez, F. J., Paredes, F., Zamora, G., Martín, F. and Bonache, J. (2012). Dual-band printed dipole antenna loaded with open complementary split-ring resonators (OCSRRs) for wireless applications, *Microw. Opt. Technol. Lett.* **54**, pp. 1014–1017.
17. Paredes, F., Zamora, G., Bonache, J. and Martín, F. (2010). Dual-band impedance matching networks based on split ring resonators for applications in radiofrequency identification (RFID), *IEEE Trans. Microw. Theory Tech.* **58**, pp. 1159–1166.
18. Mandel, C., Kubina, B., Schüßler, M. and Jakoby, R. (2011). Passive chipless wireless sensor for two-dimensional displacement measurement, *Proc. 41st European Microwave Conf.*, Manchester, UK, pp. 79–82.
19. Puentes, M., Weiss, C., Schüßler, M. and Jakoby, R. (2011). Sensor array based on split ring resonators for analysis of organic tissues, *IEEE MTT-S Int. Microw. Symp.*, Baltimore, MD, USA.
20. Withayachumnankul, W., Jaruwongrungsee, K., Tuantranont, A., Fumeaux, C. and Abbott, D. (2013). Metamaterial-based microfluidic sensor for dielectric characterization, *Sensor Actuat. A Phys.* **189**, pp. 233–237.
21. Ebrahimi, A., Withayachumnankul, W., Al-Sarawi, S. and Abbott, D. (2014). High-sensitivity metamaterial-inspired sensor for microfluidic dielectric characterization, *IEEE Sensors J.* **14**(5), pp. 1345–1351.
22. Abduljabar, A., Rowe, D., Porch, A., and Barrow, D. (2014). Novel microwave microfluidic sensor using a microstrip split-ring resonator, *IEEE Trans. Microw. Theory Tech.* **62**(3), pp. 679–688.
23. Ebrahimi, A., Withayachumnankul, W., Al-Sarawi, S. and Abbott, D. (2014). Metamaterial-inspired rotation sensor with wide dynamic range, *IEEE Sensors J.* **14**(8), pp. 2609–2614.
24. Albishi, A., Boybay, M. S. and Ramahi, O. M. (2012). Complementary split ring resonator for crack detection in metallic surfaces, *IEEE Microw. Wireless Compon. Lett.* **22**(6), pp. 330–332.

25. Boybay, M. S. and Ramahi, O. M. (2012). Material characterization using complementary split-ring resonators, *IEEE Trans. Instrum. Meas.* **61**(11), pp. 3039–3046.
26. Chen, T., Li, S. and Sun, H. (2012). Metamaterials application in sensing, *Sensors* **12**(3), pp. 2742–2765.
27. Ekmekci, E. and Turhan-Sayan, G. (2013). Multi-functional metamaterial sensor based on a broad-side coupled SRR topology with a multi-layer substrate, *App. Phys. A* **110**(1), pp. 189–197.
28. Naqui, J., Durán-Sindreu, M. and Martín, F. (2011). Novel sensors based on the symmetry properties of split ring resonators (SRRs), *Sensors* **11**(8), pp. 7545–7553.
29. Naqui, J. (2015). *Symmetry Properties in Transmission Lines Loaded with Electrically Small Resonators: Circuit Modeling and Applications* (Springer, Berlin).
30. Naqui, J., Durán-Sindreu, M. and Martín, F. (2012). On the symmetry properties of coplanar waveguides loaded with symmetric resonators: analysis and potential applications, *IEEE MTT-S Int. Microw. Symp.*, (Montreal, Canada).
31. Naqui, J., Durán-Sindreu, M. and Martín, F. (2012). Alignment and position sensors based on split ring resonators, *Sensors* **12**(9), pp. 11790–11797.
32. Naqui, J., Durán-Sindreu, M. and Martín, F. (2013). Transmission lines loaded with bisymmetric resonators and applications, *IEEE MTT-S Int. Microw. Symp.*, Seattle, WA, USA.
33. Naqui, J. and Martín, F. (2014). Mechanically reconfigurable microstrip lines loaded with stepped impedance resonators and potential applications, *Int. J. Antennas Propag.* **2014**, ID 346838, 8 pages.
34. Karami-Horestani, A., Fumeaux, C., Al-Sarawi, S. F. and Abbott, D. (2013). Displacement sensor based on diamond-shaped tapered split ring resonator, *IEEE Sensors J.* **13**(4), pp. 1153–1160.
35. Horestani, A., Abbott, D. and Fumeaux, C. (2013). Rotation sensor based on horn-shaped split ring resonator, *IEEE Sens. J.* **13**(8), pp. 3014–3015.
36. Horestani, A. K., Naqui, J., Abbott, D., Fumeaux, C. and Martín, F. (2014). Two-dimensional displacement and alignment sensor based on reflection coefficients of open microstrip lines loaded with split ring resonators, *IET Electron Lett.* **50**(8), pp. 620–622.
37. Naqui, J. and Martín, F. (2013). Transmission lines loaded with bisymmetric resonators and their application to angular displacement and velocity sensors, *IEEE Trans. Microw. Theory Tech.* **61**(12), pp. 4700–4713.
38. Naqui, J. and Martín, F. (2014). Angular displacement and velocity sensors based on electric-LC (ELC) loaded microstrip lines, *IEEE Sensors J.* **14**(4), pp. 939–940.
39. Shaterian, Z., Horestani, A. K. and Fumeaux, C. (2015). Rotation sensing based on the symmetry properties of an open-ended microstrip line loaded with a split ring resonator, *German Microwave Conference* (*GeMiC*), pp. 33–35.

40. Horestani, A. K., Naqui, J., Shaterian, Z., Abbott, D., Fumeaux, C. and Martín, F. (2014). Two-dimensional alignment and displacement sensor based on movable broadside-coupled split ring resonators, *Sensors and Actuators A* **210**, pp. 18–24.
41. Naqui, J., Damm, C., Wiens, A., Jakoby, R., Su, L. and Martín, F. (2014). Transmission lines loaded with pairs of magnetically coupled stepped impedance resonators (SIRs): Modeling and application to microwave sensors, *IEEE MTT-S Int. Microwave Symp.*, Tampa, FL, USA.
42. Pendry, J. B., Holden, A. J., Robbins, D. J. and Stewart, W. J. (1999). Magnetism from conductors and enhanced nonlinear phenomena, *IEEE Trans. Microw. Theory Tech.* **47**, pp. 2075–2084.
43. Marqués, R., Medina, F. and Rafii-El-Idrissi, R. (2002). Role of bianisotropy in negative permeability and left handed metamaterials, *Phys. Rev. B* **65**, paper 144441.
44. Aznar, F., García-García, J., Gil, M., Bonache, J. and Martín, F. (2008). Strategies for the miniaturization of metamaterial resonators, *Microw. Opt. Technol. Lett.* **50**(5), pp. 1263–1270.
45. Schurig, D., Mock, J. J. and Smith, D. R. (2006). Electric-field-coupled resonators for negative permittivity metamaterials, *Appl. Phys. Lett.* **88**, paper 041109.
46. Falcone, F., Lopetegi, T., Baena, J. D., Marqués, R., Martín, F. and Sorolla, M. (2004). Effective negative-ε stop-band microstrip lines based on complementary split ring resonators, *IEEE Microw. Wireless Compon. Lett.* **14**, pp. 280–282.
47. Hand, T., Gollub, J., Sajuyigbe, S., Smith, D. and Cummer, S. (2008). Characterization of complementary electric field coupled resonant surfaces, *Appl. Phys. Lett.* **93**(212504).
48. Naqui, J., Durán-Sindreu, M. and Martín, F. (2013). Differential and single-ended microstrip lines loaded with slotted magnetic-LC resonators, *Int. J. Antennas Propag.* **2013**, Article ID 640514.
49. Naqui, J., Durán-Sindreu, M. and Martín, F. (2014). Selective mode suppression in microstrip differential lines by means of electric-LC (ELC) and magnetic-LC (MLC) resonators, *Appl. Phys. A* **115**, pp. 637–643.
50. Naqui, J., Durán-Sindreu, M. and Martín, F. (2013). Modeling split ring resonator (SRR) and complementary split ring resonator (CSRR) loaded transmission lines exhibiting cross polarization effects, *IEEE Antenn. Wireless Propag. Lett.* **12**, pp. 178–181.
51. Rodgers, J. L. and Nicewander, W. A. (1988). Thirteen ways to look at the correlation coefficient, *American Statistician* **42**, p. 59.
52. Chretiennot, T., Dubuc, D. and Grenier, K. (2013). A microwave and microfluidic planar resonator for efficient and accurate complex permittivity characterization of aqueous solutions, *IEEE Trans. Microw. Theory Tech.* **61**(2), pp. 972–978.

53. Grenier, K., Dubuc, D., Chen, T., Artis, F., Chretiennot, T., Poupot, M. and Fournie, J. (2013). Recent advances in microwave-based dielectric spectroscopy at the cellular level for cancer investigations, *IEEE Trans. Microw. Theory Tech.* **61**(5), pp. 2023–2030.
54. Chen, T., Artis, F., Dubuc, D., Fournie, J. J., Poupot, M. and Grenier, K. (2013). Microwave biosensor dedicated to the dielectric spectroscopy of a single alive biological cell in its culture medium, *IEEE MTT-S Int. Microw. Symp.*, Seattle, WA, USA.
55. Chretiennot, T., Dubuc, D. and Grenier, K. (2013). Optimized electromagnetic interaction microwave resonator/microfluidic channel for enhanced liquid bio-sensor, *Eur. Microw. Conf.* (EuMC), pp. 464–467.
56. Artis, F., Dubuc, D., Fournie, J. J., Poupot, M. and Grenier, K. (2013). Microwave dielectric bio-sensing for precise and repetitive living cells suspension analysis, *Eur. Microw. Conf.* (EuMC), pp. 468–470.
57. Chretiennot, T., Dubuc, D. and Grenier, K. (2014). Double stub resonant biosensor for glucose concentrations quantification of multiple aqueous solutions, *IEEE MTT-S Int. Microw. Symp.*, Tampa, FL, USA.
58. Puentes, M. (2014). *Planar Metamaterial Based Microwave Sensor Arrays for Biomedical Analysis and Treatment* (Springer, Berlin).
59. Su, L., Naqui, J., Mata-Contreras, J. and Martín, F. (2015). Modeling metamaterial transmission lines loaded with pairs of coupled split ring resonators, *IEEE Antenn. Wireless Propag. Lett.* **14**, pp. 68–71.

Index

World Scientific Handbook of

Metamaterials and Plasmonics

Volume 2

Elastic, Acoustic, and Seismic Metamaterials

World Scientific Series in Nanoscience and Nanotechnology*

ISSN: 2301-301X

The Series aims to cover the new and evolving fields that cover nanoscience and nanotechnology. Each volume will cover completely a subfield, which will span materials, applications, and devices.

Published

Vol. 16 *World Scientific Handbook of Metamaterials and Plasmonics (In 4 Volumes)*
Volume 1: Electromagnetic Metamaterials
Volume 2: Elastic, Acoustic, and Seismic Metamaterials
Volume 3: Active Nanoplasmonics and Metamaterials
Volume 4: Recent Progress in the Field of Nanoplasmonics
edited by Stefan A Maier (Imperial College London, UK)

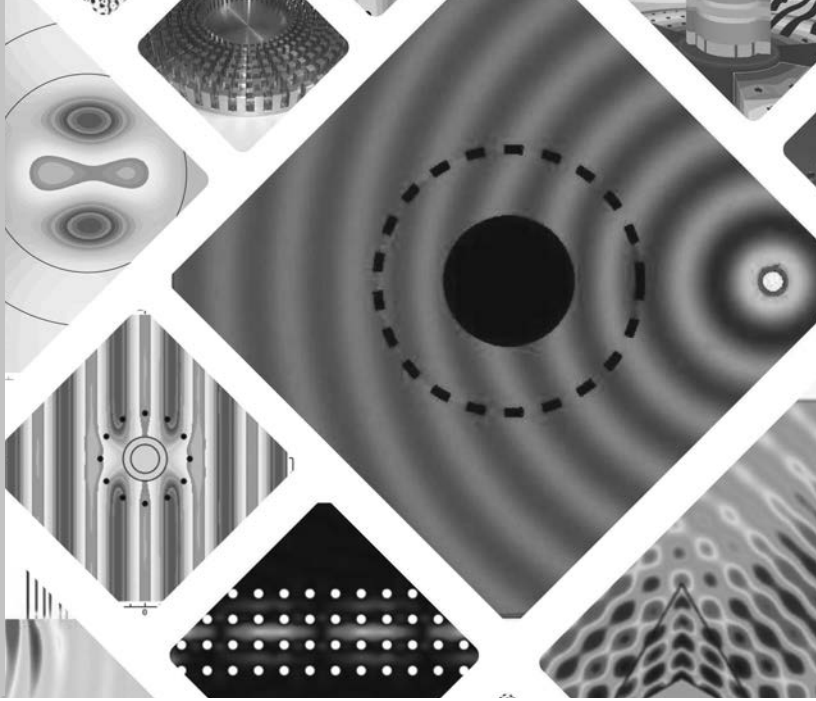
Vol. 15 *Molecular Electronics: An Introduction to Theory and Experiment Second Edition*
by Juan Carlos Cuevas (Universidad Autónoma de Madrid, Spain) and Elke Scheer (Universität Konstanz, Germany)

Vol. 14 *Synthesis and Applications of Optically Active Nanomaterials*
by Feng Bai (Henan University, China) and Hongyou Fan (Sandia National Laboratories, USA)

For further details, please visit: http://www.worldscientific.com/series/wssnn

(Continued at the end of the book)

Volume
16
World Scientific Series in
Nanoscience and Nanotechnology



World Scientific Handbook of
Metamaterials and Plasmonics

Volume 2
Elastic, Acoustic, and Seismic Metamaterials

Volume Editors

Richard Craster
Imperial College London, UK

Sébastien Guenneau
Institut Fresnel, France and Aix-Marseille Université, France

Editor

Stefan A. Maier
Imperial College London, UK

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Cataloging-in-Publication Data
Names: Maier, Stefan A., editor.
Title: World Scientific handbook of metamaterials and plasmonics / edited by:
Stefan Maier (Imperial College London, UK).
Description: Singapore ; Hackensack, NJ : World Scientific, [2017] |
Series: World Scientific series in nanoscience and nanotechnology, ISSN 2301-301X ;
volume 16 | Includes bibliographical references and index.
Identifiers: LCCN 2017029450 | ISBN 9789813227613 (set) | ISBN 9789813227620
(v. 1 ; hardcover ; alk. paper) | ISBN 9813227621 (v. 1 ; hardcover ; alk. paper) |
ISBN 9789813227637 (v. 2 ; hardcover ; alk. paper) | ISBN 981322763X
(v. 2 ; hardcover ; alk. paper) | ISBN 9789813227644 (v. 3 ; hardcover ; alk. paper) |
ISBN 9813227648 (v. 3 ; hardcover ; alk. paper) | ISBN 9789813227651
(v. 4 ; hardcover ; alk. paper) | ISBN 9813227656 (v. 4 ; hardcover ; alk. paper)
Subjects: LCSH: Metamaterials. | Plasmons (Physics)
Classification: LCC TK7871.15.M48 W67 2017 | DDC 621.3028/4--dc23
LC record available at https://lccn.loc.gov/2017029450

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Where applicable, figures in this publication will appear in color for the online version.

Desk Editor: Rhaimie Wahap

Typeset by Stallion Press
Email: enquiries@stallionpress.com

Printed in Singapore

Preface by Main Editor

It is our pleasure to present you this *Handbook of Metamaterials and Plasmonics*, charting the tremendous progress that has occurred in this exciting area of research over the last years. What continues to fascinate me about the field above all is its interdisciplinary breadth — we have arrived at a stage where metamaterials make an impact on many arrays of science where control over waves is a prominent ingredient — be they electromagnetic, acoustic, elastic, or even seismic! In these four volumes, we hence attempt to set out the richness of the field, taking metamaterials in the widest sense as artificial media with sub-wavelength structure for control over wave propagation.

Volume 1 focuses on the fundamentals of electromagnetic metamaterials in all their richness, including metasurfaces and hyperbolic metamaterials. Volume 2 widens the picture to include elastic, acoustic, and seismic systems, whereas Volume 3 presents nonlinear and active photonic metamaterials. Finally, Volume 4 includes recent progress in the field of nanoplasmonics, used extensively for the tailoring of the unit cell response of photonic metamaterials.

In its totality, we hope that this handbook will be useful for a wide spectrum of readers, from students to active researchers in industry, as well as teachers of advanced courses on wave propagation. I want to thank the volume editors Ekaterina Shamonina,

Richard Craster, Sébastien Guenneau, Ortwin Hess and Javier Aizpurua, and all the authors for their excellent and sustained work to put these four volumes together.

Stefan A. Maier
Imperial College London, UK
September 2016

Preface by Volume Editors

In December 2013, the United Nations proclaimed 2015 as the International Year of Light and Light-based Technologies. Metamaterials were amongst the most popular subjects in public lectures delivered in 2015 with topics such as invisibility cloaks attracting popular attention. However, there is more potential to metamaterials than simply one research area: correspondences between the governing equations for electromagnetic waves and acoustic, elastodynamic and hydrodynamic waves suggest that topics such as invisibility cloaks could translate into these fields and find exciting applications. For instance, a cloak for pressure waves, the sound of silence, might improve the acoustics of concert halls by acoustically concealing columns, or a seismic metamaterial provide protection from earthquakes by rerouting or diverting seismic waves. Allied to designing or creating devices, many interesting conceptual questions naturally arise, for instance one question from a public audience was about the reciprocity principle: If light cannot reach a fictional character covered by an invisibility cloak, can he or she see the outer world? In fact, he or she would be in complete darkness if the cloak would have no eyeholes; these two defects in the cloak could then be perceived by an outer observer. Similarly, in other wave systems, one has to think of the consequences of such cloaking or devices, for instance, the invisibility region should act as a quiet zone if one were

to design a cloak for acoustic waves. The present acoustic volume of the Handbook for metamaterials will present many more potential applications of acoustic, hydrodynamic and seismic metamaterials. This collection of chapters by leading international experts in the fields of acoustic and elastodynamic waves includes a survey of space–time transformations as a design tool for acoustic metamaterials underpinning intimate connections between Maxwell's equations and Einstein's theory of relativity, the book then moves on to practical applications in the control of mechanical waves, from the laboratory scale (micrometer to meter) the civil engineering (meter to decameter) and geophysics (decameters up to kilometers) scales. All these contributions promise to revolutionize ways of controlling the propagation of sound, light, and any particular form of waves at macroscopic and microscopic scales. Indeed, potential applications range from subwavelength lensing and time reversal, to underwater camouflaging and electromagnetic invisibility, to enhanced biosensors and protection from harmful physical waves (e.g., tsunamis and earthquakes). This volume covers theoretical as well experimental aspects in these different areas that include nanoscale (plasmonics) and meter-scale (geophysics) media. The outline of the chapters is as follows:

The theory of homogenization is a vast and highly interdisciplinary topic (rooted in the theory of composites, calculus of variations, asymptotic analysis etc.) that can be traced back to Lord Rayleigh's work on effective medium formulae towards the end of the nineteenth century. However, Rayleigh's work was limited to the static, infinite wavelength, case and much of the research progress in homogenization theory throughout the twentieth century has been to add dynamic corrections to Rayleigh's formulae in the quasi-static limit, with notably so-called dipole approximations in the dilute composite limit. Interestingly, for densely packed and high-contrast, composites, alternative asymptotic theories have been developed by groupings in Australia, France, USA and Russia. For wavelengths on the order of the periodicity of the composite, which is of particular physical interest as this allows for analysis of effective properties near

stop bands, progress has been made using high-frequency homogenization (HFH). A comprehensive survey of HFH and its application to problems of interest to the metamaterial community in acoustics and elastodynamics is presented by the group of Richard Craster, in Chapter 1. HFH actually encompasses long-wavelength and high-contrast homogenization theories, and only requires a priori knowledge of eigenfrequencies and associated eigenfields associated with Floquet-Bloch standing waves within an infinite periodic structure, to recover the fine features of any wave propagating in the same structure when it has some defects or when it is finite (what makes possible the study of scattering problems by finite size photonic/phononic crystals and metamaterials).

The second chapter by the group of Nicholas Fang at Massachusetts Institute of Technology is a review of inspiring lensing and cloaking results with acoustic metamaterials. The authors start with a general overview of key advances in this field, and then embarks the reader on a scientific journey where notions of effective dispersive parameters with non-local properties at low frequencies (similar to those achieved via high frequency homogenization in Chapter 1), are deduced from models Macroscopic Maxwellian acoustics and useful analogies between electromagnetic and acoustic meta-atoms. The concept of meta-atoms underpins much of the averaging theories of metamaterials and was put forward by Sir John Pendry at the turn of this millennium. Pendry's perfect lens via negative refraction is revisited in light of simultaneously negative density and compressibility modulus. Theory of transmission line network is then used to design acoustic cloaks, the touchstone being the Telegraph equation. This chapter finally discusses acoustic metasurfaces, and here again the averaging approach chosen by the authors proves very useful. Importantly, theory and numerics are nicely complemented by experimental data.

The third chapter by Patrick Sebbah at Institut Langevin in Paris and Marc Dubois at University of California Berkeley discusses theoretical and experimental results on flat lens focusing of flexural waves in a thin plate structured with air holes. A review

of all-angle-negative refraction in periodic structures is made, with analogies drawn between light and sound, in the framework of the Floquet-Bloch theory for time-harmonic waves. The authors then move on to the fascinating problem of focusing and high-resolution for time driven super oscillations, this is the concept of a convergent flat lens with subwavelength resolution without evanescent waves, which is an interesting alternative to Sir John Pendry's perfect lens.

As we have seen in the preceding chapter, control of space and time can break physical barriers. Chapter 4 on space-time cloaking starts with a survey by Martin McCall at Imperial College London and Paul Kinsler at Cockroft Institute and Lancaster University of space-time transformations as a design tool for metamaterials underpinning intimate connections between Maxwell's equations and Einstein's theory of relativity, and their extension to practical applications in moving electromagnetic media and acoustic waves. Quite amazingly, an experiment on a temporal cloak at telecommunication data rate combining a split-time lens with single-mode and dispersion compensating optical fibres has validated these theoretical concepts back in 2012, and this scientific odyssey is recalled together with a proposal for a digital signal processing circuit modified to incorporate a spacetime cloak.

Chapter 5 by the group of Mathias Fink at Institut Langevin in Paris, provides core material on acoustic metamaterials, demonstrated with a popular science device consisting of soda cans. A standard time reversal experiment, waves generated by a source are first measured by an array of antennas positioned around the source and then time reversed and simultaneously rebroadcasted by the same antenna array. Due to the time invariance of the wave process, the re-emitted energy will focus back on the original source, whatever the complexity of the propagation medium. This chapter concentrates on the application of negative refraction (associated with Bragg interferences or local resonances) and time reversal to the high-resolution focusing and manipulation of sonic waves, and as its title suggests it also has strong connections with the first chapter beyond classical homogenization.

John William Strutt, 3rd Baron Rayleigh is most famous among the wave community for Rayleigh scattering (this explains why the sky is blue) and Rayleigh waves that are, with Love waves, responsible for many of the earthquake disasters in human infrastructures. Rayleigh's textbook, "*The Theory of Sound*", has been together with Augustus Edward Hough Love's monograph "Some Problems of Geodynamics", an invaluable source of inspiration for generations of physicists, engineers, and mathematicians. The sixth chapter of this volume by the geophysics group of Philippe Roux and his collaborators in mathematics and physics fields, is a detailed analysis, based on numerical simulations, experimental observation and physical interpretation in light of effective properties near stop bands, of an extreme control of flexural (Lamb) waves in plates structured (at the centimeter scale) with a forest of thin straight rods and Rayleigh surface waves in soils structured (at the meter scale) with a forest of trees. The starting point of this work is an experiment where plate waves interact with a dense collection of sub-wavelength resonators that consist of long vertical rods attached to the plate. Wide band gaps are observed at the laboratory scale inside this locally-resonant metamaterial as well as sub-wavelength and supra-wavelength modes from which optimal sub-wavelength focusing can be achieved inside the metamaterial. When combined with numerical simulations and theoretical predictions, the full understanding of the complex wave interaction can be performed, which then leads to potential cloaking applications when the forest of vertical rods is shaped accordingly. Finally, the transition from the laboratory configuration to the geophysical scale is considered with vertical trees in a dense natural forest playing the role for seismic surface waves of the rods attached to the plate for Lamb waves. This review paper gathers results from a collection of papers published since 2013 on this plate and rods configuration and its natural generalization to geophysics scale that were interpreted through the physics of locally-resonant metamaterials.

The previous chapter can be considered as a milestone in the emerging field of seismic metamaterials, although this terminology remains debatable. Indeed, metamaterials are composites with

extraordinary properties that were introduced for electromagnetic waves by Sir John Pendry, a British physicist working at Imperial College London, and his colleagues at GEC-Marconi Materials Technology Ltd. (A. J. Holden, D. J. Robbins, W. J. Stewart) and the Defence Research Agency Holton Heath (I. Youngs) in the UK and at the University of California of San Diego and Duke University in the USA, towards the end of the twentieth century, following the advent of photonic crystals (PCs). Unlike for seismic metamaterials, that are structured on the meter scale, electromagnetic metamaterials are concerned with wave wavelengths on the order of a few hundredths of nanometers (for visible light) and their fabrication requires advanced nanotechnologies. To achieve the tour de force of an invisibility cloak for microwaves (whose wavelengths are on the order of a few centimeters), Sir John Pendry and his colleague David Smith proposed to combine electrical circuits with intertwined split-ring resonators and thin, straight wires, which exhibit, respectively, a strongly dispersive anisotropic permeability and permittivity. With a bit of imagination, one might argue that thin-straight wires are reminiscent of the forest of trees in Chapter 6, and an interesting parallel could be drawn. Chapter 7, by the civil engineering group of Stéphane Brûlé at Ménard company, physicists at Institut Fresnel and a mathematician at Imperial College, actually focuses its attention on control of soils structured by an array of boreholes (that are more akin to photonic crystals than metamaterials, as they essentially work in the Bragg regime), that have been shown to allow for shielding and focusing effects. Interestingly, the first experimental proof of negative refraction came in 2000 with the team of David Smith, inspired by Victor Veselago's 1968 proposal of a flat convergent lens via negative refraction, which was popularized by Sir John Pendry (who pointed out the theoretically infinite resolution achieved by capturing not only the propagating but also the evanescent components of the source in the image construction), and a double focusing effect of Rayleigh waves through an array of boreholes observed by Brûlé's team in a field test near Lyon in France suggests some similar ray trajectory to the Veselago-Pendry lens.

The eighth chapter by the group of Alexander Movchan at Liverpool University points out that external cloaking, an emerging topic first touched upon by Nicolae Nicorovici, Ross McPhedran and Graeme Milton in 1994, can be seen as a look-a-like contest: the goal is to make some object A surrounded by active sources appear like another object B with respect to some physical observable. Early mathematical literature has indeed spoken of the nonuniqueness of the tomography (inverse) problem with the works of Calderon (1980), Kohn and Vogelius (1984), and Greenleaf *et al.* (2003). The focus of this chapter is on acoustic and elastic plates, but as noted in the preceding two chapters, one can envisage to scale up such designs to geophysics and end up with seismic external cloaks.

The ninth chapter by Richard Porter at University of Bristol is a thorough review of various cloaking routes towards control of surface water waves. These approaches include notably the beautiful design of cloaks with spatially varying bathymetry in water depth. Porter uses not only conformal but also non-conformal mappings, in the Helmholtz and mild-slope equations and he investigates various routes that have either great potential or, equally important, are found to be cul-de-sac.

The tenth chapter on molding sound propagation and scattering with acoustic metamaterials and metasurfaces by Romain Fleury and Andrea Alù at the University of Texas at Austin aims at providing a comprehensive insight into recent developments and applications of sound manipulation via anomalous matching and tunneling effects, some analogue of plasmonic cloaking for acoustic waves, as well as impedance matching phenomena with near zero effective density (so-called acoustic DNZ which is somewhat a counterpart of ultra-refractive optics). This chapter also invites the reader to read more about active acoustic metasurfaces obeying parity-time symmetry conditions, which can induce anomalous tunneling effects with large reflection asymmetry, and the potential to relax some of the limitations associated with bulk passive 3D metamaterials, including sensitivity to losses, bandwidth of operation and imperfect isotropy. Some aspects of linear nonreciprocal acoustics are presented in light

of the by now famous acoustic circulator, which gives birth to a new field in acoustics akin to magneto-optics.

The eleventh chapter on the carpet cloak by Fu Liu and Jensen Li at Birmingham University (UK) and Jingjing Zhang at Nanyang Technical University (Singapore) concerns the observables light and sound, in the framework of quasi-conformal mappings. The concept of ground cloak was first proposed in 2008 by one of the authors in collaboration with Sir John Pendry, and has given rise to numerous theoretical and experimental works. Notably, carpets with linear boundaries lead to piecewise homogeneous and anisotropic material parameters which can then be achieved via classical homogenization formulas (for layered media).

The twelfth chapter by Andrew Norris and William Parnell on Hyperelastic cloaking: transformation elasticity with pre-stressed solids is concerned with transformational elastodynamics, a counterpart of transformational optics dealing with coordinate transforms in Navier, rather than Maxwell's, equations. including pre-stressed media. Navier equations, unlike the Helmholtz equation for acoustics of the preceding chapter, do not retain their form in general. As discovered by Graeme Milton, Marc Briane and John Willis in 2006, the former equations usually take the form of transformed Willis equations with not only an anisotropic heterogeneous rank-4 tensor of elasticity, but also rank-3 tensors of inertia and viscosity, and a rank-2 tensor of density. However, there are certain cases of transformed Navier equations, which have simpler structure such as a Navier equation that retains its form at the cost of a symmetry breaking of the elasticity tensor, or pre-stressed media, as unveiled by Norris and Parnell in 2012. The authors review this emerging field of transformational elastodynamics and make it as fun and rigorous as it possibly can be.

The thirteenth chapter on acoustic metamaterials with conical dispersion by Xueqin Huang and Meng Xiao and Che Ting Chan at Hong Kong University of Science and Technology and Fengming Liu at Hubei University of Technology (China) is concerned with graphene-like acoustic metamaterials. The physicists Sir Konstantin

Novoselov and Andre Geim shared a Nobel Prize in Physics in 2010 for 'groundbreaking experiments regarding the material graphene'. In the tracks of this revolution in condensed matter physics, the grouping of Prof. Chan has made long lasting contributions to the study of Dirac and Weyl points in periodic acoustic structures with unit cells and supercells, respectively, leading to exciting effects such as extraordinary transmission and cloaking associated with dynamic effective anisotropy at certain frequencies. Numerical simulations are complemented by analytic expressions that capture the essence of the physical mechanisms underpinning long wave high frequency effects near Dirac cones (as unveiled by the Bloch Hamiltonian which is akin to that in the electronic band structure of graphene). These beautiful results and their physical insight can be paralleled to predictions using High Frequency Homogenization as shown in Chapter 1.

The fourteenth, and last, chapter of the book on an interface model for homogenisation of acoustic metasurfaces by Agnes Maurel at Institut Langevin in Paris and Jean-Jacques Marigo at École Polytechnique in Palayseau takes us full circle as these researchers take us on a journey to the homogenization of structured interfaces for acoustic waves, and some of the asymptotic tools they introduce can be paralleled with those in the preceding chapter and chapter 1 on HFH. More precisely, Maurel and Marigo are concerned with ultra-thin metamaterials that can be replaced by anisotropic dispersive and non-local effective interfaces that encapsulate the boundary layer effects. Matched asymptotic expansions are the tools of choice to identify the limit problems, and asymptotic predictions are compared against fully numerical solutions to validate this approach.

This acoustic volume of the Handbook of Metamaterials and Nanophotonics touches upon many hot subjects in the mathematics and physics of phononic crystals and metamaterials, which were discovered less than a quarter of a century ago. Much remains to be discovered, and doubtless the future will be exciting, we hope that the present book will help foster theoretical and experimental efforts in metamaterials. We stress that all these contributions promise to revolutionize ways of controlling the propagation of sound, light, and

any particular form of waves at macroscopic and microscopic scales. Indeed, potential applications range from subwavelength lensing and time reversal, to underwater camouflaging and electromagnetic invisibility, to enhanced biosensors and protection from harmful physical waves (e.g., tsunamis and earthquakes). This volume covers theoretical as well as experimental aspects in these different areas that include nanoscale (plasmonics) and meter-scale (geophysics) media. The fourteen chapters constituting this acoustic volume give a comprehensive survey of recent advances in this field.

We would like to convey our warmest thanks to all chapter authors, who are the principal architects of this volume, for their excellent scientific contribution and their willingness to share their knowledge of acoustic metamaterials. The assistance and professionalism of the World Scientific team is also greatly acknowledged. We hope that you will enjoy reading these chapters and find these as informative as we did.

Richard Craster
London

and

Sébastien Guenneau
Marseille

September 2016

Contents

CHAPTER 1

Dynamic Homogenization of Acoustic and Elastic Metamaterials and Phononic Crystals

RICHARD CRASTER*

Imperial College London, UK

TRYFON ANTONAKAKIS

Multiwave Technologies, Switzerland

SÉBASTIEN GUENNEAU

Institut Fresnel and Aix-Marseille Université, France

1.1. A Brief History of Homogenization

It is naturally highly subjective, given the vast literature on homogenization theories, to even attempt to chart the journey from its source in static and quasi-static effective media to its current location in the area of waves through periodic media. Limitations of space mean that we will unfortunately omit some routes and hence the path we describe is a personal one and cannot be exhaustive, none the less

*Corresponding author: r.craster@imperial.ac.uk

we hope to capture a flavour of the journey. Although of a genuinely cross-disciplinary nature, homogenization theories take their root in problems of electrostatics and magnetostatics in the 19th century. The philosophy expressed by Lord Rayleigh in 1892 concerning the Lorentz-Lorenz equations acts as the foundation of homogenization theory, at least from a physical standpoint: "*In the application of our results to the electric theory of light we contemplate a medium interrupted by spherical, or cylindrical, obstacles, whose inductive capacity is different from that of the undisturbed medium. On the other hand, the magnetic constant is supposed to retain its value unbroken. This being so, the kinetic energy of the electric currents for the same total flux is the same as if there were no obstacles, at least if we regard the wave-length as infinitely great. [Strutt (1892)]*", which sets the scene for contemplating an effective medium at least in a quasi-static sense.

1.1.1. *Early history and static homogenisation*

As far as we know, the first theoretical works related to homogenization in electromagnetism began with the static effective medium models of Ottavanio Fabrizio Mossotti in 1836, and Michael Faraday in 1837, which proceeded in a similar way to analyse the distribution of electric charge on several metallic bodies embedded in a dielectric medium; Mossotti's analysis was based on earlier closely related work by Siméon Poisson on magnetic media. The latter analysed the interaction between the polarizable entities, invoking the first of many cavity considerations in the derivation of the effective field for a dielectric medium. The English version of Mossotti's paper [Mossotti (1850)] has had a long lasting influence in the theory of composites [Milton (2002)]. Historically, since Mossotti was made prisoner by the Austrians during the first war for Italian independence (1848–49), his paper first submitted in 1846 was only published in 1850. A subsequent step was taken by Ludvig Lorenz, whose 1869 and 1875 papers in Danish were succeeded by one in 1878 by Hendrik Lorentz [Born and Wolf (1970)]. In their papers, Lorenz and Lorentz assign a refractive index to the interior of the molecules, which differs from that of the surroundings and then proceed in a way akin to modern

wave propagation theories in stochastic media: in the presence of the irregularly distributed molecules Lorenz and Lorentz search for the value of the propagation constant of light such that the deviations from this well behaved *sine*-wave average out to zero, and do not build up. From a mathematical standpoint the Lorenz and Lorentz effective medium theory amounts to considering a large sphere cut out of the dielectric, centered about the molecule in consideration, to derive the effective macroscopic field. The large sphere separates the far away molecules, which can be treated as a continuum, from the close molecules which have to be taken into account more explicitly. In contrast to Lorenz, Rudolf Clausius assumed [Clausius (1858)], as did Mossotti and Faraday, that molecules are conducting spheres, and considered each molecule as contained in a small sphere tightly cut out of the continuous surrounding (perforated) medium, which is characterized by the final dielectric constant of the material.

All these works, which have a strong physical flavour, set the scene for the 1892 foundation paper of John William Strutt, the third Lord Rayleigh, who was able to mathematically solve Laplace's equation in two dimensions for rectangular arrays of cylinders, and in three-dimensions for cubic lattices of spheres. Although a mathematical *tour de force*, the original proof of Rayleigh suffers a conditionally convergent sum in order to calculate the dipolar field within the array. Many authors proposed extensions of Rayleigh's method to avoid this drawback, amongst them George Batchelor with an influential paper on micro-hydrodynamics of dilute dispersion of spheres [Batchelor (1972)]. Another limit of Rayleigh's method is that it does not hold when the volume fraction of the inclusions increases: there is a point where there exists a connected path through the material, along which a current may flow. This phenomenon is known as a percolation threshold, which has been extensively studied by Dirk Bruggeman [Bruggeman (1935)], who developed two effective medium theories for dielectric materials in 1935–1936. The first one made a distinction (like previous theories), between the inclusions and the background matrix, whereas the second one exhibited a percolation threshold; his two theories agreed well with the Lorenz-Lorentz equation for the small volume fractions.

An important act in the history of homogenization theory, as applied to waves (which propagate with a given, non-zero, frequency), takes place in the Electrical, Transport and Optical Properties in Inhomogeneous Media (ETOPIM) congress of 1977, where Ross McPhedran and David McKenzie proposed to extend Rayleigh's fundamental work of 1892 (replacing Laplace by Helmholtz equation). At this occasion further developments were discussed with Ping Sheng, David Bergman and Graeme Milton in relation with new mathematical techniques, which appeared one year after in a fundamental book by Alain Bensoussan, Jacques-Louis Lions and George Papanicolaou [Bensoussan *et al.* (1978)], which remains a treasure trove of ideas. In this book, asymptotic methods such as multiple scale expansions were applied to governing equations for conductivity, heat, electrostatics, magnetostatics, elastostatics, elastodynamics and so forth. The first mathematical results obtained in conjunction with abstract mathematical homogenization theory were in fact that of the Italian mathematician Sergio Spagnolo in 1968. Spagnolo studied elliptic and parabolic problems of diffusion, when the operators were symmetric, and he chose the term of G-convergence to refer to convergence of Green's functions [Spagnolo (1968)], which clearly is a reference to earlier work by physicists. The term homogenization itself was first introduced in mathematics by Ivo Babuska [Babuska (1975)].

In the early seventies, Luc Tartar and François Murat developed the powerful mathematical tool of H-convergence [Murat and Tartar (1978)] (a name chosen in reference to G-convergence) for mathematical homogenization theory. In 1974, Tartar described the notion of weak convergence (introducing a mathematical tool known as compensated compactness), to link microscopic and macroscopic scales [Tartar (1975)] in relation to effective properties in homogenization, which were derived heuristically by Sanchez-Palencia [Sanchez-Palencia (1972)], thanks to asymptotic expansions for periodic structures (note that Sanchez-Palencia was inspired by similar multiple scale methods used to approximate solutions of nonlinear oscillation problems in the applied mathematics and mechanical engineering communities, not for analyzing spatial oscillations

as in homogenization). In 1974, Murat referred to the method of H-convergence (H for homogenization) in a famous seminar he gave in Alger. The H-convergence approach inspired probabilistic models, such as those of George Papanicolaou and Sathamangalam R. Srinivasa Varadhan [Papanicolaou and Varadhan (1978)] and Serguei Kozlov [Kozlov (1977)]. Ten years ago, Blanc, Le Bris and Lions applied stochastic homogenization theory [Blanc *et al.* (2007)] to linear elliptic problems associated with random media characterized by a diffusion coefficient combining a periodic function with a random diffeomorphism, which models random media that are small perturbation of periodic media. This approach can also be used to model periodic media with defects. Bouchitté, Guenneau and Zolla [Bouchitté *et al.* (2010)] proposed an alternative approach to homogenization of other types of almost-periodic media based on the cut-and-projection method, which amounts to applying homogenization tools in an upper periodic dimensional space. Importantly, the homogenized problem derived in [Bouchitté *et al.* (2010)] making use of abstract functional spaces in Fourier space has been recently recast in a mathematical framework compatible with two scale asymptotic expansions [Wellander *et al.* (2017)]. This opens new avenues for asymptotic analysis of quasi-periodic structures with multiple scale methods.

Overall, a great number of papers have appeared on homogenization theory in electromagnetism in the past forty years. To choose, but a couple of directions: the extension of Rayleigh's technique to arbitrary multipole orders for cylindrical and spherical inclusions [Perrins *et al.* (1979)], to formulae for the effective conductivity and the investigation [McPhedran *et al.* (1988)] of models for arrays of circular inclusions, which were both close to touching and highly conducting. These multipole methodologies, although limited to cylinders and spheres are powerful and provide considerable insight. Connecting with a later chapter, another limiting case of great physical interest is the homogenization of gratings to get effective gratings as studied by Bouchitté and Petit about thirty years ago [Bouchitté and Petit (1985); Bouchitté (1990)], i.e. for structured

interfaces, and it is interesting to note that the last Chapter 14, of this volume by Maurel and Marigo builds upon this work to propose a new homogenization theory devoted to lower dimensional periodic structures.

1.1.2. *Quasi-static extensions and multipole methods*

Thus far we have only considered static problems, and the results can be vastly different from dynamic problems — except for very long waves relative to the microstructural lengthscale.

In the case of a dynamic problem in a periodic medium, say a periodic arrangement of rods (photonic crystal fiber, PCF [Russell (2003)]), the solutions are in general not periodic (they undergo a phase shift on opposite edges of periodic cells) and exist in modes, or energy bands, which are functions of the spatial frequency of the solution. The behaviour of waves that exist in the system is described by a dispersion diagram relating frequency to the wavenumber of the phase shift, a typical set of dispersion curves for a PCF preform, is shown in Fig. 1.1. There is a substantial literature on periodic systems and we describe some of this in the next section. The static homogenisation simply captures the behaviour of the system at the origin (see Γ point on the right panel of Fig. 1.1.) and nowhere else, clearly it makes sense to extend the static theory to describe the linear pieces of the dispersion curves near the origin and such quasi-static extensions are often made.

For instance, the so-called Generalized Rayleigh Method for circular inclusions, inspired by Rayleigh's technique for static problems [Strutt (1892)], can be utilised for a wide range of physical systems.

A dynamic theory can be used to solve Maxwell's equations in a heterogeneous periodic medium and then take the long wavelength limit in order to determine the effective properties of the structures in this long wave limit. This methodology expands the solution in terms of a set of functions that satisfy the boundary conditions exactly on the inclusion. The periodicity conditions are then satisfied approximatively, by the appropriate choice of expansion coefficients [Chin *et al.* (1994); Nicorovici *et al.* (1995); McPhedran

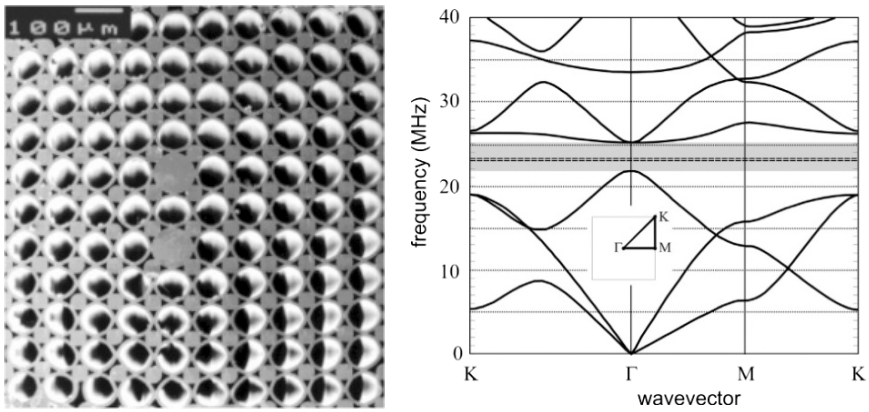


Fig. 1.1. (Left panel) Scanning electron micrograph of the preform used in a son et lumière experiment in a photonic crystal fibre preform by the group of Philip Russell at Bath University. PCF has two solid defects, an inter-hole period of 80 μm, a hole diameter of 59 μm, and an interstitial hole diameter of 8 μm. (Right panel) Band structure for in-plane mixed-polarized shear and dilatational waves in the sonic crystal depicted in left panel, with defects removed but including interstitial holes. The experimentally observed resonances (at 23 and 23.25 MHz — the dashed lines) sit near the middle of the sonic band gap, which extends from 21.8 to 25 MHz. Figure taken from [Russell *et al.* (2003)].

et al. (1997)], large matrix systems of equations ensue, but simplifications occur in the long wave limit. These quasi-static low frequency theories capture the behaviour near the Γ point in Fig. 1.1 but cannot capture any of the behaviour at higher frequencies (e.g. the nearly flat band on the upper edge of the stop band has potential applications in slow wave guiding).

The multipole method takes its root in an aforementioned seminal paper of Lord Rayleigh (1892) which is the foundation stone upon which the current edifice of homogenization theories is built. More precisely, in that paper, John William Strutt, the third Lord Rayleigh, solved Laplace's equation in two dimensions for rectangular arrays of cylinders, and in three dimensions for cubic lattices of spheres. However, a limitation of Rayleigh's algorithm is that it does not hold when the volume fraction of inclusions increases, and this pitfall has also been observed for elastic waves [Zalipaev *et al.*

(2002); McPhedran *et al.* (2009)]. Multipole methods, in conjunction with lattice sums, overcome such obstacles and lead to the Rayleigh system which is an infinite linear algebraic system; this formulation, in terms of an eigenvalue problem, facilitates the construction of dispersion curves and the study of both photonic and phononic band-gap structures. In the limit of small inclusion radii, when propagating modes are very close to plane waves, one can truncate the Rayleigh system ignoring the effect of higher multipoles, to produce a series of approximations each successively more accurate to higher values of filling fraction. At the dipole order, one is able to fit the acoustic band, which has linear behaviour in the neighbourhood of zero frequency, see Fig. 1.1 (Left panel), except of course in the singular case whereby Dirichlet data are enforced on the boundary inclusions [Poulton *et al.* (2001, 2004)] and there is a zero-frequency stop band.

1.1.3. *Photonic crystals and other advances in periodic structures*

Until relatively recently, attention in terms of homogenisation theory was mainly focussed on the so-called quasi-static limits discussed above. Recent developments have been driven by the discovery, implementation, and extensions of, exotic wave phenomena created by waves through periodic structured media.

Thirty years ago, the physicists Eli Yablonovitch and Sajeev John published two seminal papers [Yablonovitch (1987); John (1987)] in the same issue of Physical Review Letters, that reported the discovery of composites that became known as photonic crystals (PCs). PCs are periodic structures displaying a photonic band spectrum similar to that of semiconductors in electronic band diagrams, at wave frequencies away from the quasi-static limit (beyond the models of low frequency homogenization). Ultimately this has driven the interest in generating homogenisation theories or effective media models valid at high frequencies that are no longer limited by quasi-static low frequency assumptions.

There is now, a huge literature on photonic crystals, the scale of which can be seen from Jonathan Dowling's webpage at Louisiana State University on photonic and sonic band gaps and metamaterials that started in 1994 and ceased in 2008 as the number of papers exceeded 12,000. Thus we outline our own personal introduction to the topic and journey through it.

In the early nineties, in the tracks of Yablonovitch and John, Sir John Pendry and Daniel Maystre and their colleagues at the Condensed Matter Theory Group (CMTH) at the Physics Department of Imperial College London and at the Laboratoire Optique Electromagnetique (LOE) of Aix-Marseille University started to work on PCs for perfect dielectric mirrors, absorbers, and photonic band gap light guidance. Exchange of postgraduate students between the photonics groups in London and Marseille were subsequently initiated about twenty years ago. In parallel, strong links were maintained between LOE and the group of Ross McPhedran in University of Sydney, who had been working in collaboration with Maystre on the theory of gratings since the early seventies, when he was a postdoc in Marseille. The rich common history of the photonics laboratory in Marseille, nowadays known as Institut Fresnel, and the laboratory CUDOS in Sydney, led to the so-called yellow and white books. The yellow book on Theory of Gratings has been edited by Roger Petit and published by Springer in the early eighties, [Petit (1980)] and the white book on Foundations of Photonic Crystal Fibres has been edited by Frederic Zolla *et al.* and published by Imperial College Press 12 years ago [Zolla *et al.* (2005)]. One of the authors of this chapter (Richard Craster, from the Mathematics Department at Imperial College), was a critical reader of a preliminary version of the white book back in 2003–2004, and he initiated scientific collaboration on periodic acoustic waveguides with Sébastien Guenneau who was back then a research associate in Physics at Imperial. The demand for asymptotic theories devoted to Crystal Fibres for sound and light control through Brillouin scattering, an example of which is shown in Figure 1.1, was one of the physical motivations for such advanced mathematical models, see [Russell *et al.* (2003)] for more details.

Actually, two other laboratories have had a long lasting collaboration, namely the Department of Mathematical Sciences of Alexander Movchan at Liverpool University and the Royal College of Sciences of Sir John Pendry at Imperial College London. It is Ross McPhedran who initiated the exchange programs between Master and PhD students, Postdocs and permanent staff between the photonic and mathematical groups in London, Liverpool, Marseille and Sydney. During his seven years of postdoctoral studies in UK, one of us has initiated a new collaboration between Institut Fresnel and the Department of Mathematical Sciences at Imperial College London, and as a result a co-edited book on Acoustic Metamaterials was published by Springer in 2012 [Craster and Guenneau (2012)]. Electromagnetic and acoustic metamaterials are periodic structures displaying exotic effective properties such as negative refraction (with the electromagnetic paradigm of Pendry's perfect lens), cloaking with strong artificial anisotropy(see Fig. 1.2 for an application of this concept to water waves) at low frequencies and can be considered an extension of photonic crystals where subwavelength resonance becomes important.

1.1.4. *Band gap and related phenomena*

Over the last 30 years, many significant advances have created a deep understanding of the optical properties of photonic crystals (PCs), as described in [Yablonovitch (2001); Zolla *et al.* (2005); Joannopoulos *et al.* (2008)], and such periodic structures prohibit the propagation of light, or allow it only in certain directions at certain frequencies, or localize light in specified areas, and this occurs for high-frequencies [Villeneuve and Piché (1992)], and this behaviour has been also observed for acoustic waves [Sigalas (1998)]. This sort of metamaterial (using the loose terminology that this means artificial materials engineered to have desired properties that may not be found in nature, such as negative refraction, see e.g. [Pendry (2004); Ramakrishna (2005); Guenneau and Ramakrishna (2009)]) enables a marked enhancement of control over light propagation.

PCs are periodic media whose spectrum is characterized by photonic band gaps and pass bands [Conca *et al.* (1995); Gazalet *et al.*

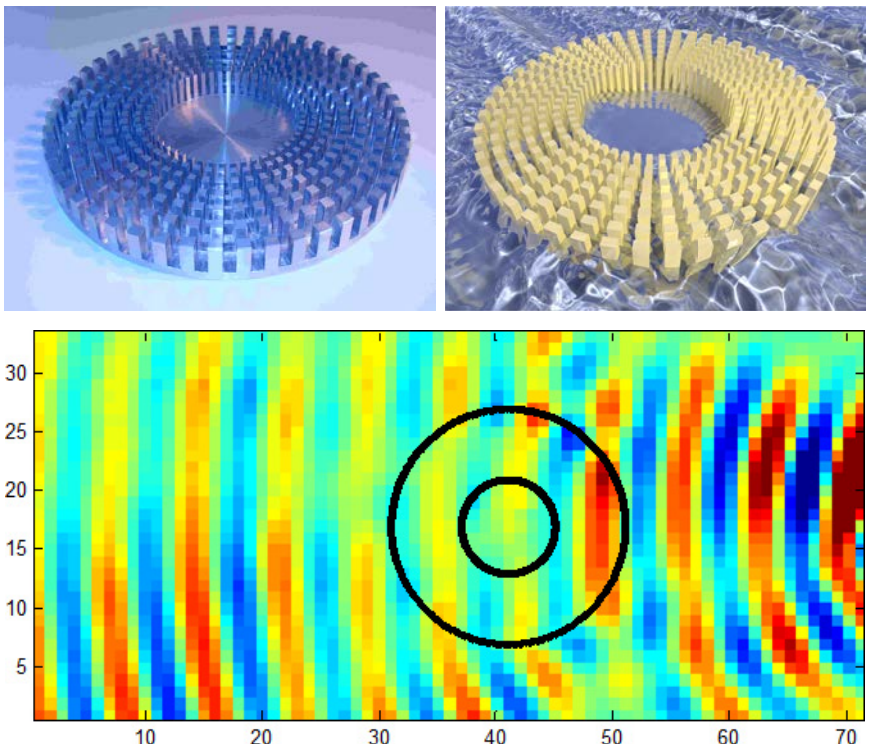


Fig. 1.2. (Upper left panel) Photo of an aluminium cloak (20 cm in diameter) working for water, sound and transverse electric waves (courtesy of Stefan Enoch at Institut Fresnel). (Upper right panel) Artistic view of a scaled-up version of the aluminium cloak immersed in ocean for protection (still water) purpose (courtesy of Martin Wegener's group at KIT). (Lower panel) Measured field map (with green color for vanishing field) of a transverse electric wave incident upon aluminium cloak (whose inner and outer boundaries are marked by black circles) at 4.3 GHz (courtesy of Redha Abdeddaim at Institut Fresnel). One notes the almost unperturbed field outside the cloak (invisibility) and the vanishing field inside its center (protection).

(2013)], just as electronic band gaps exist in semiconductors. In PCs, light propagation in particular directions can be disallowed for certain frequencies if partial band gaps occur, as shown by Yablonovitch and Gmitter [Yablonovitch and Gmitter (1989)] or even in all directions. The search for such complete stop bands is usually done by traversing the edges of the irreducible Brillouin zone, which can be misleading

[Figotin and Kuchment (1998); Dossou *et al.* (2007); Craster *et al.* (2012)] as complete stop bands very rarely turn out to be partial stop bands if the inside of the Brillouin zone is checked. The effect of photonic stop bands is well-known, with many textbooks offering comprehensive introductions to the physics and models of PCs [Movchan *et al.* (2002); Zolla *et al.* (2005); Melrose and McPhedran (2005); Joannopoulos *et al.* (2008)], and these effects form the basis of many devices, including Bragg mirrors, dielectric Fabry Perot filters and distributed feedback lasers. All of these contain low-loss dielectrics periodic in one dimension, and so are one-dimensional PCs.

Bragg mirrors are tremendously useful, but their reflecting properties critically depend upon the frequency of the incident wave with regard to its incidence. For broad frequency ranges, one wishes to reflect light of any polarization at any angle (which requires a complete photonic band gap) and for Dirichlet media (i.e. those composed with microstructure where the field is zero on the microparticles) such a gap occurs at zero frequency. This explicitly excludes dispersion curves passing through the Γ point in dispersion diagrams and precludes any long-wave homogenisation being effective. Zero-frequency band gaps are important as they mean that very long waves cannot propagate through a structure and a similar situation occurs in periodically perforated elastic plates with clamped holes: zero-frequency platonic stop bands then disallow the propagation of flexural waves at arbitrarily low frequencies [Antonakakis and Craster (2012)]. Using approximate models of thin plates to predict the propagation of surface elastic waves in structured soils [Brûlé *et al.* (2014)], it is therefore possible to create seismic shields for moderately low and even ultra-low-frequency waves [Antonakakis *et al.* (2014b); Achaoui *et al.* (2017)].

Zero-frequency band gaps are just one feature not captured with quasi-static homogenisation effects, many more exotic wave phenomena exist that require and deserve understanding through effective media: all-angle negative refraction [Yablonovitch (1987); Zengerle (1987); Notomi (2002)], ultra-refraction [Dowling and Bowden (1994); Gralak *et al.* (2000); Enoch *et al.* (2002)] and cloaking

at Dirac-like cones [Chan *et al.* (2012)]. We will demonstrate in this chapter that an effective, frequency-dependent non-local medium can be created that accurately reproduces these effects.

Given the importance of wave propagation through periodic media there has naturally been considerable recent activity and advances in numerical techniques. The main methods used are based on Fourier expansions in the vector electromagnetic Maxwell equations [Johnson and Joannopoulos (2001); Joannopoulos *et al.* (2008)], finite elements [Zolla *et al.* (2005)] and multipole and lattice sums [Movchan *et al.* (2002)] for cylinders, to name but a few: these allow one to visualize the various effects and design PCs. In practical terms these methods and commercial software such as [COMSOL (2012)] dominate the analysis of PCs and their implementation. One of us, from the spin-off company Multiwave Technology, develops alternative numerical solutions to advance research in metamaterials, such as those based on so-called High Frequency Homogenization, which is introduced in the next section.

1.1.5. *High Frequency Homogenization (HFH) theory*

Although the numerical approaches discussed briefly are efficient, they still require substantial computational effort and can obscure physical understanding and interpretation. An option to substantially reduce the numerical complexity of the wave problem is to use asymptotic analysis; this has been developed over the last 40 years by applied mathematicians primarily for solving partial differential equations (PDEs), with rapidly oscillating periodic coefficients, in the context of thermostatics, continuum mechanics or electrostatics. Indeed, the past forty years have been marked by strong research monographs on the topic by Bensoussan, Louis and Papanicolaou [Bensoussan *et al.* (1978); Sanchez-Palencia (1980)], Bakhvalov and Panasenko [Bakhvalov and Panasenko (1989)], Jikov, Kozlov and Oleinik [Jikov *et al.* (1994)] and Milton [Milton (2002)]. The available literature on such effective medium theories is vast, but only a very few groups have addressed such problems as the homogenization of media in three dimensions with moderate contrast

in the material properties, [Guenneau and Zolla (2000); Wellander and Kristensson (2003); Cioranescu *et al.* (2008)] and high-contrast two-dimensional [Zhikov (2000)] photonic crystals, that have important potential applications in photonics [Zolla *et al.* (2005)]. Besides this, the aforementioned literature does not address the challenging problem of homogenization for moderate contrast photonic crystals near stop band frequencies. Classical homogenization (by classical we mean the static and quasi-static theories) is constrained to low frequencies and long waves in moderate contrast PCs, but it nonetheless provides interesting results for effective properties of 3-D-connected and non-connected wire metamaterials [Silveirinha and Fernandes (2005)] and other metamaterials with negative parameters [Silveirinha (2007)].

So-called high-contrast homogenization captures the essence of stop bands in PCs when the permittivity inside the inclusions is much higher than that of the surrounding matrix [Bouchitté and Felbacq (2004)]. Here, the contrast is typically of the order of η^2, where η is the array pitch, which in turn is much smaller than the wavelength, and this area of homogenization theory is fuelled by interest in artificial magnetism, initiated by the work of O'Brien and Pendry [O'Brien *et al.* (2002)]. We should also point out that moderate contrast one-dimensional photonic crystals have been recently shown to display not only artificial magnetism, but also chirality, using techniques of high-order-homogenization [Liu *et al.* (2013)].

For all these reasons, there is a need for homogenization or effective medium theories of PCs that would cover the spectrum, not just at vanishing (quasi-static) frequencies, in order to grasp, and fully exploit, the rich behaviour of photonic band gap structures [Notomi (2002)]. This demand has created a suite of extended homogenization theories for periodic media called Bloch homogenization by applied mathematicians such as Conca, Allaire, Piatnitski, Birman, Suslina, Hoefer and Weinstein from the PDE analysis community [Conca *et al.* (1995); Allaire and Conca (1998); Allaire and Piatnitski (2005); Birman and Suslina (2006); Hoefer and Weinstein (2011)]. There is also a flourishing literature on developing homogenized elastic media

from the theoretical mechanics community, such as Nemat-Nasser, Willis, Auriault and Boutin, who achieve frequency-dependent effective parameters, also based upon periodic media [Nemat-Nasser *et al.* (2011); Auriault and Boutin (2012)]. There is indeed considerable interest in creating effective continuum models of microstructured media that break free from the conventional low-frequency homogenization limitations.

In 2010, noting that striking physical effects such as slow light can be achieved at arbitrarily high frequencies on edges of stop bands where the wave group velocity vanishes [Figotin and Vitebskiy (2006)], Craster, Kaplunov and Pichugin realized that they would be able to develop an asymptotic theory around these critical points in the band spectrum. They called their asymptotic method high frequency homogenization, or HFH in short [Craster *et al.* (2010a)]. The mathematical tools involved in HFH come from earlier work on effective equations in waveguide theory, in particular from high-frequency long-wave models for the vibrations of plates [Berdichevskii (1977)] or weakly curved waveguides [Gridin and Craster (2003)], or even electronic states in rings [Gridin *et al.* (2004)]; the cut-off frequencies for waveguides are the analogues of the band-gap edges and this analogy is explored in [Craster *et al.* (2014)]. The philosophy of HFH is also closely related to ideas in [Bensoussan *et al.* (1978); Birman and Suslina (2006); Allaire and Conca (1998)] from Bloch homogenisation and it applies to any of the linear PDEs in periodic continua from electromagnetism, acoustics, hydrodynamics or elasticity, the only requirement being applicability of Floquet-Bloch theory. It can be summarized schematically as in Fig. 1.3, where the two typical lengthscales $\mathbf{X}$ and $\boldsymbol{\xi}$ involved in the asymptotic model account respectively for fast oscillations of the field on the order of supercells and the much smaller cells. Importantly, unlike for classical low frequency homogenization whereby the angular wave frequency ω (resp. the wavelength λ) is rescaled as $\eta\omega$ (resp. λ/η), here we shall perturb away from a finite, possibly high, frequency ω.

In this remainder of this chapter, we survey the applications of HFH theory to the physics of acoustic and platonic metamaterials,

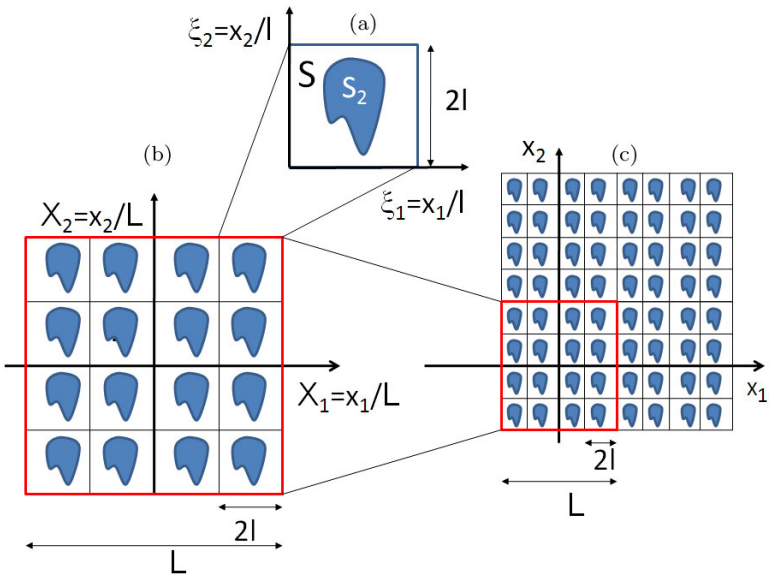


Fig. 1.3. High Frequency Homogenisation (HFH) principle: An elementary cell (a) of sidelength $2l$, modelled by a fast oscillating variable $\boldsymbol{\xi}$, is repeated periodically within a supercell (b) of sidelength L, modelled by a slow variable $\mathbf{X}$, which is itself repeated periodically in space (c). One then assumes that the parameter $\epsilon = l/L$ is small, and its vanishing limit thereafter studied. The leading order, homogenised, term of Floquet-Bloch eigenfields within the crystal is then sought as $u_0(\mathbf{X}, \boldsymbol{\xi}) = f_0(\mathbf{X})U_0(\boldsymbol{\xi}; \Omega_0)$, wherein f_0 accounts for variations of the fields on the order of the supercells, and U_0 captures their fast oscillations in the much smaller cells, when either periodic or anti-periodic conditions are enforced on the cells: Perturbing away from these standing waves of frequency Ω_0 allows for a complete reconstruction of the Bloch spectrum and associated Floquet-Bloch eigenfields, by considering larger and larger cells of sidelength $2Nl$ instead of elementary cells $2l$ [Antonakakis *et al.* (2013b)].

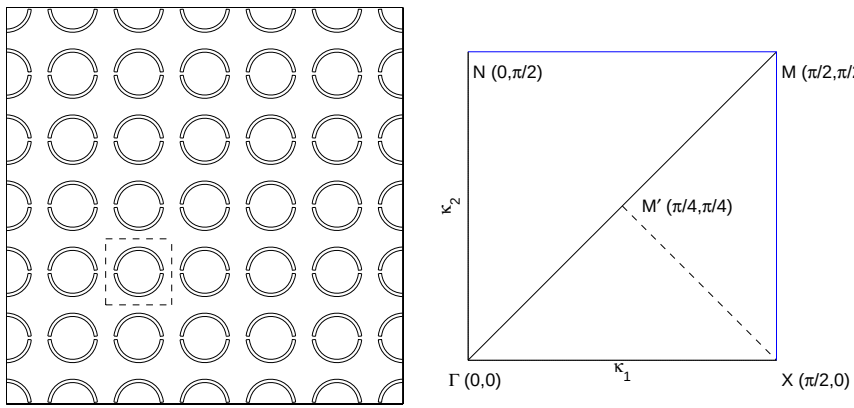
in the context of phononics for anti-plane shear waves and in-plane coupled shear and pressure waves in periodic arrays of inclusions, as well as for flexural (Lamb) waves in pinned plates. The latter analysis is made possible thanks to the extension of HFH to plate theory [Antonakakis and Craster (2012)]. Our aim is to show the universal features of stop band structures thanks to HFH, and to further exemplify their potential use in control of light and mechanical waves, with novel applications ranging from cloaking [Milton

and Nicorovici (2006); Guenneau *et al.* (2007b); Craster and Guenneau (2012); Achaoui *et al.* (2017)] to seismic shields [Antonakakis *et al.* (2014b); Brûlé *et al.* (2014)]. The need for modern asymptotic models of PCs and metamaterials is because these composites (that can be structured from the nanoscale to the meter scale, depending upon whether one would like to control light, sound, water or seismic waves) have typically tens to thousands of elementary cells and it is therefore fairly difficult to improve their designs without resorting to simplified models. For instance, seismic waves propagating in structured soil require use of super-computers if one simply attacks the complete full three-dimensional elastodynamic problem [Colombi *et al.* 2016]. Dynamic effective anisotropic effects achieved with arrays of boreholes could help prevent the disastrous impact of earthquakes in sedimentary basins and there are recent patents being filed in this direction [Achaoui *et al.* (2016)]. Similarly, hydrodynamic wave models are very demanding computationally, a full 3D Navier model for a water wave cloak like in Figure 1.2 is computationally intense and again, HFH can help simplify the model.

1.2. HFH for Acoustic Metamaterials and Phononic Crystals

1.2.1. *The scalar problem: theory*

The basic theory is more easily illustrated by considering a scalar wave equation, with time-harmonic waves assumed, and allowing for periodic variation in the material properties or geometry. For infinite perfectly periodic media, consisting of elementary cells that repeat, one focuses attention on a single elementary cell as shown in Fig. 1.4. Quasi-periodic Floquet-Bloch boundary conditions on the elementary cell describe the phase-shift as a wave moves through the material and a dispersion relation relates the Bloch wavenumber, the phase-shift, to frequency. Figure 1.4(b) shows the irreducible Brillouin zone [Brillouin (1953)] ΓMX associated with a single repeating elementary square cell containing, say, a circular hole. The dispersion diagrams we show are of the frequency versus wavenumber around



(a) (b) Brillouin zone

Fig. 1.4. Panel (a) An infinite square array of split ring resonators with the elementary cell shown as the dashed line inner square. Panel (b) shows the irreducible Brillouin zone, in wavenumber space, used for square arrays in perfectly periodic media based around the elementary cell shown of length $2l$. Figure taken from [Antonakakis *et al.* (2013a)].

the edges of the Brillouin zone as is traditional in solid state physics. There are occasions upon which doing this misses interesting details [Craster *et al.* (2012)] where the path $M'X$, see Fig. 1.4(b), yielded important dispersion properties. We also note that the symmetry of the hole is important, and for the two thin ligament SRR of Fig. 1.4(a) one should use the square ΓXMN.

The eigensolutions that emerge are the Bloch modes at the edges of the Brillouin zone, and when these eigensolutions are perfectly in-phase or out-of-phase across the cell then standing waves exist and there are standing wave frequencies (whose frequencies can be arbitrarily high). Asymptotic techniques based around high frequency long wave asymptotics have been developed [Craster *et al.* (2010a)] and Schrödinger ordinary differential equations in 1D periodic media (or partial differential equations in 2D) emerge; this approach also works for microstructured discrete [Craster *et al.* (2010b)] or frame-like media [Nolde *et al.* (2011)]. These recent theories avoid the issue of holes with various boundary conditions

and have the material properties varying periodically on the scale of the elementary cell and only treat model problems, mainly in one dimension, for which completely analytic progress can be made. The key idea for periodic media is to replace the complicated microstructured medium with an effective (possibly anisotropic) continuum on a macroscale, that is, one wishes to homogenize the medium even when the wavelength and microstructure may be of similar scales.

The theory is ultimately not limited to just reproducing dispersion curves asymptotically, it can be used to obtain envelope functions of the resulting fields or it can be adjusted to treat localized defect modes and other features due to local non-periodic material changes or boundaries, with these effects coming through in extra forcing terms within the continuum partial differential equations and effective boundary conditions.

We begin with a two dimensional structure composed of a square lattice geometry of identical cells with identical holes inside each of them. The side length of the direct lattice base vectors, i.e. the side of each square cell, is taken as $2l$. Note that for simplicity equal length lattice vectors and a square lattice are assumed and both assumptions could be relaxed. These elementary cells define a lengthscale which is the micro-scale of the structure. As noted above, real structures could be created from many hundreds or thousands of such elementary cells and we introduce a macro-scale length denoted by L that could be viewed as a characteristic overall dimension of the structure: The ratio of these scales, $\eta \equiv l/L$, is assumed small, see also Fig. 1.3. Each cell is identical in geometry and the material within each cell is characterized by two periodic functions, in $\boldsymbol{\xi} \equiv (x_1/l, x_2/l)$, namely $a(\boldsymbol{\xi})$ and $\rho(\boldsymbol{\xi})$; these are stiffness and density for pressure waves.

A time harmonic dependence of propagation $\exp(-i\omega t)$, with angular wave frequency ω, is assumed throughout, and henceforth suppressed, together with a non-dimensionalization of physical parameters by setting, $a \equiv a_0\hat{a}(\boldsymbol{\xi})$, $\rho \equiv \rho_0\hat{\rho}(\boldsymbol{\xi})$ and $\mathbf{x} = l\hat{\mathbf{x}}$ where $c_0 = \sqrt{a_0/\rho_0}$ is the characteristic wave speed. This leads to the resulting equation of study where the hat decoration has been

removed:

$$l^2\nabla_{\boldsymbol{x}}\cdot[a(\boldsymbol{\xi})\nabla_{\boldsymbol{x}}u(\boldsymbol{x})]+\Omega^2\rho(\boldsymbol{\xi})u(\boldsymbol{x})=0,\quad \text{with } \Omega=\frac{\omega l}{c_0}, \tag{1.1}$$

on $-\infty<x_1,x_2<\infty$, Ω is the non-dimensional frequency and u is the out-of-plane displacement in elasticity.

The two scale nature of the problem is incorporated using the small and large length scales to define two new independent coordinates namely $\mathbf{X}=\boldsymbol{x}/L$, and $\boldsymbol{\xi}=\boldsymbol{x}/l$, see Figure 1.3. Equation (1.1) then becomes,

$$\begin{aligned}\nabla_{\boldsymbol{\xi}}\cdot[a(\boldsymbol{\xi})\nabla_{\boldsymbol{\xi}}u(\mathbf{X},\boldsymbol{\xi})]+\Omega^2\rho(\boldsymbol{\xi})u(\mathbf{X},\boldsymbol{\xi})\\ +\eta[2a(\boldsymbol{\xi})\nabla_{\boldsymbol{\xi}}+\nabla_{\boldsymbol{\xi}}a(\boldsymbol{\xi})]\cdot\nabla_{\mathbf{X}}u(\mathbf{X},\boldsymbol{\xi})+\eta^2a(\boldsymbol{\xi})\nabla^2_{\mathbf{X}}u(\mathbf{X},\boldsymbol{\xi})=0.\end{aligned} \tag{1.2}$$

Standing waves occur when there are periodic (or anti-periodic) boundary conditions across the elementary cell (in the $\boldsymbol{\xi}$ coordinates) and these standing waves encode the local information about the multiple scattering that occurs due to the neighbouring cells. The asymptotic technique is then a perturbation about these standing wave solutions. As these are associated with periodic and anti-periodic boundary conditions, which are respectively in-phase and out-of-phase waves across the cell, the conditions in $\boldsymbol{\xi}$ on the edges of the cell, ∂S_1, are known:

$$u|_{\xi_i=1}=\pm u|_{\xi_i=-1}\quad\text{and}\quad u_{,\xi_i}|_{\xi_i=1}=\pm u_{,\xi_i}|_{\xi_i=-1}, \tag{1.3}$$

where the subscript, $\boldsymbol{\xi}$ denotes partial derivatives with respect to the variable $\boldsymbol{\xi}$ and with the $+,-$ for periodic or anti-periodic cases respectively. We now pose an ansatz for the field and the frequency,

$$\begin{aligned}u(\mathbf{X},\boldsymbol{\xi})=u_0(\mathbf{X},\boldsymbol{\xi})+\eta u_1(\mathbf{X},\boldsymbol{\xi})+\eta^2u_2(\mathbf{X},\boldsymbol{\xi})+\cdots,\\ \Omega^2=\Omega_0^2+\eta\Omega_1^2+\eta^2\Omega_2^2+\cdots\end{aligned} \tag{1.4}$$

The $u_i(\mathbf{X},\boldsymbol{\xi})$'s adopt the boundary conditions (1.3) on the edge of the cell. An ordered set of equations emerge indexed with their respective

power of η, and are treated in turn

$$(au_{0,\xi_i})_{,\xi_i} + \Omega_0^2 \rho u_0 = 0, \tag{1.5}$$

$$(au_{1,\xi_i})_{,\xi_i} + \Omega_0^2 \rho u_1 = -(2au_{0,\xi_i} + a_{,\xi_i} u_0)_{,X_i} - \Omega_1^2 \rho u_0, \tag{1.6}$$

$$\begin{aligned}(au_{2,\xi_i})_{,\xi_i} + \Omega_0^2 \rho u_2 \\ = -au_{0,X_iX_i} - (2au_{1,\xi_i} + a_{,\xi_i} u_1)_{,X_i} - \Omega_1^2 \rho u_1 - \Omega_2^2 \rho u_0.\end{aligned} \tag{1.7}$$

The leading order Eq. (1.5) is independent of the longscale $\mathbf{X}$ and is a standing wave on the elementary cell excited at a specific eigenfrequency Ω_0 and associated eigenmode $U_0(\boldsymbol{\xi}; \Omega_0)$, modulated by a long scale function $f_0(\mathbf{X})$ and so

$$u_0(\mathbf{X}, \boldsymbol{\xi}) = f_0(\mathbf{X}) U_0(\boldsymbol{\xi}; \Omega_0). \tag{1.8}$$

At this point we will assume isolated eigenfrequencies, but repeated eigenvalues arise and are discussed later. The entire aim is to arrive at a PDE with a matrix valued coefficient for f_0 posed entirely upon the longscale, but with the microscale incorporated through coefficients that are integrated, not necessarily averaged, quantities.

1.2.1.1. *The Neumann problem*

We define the Neumann boundary conditions on the holes, ∂S_2,

$$\frac{\partial u}{\partial \mathbf{n}} = u_{,x_i} n_i|_{\partial S_2} = 0, \tag{1.9}$$

where $\mathbf{n} = (n_1, n_2)$ is the unit outward normal to ∂S_2 and which, in terms of the two-scales and $u_i(\mathbf{X}, \boldsymbol{\xi})$, become

$$U_{0,\xi_i} n_i = 0, \quad (U_0 f_{0,X_i} + u_{1,\xi_i}) n_i = 0, \quad (u_{1,X_i} + u_{2,\xi_i}) n_i = 0. \tag{1.10}$$

The leading order eigenfunction $U_0(\boldsymbol{\xi}; \Omega_0)$ must satisfy the first of these conditions. Moving to the first order equation (1.6) we invoke a solvability condition by integrating over the cell the product of equation (1.6) and U_0 minus the product of equation (1.5) and $u_1/f_0(\mathbf{X})$. The eigenvalue Ω_1 is zero and we can solve for $u_1 = f_{0,X_i} U_{1_i}(\boldsymbol{\xi})$, so $\mathbf{U}_1$ is a vector field and $\mathbf{U}_{1_i}$ its ith component. By re-invoking a

similar solvability condition for equation (1.7) we obtain the desired partial differential equation for f_0

$$T_{ij} f_{0,X_i X_j} + \Omega_2^2 f_0 = 0 \quad \text{where,}$$
$$T_{ij} = \frac{t_{ij}}{\iint_S \rho U_0^2 dS} \quad \text{for} \quad i,j = 1,2 \tag{1.11}$$

posed entirely on the longscale $\mathbf{X}$. The tensor t_{ij} consists of integrals over the microcell in $\boldsymbol{\xi}$ and is ultimately independent of $\boldsymbol{\xi}$. The formulations for t_{ij}'s read,

$$t_{ii} = \iint_S a U_0^2 dS + \iint_S a(U_{1_i,\xi_i} U_0 - U_{1_i} U_{0,\xi_i}) dS \quad \text{for } i = 1,2, \tag{1.12}$$

$$t_{ij} = \iint_S a(U_{1_j,\xi_i} U_0 - U_{1_j} U_{0,\xi_i}) dS \quad \text{for } i \neq j. \tag{1.13}$$

There is no summation over repeated suffices for the t_{ii}. The numerical solutions of U_0, and subsequently U_{1_j}, are computed using a standard finite element package [COMSOL (2012)], although many other numerical methods could be used instead, thereby allowing us to treat general geometries. Besides this we note that a could be a spatially varying function.

Floquet Bloch boundary conditions on the cell imply $f_0(\mathbf{X}) = \exp(i\kappa_j X_j/\eta)$. In this notation $\kappa_j = K_j - d_j$ and $d_j = 0, \pi/2, -\pi/2$ depending on the location in the Brillouin zone with $K_j \in [0, \pi/2]$. Equation (1.11) and the frequency expansion of Eq. (1.4) lead to,

$$\Omega \sim \Omega_0 + \frac{T_{ij}}{2\Omega_0} \kappa_i \kappa_j \tag{1.14}$$

and thus one can compare directly with the full numerics. It is worthwhile noting the utility of T_{ij} coefficients as their sign and absolute value give information about the group velocity for the specified frequencies and locations of the Brillouin zone. We have glossed over some tedious details here, in particular how to deal with repeated roots for which coupled PDEs emerge for the, now multiple, f_0's [Antonakakis *et al.* (2013a)].

1.2.1.2. *The Dirichlet problem*

If boundary conditions on the inside boundary of the cell (∂S_2) change to Dirichlet then

$$u(\mathbf{X}, \boldsymbol{\xi})|_{\partial S_2} = 0 \iff u_i(\mathbf{X}, \boldsymbol{\xi})|_{\partial S_2} = 0, \quad i \in \mathbb{N} \tag{1.15}$$

and are set in the short-scale $\boldsymbol{\xi}$ so, for $i = 0$, $U_0(\boldsymbol{\xi}; \Omega_0)|_{\partial S_2} = 0$. Using the same solvability conditions as in section 1.2.1.1 we obtain the following equation on the long scale,

$$T_{ij} f_{0,X_i X_j} + \Omega_2^2 f_0 = 0, \quad \text{with} \quad T_{ij} = \frac{t_{ij}}{\iint_S \rho U_0^2 dS} \quad \text{for} \quad i, j = 1, 2. \tag{1.16}$$

In particular, the coefficients T_{ij} encode the anisotropy at a specific frequency and the t_{ij}'s are integrals over the small-scale cell

$$t_{ii} = \iint_S a U_0^2 dS + 2 \iint_S a U_{1_i,\xi_i} U_0 dS + \iint_S a_{,\xi_i} U_{1_i} U_0 dS, \tag{1.17}$$

$$t_{ij} = 2 \iint_S a U_{1_j,\xi_i} U_0 dS + \iint_S a_{,\xi_i} U_{1_j} U_0 dS \quad \text{for} \quad i \neq j, \tag{1.18}$$

where U_{1_i} is the ith component of vector function $\mathbf{U}_1$. Note that there is no summation over repeated indexes for t_{ii}.

The PDE for f_0, equations (1.16) or (1.11), is crucial as the local microstructure is completely encapsulated within the tensor T_{ij} (physically a hallmark of effective anisotropy); these are, for a specific structure at an Ω_0, just numerical values. Notably the tensor can have negative values, or components, and it allows one to interpret and, even more importantly, predict changes in behaviour or when specific effects occur. The structure of the tensor depends upon the boundary conditions of the holes, the results for the Dirichlet problem, for instance, are different from those of the Neumann case in section 1.2.1.1. Another key point is that numerically the short scale is no longer present and the PDE (1.16) is simple and quick to solve numerically, or even by hand.

1.2.2. *Effective dispersive media*

The major goal of HFH is to represent the periodic medium of finite extent using an effective homogeneous medium.

Transforming equation (1.16) back to the original $x_i = X_i L$ coordinates and using the solution of Ω_2 we obtain an effective medium equation for $\hat{f}_0$ that is,

$$T_{ij}\hat{f}_{0,x_i x_j}(\mathbf{x}) + \frac{\Omega^2 - \Omega_0^2}{l^2}\hat{f}_0(\mathbf{x}) = 0. \tag{1.19}$$

The nature of equation (1.19) is not necessarily elliptic since T_{11} and T_{22} can take different values and/or different signs. In the illustrations herein the $T_{ij} = 0$ for $i \neq j$. The hyperbolic behaviour of equation (1.19) yields asymptotic solutions that describe the endoscope effects where, if one of the T_{ii} coefficients is zero as often happens near the point M of the Brillouin zone, waves propagate only in one direction. Note that when the cell has the adequate symmetries the dispersion relation near point $N(0, \pi/2)$ is identical to that near point $X(\pi/2, 0)$, which explains the existence of two orthogonal directions of propagation instead of only one.

At Dirac-like cones that correspond to the coalescence of three dispersion curves as would happen in Fig. 1.5(b) if the radius decreases toward zero, the linear behaviour of the effective medium yields an equation slightly different from (1.19). Typically it is represented by a system of three coupled PDEs that uncouple to yield one identical PDE for all $f_0^{(i)}$s that is of the form,

$$\hat{f}_{0,x_i x_i}(\mathbf{x}) + \beta\frac{(\Omega^2 - \Omega_0^2)^2}{l^2}\hat{f}_0(\mathbf{x}) = 0, \tag{1.20}$$

where β is a coefficient equivalent of the T_{ij} but this time is the same for all combinations of i and j. The quadratic mode that emerges in the middle of the linear ones follows Eq. (1.19).

1.2.3. *Dispersion curves and asymptotics*

To illustrate the results we consider the dispersion curves for a system of split-ring resonators, as shown in Fig. 1.6; there have been many

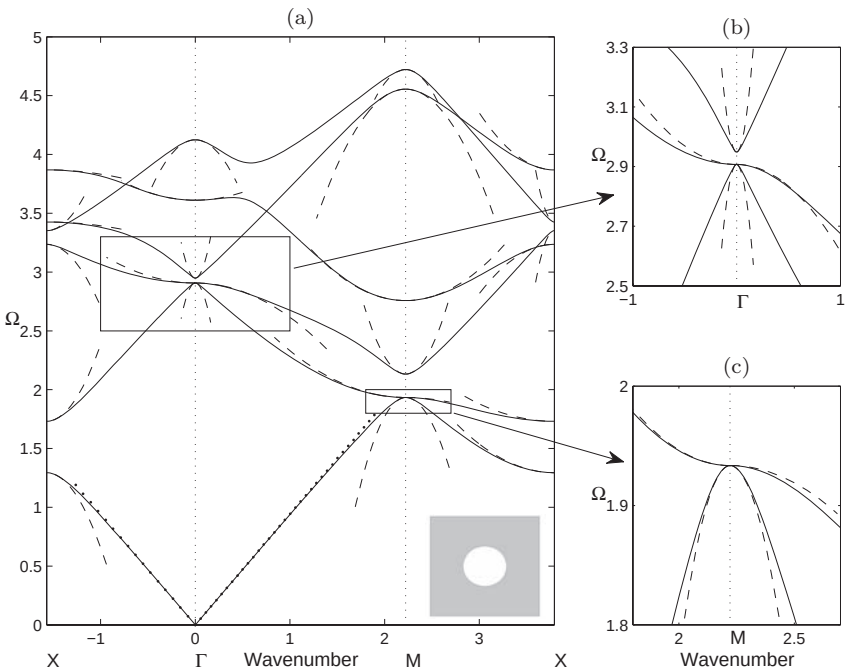


Fig. 1.5. The dispersion diagram for an array of square cells (side 2) with circular inclusions of radius 0.4. Panel (a) shows dispersion curves from finite element simulations (solid lines), the asymptotic solutions from HFH theory (dashed lines) and the linear long wave classical homogenization asymptotics are the dotted lines emerging from the origin. Panels (b) and (c) show enlargements near repeated eigenvalues where the asymptotics from repeated roots are used. Taken from [Antonakakis *et al.* (2013a)].

numerical studies for dispersion diagrams for instance in [Guenneau *et al.* (2007a)] and semi-analytical work for narrow gaps as in [Llewellyn-Smith and Davis (2010)]. Figures 1.6(a, b, c) show the dispersion curves together with the HFH asymptotics. Notable here is the appearance of a low-frequency stop band (at roughly $\Omega \sim 1$) below the Bragg stop band (at roughly $\Omega \sim 1.75$), whose upper edge remains virtually unaffected by the insertion of the resonator in each circular inclusion of the array. The low frequency stop band is associated with a localized mode upon resonance of the resonator; the inner cylinder of material acts as a resonator attached by thin

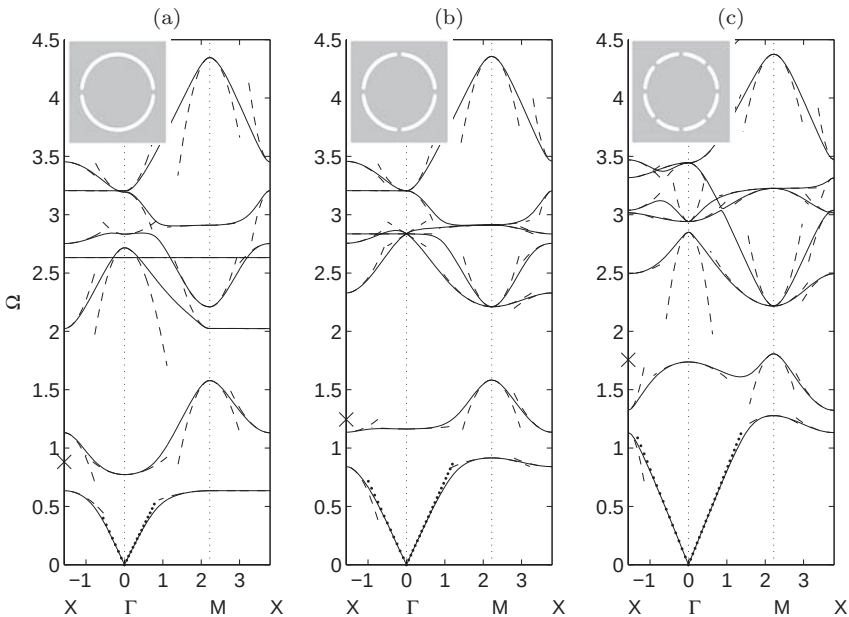


Fig. 1.6. The dispersion diagram shown for an array of square cells (side 2) containing circular holes of radius 0.8, and with circular inclusions of radius 0.7, attached to the rest of the cell by 2, 4 and 8 thin ligaments in panels (a, b) and (c) respectively. The dispersion curves from numerics are shown as solid lines, the asymptotic HFH results are shown as dashed lines with the low frequency linear classical homogenization shown dotted emerging from the origin. Taken from [Antonakakis *et al.* (2013a)].

ligaments to the outer material. The low frequency stop band appears at frequencies already beyond the scope of classical homogenization, but HFH captures its finer details, as in Figs. 1.6(a, b, c).

The flat band along the MX path is captured by HFH, and this band is associated with a localized mode in the SRR (which is therefore insensitive to any variation of the Floquet-Bloch phase-shift across the unit cell along this path). The highly dispersive physics of the low frequency stop band can lead to ultra-refraction and all-angle-negative-refraction effects.

Figures 1.6(a, b, c) also show changes of curvature, in say, the second dispersion curve around Γ point, which flattens for 4 holes,

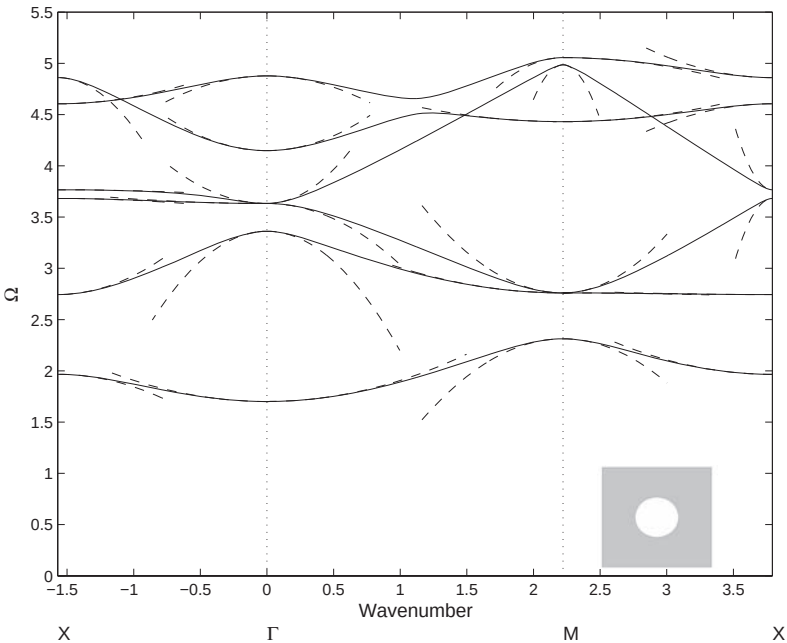


Fig. 1.7. The dispersion diagram for a doubly periodic array of square cells with circular inclusions, of radius 0.4, fixed at their inner boundaries shown for the irreducible Brillouin zone of Fig. 1.4. The dispersion curves are shown in solid lines and the asymptotic solutions from HFH are shown in dashed lines. Taken from [Antonakakis *et al.* (2013b)].

and the curvature actually changes sign when comparing 2 and 8 ligaments; the HFH asymptotics capture this behaviour. A minor note of warning is worth sounding regarding the irreducible Brillouin zone, it is all too easy to overlook the fact that the two ligament SRR does not have the appropriate symmetries such that one can use just the triangle ΓMX as the irreducible Brillouin zone, but here one loses nothing in the physical interpretation.

As noted earlier a commonly arising issue for Dirichlet (clamped/fixed) inclusion is that a zero-frequency band gap occurs, see Fig. 1.7, which illustrates the typical dispersion curves versus the asymptotics for wavenumbers, on the specified path of Fig. 1.4, for the example of a hole with radius 0.4. There is, as expected, a wide zero-frequency

Table 1.1. The first six standing wave frequencies for in-phase waves at Γ, cf. Fig. 1.7, together with associated values for T_{11} and T_{22}. Symmetry between T_{11} and T_{22} is breaking when the multiplicity of the eigenvalue is greater than two. Negative group velocity is demonstrated by the negative sign of both T_{ij}'s.

T_{11}	T_{22}	Ω_0
0.6988	0.6988	1.7009
7.8677	7.8677	3.3616
0.3230	−8.9980	3.6327
4.9681	14.2899	3.6327
4.7755	4.7755	4.1487
−4.2798	−4.2798	4.8770

stop-band and the lowest branch is isolated with a full stop-band also lying above it. It is not possible to homogenize this using quasi-static theories, but the HFH mehodology captures the behaviour perfectly. At some standing wave frequencies two modes share the same frequency, for instance the third and fourth modes at Γ; these modes are asymptotically quadratic in the local wavenumber. Table 1.1 shows the T_{11}, T_{22} values for point Γ of the Brillouin zone. Note how $T_{11} = T_{22}$ for all single eigenfrequencies, but that symmetry breaks for the double roots. Moreover the signs of the T_{11} and T_{22} naturally inform one of the local curvatures near Γ. Physically, this tells us the sign of group velocity of waves with small phase-shift across the unit cells, and thus whether or not they undergo backward propagation, which is one of the hallmarks of negative refraction.

In this section, for simplicity, we assumed Dirichlet and Neumann data on inclusions' boundaries, but the HFH theory also works for transmission boundary conditions (e.g. continuity out-of-plane displacement and its flux). It is also well-known that the Helmholtz equation models transverse electromagnetic waves, and HFH successfully predicted parabolic and hyperbolic type features in the microwave regime for an array of parallel dielectric rods [Ceresoli *et al.* (2015)].

1.3. HFH for Platonic Crystals

The presentation above was for the simple scalar Helmholtz equation, but the methodology works for all other linear continuum equations. We now turn our attention to thin elastic plates, which support flexural waves and have been much studied particularly in connection with the dynamics of shells [Kaplunov *et al.* (1998)] and the vibration of periodic engineering structures [Mead (1996)]. More recently the subject has undergone a renaissance with the analysis of doubly-periodic arrays of supports, [Mace (1996); Langley (1997)], revisited from the point of view of phononic crystals [Movchan *et al.* (2007)] with these structured elastic plates being dubbed platonic crystals [McPhedran *et al.* (2009)]; the aim being to manipulate and re-direct vibrational wave fields through negative refraction and ultra-refraction [Farhat *et al.* (2010a,b)], Dirac-like point effects [Torrent *et al.* (2013); Haslinger *et al.* (2016)], highly-directive dynamic anisotropy [Langley (1997); Antonakakis *et al.* (2014b)] and these ideas use Bragg scattering and interactions due to the periodic structuring.

Elastic plates and beams are not governed by a simple wave equation and instead classical Bernoulli-Euler beam or plate theory [Graff (1975)] allows for bending moments and transverse shear forces to be present; a fourth order equation emerges for the out-of-plane plate displacement u as

$$\rho A \frac{\partial^2 u}{\partial t^2} + \nabla^2 \left(EI \nabla^2 u \right) = 0. \tag{1.21}$$

Here ρ, A, E, I are density, cross-sectional area, Young's modulus and cross-section moment of inertia, respectively.

We operate in non-dimensional variables that is we choose constant reference values for the physical parameters $E_0, I_0, \rho_0, \mu_0, A_0$ and incorporate any spatial variation through $\hat{\beta}, \hat{\mu}$ defined through $EI = E_0 I_0 \hat{\beta}(\hat{\mathbf{x}})$ and $\rho A = \rho_0 A_0 \hat{\mu}(\hat{\mathbf{x}})$. The spatial coordinate $\mathbf{x} = l\hat{\mathbf{x}}$ and, assuming time harmonic dependence $u = u_0 \hat{u} \exp(-i\Omega t)$ (henceforth suppressed) leads to a non-dimensional frequency $\hat{\Omega}$ from

$$\hat{\Omega}^2 = \frac{\rho_0 A_0 l^4 \Omega^2}{E_0 I_0}.$$

The governing equation (1.21) is now

$$\nabla^2_{\hat{\mathbf{x}}}\left(\hat{\beta}\nabla^2_{\hat{\mathbf{x}}}\hat{u}\right) - \hat{\mu}\hat{\Omega}^2\hat{u} = 0 \tag{1.22}$$

The subscript on ∇ denotes that differentiation is now with respect to $\hat{\mathbf{x}}$. Henceforth we operate in the non-dimensional setting and drop the hat decoration, and the subscript for ∇.

Following the same method as for the Helmholtz equation, with the multiple scales and leading order solution as in Eq. (1.8), we obtain

$$\begin{aligned} f_{0,X_iX_j}\int_{-1}^{1}\int_{-1}^{1} & 2((\beta(V_{1(i)}-\xi_iU_0)_{,\xi_k\xi_k})_{,\xi_j}U_0 \\ & + (\beta(V_{1(i)}-\xi_iU_0)_{,\xi_j})_{,\xi_k\xi_k}U_0)d\xi_1d\xi_2 \\ & + f_{0,X_jX_j}\int_{-1}^{1}\int_{-1}^{1}(\beta U_{0,\xi_i\xi_i}U_0 + (\beta U_0)_{,\xi_i\xi_i}U_0)\,d\xi_1d\xi_2 \\ & + f_{0,X_iX_j}\int_{-1}^{1}\int_{-1}^{1}4(\beta U_{0,\xi_i})_{,\xi_j}U_0d\xi_1d\xi_2 \\ & - \Omega_2^2 f_0\int_{-1}^{1}\int_{-1}^{1}\mu U_0^2d\xi_1d\xi_2 = 0. \end{aligned} \tag{1.23}$$

Note that Eq. (1.23) is like Eq. (1.11) in the form of a second order PDE,

$$T_{ij}f_{0,X_iX_j} - \Omega_2^2 f_0 = 0, \tag{1.24}$$

where T_{ij} are the respective coefficients in front of the double integrals normalised by double integral coefficient in front of the last term of Eq. (1.23). Invoking the Bloch conditions for, in and out of phase waves, in both directions we obtain a relation for $f_0(\mathbf{X}) = \exp(i\kappa_jX_j/\eta)$ where $\kappa_j = K_j - d_j$ and $d_j = 0, \pi/2, -\pi/2$ depending on the location we refer to in the Brillouin zone. The final asymptotic

dispersion relation reads as

$$\Omega \sim \Omega_{\hat{0}} - \frac{T_{ij}}{2\Omega_{\hat{0}}}\kappa_i\kappa_j, \tag{1.25}$$

repeated indexes imply summation. Note that the generic $\Omega_{\hat{0}}$ denotes the standing wave frequency for the respective location in the Brillouin zone that is defined by the parameters d_i and d_j. For the case of triple modes originating from the same point, as for example the Dirac-like cones, Eq. (1.24) is no longer valid and one obtains three coupled equations for $f_0^{(i)}$ for $i = 1, \ldots, 3$.

We stress that Eq. (1.24) is of second order and not fourth. The fourth order plate equation yields two types of solutions namely propagating and exponentially decaying. Indeed in the long scale, solutions follow a second order PDE and the remaining information of the fourth order problem is enclosed in the T_{ij} integrated quantities. This homogenization theory is not limited to long-waves relative to the microstructure, one apparent pitfall is that the asymptotics appear to be only valid near the band edge frequencies but further refinements are possible, using foldings of the Brillouin zone (by considering a supercell instead of an elementary cell in Fig. 1.3), that extend the theory to provide complete coverage of the dispersion curves (for a large enough supercell) and provide useful information at all frequencies and Bloch wavevectors [Antonakakis *et al.* (2013b)].

1.3.1. *Applications in platonics*

Generating asymptotic dispersion curves is a nice verification of the theory, but naturally one wants to do more with it and actually model real effects. Let us now move to the fascinating area of platonics, plates structured in a periodic way so that one can achieve an extreme control of flexural waves via dynamic effective anisotropy, see Fig. 1.8 (left), wherein the HFH is required to predict frequencies at which some ultra-directive, or lensing, effects can occur. Platonics now encompass the case of plates structured periodically, with the paradigm of invisibility cloak for flexural waves, see Fig. 1.8 (right), wherein the required shape and spacing of stress-free inclusions was

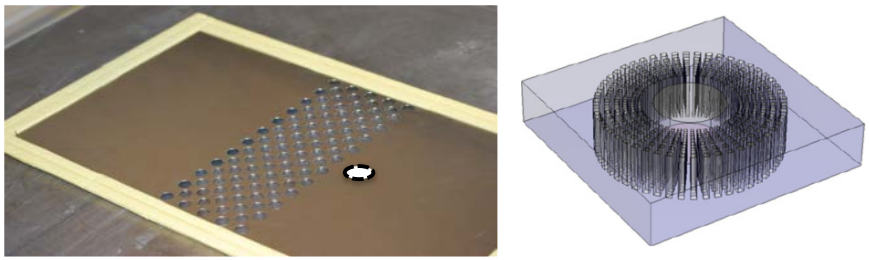


Fig. 1.8. (Left panel) Photo of a duraluminium plate (group of Patrick Sebbah, Institut Langevin, Paris) structured with a periodic array of boreholes (stress-free inclusions) that allows for focussing and directive effects of flexural waves via dynamic effective anisotropy (high-frequency homogenization). (Right panel) Schematic view of a thin plate structured with radially symmetric stress-free inclusions of increasing area with distance from the center of the cloak that achieves the required spatially varying effective anisotropy (low-frequency homogenisation).

deduced from low frequency homogenization techniques (a limit case of HFH when the frequency goes to zero).

In this section, we shall focus on the periodic case (Fig. 1.8 (left)), but HFH indeed encompasses the aperiodic case, when the unit cell size varies from one row to the next as in Fig. 1.8 (right).

1.3.1.1. *Guiding bending waves in platonic crystals*

A useful special case is that of pinned points for which there is an explicit solution [Mace (1996)], for this we take a doubly periodic array of points at $x_1 = 2n_1$, $x_2 = 2n_2$ where $u = 0$ and so the elementary cell is one in $|x_1| < 1, |x_1| < 1$ with $u = 0$ at the origin. Floquet-Bloch conditions are applied at the edges of the cell. The PDE problem is

$$\nabla^4 u - \Omega^2 u = -\sum_{n_1,n_2} F_{n_1 n_2} \delta(x_1 - 2n_1)\delta(x_2 - 2n_2). \tag{1.26}$$

Solving for the unit cell with Bloch conditions a dispersion relation emerges,

$$D(\kappa_1, \kappa_2, \Omega) = \sum_{n_1,n_2} \frac{1}{[(\pi n_1 - \kappa_1)^2 + (\pi n_2 - \kappa_2)^2]^2 - \Omega^2} \tag{1.27}$$

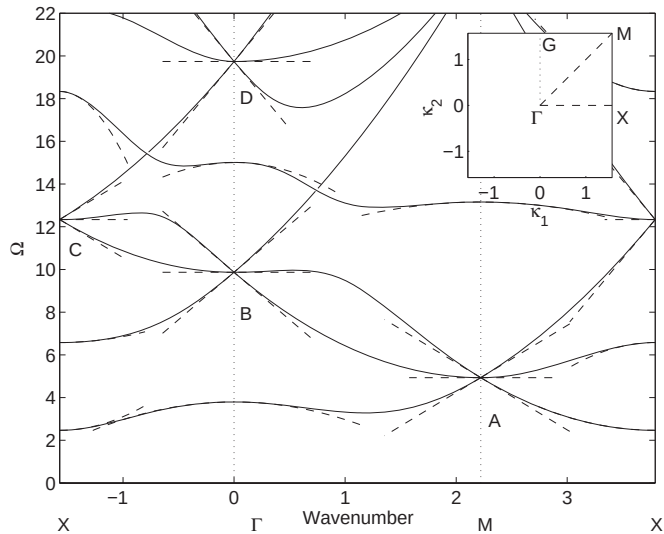


Fig. 1.9. Bloch dispersion curves around the edges of the Brillouin zone ΓXM (inset) for an array of clamped circles (radius 0.01) of pitch 2. One notes the stop band for $\Omega \in [0, \pi^2/4]$ and four Dirac-like cones (so-named by analogy with gapless dispersion curves in graphene electronic band structure wherein the energy of electrons is a linear function of momentum k) [Novoselov *et al.* (2004)] at frequencies $\Omega = \pi^2/2$ (marked A), $\Omega = \pi^2$ and $2\pi^2$ (marked B and D) and $\Omega = 5\pi^2/4$ (marked C). Solid curves are computed with finite element simulations, and dashed curves are from HFH.

which converges as $1/n_i^4$, in n_1 and n_2. Note that when $\boldsymbol{\kappa} = \mathbf{0}$ it is immediate that $\Omega = 0$ is not a solution and thus that a zero-frequency stop-band ensues. The dispersion curves, actually generated by finite elements for a very small clamped circle but indistinguishable from this exact solution, are given in Fig. 1.9 and show a zero frequency stop band, Dirac-like points, and many other interesting features.

We illustrate potential applications of platonic crystals whose dynamics are described by Fig. 1.9, very small clamped holes, namely a shielding effect, a directive antenna and focusing effects. A finite array of these small clamped holes is embedded within an infinite elastic plate and the full finite element simulations (making use of specially designed perfectly matched layers [Farhat *et al.* (2011)])

are compared to solutions constructed using the continuum long-scale HFH theory to replace the finite array; standard continuity conditions then join this HFH material to the surrounding infinite elastic plate.

This theory captures a rich array of behaviours, for instance the material can be used to display the shielding effect of the zero frequency stop band or strong anisotropy with propagation along characteristics, that is well captured by HFH. When the coefficients T_{ij} of the HFH theory are of opposite sign the theory captures the strong anisotropy of the medium and the effective material becomes hyperbolic rather than elliptic, which unveils the similar lensing effect, shown in Fig. 1.10(c), without negative refractive index similarly to that observed in [Luo *et al.* (2002)]. Qualitatively one can notice the opposite slope of the dispersion curves at the second standing wave frequency on point X that is responsible for such effects. Quantitatively the difference in group velocities in the two orthogonal directions is determined with coefficients T_{11} and T_{22} and cannot be evaluated by only looking at the band diagram.

Figures 1.10(c) and (d) show a focusing effect wherein a point source located close to a tilted array leads to a localised beam. The behaviour near point M of the Brillouin zone is responsible for the focussing effects of Fig. 1.10(c) where the group velocities are of opposite sign for each of the two directions and waves are therefore directed along the diagonals of the cells, what a physicist would consider as a hyperbolic type metamaterial effect.

1.3.1.2. *Cloaking with a pinned elastic plate*

Last, but not least, the triple crossings in the dispersion diagram in Fig. 1.9 comprise Dirac-like cones with a flat mode passing through the vertex. This is highly interesting given that Dirac-like cones are normally limited to graphene-like hexagonal structures [Neto *et al.* (2009)]; the current situation involves a square lattice is akin to the Dirac-like cones for photonic crystals recently described in [Huang *et al.* (2011); Liu *et al.* (2011)]. One can use the properties of Dirac cones-like to reduce the scattering of a clamped obstacle placed

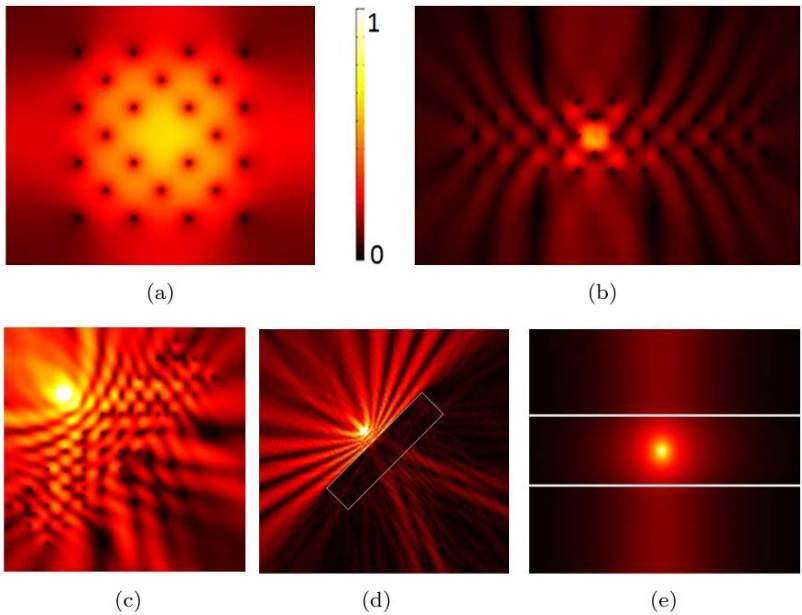


Fig. 1.10. A bending wave excited by a forcing of normalized frequency $\Omega = 3.7952$ (that of vanishing group velocity at Γ point in Fig. 1.9) inside an array of clamped circles (pitch 2, radius 0.01) tilted through an angle $\pi/4$ gives rise to respectively four (a) and two (b) highly-directed beams outside 24 clamped circles making a square (a) and 47 clamped circles making a rectangle (b). Focusing through a slab of 48 clamped circles is shown in (c) with its HFH equivalent in (d) for a point forcing at normalized frequency of $\Omega = 6.58$. The focusing (d) and antenna (e) effects are simulations of the continuum PDEs generated by HFH with respective (T_{11}, T_{22}) coefficients of $(25.65, -11.18)$ at point X and $(6.2524, 6.2524)$ at point Γ, where the effective media are highlighted by white lines. Panels (a) and (b) use the array in the ΓM symmetry direction and panel (c) uses the ΓX symmetry direction. Figure taken from [Antonakakis *et al.* (2014b)].

within a periodic array, as demonstrated in Fig. 1.11(a),(b). This behaviour is also captured by HFH asymptotics: In this particular case three coupled equations emerge with variables $f_0^{(i)}$ for $i = 1, 2, 3$. The system decouples to yield the same governing equation for all three functions $f_0^{(i)}$ as,

$$8\pi^6 \frac{\partial^2 f_0}{\partial x_i^2} + \frac{(\Omega^2 - \Omega_0^2)^2}{l^2} f_0 = 0, \tag{1.28}$$

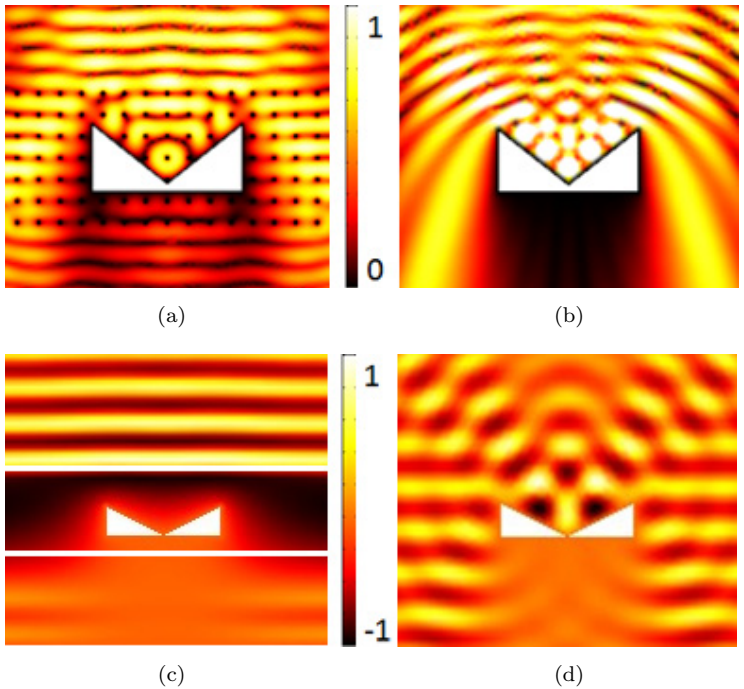


Fig. 1.11. A plane bending wave of normalized frequency $\Omega = 9.7$ (just below the first Dirac-like cone at Γ point in Fig. 1.9) incident from the top on an array of clamped circles (pitch 2, radius 0.01) undergoes considerably less scattering (a) than by a clamped obstacle on its own (b); panels (a, b) are from finite element simulation. The asymptotic HFH PDEs from equation (1.28) capture the essence of physics, and the equivalent results are shown for cloaking (c) and scattering (d) by the same clamped obstacle. Taken from [Antonakakis *et al.* (2013c, 2014b)].

where the coefficient in front of f_0 comes from the first order correction Ω_1 and the change into the original coordinates; the numerics using HFH are shown in Fig. 1.11(c). The front and back scattering do not depend on the shape of the obstacle but mostly on its area. Figure 1.12(c) and (d) show the scattering of three different obstacles (a square, a tilted square and a circle) are virtually indistinguishable, whereas in panels (a) and (b) the three curves can be clearly told apart. The platonic crystal is used to destroy any one-to-one correspondence between the scattered field and the shape of

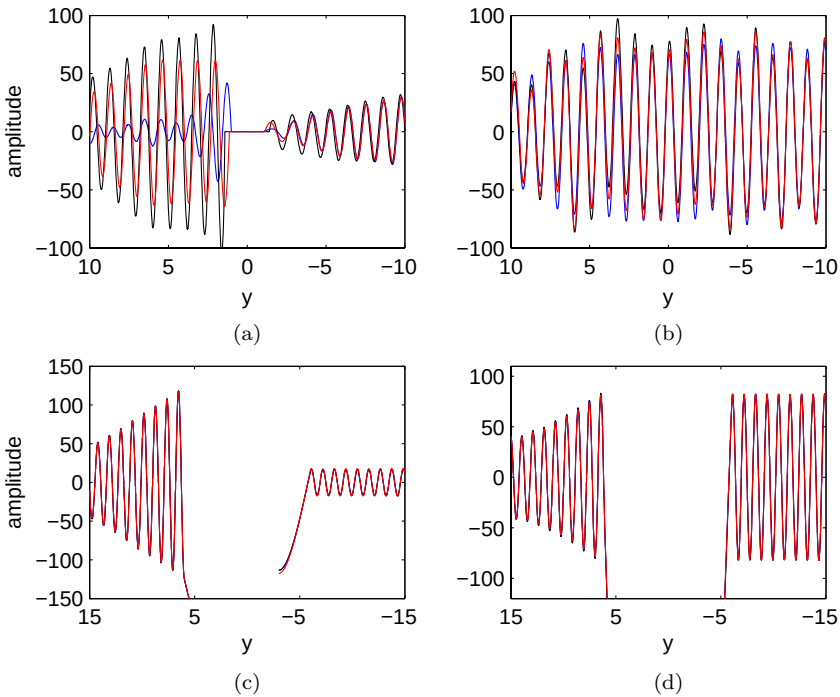


Fig. 1.12. Shape independence for three obstacles of same area ~ 4 centered at $(0, 0)$, namely a square in blue, $\pi/4$ tilted square in black and a circle in red, in the same setting as in Fig. 1.11 but with a bending wave's normalized frequency of $\Omega = 5.75$. (a) and (b) plot forward and back scattering on the respective paths $x = 0$ and $x = 15$ (outside the obstacle abscissa) for three different obstacle shapes without any periodic medium. (c) and (d) plot forward and back scattering on the same paths only this time with an effective medium that surrounds the obstacle just like in Fig. 1.11(c) [Antonakakis *et al.* (2013c)].

an obstacle, which is the essence of cloaking in impedance tomography [Greenleaf *et al.* (2003); Kohn *et al.* (2008)]. However, we note that the present cloaking is constrained to normal incidence and is reminiscent of [Urzhumov and Smith (2010)] which is a more elaborate type of cloaking somewhat constrained to the eikonal limit of transformation optics with photonic band gap media [Liang and Li (2011)].

We hope that our results will foster experimental effort towards Dirac-cone platonic cloaking. Recent experiments have already

validated HFH prediction of dynamic anisotropy in platonic crystals [Lefebvre *et al.* (2016)]. However, to convince experimentalists to engage into the more challenging case of cloaking might require further theoretical and numerical efforts using the more elaborate, full elastic, Navier system, instead of the Kirchhoff-Love theory. In actuality, HFH theory has already been applied to elastic wave equations, and we now pursue our journey in this framework.

1.4. HFH for Elastic Metamaterials

Full vector elasticity is clearly more complicated, however the methodology carries through again. The elastic medium is taken to be linear and isotropic with the stress related to the elastic displacements via the constitutive relation

$$\sigma_{ij} = \mu(u_{i,x_j} + u_{j,x_i}) + \lambda\delta_{ij}u_{k,x_k}. \tag{1.29}$$

The usual Einstein summation convention is adopted and $\mathbf{x} = (x_1, x_2)$ are a Cartesian coordinate system with μ, λ the Lamé parameters and, x_j denoting differentiation with respect to variable x_j. The equation of state, with $\exp(-i\Omega t)$ time dependence assumed and suppressed hereafter, is

$$\sigma_{ij,x_j} + \rho\Omega^2 u_i = 0, \tag{1.30}$$

and ρ is the density. To generate the effective description of the medium we follow the methodology given in [Antonakakis *et al.* (2014a)] with the separation of scales and HFH approach giving the leading to the order solution

$$u_{0_i} = f_0(\mathbf{X})U_{0_i}(\boldsymbol{\xi}; \Omega_0). \tag{1.31}$$

Where U_{0_i} is the ith component of the short scale vector dispalcement field $\mathbf{U}_0$ and it is entirely on the short-scale and its behaviour is modulated by a long-scale, unknown, function $f_0(\mathbf{X})$. The short-scale displacement field is different for each standing wave frequency Ω_0 and we emphasise this by including it in the argument of the U_{0_i}. It is worth dwelling on (1.31) and noting that one could envisage having

two different scalar functions f_{0_1}, f_{0_2} multiplying each displacement field or, if one used a potential formulation, the shear and compressional potentials. However, ultimately for self-consistency one is drawn to conclude that $f_{0_1} = f_{0_2}$ and unexpectedly a single scalar function f_0 emerges on the long-scale.

Associated with the ith component of the short-scale displacement, U_{0_i}, is a stress field, $S_{0_{ij}}(\boldsymbol{\xi}; \Omega_0)$, defined as

$$S_{0_{ij}} = \mu(U_{0_i,\xi_j} + U_{0_j,\xi_i}) + \lambda\delta_{ij}U_{0_k,\xi_k}. \tag{1.32}$$

The leading order equation, for the short scale, is then

$$S_{0_{ij},\xi_j} + \rho\Omega_0^2 U_{0_i} = 0 \tag{1.33}$$

and this can be solved, subject to the boundary conditions, to identify the displacements U_{0i} and stresses S_{0ij}; in the examples treated later these solutions are found numerically.

As in the earlier sections, one moves up the hierarchy of equations arriving at a solvability equation for the longscale function as

$$\begin{aligned} f_{0,X_kX_j} \int_S \mu(U_{0_i}(U_{1_{ik},\xi_j} + U_{1_{jk},\xi_i}) - U_{1_{ik}}(U_{0_i,\xi_j} + U_{0_j,\xi_i}) \\ + \lambda(U_{1_{ik},\xi_i}U_{0_j} + U_{0_k}U_{0_j} - U_{1_{kj}}U_{0_i,\xi_i}))dS \\ + f_{0,X_jX_j} \int_S \mu U_{0_i}U_{0_i}dS + f_{0,X_iX_j} \int_S \mu U_{0_i}U_{0_j}dS \\ + f_0\Omega_2^2 \int_S \rho U_{0_i}U_{0_i}dS = 0. \end{aligned} \tag{1.34}$$

The important point about this equation is that the short-scale is completely absent, it has been integrated out and is encapsulated within coefficients of a rank-2 tensor T_{ij} as (1.34) can be written as

$$T_{ij}\frac{\partial^2 f_0}{\partial X_i \partial X_j} + \Omega_2^2 f_0 = 0. \tag{1.35}$$

We will verify this using non-trivial examples later where we use equation (1.35) to find asymptotic dispersion curves for perfectly periodic Floquet-Bloch problems and compare with full numerical

solutions. We stress that we have reduced the complexity of the elasticity problem described by (1.29) and (1.30) to solving the scalar PDE (1.35), which has the exact same form as the effective equation (1.11) that emerged from the HFH of Helmholtz case. If we were to perform our analysis in a three-dimensional elastic periodic setting (e.g. a cubic array of rigid or traction-free spheres), the derivation of (1.35) would apply mutatis mutandis with i, j ranging from 1 to 3 and the Bloch vector now describing the edges of a tetrahedron. We finally note that Auriault and Boutin [Auriault and Boutin (2012); Boutin *et al.* (2014)] independently developed a theory for long wave wavelength near cut-off frequencies in elastic composite materials, which also produces dynamic effective elastic parameters at high-frequencies.

1.4.1. *Long wave zero frequency limit*

Under traction free boundary conditions on the holes the medium permits wave propagation at frequencies of the order of $\Omega^2 \sim O(\eta^2)$; the case of clamped boundary conditions is fundamentally different and the material exhibits a zero-frequency stop-band and no long-wave low frequency solution, see the dispersion curves of Fig. 1.13. Note this can have potential applications in shielding elastic waves of very long wavelengths (e.g. seismic waves).

For the traction-free case, the leading order equation is $\sigma_{ij,\xi_j} = 0$ and admits two independent solutions for $u_i = f_0^{(l)}(\mathbf{X})U_{0_i}^{(l)}$, where $U_{0_i}^{(l)}$ are constants. Constant U_0 solutions are inserted in the first order equation to yield $\Omega_1 = 0$. Then we solve for $u_{1_i} = f_{0,X_j}^{(l)} U_{1_{ij}}^{(l)}$ where the first order equation is the same as the leading order but the boundary conditions are $\sigma_{1_{ij}} n_j = 0$. Finally Ω_2 is obtained and using that $\Omega_0 = \eta\Omega_2$ gives the linear asymptotics.

1.4.2. *Effective equation*

Changing back to normal coordinates equation (1.34) yields the effective medium equation which is scalar and entails a bulk envelope wave

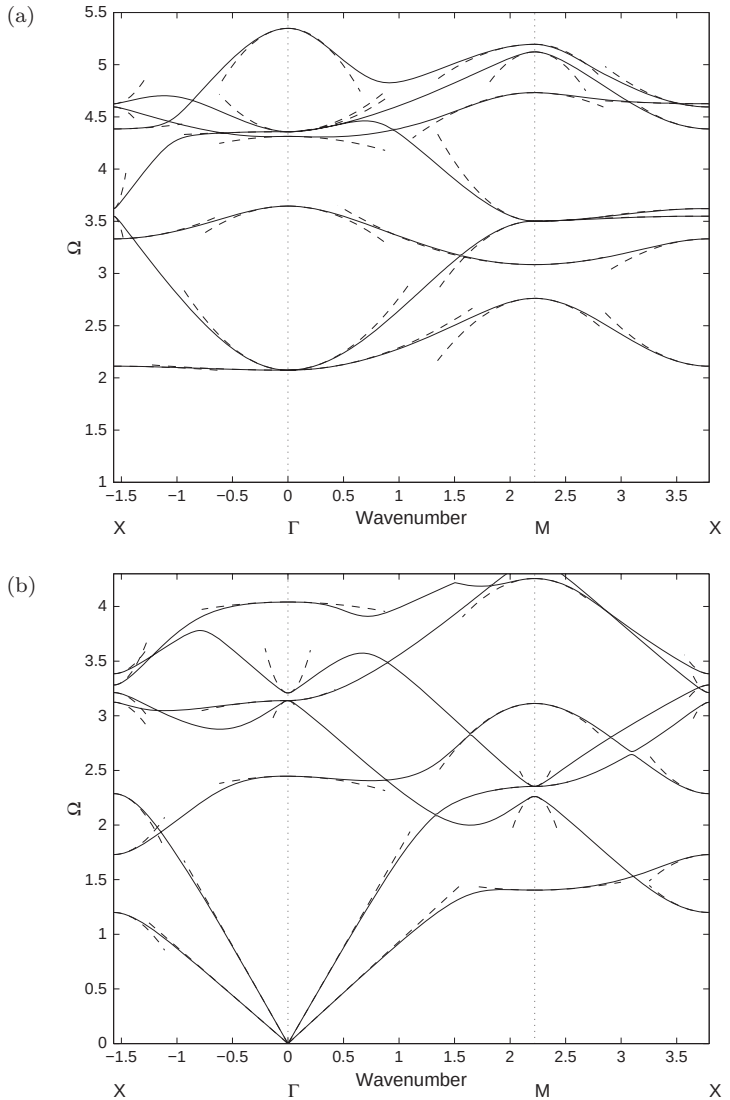


Fig. 1.13. Bloch dispersion curves and HFH asymptotics for cylinders of radius 0.4 in an elementary cell of side 2: (a) Clamped and (b) traction free boundary conditions. The solid lines are from finite element numerical simulations and the dashed lines are from the HFH. Taken from [Antonakakis *et al.* (2014a)].

propagation law even in vector elasticity.

$$T_{ij} f_{0,x_i x_j} + \frac{\Omega^2 - \Omega_0^2}{l^2} f_0 = 0, \tag{1.36}$$

where $l = 1$ is the short characteristic length as half the length of a square cell. By reordering terms in equation (1.36) the effective wave velocity is given by the tensor,

$$c_{ij} = \frac{\Omega^2 l^2}{\Omega^2 - \Omega_0^2} T_{ij}, \tag{1.37}$$

to yield,

$$c_{ij} f_{0,x_i x_j} + \Omega^2 f_0 = 0 \tag{1.38}$$

Equation (1.37) depicts the dependence of the effective velocity on the T_{ij} coefficients as well as on the perturbation away from the standing wave point. The effective medium is often anisotropic elliptic or even hyperbolic causing the occurence of highly directive effects as will be seen in the following illustrations.

1.4.3. *Bloch dispersion curves*

We begin by verifying the accuracy of the asymptotic technique by creating asymptotic dispersion curves from HFH and comparing them to full numerical simulations performed using finite elements, before then moving on to investigate some of the broad features and wave phenomena that can arise in elasticity. For definiteness we fix the normalised elastic Lamé parameters λ, μ as $2.3, 1$ respectively, and similarly for the density $\rho = 1$, in all of the following computations as typical material values. These normalised parameters can be used for elastic media with a Poisson ratio close to 0.35, for instance fused silica, which is drilled to fabricate photonic crystal fibres (in dimensional terms $\rho = 2.2 \times 10^3$ kg m^3, $\lambda = 31.15 \times 10^9$ Pa, $\mu = 16.05 \times 10^9$ Pa).

Typical Bloch dispersion curves, with both the HFH asymptotics (dashed lines) and numerical simulations from finite elements (solid lines) shown are plotted in Fig 1.13 for cylindrical inclusions of radius 0.01 in a square array of side 2. This figure shows the

clamped inclusions in panels (a) and traction free holes in panels (b); an immediate conclusion is that clamped inclusions have a zero frequency stop-band, in common with Dirichlet inclusions in the acoustic/polarised electromagnetic analogues, c.f. [Antonakakis *et al.* (2014b)], whereas traction-free inclusions share features with the analogous Neumann case at low frequencies, i.e., there is a low-frequency linear response. The zero-frequency stop-band behaviour of the clamped inclusions completely derails the classical homogenisation approach of effective media that requires low frequencies and long waves and the microstructured medium supports no waves in this limit: nonetheless the HFH asymptotics accurately represent the dispersion curves close to the edges of the Brillouin zone for both clamped and traction-free cases.

It is satisfying to note that the asymptotics based upon the scalar f_0 equation capture the detailed behaviour of the dispersion curves and this is strong verification of the approach we have employed. We now move on to using the asymptotics to predict and explain specific phenomena. The power of the asymptotic technique, is that it allows us to replace a microstructured medium by an effective dispersive medium whose elastic properties are encapsulated within the coefficients T_{ij} in the PDE for f_0, or by the coefficients in the coupled cases of repeated roots. The form of the f_0 equation then guides one to anticipate or predict particular features.

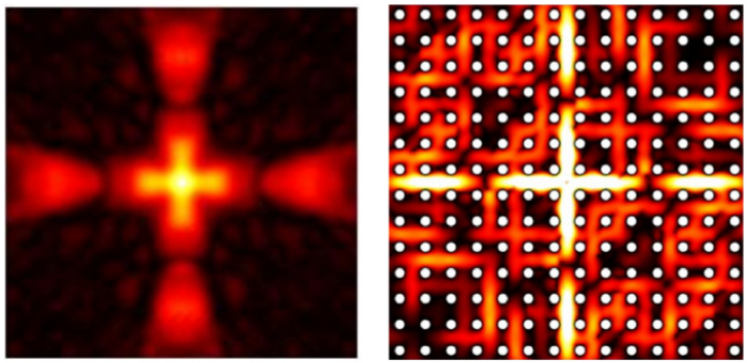
1.4.4. *Elastic metamaterials applications*

1.4.4.1. *Highly-directional standing waves*

Highly-directional waves are shown in Fig. 1.14 and the cross effect is predicted by the T_{ij} coefficients where one is of order one while the other is nearly null and vice versa, depending on which point of the Brillouin zone is considered, $X(\pi/2, 0)$ or $G(0, \pi/2)$.

1.4.4.2. *Unidirectivity*

Another application of the unidirectivity of the effective medium is shown in Fig. 1.15 where the source's position and direction is defined by the arrows. In all four panels the excitation frequency is $\Omega = 2.19$



(a) (b)

Fig. 1.14. Directive emission in horizontal and vertical directions, caused by a point compressive source at frequency $\Omega = 2.19$, in a doubly periodic array of square cells with circular holes of radius $r = 0.4$, clamped at their surface. c.f Fig. 1.13(a) first mode at point X. Panel (b) shows FEM calculations and panel (a) reproduces the effects by HFH with $T_{11} = -0.18255$ and $T_{22} = 2.46973$ (see Table 1.2). Taken from [Antonakakis *et al.* (2014a)].

Table 1.2. The first seven standing wave frequencies for a square cell with clamped holes of radius $r = 0.4$ at wavenumber $\boldsymbol{\kappa} = (\pi/2, 0)$ at X, c.f. Fig. 1.13(a), together with associated values for T_{11} and T_{22}. Opposite signs of T_{11}, T_{22} together with absolute values of similar orders of magnitude yield characteristic type of propagation. If one coefficient is much less in magnitude than the other, e.g. $|T_{11}| \ll |T_{22}|$, then directive propagation emerges in the horizontal and vertical directions. The above coefficients are used in Fig. 1.14.

T_{11}	T_{22}	Ω_0
−0.1826	2.470	2.112
1.6246	−2.4885	3.3317
−196.4900	−0.1476	3.5485
200.6397	−0.6115	3.6216
1.0742	4.9378	4.3847
−28.7271	4.0614	4.5952
27.4797	0.2967	4.6258

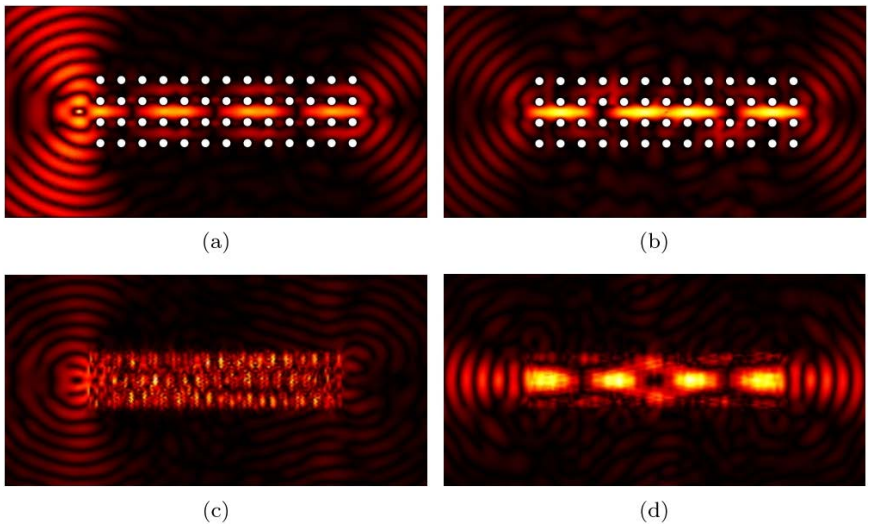


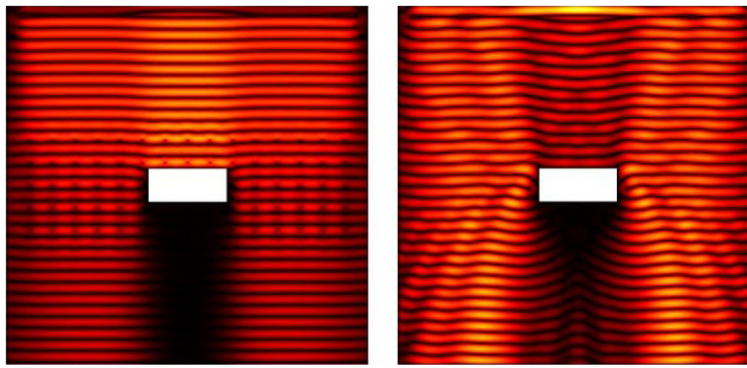
Fig. 1.15. A doubly periodic rectangular array of 52 clamped circular holes of radius $r = 0.4$, is excited at a frequency of $\Omega = 2.19$ c.f. Fig. 1.13(a) and with a prescribed direction of $(1, 0)$ and $(1, 1)$ for panels (a) and (b) respectively. Panels (c) and (d) show the equivalent effects by replacing the periodic array by an effective medium of material values prescribed by HFH. The preferred direction of propagation is in the horizontal direction. The effect is magnified when the source is placed inside the array as shown in panels (b) and (d). The HFH coefficients of interest are $T_{11} = 2.4697$ and $T_{22} = -0.1826$ and at the excited frequency there is a superposition of two effective media each with interchanged values of T_{11} and T_{22}.

and is close to the standing wave frequency $\Omega_0 = 2.1120$, where the effective medium admits exactly one $T_{ii} \approx 0$ coefficient. This is near the first mode of Fig. 1.13(a) where $T_{11} = 2.4697$ and $T_{22} = -0.1826$. The prefered directivity of the source is in the horizontal direction due to the large length to height ratio of the crystal. The two dimensional effective medium is then governed by equation 1.36,

$$2.4697 f_{0,x_1x_1} - 0.1826 f_{0,x_2x_2} + 0.3440 f_0 = 0 \tag{1.39}$$

1.4.4.3. *Dirac cone*

A cloaking application is observed in Fig. 1.16 where an obstacle is placed in a photonic crystal slab composed of a doubly periodic



(a) (b)

Fig. 1.16. An elastic slab clamped at a doubly periodic array of 72 points with a rectangular hole of width 7 and height 3 is placed in between a homogeneous elastic material. In panel (a) a normal incident wave, at frequency $\Omega = 3.14$ where the dispersion curves are reminiscent of a Dirac-like cone, coming from above travels through the crystal which removes all detail in the front and back scattering fields when compared to the same obstacle placed alone in the homogeneous medium shown in panel (b). Although the amplitude of the wave is damped behind the obstacle in panel (a), the wavefronts are neatly flattened compared to panel (b).

array of very small clamped holes ($r = 0.01$) with the dispersion diagram given in [Antonakakis *et al.* (2014a)]. This is not the typical definition of cloaking as the transmitted field in panel (a) contains a shadow of the obstacle but the wavefront appears to stay parallel to that of the incident one and so removes all detail of the hidden obstacle. On the contrary in panel (b) the wavefront is distorted which reveals information about the obstacle.

1.4.5. *HFH for periodic bianisotropic media*

As a further note, we can apply this approach very widely, for instance there is currently a surge for effective properties of advanced designs of photonic crystals and electromagnetic metamaterials. In most cases electric and magnetic fields are inherently coupled, even in the transverse case, and so called bianisotropy comes in to play. In this section, we would like to make a note of a useful correspondence between in-plane elastodynamic and electromagnetic wave

equations, which can be used to translate HFH results of section 1.4 to bianisotropic media. Let us first recall that an electromagnetic field $(\mathbf{E}, \mathbf{H})$ is solution to the so-called time harmonic Maxwell-Tellegen equations in a bianisotropic media:

$$\nabla \times \mathbf{E} = \omega\kappa_H \mathbf{E} + i\omega\mu \mathbf{H}, \quad \nabla \times \mathbf{H} = -i\omega\varepsilon \mathbf{E} + \omega\kappa_E \mathbf{H}, \tag{1.40}$$

where ω is the angular wave frequency, ε the permittivity, μ the permeability and κ the magneto-electric coupling parameter.

If we assume that the material parameters and the electromagnetic field are independent of the longitudinal coordinate x_3 (i.e. we consider a bianisotropic medium invariant along x_3), we can write (1.40) as six coupled PDEs:

$$\begin{aligned}
&\partial_{x_2} E_3 = \omega\kappa_H (E_1 + E_2) + i\omega\mu (H_1 + H_2),\\
&-\partial_{x_1} E_3 = \omega\kappa_H (E_1 + E_2) + i\omega\mu (H_1 + H_2),\\
&\partial_{x_1} E_2 - \partial_{x_2} E_1 = \omega\xi E_3 + i\omega\mu H_3,\\
&\partial_{x_2} H_3 = \omega\kappa_E (H_1 + H_2) - i\omega\varepsilon (E_1 + E_2),\\
&-\partial_{x_1} H_3 = \omega\xi (E_1 + E_2) - i\omega\mu (H_1 + H_2),\\
&\partial_{x_1} H_2 - \partial_{x_2} H_1 = \omega\kappa_E H_3 - i\omega\varepsilon E_3.
\end{aligned}$$

Noting that one can express the above equation only in terms of the longitudinal components of the electric and magnetic fields, then one can recast (1.40) as follows [Liu *et al.* (2016)]

$$\nabla \cdot (\mathbf{C} : \nabla \mathbf{u}) + \rho\omega^2 \mathbf{u} = \mathbf{0}\,, \tag{1.41}$$

where $\mathbf{C}$ is the rank-4 elasticity tensor with the following non-zero entries:

$$\begin{aligned}
&C_{1111} = C_{2222} = -\tfrac{\varepsilon}{\mu\varepsilon - \kappa_E\kappa_H}, \quad C_{1211} = C_{1222} = \tfrac{i\kappa_E}{\mu\varepsilon - \kappa_E\kappa_H},\\
&C_{2111} = C_{2122} = \tfrac{i\kappa_H}{\mu\varepsilon - \kappa_E\kappa_H}, \qquad C_{2211} = C_{2222} = \tfrac{\mu}{\mu\varepsilon - \kappa_E\kappa_H}
\end{aligned}$$

where $\varepsilon\mu - \kappa_E\kappa_H \neq 0$ and the displacement field $\mathbf{u} = (E_3, H_3)$. Besides from that, ρ is a rank-2 density tensor whose entries are

given by:

$$\rho_{11} = \varepsilon, \quad \rho_{12} = i\kappa_E, \quad \rho_{21} = i\kappa_H, \quad \rho_{22} = -\mu. \tag{1.42}$$

One can then straightforwardly apply the homogenization results of elastic crystals to bianisotropic crystals. However, this correspondence does not hold for three-dimensional periodic structures, in which case one need generalize high-frequency homogenization of the Maxwell's system [Maling *et al.* (2016)] to the Maxwell-Tellegen's system. Taking the low frequency limit in HFH, one recovers result of [Guenneau and Zolla (2007)].

1.5. Concluding Remarks

In this chapter, we reviewed the key features of HFH which is a universal approach for any wave equation in physics. HFH can also be applied to diffusion problems, such as heat or mass diffusion in thermal and biological metamaterials. A key feature of HFH is that it identifies simplified, limit, equations that unveil deep analogies between quite different physics areas. For instance, anti-plane shear/pressure/transverse electric waves in cylindrical media (all of which are solutions of a Helmholtz equation) and flexural waves in thin plates (solutions of a Kirchhoff-Love equation) are governed by the same second order limit PDE unveiled by HFH. This makes possible mathematical correspondences between different wave physics, and in turn facilitate designs of mechanical metamaterials by analogy with electromagnetic metamaterials at least in a qualitative sense. Another important point is the flexibility of the HFH approach, it can be simply accommodated with most linear PDEs, for instance in water wave theory one might wish to use the mild-slope equation instead of the Helmholtz equation to model water waves propagating in structured liquids with a varying depth. In fact, applicability of HFH is virtually unlimited. We believe the future of HFH is bright as many areas of physics of metamaterials remain unexplored, notably three-dimensional bianisotropic or high-contrast media that have considerable interest for negative refraction, and non-linear and dispersive media in general.

Acknowledgments

RVC thanks the EPSRC (UK) for their support through research grants EP/I018948/1 & EP/J009636/1 and Mathematics Platform grant EP/I019111/1. S.G. would like to acknowledge a European Research Council funding (ERC grant ANAMORPHISM) that helped strengthening links with the group of R.V.C. at Imperial College in 2011–2016.

Bibliography

[1] Achaoui, Y., Antonakakis, T., Brul, S., Craster, R., Enoch, S. and Guenneau, S. (2016). Seismic defence structures, Patent.

[2] Achaoui Y., Antonakakis T., Brule S., Craster R.V., Enoch S., Guenneau S. (2017). Clamped seismic metamaterials: Ultra-low broad frequency stop-bands, arXiv preprint arXiv:1701.08841.

[3] Allaire, G. and Conca, C. (1998). Bloch wave homogenization and spectral asymptotic analysis, *Journal de Mathématiques Pures et Appliquées* **77**, pp. 153–208.

[4] Allaire, G. and Piatnitski, A. (2005). Homogenisation of the Schrödinger equation and effective mass theorems, *Commun. Math. Phys.* **258**, pp. 1–22.

[5] Antonakakis, T., Craster, R. and Guenneau, S. (2014a). Homogenisation for elastic photonic crystals and dynamic anisotropy, *Journal of the Mechanics and Physics of Solids* **71**, pp. 84–96.

[6] Antonakakis, T. and Craster, R. V. (2012). High frequency asymptotics for microstructured thin elastic plates and platonics, *Proc. R. Soc. Lond. A* **468**, pp. 1408–1427.

[7] Antonakakis, T., Craster, R. V. and Guenneau, S. (2013a). Asymptotics for metamaterials and photonic crystals, *Proc. R. Soc. Lond. A* **469**, p. 20120533.

[8] Antonakakis, T., Craster, S. and Guenneau, S. (2013c). Moulding flexural waves in elastic plates lying atop a Faqir's bed of nails, arXiv:1301.7653v2.

[9] Antonakakis, T., Craster, R. V. and Guenneau, S. (2013b). High-frequency homogenization of zero frequency stop band photonic and phononic crystals, *New Journal of Physics* **15**, p. 103014.

[10] Antonakakis, T., Craster, R. V. and Guenneau, S. (2014b). Moulding and shielding flexural waves in elastic plates, *Euro. Phys. Lett.* **105**, p. 54004.

[11] Achaoui Y., Antonakakis T., Brule S., Craster R.V., Enoch S., and Guenneau S. (2017). Clamped seismic metamaterials: Ultra-low broad frequency stop-bands, *New J. Phys.* **19**, p. 063022.

[12] Auriault, J. and Boutin, C. (2012). Long wavelength inner-resonance cut-off frequencies in elastic composite materials. *International Journal of Solids and Structures* **49**, p. 32693281.
[13] Babuska, I. (1975). Homogenization and its application. Mathematical and computational problems, *Numerical solution of Partial differential Equations, III, Proc. Third Sympos. Synspade, Univ. Maryland, college Park, Md.*
[14] Bakhvalov, N. and Panasenko, G. (1989). *Homogenization: Averaging Processes in Periodic Media* (Kluwer, Amsterdam).
[15] Batchelor, G. (1972). Sedimentation in a dilute dispersion of spheres, *Journal of fluid Mechanics* **52**, pp. 245–268.
[16] Bensoussan, A., Lions, J. and Papanicolaou, G. (1978). *Asymptotic analysis for periodic structures* (North-Holland, Amsterdam).
[17] Berdichevskii, V. L. (1977). Equations of high-frequency long-wave vibrations of plates, *Doklady Akad. Nauk SSSR (Soviet Phys. Dokl.)* **236**, pp. 1319–1322, in Russian.
[18] Birman, M. S. and Suslina, T. A. (2006). Homogenization of a multidimensional periodic elliptic operator in a neighborhood of the edge of an internal gap, *Journal of Mathematical Sciences* **136**, pp. 3682–3690.
[19] Blanc, X., Bris, C. L. and Lions, P. (2007). Stochastic homogenization and random lattices, *J. Math. Pures Appl.* **88**, pp. 34–63.
[20] Born, M. and Wolf, E. (1970). *Principles of Optics* (Pergamon Press).
[21] Bouchitté, G. (1990). Analyse limite de la diffraction d'ondes electomagnetiques par une structure mince, *Compres Rendus Academie des Sciences Serie II* **311**, pp. 51–56.
[22] Bouchitté, G. and Felbacq, D. (2004). Homogenization near resonances and artificial magnetism from dielectrics, *Comptes Rendus Mathématique* **339**, pp. 377–382.
[23] Bouchitté, G., Guenneau, S. and Zolla, F. (2010). Homogenization of dielectric photonic crystals, *Multiscale Modeling and Simulation* **8**, pp. 1862–1881.
[24] Bouchitté, G. and Petit, R. (1985). Homogenization techniques as applied in the electromagnetic theory of gratings. *Electromagnetics* **5**, pp. 17–36.
[25] Boutin, C., Rallu, A. and Hans, S. (2014). Large scale modulation of high frequency waves in periodic elastic composites, *Journal of the Mechanics and Physics of Solids* **70**, pp. 362–381.
[26] Brillouin, L. (1953). *Wave propagation in periodic structures: electric filters and crystal lattices*, 2nd edn. (Dover, New York).
[27] Bruggeman, D. (1935). Berechnung verschiedener physikalischer konstanten von heterogenen substanzen. i. dielektrizitatskonstanten und leitfahigkeiten der mischkorper aus istropen substanzen, *Annalen der Physik* **416**, pp. 636–664.
[28] Brûlé, S., Javelaud, E. H., Enoch, S. and Guenneau, S. (2014). Experiments on seismic metamaterials: Molding surface waves, *Phys. Rev. Lett.* **112**, p. 133901.

[29] Ceresoli, L., Abdeddaim, R., Antonakakis, T., Maling, B., Chmiaa, M., Sabouroux, P., Tayeb, G., Enoch, S., Craster, R. and Guenneau, S. (2015). Dynamic effective anisotropy: Asymptotics, simulations, and microwave experiments with dielectric fibers, *Phys. Rev. B* **92**, p. 174307.

[30] Chan, C. T., Huang, X., Liu, F. and Hang, Z. H. (2012). Dirac dispersion and zero-index in two dimensional and three dimensional photonic and phononic systems, *Progress In Electromagnetics Research B* **44**, pp. 163–190.

[31] Chin, S. K., Nicorovici, N. A. and McPhedran, R. C. (1994). Green's function and lattice sums for electromagnetic scattering by a square array of cylinders, *Phys. Rev. E* **49**, 5, pp. 4590–4602, doi:10.1103/PhysRevE.49.4590.

[32] Cioranescu, D., Damlamian, A. and Griso, G. (2008). The periodic unfolding method in homogenization, *SIAM Journal on Mathematical Analysis* **40**, pp. 1585–1620.

[33] Colombi, A., Guenneau, S., Roux, P., Craster, R. V. (2016). Transformation seismology: composite soil lenses for steering surface elastic Rayleigh waves, *Sci. Rep.* **6**, 25320.

[34] Clausius, R. (1858). *Die Mechanische Behandlung der Elektricitat* (Vieweg, Braunschweig).

[35] COMSOL (2012). *www.comsol.com*.

[36] Conca, C., Planchard, J. and Vanninathan, M. (1995). *Fluids and Periodic structures* (Res. Appl. Math., Masson, Paris).

[37] Craster, R. V., Antonakakis, T., Makwana, M. and Guenneau, S. (2012). Dangers of using the edges of the Brillouin zone, *Phys. Rev. B* **86**, p. 115130.

[38] Craster, R. V. and Guenneau, S. (eds.) (2012). *Acoustic Metamaterials* (Springer-Verlag).

[39] Craster, R. V., Joseph, L. M. and Kaplunov, J. (2014). Long-wave asymptotic theories: The connection between functionally graded waveguides and periodic media, *Wave Motion* **51**, pp. 581–588.

[40] Craster, R. V., Kaplunov, J. and Pichugin, A. V. (2010a). High frequency homogenization for periodic media, *Proc. R. Soc. Lond. A* **466**, pp. 2341–2362.

[41] Craster, R. V., Kaplunov, J. and Postnova, J. (2010b). High frequency asymptotics, homogenization and localization for lattices, *Q. Jl. Mech. Appl. Math.* **63**, pp. 497–519.

[42] Dossou, K. B., Botten, L. C., McPhedran, R. C., Asatryan, A. A. and Martijn de Sterke, C. (2007). Gap-edge asymptotics of defect modes in two-dimensional photonic crystals, *Optics Express* **15**, pp. 4753–4762.

[43] Dowling, J. P. and Bowden, C. M. (1994). Anomalous index of refraction in photonic bandgap materials, *J. Mod. Optics* **41**, pp. 345–351.

[44] Enoch, S., Tayeb, G., Sabouroux, P., Guérin, N. and Vincent, P. (2002). A metamaterial for directive emission, *Phys. Rev. Lett.* **89**, p. 213902.

[45] Farhat, M., Guenneau, S. and Enoch, S. (2010a). High-directivity and cofinement of flexural waves through ultrarefraction in thin perforated plates, *European Physics Letters* **91**, p. 54003.
[46] Farhat, M., Guenneau, S. and Enoch, S. (2011). Finite elements modeling of scattering problems for flexural waves in thin plates: Application to elliptic invisibility cloaks, rotators and the mirage effect, *Journal of Computational Physics* **230**, pp. 2237–2245.
[47] Farhat, M., Guenneau, S., Enoch, S., Movchan, A. and Petursson, G. (2010b). Focussing bending waves via negative refraction in perforated thin plates, *Appl. Phys. Lett.* **96**, p. 081909.
[48] Figotin, A. and Kuchment, P. (1998). Spectral properties of classical waves in high contrast periodic media, *SIAM J. Appl. Math.* **58**, pp. 683–702.
[49] Figotin, A. and Vitebskiy, I. (2006). Slow light in photonic crystals, *Waves Random Complex Media* **16**, pp. 293–392.
[50] Gazalet, J., Dupont, S., Kastelik, J., Rolland, Q. and Djafari-Rouhani, B. (2013). A tutorial survey on waves propagating in periodic media: Electronic, photonic and phononic crystals. perception of the bloch theorem in both real and fourier domains. *Wave Motion* **50**, p. 619654.
[51] Graff, K. F. (1975). *Wave motion in elastic solids* (Oxford University Press).
[52] Gralak, B., Enoch, S. and Tayeb, G. (2000). Anomalous refractive properties of photonic crystals, *J. Opt. Soc. Am. A* **17**, pp. 1012–1020.
[53] Greenleaf, A., Lassas, M. and Uhlmann, G. (2003). On nonuniqueness for calderon's inverse problem, *Mathematical Research Letters* **10**, pp. 685–693.
[54] Gridin, D., Adamou, A. T. I. and Craster, R. V. (2004). Electronic eigenstates in quantum rings: Asymptotics and numerics, *Phys. Rev. B* **69**, p. 155317.
[55] Gridin, D. and Craster, R. V. (2003). Quasi-modes of a weakly curved waveguide, *Proc. R. Soc. Lond. A* **459**, pp. 2909–2931.
[56] Guenneau, S., Movchan, A. B., Petursson, G. and Ramakrishna, S. A. (2007a). Acoustic metamaterials for sound focusing and confinement (doi:10.1088/1367-2630/9/11/399), *New J. Phys.* **9**, p. 399.
[57] Guenneau, S. and Ramakrishna, S. A. (2009). Negative refractive index, perfect lenses and checkerboards: Trapping and imaging effects in folded optical spaces, *C. R. Physique* **10**, pp. 352–378.
[58] Guenneau, S., Ramakrishna, S. A., Enoch, S., Chakrabarti, S., Tayeb, G. and Gralak, B. (2007b). Cloaking and imaging effects in plasmonic checkerboards of negative epsilon and mu and dielectric photonic crystal checkerboards, *Photonics and Nanostructures Fundamentals and Applications* **5**, pp. 63–72.
[59] Guenneau, S. and Zolla, F. (2000). Homogenization of three-dimensional finite photonic crystals, *Progress In Electromagnetic Research* **27**, pp. 91–127.

[60] Guenneau, S. and Zolla, F. (2007). Homogenization of 3d finite chiral photonic crystals, *Physica B: Condensed Matter* **394**, pp. 145–147.
[61] Haslinger, S. G., Craster, R. V., Movchan, A. B., Movchan, N. V. and Jones, I. S. (2016). Dynamic interfacial trapping of flexural waves in structured plates, *Proc. R. Soc. Lond. A* **472**, p. 20150658.
[62] Hoefer, M. A. and Weinstein, M. I. (2011). Defect modes and homogenization of periodic Schrödinger operators, *SIAM J. Math. Anal.* **43**, pp. 971–996.
[63] Huang, X., Lai, Y., Hang, Z. H., Zheng, H. and Chan, C. T. (2011). Dirac cones induced by accidental degeneracy in photonic crystals and zero-refractive-index materials, *Nature Materials* **10**, pp. 582–586.
[64] Jikov, V., Kozlov, S. and Oleinik, O. (1994). *Homogenization of differential operators and integral functionals* (Springer-Verlag).
[65] Joannopoulos, J. D., Johnson, S. G., Winn, J. N. and Meade, R. D. (2008). *Photonic Crystals, Molding the Flow of Light*, 2nd edn. (Princeton University Press, Princeton).
[66] John, S. (1987). Strong localization of photons in certain disordered dielectric superlattices, *Phys. Rev. Lett.* **58**, pp. 2486–2489.
[67] Johnson, S. G. and Joannopoulos, J. D. (2001). Bock-iterative frequency-domain methods for Maxwell's equations in a planewave basis, *Optics Express* **8**, pp. 173–190.
[68] Kaplunov, J. D., Kossovich, L. Yu. and Nolde, E. V. (1998). *Dynamics of Thin Walled Elastic Bodies* (Academic Press, New York).
[69] Kozlov, S. M. (1977). Averaging of differential operators with almost periodic, fast oscillating coefficients, *Dokl. Akad. Nauk SSSR* **236**, 1068–1071.
[70] Kohn, R., Shen, H., Vogelius, M. and Weinstein, M. (2008). Cloaking via change of variables in electric impedance tomography, *Inverse Problems* **24**, p. 015016.
[71] Langley, R. S. (1997). The response of two-dimensional periodic structures to point harmonic forcing, *J. Sound Vib.* **197**, pp. 447–469.
[72] Lefebvre, G., Antonakakis, T., Achaoui, Y., Craster, R., Guenneau, S. and Sebbah, P. (2016). Unveiling extreme anisotropy in elastic structured media, *arXiv:1610.04884*.
[73] Liang, Z. and Li, J. (2011). Scaling two-dimensional photonic crystals for transformation optics, *Opt. Express* **19**, p. 16821.
[74] Liu, F., Lai, Y., Huang, X. and Chan, C. T. (2011). Dirac cones at $\overline{k} = 0$ in photonic crystals, *Phys. Rev. B* **84**, p. 224113.
[75] Liu, Y., Gralak, B. and Guenneau, S. (2016). Finite element analysis of electromagnetic waves in two-dimensional transformed bianisotropic media, *Optics Express* **24**, p. 26479.
[76] Liu, Y., Guenneau, S. and Gralak, B. (2013). Artificial dispersion via high-order homogenization: magnetoelectric coupling and magnetism from dielectric layers, *Proc. R. Soc. A* **469**, p. 20130240.

[77] Llewellyn-Smith, S. G. and Davis, A. M. J. (2010). The split ring resonator, *Proc. R. Soc. Lond. A* **466**, pp. 3117–3134.
[78] Luo, C., Johnson, S. G. and Joannopoulos, J. D. (2002). All-angle negative refraction without negative effective index, *Physical Review B* **65**, p. 201104.
[79] Mace, B. R. (1996). The vibration of plates on two-dimensionally periodic point supports, *J. Sound Vib.* **192**, pp. 629–643.
[80] Maling, B., Colquitt, D. and Craster, R. (2016). Dynamic homogenisation of maxwells equations with applications to photonic crystals and localised waveforms on gratings, *Wave motion* **69**, pp. 35–49.
[81] McPhedran, R., Poladian, L. and Milton, G. (1988). Asymptotic studies of closely spaced, highly conducting cylinders, *Proc. R. Soc. Lond. A* **415**, pp. 185–196.
[82] McPhedran, R. C., Movchan, A. B. and Movchan, N. V. (2009). Platonic crystals: Bloch bands, neutrality and defects, *Mechanics of Materials* **41**, pp. 356–363.
[83] McPhedran, R. C., Nicorovici, N. A. and Botten, L. C. (1997). The TEM mode and homogenization of doubly periodic structures, *J. Elect. Waves and Appl.* **11**, pp. 981–1012.
[84] Mead, D. J. (1996). Wave propagation in continuous periodic structures: research contributions from Southampton 1964–1995, *J. Sound Vib.* **190**, pp. 495–524.
[85] Melrose, D. B. and McPhedran, R. C. (2005). *Electromagnetic processes in dispersive media* (Cambridge University Press, Cambridge).
[86] Milton, G. W. (2002). *The Theory of Composites* (Cambridge University Press, Cambridge).
[87] Milton, G. W. and Nicorovici, N. A. (2006). On the cloaking effects associated with anomalous localised resonance, *Proc. Roy. Lond. A* **462**, p. 3027.
[88] Mossotti, O. (1850). *Memorie di Matematica e di Fisica della Societa Italiana delle Scienze Residente in Modena, part 2* **24**, pp. 49–74.
[89] Movchan, A. B., Movchan, N. V. and McPhedran, R. C. (2007). Bloch-Floquet bending waves in perforated thin plates, *Proc. R. Soc. Lond. A* **463**, pp. 2505–2518.
[90] Movchan, A. B., Movchan, N. V. and Poulton, C. G. (2002). *Asymptotic Models of Fields in Dilute and Densely Packed Composites* (ICP Press, London).
[91] Murat, F. and Tartar, L. (1978). H-convergence, *Seminaire d'Analyse fonctionnelle et numerique de l'Université d'Alger* **107**.
[92] Nemat-Nasser, S., Willis, J. R., Srivastava, A. and Amirkhizi, A. V. (2011). Homogenization of periodic elastic composites and locally resonant sonic materials, *Phys. Rev. B* **83**, p. 104103.
[93] Neto, A. H. C., Guinea, F., Peres, N. M. R., Novoselov, K. S. and Geim, A. K. (2009). The electronic properties of graphene, *Rev. Modern Phys.* **81**, pp. 109–162.

[94] Nicorovici, N. A., McPhedran, R. C. and Botten, L. C. (1995). Photonic band gaps: Non-commuting limits and the 'acoustic band', *Phys. Rev. Lett.* **75**, pp. 1507–1510.

[95] Nolde, E., Craster, R. V. and Kaplunov, J. (2011). High frequency homogenization for structural mechanics, *J. Mech. Phys. Solids* **59**, pp. 651–671.

[96] Notomi, M. (2002). Negative refraction in photonic crystals, *Opt. and Quant. Elec.* **34**, pp. 133–143.

[97] Novoselov, K., Geim, A., Morosov, S., Jiang, D., Zhang, Y., Grigorieva, S. D. I. and Firsov, A. (2004). Electric field effect in atomically thin carbon films, *Science* **306**, pp. 666–669.

[98] O'Brien, S., and Pendry, J. B. (2002). Photonic band gap effects and magnetic activity in dielectric composites, *J. Phys. Cond. Mat.* **14**, pp. 4035–4044.

[99] Papanicolaou, G. and Varadhan, S. (1978). Diffusion in regions with many small holes, *Stochastic differential systems, proc. IFIP-WG 7/1 Working Conf., Vilnius*, pp. 190–206.

[100] Pendry, J. B. (2004). Negative refraction, *Contemp. Phys.* **45**, pp. 191–202.

[101] Perrins, W., McKenzie, D. and McPhedran, R. (1979). Transport properties of regular arrays of cylinders, *Proc. R. Soc. Lond. A* **369**, pp. 207–225.

[102] Petit, R. (1980). *Electromagnetic theory of gratings, Topics in current physics* (Springer-Verlag, Berlin).

[103] Poulton, C. G., Botten, L. C., McPhedran, R. C., Nicorovici, N. A. and Movchan, A. B. (2001). Non-commuting limits in electromagnetic scattering: asymptotic analysis for an array of highly conducting inclusions, *SIAM J. Appl. Math.* **61**, pp. 1706–1730.

[104] Poulton, C. G., Guenneau, S. and Movchan, A. B. (2004). Noncommuting limits and effective properties for oblique propagation of electromagnetic waves through an array of aligned fibres, *Phys. Rev. B* **69**, p. 195112.

[105] Ramakrishna, S. A. (2005). Physics of negative refractive index materials, *Rep. Prog. Phys.* **68**, pp. 449–521.

[106] Russell, P. S., Marin, E., Diez, A., Guenneau, S. and Movchan, A. B. (2003). Sonic band gap PCF preforms: enhancing the interaction of sound and light, *Opt. Exp.* **11**, p. 2555.

[107] Sanchez-Palencia, E. (1972). équations aux dérivées partielles dans un type de milieux hétérogènes, *Comptes Rendus Académie Sciences Paris, Série A-B*, pp. 395–398.

[108] Sigalas, M. M. (1998). Defect states of acoustic waves in a two-dimensional lattice of solid cylinders, *J. Appl. Phys.* **84**, pp. 3026–3030.

[109] Sanchez-Palencia, E. (1980), Non-homogeneous media and vibration theory, Lecture Notes in Physics 127 (Springer-Verlag, Berlin Heidelberg).

[110] Silveirinha, M. (2007). Metamaterial homogenization approach with application to the characterization of microstructured composites with negative parameters, *Physical Review B* **75**, p. 115104.

[111] Silveirinha, M. and Fernandes, C. (2005). Homogenization of 3-d-connected and nonconnected wire metamaterials, *IEEE transactions on microwave theory and techniques* **53**, pp. 1418–1430.
[112] Spagnolo, S. (1968). Sulla convergenza di soluzioni di equazioni paraboliche ed ellitiche, *Ann. Scuola Norm. Sup. Pisa* **22**, pp. 571–597.
[113] Strutt, J. (1892). On the influence of obstacles arranged in rectangular order upon the properties of a medium. *Phil. Mag.* **34**, pp. 481–502.
[114] Tartar, L. (1975). Problèmes de contrôle des coefficients dans des équations aux dérivées partielles, *Lectures Notes in Econom. and Math. Systems, Springer, Berlin* **107**.
[115] Torrent, D., Mayou, D. and Sanchez-Dehesa, J. (2013). Elastic analog of graphene: Dirac cones and edge states for flexural waves in thin plates, *Phys. Rev. B* **87**, p. 115143.
[116] Urzhumov, Y. and Smith, D. (2010). Transformation optics with photonic band gap media, *Phys. Rev. Lett.* **105**, p. 163901.
[117] Villeneuve, P. R. and Piché, M. (1992). Photonic band gaps in two-dimensional square and hexagonal lattices, *Phys. Rev. B* **46**, pp. 4969–4972.
[118] Wellander, N. and Kristensson, G. (2003). Homogenization of the Maxwell equations at fixed frequency, *SIAM Journal on Applied Mathematics* **64**, pp. 170–195.
[119] Wellander, N., Guenneau, S. and Cherkaev, E. (2017). Two-scale cut-and-projection convergence; homogenization of quasiperiodic structures, *Mathematical Methods in the Applied Sciences*, DOI: 10.1002/mma.4345.
[120] Yablonovitch, E. (1987). Inhibited spontaneous emission in solid-state physics and electronics, *Phys. Rev. Lett.* **58**, p. 2059.
[121] Yablonovitch, E. (2001). Photonic crystals: semiconductors of light, *Scientific American* **285**, pp. 34–41.
[122] Yablonovitch, E. and Gmitter, T. J. (1989). Photonic band structure: The face-centered-cubic case, *Phys. Rev. Lett.*
[123] Zalipaev, V. V., Movchan, A. B., Poulton, C. G. and McPhedran, R. C. (2002). Elastic waves and homogenization in oblique periodic structures, *Proc. R. Soc. Lond. A* **458**, pp. 1887–1912.
[124] Zengerle, R. (1987). Light propagation in singly and doubly periodic waveguides, *J. Mod. Opt.* **34**, pp. 1589–1617.
[125] Zhikov, V. V. (2000). On an extension of the method of two-scale convergence and its applications, *Sb. Math.* **191**, pp. 973–1014.
[126] Zolla, F., Renversez, G., Nicolet, A., Kuhlmey, B., Guenneau, S. and Felbacq, D. (2005). *Foundations of photonic crystal fibres* (Imperial College Press, London).

CHAPTER 2

Acoustic Metamaterial

JUN XU, NAVID NEMATI, NICOLAS VIARD and NICHOLAS FANG*

Massachussetts Institute of Technology, USA

DENIS LAFARGE

Laboratoire d'Acoustique de l'Université du Maine, France

2.1. Introduction

Over the past decade, electromagnetic metamaterials have shown tremendous potential in many disciplines of science and technology, and have led to a broad interest in devising techniques for manipulating different classes of waves with unconventional materials. Among them, acoustic metamaterials show the greatest promise for manipulating acoustic waves in a novel fashion for imaging, communication, detection and sound protection applications. Because of similar governing equations of electromagnetic waves and acoustic waves, which are the Maxwell equations and Helmholtz equations, respectively, the successes of metamaterials study in electromagnetic wave can be extended to the acoustic community. Permittivity and permeability of materials are used to describe electromagnetic wave properties, while acoustic wave properties are determined by

*Corresponding author: nicfang@mit.edu

Table 2.1. Analogy between electromagnetic and acoustic variables and material characteristics.

Electromagnetism (TMz)	Acoustics	Analogy
$\frac{\partial E_z}{\partial x} = -i\omega\mu_y H_y$	$\frac{\partial P}{\partial x} = -i\omega\rho_x v_x$	
$\frac{\partial E_z}{\partial y} = i\omega\mu_x H_x$	$\frac{\partial P}{\partial y} = -i\omega\rho_y v_y$	
$\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = -i\omega\epsilon_z E_z$	$\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} = -i\omega\chi P$	
Electric field E_z	Acoustic pressure P	$P \leftrightarrow -E_z$
Magnetic field $H_x,\ H_y$	Particle velocity $v_x,\ v_y$	$v_x \leftrightarrow -H_y,\ v_y \leftrightarrow H_x$
Permittivity ϵ_z	Dynamic compressibility χ	$\chi \leftrightarrow \epsilon_z$
Permeability $\mu_x,\ \mu_y$	Dynamic density $\rho_x,\ \rho_y$	$\rho_x \leftrightarrow \mu_y,\ \rho_y \leftrightarrow \mu_x$

the mass density and compressibility of materials. Moreover, in a two-dimensional (2D) case, when there is only one polarization mode, the electromagnetic wave has scalar wave formulation. The two sets of equations for the electromagnetic and the acoustic waves in isotropic media are dual of each other by the replacement as shown in Table 2.1 and this isomorphism holds for anisotropic medium as well. Table 2.1 presents the analogy between acoustic and transverse magnetic fields in 2D under harmonic excitation. It is noted that this analogy between the electromagnetic and acoustic waves is not unique. We can also have the one-to-one mapping between the acoustic and transverse electric field. The latter analogy, which is presented in Table 2.2 for 1D propagation case (here, according to x-axis), is used, in particular, within the nonlocal description of acoustic wave propagation (see Section 2.2). From the equivalence shown in Tables 2.1 and 2.2, the desirable effective density and compressibility need to be established by structured material to realize unusual sound wave properties. Therefore, the basic question in acoustic metamaterials is to build "acoustic artificial" atoms, on a scale much smaller than

Table 2.2. Analogy between macroscopic electromagnetics and macroscopic acoustics in 1D, including temporal and spatial dispersion. Macroscopic acoustic variables are described in page 75.

Electromagnetism	Acoustics	Analogy
$\frac{\partial B_z}{\partial t} = -\frac{\partial E_y}{\partial x}$	$\frac{\partial B}{\partial t} = -\frac{\partial V_x}{\partial x}$	
$\frac{\partial D_y}{\partial t} = -\frac{\partial H_z}{\partial x}$	$\frac{\partial D_x}{\partial t} = -\frac{\partial H}{\partial x}$	
$D_y(\omega, k) = \epsilon(\omega, k)\ E_y(\omega, k)$	$D_x(\omega, k) = \rho(\omega, k)\ V_x(\omega, k)$	
$H_z(\omega, k) = \mu^{-1}(\omega, k)\ B_z(\omega, k)$	$H(\omega, k) = \chi^{-1}(\omega, k)\ B(\omega, k)$	
$S_x = E_y\ H_z$	$S_x = H\ V_x$	
Electric field E_y	Velocity field V_x	$V_x \leftrightarrow E_y$
Electric displacement field D_y	Acoustic D_x field (momentum density)	$D_x \leftrightarrow D_y$
Magnetic B_z field	Acoustic B field (condensation)	$B \leftrightarrow B_z$
Magnetic H_z field	Acoustic H field (pressure)	$H \leftrightarrow H_z$
Pointing vector (x-component) S_x	Acoustic part of energy current density S_x	$S_x \leftrightarrow S_x$
Effective permittivity $\epsilon(\omega, k)$	Effective density $\rho(\omega, k)$	$\rho \leftrightarrow \epsilon$
Effective permeability $\mu(\omega, k)$	Effective compressibility $\chi(\omega, k)$	$\chi \leftrightarrow \mu$

the relevant wavelength, to achieve effective parameters beyond the properties existing in nature.

There are several ways to design acoustic "atoms" in acoustic metamaterials. One option is to use mechanical oscillation to build deep subwavelength local resonators. Due to the resonant properties of the artificial structures, they can exhibit novel effective parameters for acoustic waves (i.e. singly or simultaneously negative mass

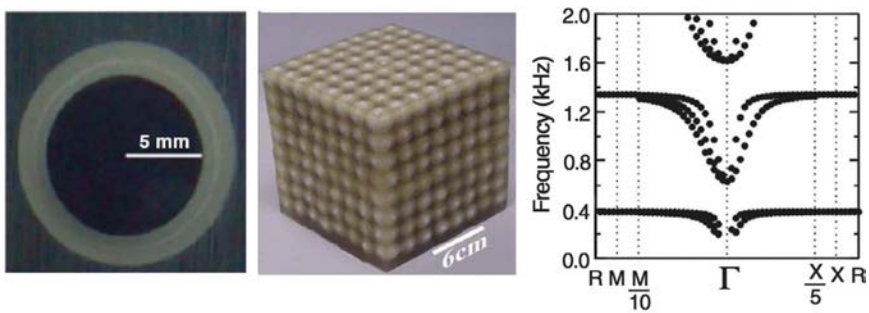


Fig. 2.1. (a) Cross section of a coated lead sphere that forms the basic man-made "atom" with local resonance; (b) Unit cells arranged in an $8 \times 8 \times 8$ sonic crystal, which is the block of acoustic metamaterial that was experimentally tested; (c) A simple cubic structure of coated spheres from 200 to 2000 Hz. Three modes (two transverse and one longitudinal) are distinguishable in the [110] direction, to the left of the Γ point. The two transverse modes are degenerate along the [100] direction, to the right of the Γ point. Note the expanded scale near the Γ point.[1]

density and compressibility) in a certain frequency range.[1–7] The effective parameters of the acoustic metamaterials can be retrieved by analyzing the acoustic wave scattering coefficients. Further studying the oscillation modes of the resonators, we can find either monopolar or dipolar resonances, which lead to negative compressibility or mass density. Some analytical analysis have been applied on simple geometries, e.g. cylindrical or spherical resonators.[8,9] Experimentally, the negative mass density has been demonstrated by immersing core-shell structures (lead core in rubber shell) in epoxy matrix,[1] as shown in Fig. 2.1. In addition, the structure of air bubble in liquid has also been explored because of the different orders of Mie scattering, which lead to monopolar or dipolar resonance at different frequencies to achieve negative compressibility or negative mass density.[4] Recently, another type of resonators — Helmholtz resonators — has been studied to demonstrate negative compressibility.[3,4,7,10,11] A typical Helmholtz resonator (as shown in Section 2.2.2, Fig. 2.3(a)) can be presented as a series of inductance and capacitance. The fluid inside the cavity is much easier to be compressed compared with that in the neck part.

Moreover, the pressure gradient along the open neck is much greater than that inside the large cavity. Therefore the cavity displays capacitive property and leaves the smaller neck as an acoustic inductor. According to the transmission line method[12] (see Section 2.2.1), the negative properties of the acoustic metamaterials can be designed by an array of subwavelength Helmholtz resonators. Experimentally, the negative compressibility was achieved in 1D Helmholtz resonator array,[3] and focusing effect by negative index acoustic metamaterial (mass density and compressibility are negative simultaneously) was demonstrated in 2D Helmholtz resonator array.[13] In addition, using the long-wavelength approximation, the effective mass density and compressibility can also be retrieved from reflection and transmission coefficients.[14,15]

Recently, a new design paradigm called conformal mapping and coordinate transformation has inspired a series of key explorations to manipulate, store and control the flow of energy, in the form of either sound, elastic waves or light radiation. In electromagnetism, because of the coordinate invariance of Maxwell's equations, the space for light can be bent in almost any arbitrary ways by providing a desired spatial distribution of electric permittivity ϵ and magnetic permeability μ.[16,17] A set of novel optical devices were proposed based on transformation optics;[18–21] they usually call for complicated medium with anisotropic and spatially varying permittivity and permeability tensors to accomplish the desired functionality. Recent advances in synthetic structured metamaterials[22,23] whose properties are determined by subwavelength structures, offer the potential to physically implement these complicated media. By modifying the shape and arrangement of these subwavelength constituent elements, anisotropy and spatial variation can be achieved in the artificial metamaterials. Due to the coordinate invariance of Helmholtz equations, the concept can be extended to the acoustic waves as well, which is called transformation acoustics. Similar to the optical metamaterials, generally, we need spatially varying mass density and compressibility tensors to manipulate the acoustic wave in the desired fashion.

The central theme of this chapter is to introduce the design and realization of acoustic metamaterials aimed at controlling and guiding acoustic waves in applications such as acoustic imaging or communication. We will first describe the approach to build an acoustic metamaterial unit cell — Helmholtz resonator — based on the transmission line model. The basic concept and derivation of lumped acoustic circuit will be introduced (Section 2.2.1). The realization of negative dynamic compressibility for the acoustic wave using Helmholtz resonator array will be demonstrated (Section 2.2.2). Newly developed nonlocal theory of the acoustic wave propagation in Helmholtz resonator array will be discussed (Section 2.2.3). By this theory, the effective medium parameters are calculated, and the behavior of the effective modulus which appears negative near the resonance will be shown to be a consequence of the nonlocal effects in this material exhibiting local resonance phenomenon. As a promising application in acoustic imaging, the lens with negative index by acoustic metamaterials will be presented (Section 2.2.4). Both the theoretical and experimental studies demonstrate the focusing of ultrasound waves through the negative index lens, which is implemented by a two-dimensional array of subwavelength Helmholtz resonators. We will also introduce a subwavelength focusing using two-dimensional Helmholtz resonator array based on strong local resonance mode (Section 2.2.5). In the next section (Section 2.3), we will focus on the realization of acoustic cloak in various ways. The design of acoustic cloak is based on transformation acoustics. The mathematical background, numerical simulation results, and experimental characterization results will be given then.

2.2. Acoustic Metamaterials Made of Helmholtz Resonators

In this section, we first review the simple principles of the acoustic transmission line method, followed by the application of the method to design metamaterials with Helmholtz resonator as the material building block. In Section 2.2.2, we describe the first application related to the one-dimensional subwavelength metamaterial, whose elastic modulus is found to be negative experimentally near

resonance frequency.[3] This metamaterial, consisting of an array of Helmhotz resonators, exhibits local resonance phenomena in the long-wavelength limit where the wavelength is much larger than the unit cell composed of one resonator. This medium property which leads to negative modulus is described precisely in Section 2.2.3, by a nonlocal theory of sound propagation taking into account the effects due to spatial dispersion and also viscothermal dissipation.[24] In Section 2.2.4, we present the focusing of a point source from a designed ultrasonic metamaterial consisting of a planar network of subwavelength Helmholtz resonators which has been experimentally investigated in Ref. 13. Transmission line method has been employed to guide us for designing this material and analysing the experimental results. In Section 2.2.5, we demonstrate the extraordinary focusing in an array of Helmholtz resonator.[25] We explain the experimental observations based on an analytical effective-medium model as well as finite-element calculations.

2.2.1. *Transmission line method*

We will review in the following an analogy which can be established between acoustic systems and electrical circuits. This analogy is based on the equivalence of the fluid motion in the acoustic system and the current in the electrical circuit. In the acoustic counterpart of the electrical circuit, acoustic capacitance and inductance can be defined in the framework of a parameter model. This model is valid when the length of the acoustic medium is much smaller than the wavelength. The effective density and compressibility of the network structure are found to be related to the capacitance and inductance in this lumped circuit.

Acoustic impedance of a tube

We consider a hollow cylindrical tube of length l and cross section area S, open at one end and closed at the other end with impedance Z. The origin of coordinates is chosen to be coinciding with the position of the open end of the tube. We assume that the diameter of the tube is sufficiently small so that the waves travel down the tube

with plane wave fronts. For an initial pressure wave traveling in the positive x direction, $p^+ = \tilde{p}^+ \exp(ikx - i\omega t)$, the reflected wave at, traveling in the negative x direction writes as $p^- = \tilde{p}^- \exp(-ikx - i\omega t)$. The corresponding particle velocity can be written as $v^+ = (\tilde{p}^+/Z_c)\exp(ikx - i\omega t)$, and $v^- = -(\tilde{p}^-/Z_c)\exp(-ikx - i\omega t)$, where $Z_c = \rho_0 c_0$ is the characteristic impedance of the fluid in the tube, and ρ_0, and c_0 are the density and sound velocity in the fluid, respectively. The total pressure and total velocity at a given position and time are $p(x,t) = p^+(x,t) + p^-(x,t)$, and $v(x,t) = v^+(x,t) + v^-(x,t)$, respectively. This yields the impedance at x

$$Z_x = \frac{p}{Sv} = \frac{Z_c}{S}\frac{\tilde{p}^+ e^{ikx} + \tilde{p}^- e^{-ikx}}{\tilde{p}^+ e^{ikx} - \tilde{p}^- e^{-ikx}} \tag{2.1}$$

We can immediately obtain the impedances Z_0 at the open end $x = 0$, and Z_l at $x = l$. This gives Z_0 as a function of Z_l

$$Z_0 = \frac{Z_c}{S}\frac{Z_l - i\frac{Z_c}{S}\tan kl}{\frac{Z_c}{S} - iZ_l \tan kl} \tag{2.2}$$

Acoustic inductance

We consider that the tube is acoustically rigid and open on both ends [Fig. 2.2(a)]. Since all quantities are in phase when the dimension of the tube is much smaller than the corresponding wavelength, the fluid moves as a whole under the action of an unbalanced force. The

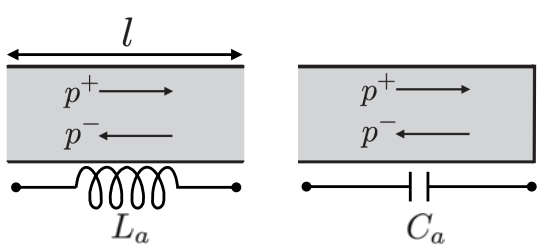


Fig. 2.2. A pipe with (a) open and (b) rigid ends is analogous to an acoustic inductor and an acoustic capacitor, respectively.

whole fluid moves without appreciable compression because of the open ends. We substitute $Z_l = 0$ in (2.2)

$$Z_0 = -i\frac{Z_c}{S}\tan kl \tag{2.3}$$

Since l is much smaller than the wavelength, kl is a small value. Therefore we can use the Taylor series to write $\tan kl = kl + \frac{(kl)^3}{3} + 2\frac{(kl)^5}{15} + \cdots$. Then (2.3) becomes $Z_0 = -i\omega\frac{\rho_0 l}{S} - i\omega^3\frac{\rho_0 l^3}{3Sc_0^2} + \cdots$. When $l < \lambda/16$, we can keep only the first term and neglect the higher order terms within about 5% error. Thus, the acoustic inductance for an open end tube can be written as

$$L_A \simeq \frac{\rho_0 l}{S} \tag{2.4}$$

Taking into account the radiation impedance, l should be replaced by an effective length of the tube l', and augmented with a correction factor: $l' = l + 8r/3\pi = l + 0.85r$, where r is the radius of the tube.

Acoustic capacitance

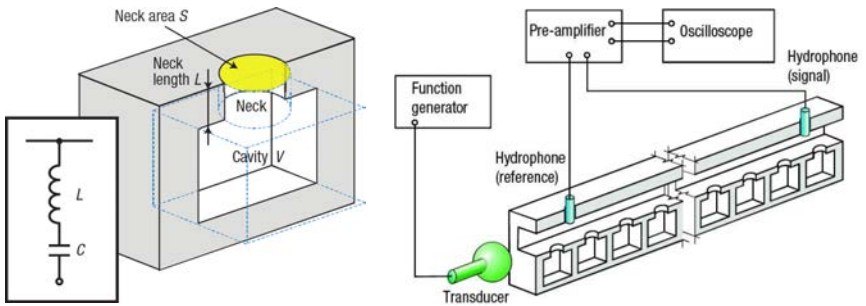
If the tube is rigidely closed at one end [Fig. 2.2(b)], we substitute $Z_l = \infty$ in (2.2)

$$Z_0 = i\frac{Z_c}{S}\cot kl \tag{2.5}$$

For a small value of kl, $\cot kl = \frac{1}{kl} - \frac{kl}{3} - \frac{(kl)^3}{45}$, and consequently (2.5) becomes $Z_0 = \frac{i}{\omega}\frac{1}{(V/\rho_0 c_0^2)} - i\omega\frac{l\rho_0}{3S} + \cdots$. This impedance is valid within 5% for l up to $\lambda/8$ as a combination of acoustic inductance and capacitance in series. Furthermore and as before, when $l < \lambda/16$, we can keep only the first term and neglect the higher order inductance term within about 5% error, such that the impedance

$$C_A \simeq \frac{V}{\rho_0 c_0^2} \tag{2.6}$$

where V is the volume of the tube.



(a) (b)

Fig. 2.3. (a) Schematic cross-sectional view of a Helmholtz resonator. The sample is made of aluminum, consisting of a rectangular cavity of $3.14 \times 4 \times 5$ mm^3, and a cylindrical neck 1-mm long and 1-mm in diameter. The cavity and neck are filled with water, and are connected at the same side to a square water duct with a 4×4 mm^2 opening. The resonators are placed in a periodicity of 9.2 mm. The inset illustrates the analogy between a Helmholtz resonator and an inductor-capacitor circuit, showing the fluidic inductance due to the neck, and the acoustic capacitance due to the cavity. (b) Illustration of the setup of the ultrasonic transmission experiment.[3]

Helmholtz resonator

A typical Helmholtz resonator as in Fig. 2.3(a) can be presented as a combination of inductance and capacitance. The fluid inside the cavity is much easier to compress compared with that in the neck part. Moreover, the pressure gradient along the open neck is much greater than that inside the large cavity. Therefore, the cavity displays capacitive property and leaves the smaller neck as an acoustic inductor.

2.2.2. *Negative bulk modulus*

A new class of ultrasonic metamaterials is presented in Ref. 3 with subwavelength resonant structural units, which leads to the negative bulk modulus near the resonance. The building block of this ultrasonic metamaterial, the Helmholtz resonator, consists of a cavity of known volume with rigid walls and a small hole in one side [Fig. 2.3(a)]. The material is embedded in water, and the periodicity

is considerably smaller than the corresponding longitudinal wavelength in water ($d \simeq \lambda/5$). A pressure variation in the channel causes the plug of fluid in the hole to oscillate in and out, producing adiabatic compression and rarefaction of the liquid enclosed in the cavity. Such a resonator, as it has been mentioned above, is analogous to an inductor-capacitor circuit [Fig. 2.3(a)], with the enclosed cavity acting as the capacitor with capacitance $C \simeq V/\rho_0 c_0^2$, and the neck acting as the inductor ($L \simeq \rho_0(L'/S)$), where V is the volume of the cavity, ρ_0 is the density of water, c_0 is the sound speed in the water, L' is the effective length of the neck, and S is the cross-sectional area of the neck. Because the Helmholtz resonator does not use typical standing waves to create a resonance, the dimension of each element can be made much smaller than the acoustic wavelength (at 33 kHz, $\lambda = 4.4$ cm in water).

Following the formalism of electromagnetic response in metamaterials,[26,27] the combination of many Helmholtz resonators into a periodic array allows the material to behave as a medium with an effective modulus $\chi^{-1}(\omega)$ that can be expressed in the form

$$\chi^{-1}(\omega) = \chi_0^{-1}\left[1 - \frac{F\omega_H^2}{\omega^2 - \omega_H^2 + i\Gamma\omega}\right] \tag{2.7}$$

where $F = V/S_c d$ is a geometrical factor, S_c the cross-sectional area of the main conduit, $\omega_H = c_0\sqrt{S/L'V}$ is the resonant angular frequency, and Γ is the dissipation loss in the resonating Helmholtz elements due to viscous effects. In Fig. 2.4(a), the imaginary part of the modulus has a negative sign because the acoustic analogue of permeability corresponds to χ. This frequency dependent response is essential to the negative modulus over a range of frequencies. At frequencies near resonance, the induced displacement in the neck becomes very large, as is typical in resonance phenomena. The large response represents accumulation of energy over many cycles, such that a considerable amount of energy is stored in the resonator relative to the driving field. This stored energy is significant to maintain the sequence of displacement near resonance even when the excitation field changes the sign. That is, as the frequency of the driving

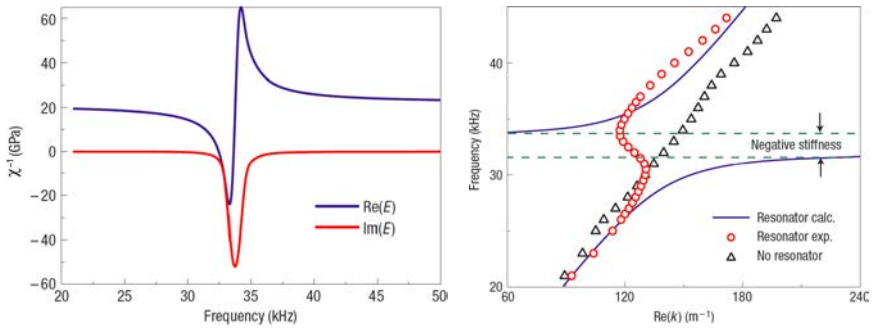


Fig. 2.4. (a) The calculated effective bulk modulus in the one-dimensional subwavelength Helmholtz resonators. (b) Measured and calculated dispersion of ultrasonic metamaterial. The red circles represent data measured from the periodical array of Helmholtz resonators in the duct, the black triangles represent the data in the duct without Helmholtz resonators, and the blue solid line is calculated using a sum of lossless Bloch waves.[3]

pressure field is swept through the resonance, the instantaneous velocity of the mass centre in the unit cell flips from in-phase to out-of-phase with the driving field, and the material shows a negative response. Similarly, a polariton effect is also observed in the electromagnetic response of metamaterials, where a negative permittivity or permeability (generally on the higher frequency side of the resonance) implies a purely imaginary wavevector in the bulk medium. Here, this idea has been implemented in the context of elastic composites at ultrasonic frequencies. By varying the size and geometry of the structural unit, we can tune the effective elastic moduli to negative values at desired frequency ranges.

Theory on the lossless resonators predicts that a full bandgap opens up between 32 and 34 kHz, whereas away from this dip the dispersion behaves linearly. However, experimental data show that possible propagation modes can exist in the bandgap with a backbending of the dispersion curve, which suggests an antiparallel relation between group and phase velocities. This is a direct result of the loss in the system. When ultrasonic metamaterials approach resonance, the complex modulus $\chi^{-1} = -|\Re(\chi^{-1})| + i\Im(\chi^{-1}) =$

$-\alpha - i\beta$ is expected in the spectral dip as a result of friction dissipation ($\alpha, \beta > 0$ as shown in Fig. 2.4(b)), where $\Im(\chi^{-1})$ indicates the imaginary part, and α, β are parameters corresponding to the real and imaginary parts of the complex modulus. It is straightforward to write the propagation constant in the system as $k = (-\alpha + i\beta)^{1/2}\omega\sqrt{\rho_0/(\alpha^2+\beta^2)}$, with a small real component

$$\Re(k) = -\frac{\omega}{2}\sqrt{\frac{\rho_0}{\alpha^2+\beta^2}}\left\{\sqrt{\alpha^2+\beta^2} - \alpha\right\}^{1/2} \tag{2.8}$$

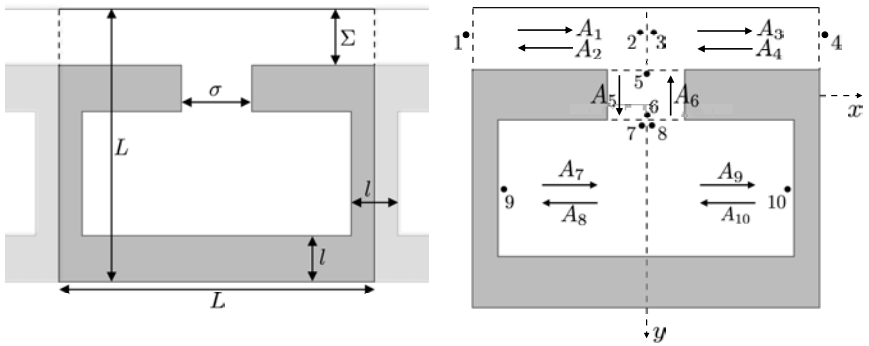
characterizing a propagating mode in the bandgap, whereas the imaginary part

$$\Im(k) = +\frac{\omega}{2}\sqrt{\frac{\rho_0}{\alpha^2+\beta^2}}\left\{\sqrt{\alpha^2+\beta^2} + \alpha\right\}^{1/2} \tag{2.9}$$

describes the decaying length of the pulse. The dispersion of these complex wavevectors can be well captured in our experiment by sweeping in real frequencies. Taking into account a small propagating component as an effect of resonant re-emission in parallel to the dominant tunnelling process in the transmission dip, the dispersion relationship can be characterized in the experimental results [Fig. 2.4(b)]. In addition, as frequency increases above the band edge, the attenuation (β) increases leading to the reduction of the real wavevector $\Re(k)$, resulting in the back-bending of the dispersion curve observed in Fig. 2.4(b). The loss term ($\Gamma = 2\pi \times 400$ Hz) is determined empirically by fitting the calculated transmission data along the edges of the experimental expectral dip.[3]

2.2.3. *Nonlocal description of sound propagation*

We employ here a generalized macroscopic nonlocal theory of sound propagation in rigid-framed porous media saturated with a viscothermal fluid[28] to describe the behavior of an acoustic metamaterial made of an array of Helmholtz resonators filled with air [see Fig. 2.5(a)].[24] Inspired by the electromagnetic theory and a thermodynamic consideration relating to the concept of acoustic part of energy current density, this macroscopic theory allows us to go beyond the limits of



(a) (b)

Fig. 2.5. (a): A unit cell in a 2D array of Helmholtz resonators. The dimensions are $L = 1\ cm$, $\Sigma = 0.2\ L$, $\sigma = 0.015\ L$, and $l = 0.15\ L$. (b): Illustration of slit portions and plane waves propagating in different parts of the resonator. Different positions are indicated by m, and different amplitudes by A_m, $m = 1, \ldots, 10$.

the classical local theory and within the limits of linear theory, to take into account temporal and spatial dispersion. By *macroscopic theory* we mean that the theory is concerned with averaged fields only. Assuming that there is a suitable ensemble of realizations of the medium, the macroscopic theory then is developed to describe the dynamics of the ensemble-averaged fields. A special case will be that of a periodic medium. The ensemble will be the collection of configurations generated by random translations of a single sample, and the ensemble average will be related to cell average of one sample. In the framework of the new approach, a homogenization procedure is proposed, through solving two independent microscopic action-response problems each of which related to the effective density and effective bulk modulus of the material. Contrary to the classical (two-scale asymptotic) method of homogenization, no asymptotic approach has been employed and there is no length-constraint to be considered within the development of the new method. Thus, there would be no frequency limit for the medium effective properties to be valid; in addition, materials with different length scales can be treated. The homogenization procedure offers a systematic way of obtaining the effective properties of the materials, regardless of their geometries.

These characteristics of the nonlocal approach permits the description of the porous media with specific geometries causing metamaterial behavior. A metamaterial with periodic structure will be studied: two-dimensional array of Helmholtz resonators connected in series.

By local theory, we refer to space locality. Nonlocality in time, or temporal dispersion, has been already taken into account through models for wave propagation in porous media.[29–32] That is, in Fourier space the effective density and bulk modulus depend on the frequency ω. In other words, the field dynamics at one location retains a memory of the field values at this location but is not affected by the neighboring values. The local description is usually based on retaining only the leading order terms in the two-scale homogenization method.[32–38] An asymptotic two-scale approach is applied in terms of a characteristic length of the medium, the period L in periodic media, which is assumed to be much smaller than the wavelength λ.[39,40] Efforts have been performed to extend the asymptotic method of homogenization to higher frequencies for the periodic composite materials[41,42] and rigid porous media[43] by introducing another type of scale separation to which the asymptotic multi-scale procedure applies. An enhanced asymptotic method has been adapted to describe sound propagation in rigid porous media with embedded damped Helmholtz resonators[44] exhibiting scattering different from Bragg scattering at high frequency in periodic media.

An effective medium approach has been proposed for periodic elastic composites based on surface responses of a structural unit of the material,[45] which can describe the macroscopic parameters beyond the frequencies within the long wavelength limit. Unlike the classical methods, based on the introduction of two-scale asymptotic expansions, or coherent potential approximation[46] based on the effective-medium parameters minimizing scatterings in the long-wavelength limit, the homogenization scheme presented in Ref. 45 uses matching the lowest-order scattering amplitudes arising from a periodic unit cell of the metamaterial with that of a homogenized material. As such, local resonant scattering can be captured as well

by the latter method in the elastic metamaterials. The asymptotic method of homogenization has been enhanced to provide the weak nonlocal effects as a small correction to the local behavior.[47] An approach has been presented[48] for random elastic composites based on ensemble averaging of the material responses to a body force, giving rise to effective parameters of the medium depending on frequency and wavenumber. By this method, the case of periodic media can be treated as well.

The nonlocal theory we use here takes fully the temporal dispersion and spatial dispersion into account. The medium is assumed unbounded and homogeneous in the stationary random statistical sense; therefore, the spatial dispersion refers only to the dependence of the permittivities, i.e. effective density and bulk modulus, on the Fourier wavenumbers $\boldsymbol{k}$ present in the macroscopic fields.[49] As mentioned above, the theory can be applied with certain considerations to a periodic medium; in particular it gives the Bloch wavenumbers and defines Bloch impedances. The materials susceptible to showing the nonlocal behavior may be classified into two main groups regarding their microgeometry. The first comprises the materials which exhibit this behavior in sufficiently high frequency regime. The second one concerns materials with microgeometry constituting the resonators, which exhibit spatial dispersion phenomena even at not very high frequencies; the resonance phenomena act as a source generating nonlocal behavior. In this article, we investigate the second type of these geometries in the form of daisy-chained Helmholtz resonators. A material made of an array of Helmholtz resonators filled by water has been studied experimentally, and has been found to show negative bulk modulus in the resonance frequency range[3] (see Section 2.2.2). Later, Helmhotz resonators as structural units were used to design novel metamaterials for focusing ultrasound waves[13] (see Section 2.2.4 and broadband acoustic cloaking[50]).

Here, we apply the nonlocal theory to quantatively describe the macroscopic dynamics of such a metamaterial filled with air as a viscothermal fluid, in 2D. Using a simplified analytical solution of the complete equations, we present the method of obtaining the

nonlocal efffective density and effective bulk modulus. When these effective parameters satisfy the dispersion equation based on the nonlocal theory, we can compute the wavenumber of the least attenuated mode, among other modes. We can then check that the wavenumber resulting from the macroscopic nonlocal theory coincides with the wavenumber associated with the Bloch wave propagating and attenuating in the medium. The Bloch solution is determined using the same simplifying way of solving as in the nonlocal modeling. Thus the results based on the two calculations should be comparable. Finally, as a check of the validity of the simplifying assumptions introduced in our modeling calculations, we have performed direct Finite Element Method (FEM) computations based on the exact equations in the framework of nonlocal homogenization.

First, we review briefly the general framework of the nonlocal theory which is used in this section. The microscopic equations governing sound propagation in a rigid porous medium are summarized, before mentioning the macroscopic Maxwellian equations describing the macroscopic nonlocal dynamics of the homogenized equivalent fluid. Then, we will see the nonlocal modeling allowing the calculation of the effective parameters and the wavenumber of the least attenuated wave in the medium. The direct calculation of the Bloch wavenumber, using similar simplifications, is presented next, followed by the results and analysis of the three different calculations.

General Framework of the Nonlocal Theory

In the following, we state the microscopic equations applied at the pore level, and the nonlocal Maxwellian macroscopic equations that describe the dynamics of the material as a homogeneous equivalent fluid medium. Then, we recall briefly the upscaling procedures allowing to obtain the frequency and wavenumber dependent effective parameters of the macroscopic equivalent fluid medium, i.e. effective density and effective bulk modulus. This section is a summary of the results which have been discussed in detail in Ref. 28. Hence, we will frequently refer to Ref. 28, for the in-depth explanations.

Microscopic Equations

The dynamics of a small amplitude perturbation in a rigid-framed porous material filled with a viscothermal fluid is governed by the linearized equations of the mass, momentum, and energy balance, and a general fluid state equation as follows: in the fluid region $\mathcal{V}^f$

$$\rho_0 \frac{\partial \boldsymbol{v}}{\partial t} = -\boldsymbol{\nabla} p + \eta \boldsymbol{\nabla}^2 \boldsymbol{v} + \left(\zeta + \frac{\eta}{3}\right) \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{v}) \tag{2.10a}$$

$$\frac{\partial b}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{v} = 0 \tag{2.10b}$$

$$\gamma \chi_0 p = b + \beta_0 \tau \tag{2.10c}$$

$$\rho_0 c_p \frac{\partial \tau}{\partial t} = \beta_0 T_0 \frac{\partial p}{\partial t} + \kappa \boldsymbol{\nabla}^2 \tau \tag{2.10d}$$

where $\boldsymbol{v}$, $b \equiv \rho/\rho_0$, p and τ, are the fluid velocity, excess condensation, thermodynamic excess pressure, excess temperature, respectively, and ρ is the excess density. The fluid constants ρ_0, η, ζ, γ, χ_0, β_0, c_p, T_0, κ, represent the ambient density, first viscosity, second viscosity, ratio of the heat capacity at constant pressure to heat capacity at constant volume c_p/c_v, adiabatic compressibility, coefficient of thermal expansion, specific heat capacity per unit mass at constant pressure, ambient temperature, and coefficient of thermal conduction, respectively.

In the (rigid) solid phase region $\mathcal{V}^s$, energy balance equation is reduced to $\rho^s c_p^s (\partial \tau^s/\partial t) = \kappa^s \boldsymbol{\nabla}^2 \tau^s$, where ρ^s is the constant solid density, τ^s solid excess temperature, and κ^s solid coefficient of thermal conductivity. On the fluid/solid interface $\partial\mathcal{V}$, we have the conditions of continuity of the excess temperature $\tau = \tau^s$ and the heat flux $\kappa \boldsymbol{\nabla} \tau = \kappa^s \boldsymbol{\nabla} \tau^s$. We admit that the coefficient of thermal conductivity of the solid is much larger than that of the fluid $\kappa^s \gg \kappa$, and the heat capacity at constant pressure of the solid part is much larger than that of the fluid part, i.e., $(1-\phi)\rho^s c_p^s \gg \phi \rho_0 c_p$; where ϕ is the fluid volume fraction (porosity). The latter assumptions combined with the Fourier heat diffusion in the solid, and the temperature and heat flux continuity relations, generally result in the vanishing of the

fluid excess temperature at the fluid/solid boundaries. In addition, we assume no-slip condition on the fluid/(rigid) solid interface. The boundary conditions for the velocity and excess temperature on $\partial\mathcal{V}$ are finally written as

$$\boldsymbol{v} = 0, \quad \tau = 0 \tag{2.11}$$

Macroscopic Maxwellian Acoustics

Before going through the macroscopic equations for sound propagation in rigid-framed porous media, and the homogenization procedure, we will precise the notion of field averaging in the nonlocal approach.

Averaging: The present macroscopic theory is statistical in nature and has been developed in principle for fluid-saturated rigid-framed media which are homogeneous in an ensemble-averaged sense; this is the case of stationnary random media. The macroscopic properties represented in the theory refer to the ensemble of realizations. Thus, for example, the propagation constants of the medium would refer to the propagation constant of *coherent waves* in multiple-scattering theory. Here, the material we wish to study is not defined by stationary random realizations. It belongs to the important class of periodic materials. The macroscopic theory can still be applied by considering the ensemble obtained through random translation of one sample. It turns out that the ensemble-average $\langle\ \rangle$ properties of the space are, in this case, precisely computable by spatial averaging over a periodic cell in a single realization. This, in a sense, reminds of *ergodicity* in stationary random media.

The macroscopic condensation and velocity are defined as the average of pore scale microscopic fields: $\boldsymbol{V} \equiv \langle \boldsymbol{v} \rangle$, and $B \equiv \langle b \rangle$; average over the periodic cell in the case of the periodic media. A macroscopic equation can be obtained directly by averaging Eq. (2.10b), using the commutation relation $\langle \boldsymbol{\nabla}.\boldsymbol{v} \rangle = \boldsymbol{\nabla}.\langle \boldsymbol{v} \rangle$ which is automatically satisfied owing to (2.11) (see Eq. (56) in Ref. 28). The second macroscopic field equation, as well as the macroscopic constitutive relations, are written using the electromagnetic analogy. This analogy suggests that the system of macroscopic equations can be carried

through by introducing new Maxwellian fields H and $\boldsymbol{D}$, as well as linear operators $\hat{\rho}$ and $\hat{\chi}^{-1}$. The field equations and constitutive relations are written as (see Section 3.3 in Ref. 28)

$$\textbf{Field equations:} \qquad \frac{\partial B}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{V} = 0, \qquad \frac{\partial \boldsymbol{D}}{\partial t} = -\boldsymbol{\nabla} H \tag{2.12}$$

$$\textbf{Constitutive relations:} \quad \boldsymbol{D} = \hat{\rho}\boldsymbol{V}, \qquad H = \hat{\chi}^{-1} B \tag{2.13}$$

where the integral operators of density $\hat{\rho}$ and bulk modulus $\hat{\chi}^{-1}$ are such that

$$\boldsymbol{D}(t, \boldsymbol{r}) = \int_{-\infty}^{t} dt' \int d\boldsymbol{r}' \rho(t - t', \boldsymbol{r} - \boldsymbol{r}') \boldsymbol{V}(t', \boldsymbol{r}') \tag{2.14a}$$

$$H(t, \boldsymbol{r}) = \int_{-\infty}^{t} dt' \int d\boldsymbol{r}' \chi^{-1}(t - t', \boldsymbol{r} - \boldsymbol{r}') B(t', \boldsymbol{r}') \tag{2.14b}$$

We notice that the kernels ρ and χ^{-1} depend on the difference $t - t'$ and $\boldsymbol{r} - \boldsymbol{r}'$, which is due to the homogeneity of the medium with respect to time and space. Therefore, we can write (2.14a) and (2.14b) in the Fourier space, respectively, as

$$\boldsymbol{D}(\omega, \boldsymbol{k}) = \rho(\omega, \boldsymbol{k}) \boldsymbol{V}(\omega, \boldsymbol{k}), \quad H(\omega, \boldsymbol{k}) = \chi^{-1}(\omega, \boldsymbol{k}) B(\omega, \boldsymbol{k}) \tag{2.15}$$

provided that

$$\rho(t - t', \boldsymbol{r} - \boldsymbol{r}') = \int \frac{d\omega}{2\pi} \frac{d\boldsymbol{k}}{(2\pi)^3} \, \rho(\omega, \boldsymbol{k}) \, e^{-i\omega(t-t') + i\boldsymbol{k}.(\boldsymbol{r}-\boldsymbol{r}')}$$

$$\chi^{-1}(t - t', \boldsymbol{r} - \boldsymbol{r}') = \int \frac{d\omega}{2\pi} \frac{d\boldsymbol{k}}{(2\pi)^3} \, \chi^{-1}(\omega, \boldsymbol{k}) \, e^{-i\omega(t-t') + i\boldsymbol{k}.(\boldsymbol{r}-\boldsymbol{r}')}$$

In other words, because of the medium homogeneity with respect to time and space, $\boldsymbol{D}(\omega, \boldsymbol{k})$ is related to $\boldsymbol{V}(\omega, \boldsymbol{k})$, and $H(\omega, \boldsymbol{k})$ is related to $B(\omega, \boldsymbol{k})$, for the same values of ω and $\boldsymbol{k}$.

In nonlocal theory, the macroscopic H field is defined through the Poynting-Schoch condition of *acoustic part of energy current density*[28,51] which is postulated as (see Section 3.4 in Ref. 28)

$$S = H\boldsymbol{V} = \langle p\boldsymbol{v} \rangle \tag{2.16}$$

As a result of this definition, the density and bulk modulus operators become susceptibility functions determinable, in principle, through independent action-response problems (see Section 2.4 in Ref. 28). Regarding Eqs. (2.14) and (2.15), it is visible that the theory allows for both temporal dispersion, shown by integration over time variable t' in physical space and frequency dependence in Fourier space, and spatial dispersion, shown by integration over space coordinates $\boldsymbol{r}'$ and wavenumber dependence in Fourier space. We will recognize the quantities in physical space $(t, \boldsymbol{r})$ and Fourier space $(\omega, \boldsymbol{k})$ by their arguments. Now, in order to clarify the relationship between constitutive operators and microgeometry, the kernel functions $\rho(\omega, \boldsymbol{k})$ and $\chi^{-1}(\omega, \boldsymbol{k})$ are needed to be determined, by introducing action-response procedures coarse-graining the dissipative fluid dynamics of the pore scale.

Procedures to Compute Effective Density and Bulk Modulus

In the 1D case of macroscopic propagation along a symmetry axis, for instance x-axis with the unit vector $\hat{\boldsymbol{x}}$, we will have $\boldsymbol{D} = D\hat{\boldsymbol{x}}$ and $\boldsymbol{V} = V\hat{\boldsymbol{x}}$, $\boldsymbol{r} = x\hat{\boldsymbol{x}}$, and $\boldsymbol{k} = k\hat{\boldsymbol{x}}$ in the above equations (2.12–2.16). To determine the Fourier functions $\rho(\omega, k)$ and $\chi^{-1}(\omega, k)$ for the 1D acoustic propagation in a medium with porosity ϕ, we solve two independent action-response problems (see Section 4 in Ref. 28). For computing the effective density we consider the macroscopic response of the fluid subject to a single-component (ω, k) Fourier bulk force. The effective bulk modulus is related to the response of the fluid subject to a single-component Fourier rate of heat supply.

Two sets of equations to be solved

The two systems of equations to be solved are written as

In the fluid region $\mathcal{V}_f$:

$$\frac{\partial b}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{v} = 0 \tag{2.17a}$$

$$\rho_0 \frac{\partial \boldsymbol{v}}{\partial t} = -\boldsymbol{\nabla} p + \eta \boldsymbol{\nabla}^2 \boldsymbol{v} + \left(\zeta + \frac{1}{3}\eta\right) \boldsymbol{\nabla} (\boldsymbol{\nabla} \cdot \boldsymbol{v}) + \underbrace{\boldsymbol{F} e^{-i\omega t + ikx}}_{\text{Added for determination of density}} \tag{2.17b}$$

$$\rho_0 c_p \frac{\partial \tau}{\partial t} = \beta_0 T_0 \frac{\partial p}{\partial t} + \kappa \boldsymbol{\nabla}^2 \tau + \underbrace{\dot{Q} e^{-i\omega t + ikx}}_{\text{Added for determination of bulk modulus}} \tag{2.17c}$$

$$\gamma \chi_0 p = b + \beta_0 \tau \tag{2.17d}$$

On the fluid/solid interface $\partial\mathcal{V}$:

$$\boldsymbol{v} = 0, \quad \tau = 0 \tag{2.18}$$

For convenience the excitation amplitudes are written as: $\dot{Q} e^{-i\omega t + ikx} = \beta_0 T_0 (\partial/\partial t)(\mathcal{P} e^{-i\omega t + ikx})$, and $\boldsymbol{F} e^{-i\omega t + ikx} = -\boldsymbol{\nabla}(\mathcal{P} e^{-i\omega t + ikx})$. Here, it is important to note that the excitation variables ω and k are set as independent variables. The solutions to the above systems for the fields p, b, τ, and components of $\boldsymbol{v}$ take the form $p(t, \boldsymbol{r}) = p(\omega, k, \boldsymbol{r}) e^{-i\omega t + ikx}$, and so on. Recall that the theory is formulated for a geometry that is stationary random, and the averaging operator $\langle\ \rangle$ refers to the ensemble averaging. Thus, here, the amplitude fields $\boldsymbol{v}(\omega, k, \boldsymbol{r})$, $p(\omega, k, \boldsymbol{r})$, $b(\omega, k, \boldsymbol{r})$, and $\tau(\omega, k, \boldsymbol{r})$, are stationary random functions of $\boldsymbol{r}$. Passing to the case of periodic geometry, we can limit ourselves to considering one periodic sample. The fields become periodic functions over a cell, and $\langle\ \rangle$ is interpreted as a volume average over a cell.

Frequency and Wavenumber Dependent Effective Density and Bulk Modulus

Once the two systems of equations are solved independently, using the right hand Maxwellian macroscopic equations in (2.12) and

(2.13), we arrive at the following expressions for the nonlocal effective density and bulk modulus

$$\rho(\omega,k)=\frac{k\left(\mathcal{P}+P(\omega,k)\right)}{\omega\left\langle v(\omega,k,\boldsymbol{r})\right\rangle} \tag{2.19a}$$

$$\chi^{-1}(\omega,k)=\frac{P(\omega,k)+\mathcal{P}}{\langle b(\omega,k,\boldsymbol{r})\rangle+\phi\gamma\chi_0\mathcal{P}} \tag{2.19b}$$

where $P\langle\boldsymbol{v}\rangle=\langle p\boldsymbol{v}\rangle$, which has been inspired by (2.16).

Wavenumbers, Constants of the Medium

Contrary to the case of local theory, here, since we fully take into account spatial dispersion, several normal mode solutions might exist, with fields varying as $e^{-i\omega t+ikx}$. Solutions should satisfy the following dispersion equation

$$\rho(\omega,k)\chi(\omega,k)\omega^2=k^2 \tag{2.20}$$

which is easily derived from the Maxwellian macroscopic equations. With each frequency ω, several complex wavenumbers $k_n(\omega)$, $\Im(k_n)>0$, $n=1,2,\ldots$, may be associated.

Frequency Dependent Effective Parameters

With each wavenumber k_n solution of the nonlocal dispersion equation (2.20) are associated a frequency-dependent density and bulk-modulus, such that

$$\rho_n(\omega)=\rho_n(\omega,k_n(\omega)) \tag{2.21a}$$

$$\chi_n^{-1}(\omega)=\chi_n^{-1}(\omega,k_n(\omega)) \tag{2.21b}$$

The fact that at each frequency ω, we can obtain several modes propagating and attenuating in the medium, with wavenumbers $k_n(\omega)$ constants of the medium, and effective parameters $\rho_n(\omega)$ and $\chi_n^{-1}(\omega)$, is a direct consequence of the nonlocal description. In other words, the existence of multiple wavenumbers at each frequency, associated with unique effective parameters, is a signature of the nonlocal effects or spatial dispersion in the medium. In what

follows, with the aim of obtaining the nonlocal effective density $\rho_1(\omega)$, effective bulk modulus $\chi_1^{-1}(\omega)$, and wavenumber k_1 of the least attenuated mode, we will apply this theoretical framework in analytical simplified manner, to a 2D array of Helmholtz resonators, illustrated in Fig. 2.5(a). Sound propagation through this material exhibits resonance phenomena resulting in metamaterial behavior.

Nonlocal Modelling

For 2D structures, we proceed to determine the functions $\rho(\omega, k)$ and $\chi^{-1}(\omega, k)$ sufficiently precise to give an appropriate modeling of the least attenuated mode, which results then in purely frequency dependent functions $\rho(\omega)$ and $\chi^{-1}(\omega)$. For this purpose, we do not need to consider in full detail the microscopic fields $\boldsymbol{v}$ and p. In the waveguide t and cavity c, instead of the microscopic fields, we can use the mean values $V_{t(c)} = \langle \boldsymbol{v} \rangle_S \cdot \hat{\boldsymbol{x}}$ and $P_{t(c)} = \langle p \rangle_S$, where $\langle\ \rangle_S$ denotes the average at a given x over the waveguide or the cavity width; and in the neck n, we can use the mean values $V_n = \langle \boldsymbol{v} \rangle_S \cdot \hat{\boldsymbol{y}}$ and $P_n = \langle p \rangle_S$, where $\langle\ \rangle_S$ denotes the average at a given y over the neck width, and $\hat{\boldsymbol{y}}$ is the unit vector in the y direction. At the same time, we make some simplifications consistent with describing the propagation of these averaged quantities in terms of the Zwikker and Kosten densities $\rho(\omega)$ and bulk moduli $\chi^{-1}(\omega)$, in the different slit portions. These depend only on the slit half-widths, which we shall denote by s_t, s_n, and s_c, in the tube, neck, and cavity, respectively. The different slit-like tube portions are illustrated in Fig. 2.5(b). The main tube t is divided in two Zwikker and Kosten ducts, a left duct, and a right duct, oriented in the x direction. The same separation is made for the cavity c, whereas the neck n is not divided but viewed as one Zwikker and Kosten duct oriented in y direction.

Determination of nonlocal effective density

Considering the periodic cell of Fig. 2.5(a), and the corresponding cell average operation $\langle\ \rangle$, we look for the response of the fluid when

a harmonic driving force $f(t,x) = fe^{-i\omega t+ikx}$ in the direction of $\hat{\boldsymbol{x}}$ is applied. If we can determine the microscopic response velocity and pressure fields $\boldsymbol{v}$, p, then we will have the function $\rho(\omega,k)$ through the relation (see Eq. (2.19a))

$$\rho(\omega,k) = \frac{f - ik\mathtt{P}(\omega,\mathtt{k})}{-i\omega\langle v(\omega,k,\boldsymbol{r})\rangle} \tag{2.22}$$

with $\mathtt{P}(\omega,k) = \langle pv\rangle/\langle v\rangle$, where the v is the x-component of the microscopic velocity $\boldsymbol{v}$.

In [51, Appendix], the Zwikker and Kosten local theory is expressed for tubes of circular cross-section. For 2D slits, exactly the same general principles of modeling may be used; only some details of the calculations are changed. In particular, the Bessel functions J_0 and J_1 are replaced by cosh and sinh functions. Zwikker and Kosten's effective densities $\rho_\alpha(\omega)$ and bulk modulii $\chi_\alpha^{-1}(\omega)$ in the guide, neck and cavity, will be[52]

$$\rho_\alpha(\omega) = \rho_0\left[1 - \frac{\tanh\left(\sqrt{-i\omega\rho_0 s_\alpha^2/\eta}\right)}{\sqrt{-i\omega\rho_0 s_\alpha^2/\eta}}\right]^{-1} \tag{2.23a}$$

$$\chi_\alpha^{-1}(\omega) = \gamma P_0\left[1 + (\gamma-1)\frac{\tanh\left(\sqrt{-i\omega\rho_0 c_p s_\alpha^2/\kappa}\right)}{\sqrt{-i\omega\rho_0 c_p s_\alpha^2/\kappa}}\right]^{-1} \tag{2.23b}$$

for $\alpha = t,n,c$, where the indexes t, n, and c are related to the tube, neck, and cavity, respectively; P_0 the fluid pressure at rest. The corresponding wavenumbers $k_\alpha(\omega)$ and characteristic admittances $Y_\alpha(\omega)$ are expressed as $k_\alpha = \omega/c_\alpha$, and $Y_\alpha(\omega) = 2s_\alpha/(\rho_\alpha c_\alpha)$, for $\alpha = t,n,c$, where $c_\alpha = 1/\sqrt{\rho_\alpha\chi_\alpha}$, is the corresponding Zwikker and Kosten's phase velocity. Notice that we include the slit width $2s_\alpha$ (resp. Σ,σ, and $L-\Sigma-2l$ in the resonator, see Fig. 2.5(a)) in the definition of the characteristic admittance, because it simplifies the subsequent writing of continuity conditions.

We start writing the Zwikker and Kosten's equations in the different parts of the periodic cell. For the tube and the cavity, i.e., $\alpha = t, c$, we have

$$-i\omega\frac{\rho_\alpha(\omega)}{S_\alpha}V_\alpha = -\frac{\partial P_\alpha}{\partial x} + fe^{ikx} \tag{2.24a}$$

$$i\omega S_\alpha \chi_\alpha(\omega) P_\alpha = \frac{\partial V_\alpha}{\partial x} \tag{2.24b}$$

where, $V_\alpha = V_x S_\alpha$ is the flow rate field across the cross section S_α, with V_x the x-component of the velocity in the sense of Zwikker and Kosten (averaged over the section), and P_α is the Zwikker and Kosten's pressure. In the neck, the external excitation having no y-component, we have

$$i\omega\frac{\rho_n(\omega)}{\sigma}V_n = \frac{\partial P_n}{\partial y} \tag{2.25a}$$

$$i\omega\sigma\chi_n(\omega)P_n = \frac{\partial V_n}{\partial y} \tag{2.25b}$$

where, $V_n = V_y\sigma$ is the flow rate, with V_y the y-component of the velocity, and P_n is the Zwikker and Kosten's pressure in the neck.

The general solution of the non homogeneous equations in the tube and the cavity, (P_α, V_α), $\alpha = t, c$, is written as the sum of the general solution $(P_{\alpha,h}, V_{\alpha,h})$ of the homogeneous equations and a particular solution $(P_{\alpha,p}, V_{\alpha,p})$ of the non homogeneous equations. A general solution of the homogeneous equations (2.24) is written as

$$\begin{pmatrix} P_{\alpha,h} \\ V_{\alpha,h} \end{pmatrix} = \begin{pmatrix} 1 \\ Y_\alpha \end{pmatrix} A^+ e^{ik_\alpha x} + \begin{pmatrix} 1 \\ -Y_\alpha \end{pmatrix} A^- e^{-ik_\alpha x} \tag{2.26}$$

where A^+ and A^- are the amplitudes of the plane waves in direction of the positive x-axis and negative x-axis, respectively. The following particular solution can be considered

$$\begin{pmatrix} P_{\alpha,p} \\ V_{\alpha,p} \end{pmatrix} = \begin{pmatrix} B_\alpha \\ C_\alpha \end{pmatrix} fe^{ikx} \tag{2.27}$$

where B_α and C_α represent four constants (for each ω) to be determined. Substituting (2.27) in (2.24) gives the four constants $B_t =$

$ik/(\omega^2\rho_t\chi_t - k^2)$, $C_t = i\omega\chi_t\Sigma/(\omega^2\rho_t\chi_t - k^2)$, $B_c = ik/(\omega^2\rho_c\chi_c - k^2)$, and $C_c = i\omega\chi_c(L - \Sigma - 2l)/(\omega^2\rho_c\chi_c - k^2)$. The particular solution is the same in the left and right portions of the tube and the cavity. On the contrary and because of the presence of the neck, the general solution will have different amplitude constants in the left and right portions. Thus, the general solution of Eqs. (2.24) can be written as

$$\begin{pmatrix} P_t \\ V_t \end{pmatrix} = \begin{pmatrix} 1 \\ Y_t \end{pmatrix} A_{1,3} f e^{ik_t x} + \begin{pmatrix} 1 \\ -Y_t \end{pmatrix} A_{2,4} f e^{-ik_t x} + \begin{pmatrix} B_t \\ C_t \end{pmatrix} f e^{ikx} \tag{2.28a}$$

$$\begin{pmatrix} P_c \\ V_c \end{pmatrix} = \begin{pmatrix} 1 \\ Y_c \end{pmatrix} A_{7,9} f e^{ik_c x} + \begin{pmatrix} 1 \\ -Y_c \end{pmatrix} A_{8,10} f e^{-ik_c x} + \begin{pmatrix} B_c \\ C_c \end{pmatrix} f e^{ikx} \tag{2.28b}$$

where (2.28a) with amplitudes A_1 and A_2 corresponds to the left part of the tube, and with amplitudes A_3 and A_4 to the right part [Fig. 2.5(b)]; similarly for (2.28b): A_7 and A_8 for the left part of the cavity, and A_9 and A_{10} for the right part [Fig. 2.5(b)]. These eight amplitudes are to be determined. The general solution of Eqs. (2.25), (P_n, V_n) has the form

$$\begin{pmatrix} P_n \\ V_n \end{pmatrix} = \begin{pmatrix} 1 \\ Y_n \end{pmatrix} A_5 f e^{ik_n y} + \begin{pmatrix} 1 \\ -Y_n \end{pmatrix} A_6 f e^{-ik_n y} \tag{2.29}$$

where A_5 and A_6 are the neck amplitude-relating constants to be determined (Fig. 2.5(b)).

Indeed, in the framework of our simple plane-wave modeling, there are 10 relations concerning the flow rate and pressure, which are assumed to be verified. These continuity relations involve the values of the fields at different locations indicated by numbers $m = 1, \ldots, 10$, in Fig. 2.5(b). As such, we have 10 equations for 10 unknown amplitudes $A_1, \ldots, A_{10}$.[24] Once these are determined, we will have all the Zwikker and Kosten's fields through Eqs. (2.28) and (2.29). At this point, we can easily obtain the cell averages $\langle v \rangle$ and $\langle pv \rangle$, regarding the fact that the Zwikker and Kosten's flow rate has no component along the y-axis.[24] Subsequently, we can obtain

explicitly the effective density function $\rho(\omega, k)$ through Eq. (2.22). In the next section, the effective bulk modulus is computed in a similar way but with a different excitation term, and with exactly the same conditions on the flow rate and pressure fields at different junctions.

Determination of nonlocal effective bulk modulus

Considering the periodic cell [Fig. 2.5(b)], when a harmonic heating $\dot{Q}(t,x) = \dot{Q}_0 e^{-i\omega t+ikx} = -i\omega\beta_0 T_0 \mathcal{P} e^{-i\omega t+ikx}$ is applied in the medium, we write the Zwikker and Kosten's equations, in each part of the resonator: tube, neck, and cavity. The aim is to obtain the function $\chi^{-1}(\omega, k)$ as it is indicated in Eq. (2.19b). In the main tube and the cavity, for $\alpha = t, c$, we write

$$-i\omega\frac{\rho_\alpha(\omega)}{S_\alpha}V_\alpha = -\frac{\partial P_\alpha}{\partial x} \tag{2.30a}$$

$$i\omega S_\alpha\chi_\alpha(\omega)P_\alpha + i\omega S_\alpha\left(\chi_\alpha(\omega) - \gamma\chi_0\right)\mathcal{P} = \frac{\partial V_\alpha}{\partial x} \tag{2.30b}$$

The second term in the second equation might not seem to be obvious but follows the very procedure of obtaining (2.19b). In the neck, the equations are written as

$$i\omega\frac{\rho_n(\omega)}{\sigma}V_n = \frac{\partial P_n}{\partial y} \tag{2.31a}$$

$$i\omega\sigma\chi_n(\omega)P_n + i\omega\sigma\left(\chi_n(\omega) - \gamma\chi_0\right)\mathcal{P}\left\langle e^{ikx}\right\rangle_\sigma = \frac{\partial V_n}{\partial y} \tag{2.31b}$$

where the term $\mathcal{P}\langle e^{ikx}\rangle_\sigma$ comes from the averaging of $\dot{Q}$ over the neck cross section. Here also, the second equation might not appear obvious, but follows the procedure of the determination of (2.19b) in nonlocal theory.[28]

As before, the general solution of the non homogeneous equations (2.30) in the right or left part of the tube and the cavity, is written as the sum of the general solution $(P_{\alpha,h}, V_{\alpha,h})$ of the homogeneous equations and a particular solution $(P_{\alpha,p}, V_{\alpha,p})$ of the non homogeneous equations. A general solution of the homogeneous equations

(2.30) is written as Eq. (2.26). The following particular solution can be considered

$$\begin{pmatrix} P_{\alpha,p} \\ V_{\alpha,p} \end{pmatrix} = \begin{pmatrix} B_\alpha \\ C_\alpha \end{pmatrix} \mathcal{P} e^{ikx} \tag{2.32}$$

where B_α and C_α are four constants to be determined. Substituting (2.32) in (2.30) gives the four constants $B_t = \omega^2\rho_t(\chi_t - \gamma\chi_0)/(k^2 - \omega^2\rho_t\chi_t)$, $C_t = \omega k(\chi_t - \gamma\chi_0)\Sigma/(k^2 - \omega^2\rho_t\chi_t)$, $B_c = \omega^2\rho_c(\chi_c - \gamma\chi_0)/(k^2-\omega^2\rho_c\chi_c)$, and $C_c = \omega k(\chi_t-\gamma\chi_0)(L-\Sigma-2L)/(k^2-\omega^2\rho_c\chi_c)$. Thus, the general solution of Eqs. (2.30) can be written as Eqs. (2.28), replacing f with $\mathcal{P}$. The amplitudes A_1, A_2, A_3, A_4, A_7, A_8, A_9, and A_{10} (Fig. 2.5(b)) are to be determined.

As for the tube and the cavity, the general solution of the non homogeneous equations (2.31) in the neck, is written as the sum of the general solution $(P_{n,h}, V_{n,h})$ of the homogeneous equations and a particular solution $(P_{n,p}, V_{n,p})$ of the non homogeneous equations. We can find a particular solution in the following form

$$\begin{pmatrix} P_{n,p} \\ V_{n,p} \end{pmatrix} = \begin{pmatrix} B_n \\ C_n \end{pmatrix} \mathcal{P} \tag{2.33}$$

where B_n and C_n are two constants which will be determined by substituting (2.33) in (2.31):

$B_n = (2/k\sigma)(\gamma\chi_0/\chi_n - 1)\sin(k\sigma/2)$, and $C_n = 0$. To obtain the above expression for B_n, the average $\langle e^{ikx}\rangle_\sigma$ can be easily calculated.[24] Therefore, the general solution of Eq. (2.31) in the neck can be written as

$$\begin{pmatrix} P_n \\ V_n \end{pmatrix} = \begin{pmatrix} 1 \\ Y_n \end{pmatrix} A_5\mathcal{P}e^{ik_n y} + \begin{pmatrix} 1 \\ -Y_c \end{pmatrix} A_6\mathcal{P}e^{-ik_n y} + \begin{pmatrix} B_n \\ 0 \end{pmatrix} \mathcal{P} \tag{2.34}$$

where A_5 and A_6 are amplitude-relating constants to be determined [Fig. 2.5(b)].

As noted previously, in the framework of our modeling, there are 10 relations which are assumed to be verified, allowing to relate the flow rates and pressures at different indicated points in Fig. 2.5(b).

These relations result in 10 equations by which we can compute the amplitudes $A_1, \ldots, A_{10}$.[24] Consequently, all Zwikker and Kosten's fields will be found. After the averages $\langle v \rangle$, $\langle pv \rangle$, and $\langle b \rangle$ are calculated[24] for the actual fields, the expression for $\chi^{-1}(\omega, k)$ will be obtained.

Bloch Wave Modeling

In this section, without using the principles of the nonlocal macroscopic theory but within the same plane wave modeling, we directly seek the macroscopic Bloch wavenumber k_B of the least attenuated wave propagating in the direction of positive x-axis, such that

$$\begin{pmatrix} P_t^{(4)} \\ V_t^{(4)} \end{pmatrix} = e^{ik_B L} \begin{pmatrix} P_t^{(1)} \\ V_t^{(1)} \end{pmatrix} \tag{2.35}$$

with the field constituted of 10 Zwikker and Kosten's slit waves, as illustrated in Fig. 2.5(b), are associated 10 complex amplitudes $A_1, \ldots, A_{10}$. As before, between these 10 amplitudes there are a set of 10 relations; where two of them express the Bloch condition (2.35), and 8 relations are based on the continuity equations. Here, all these relations are homogeneous relations, so that nontrivial solutions will be obtained only if the determinant of the coefficient matrix vanishes. This condition will give the analytical expression for Bloch wavenumber k_B, as follows

$$k_B = -\frac{i}{L} \ln \left(\frac{D}{2} \pm \sqrt{\frac{D^2}{4} - 1} \right) \tag{2.36}$$

where $D = 2\cos k_t L - i(Y_r/Y_t) \sin k_t L$, $Y_r = V_n^{(5)}/P_n^{(5)}$ is the entrance admittance of the resonator (see Eq. (31) in Ref. 24).

2D structure filled with air

For the geometry considered in Fig. 2.5(a), to perform the computations, we have set $L = 1\ cm$, $\Sigma = 0.2\ L$, and $\sigma = 0.015\ L$. The functions $\rho(\omega, k)$ and $\chi^{-1}(\omega, k)$ are first determined within the approximations of our nonlocal modeling. Given these expressions,

we know that according to nonlocal theory the possible wavenumbers in the medium will be the solutions of the dispersion relation (2.20). Solving the equation (2.20) by a Newton-Raphson scheme, we have checked that the obtained expressions for $\rho(\omega, k)$ and $\chi^{-1}(\omega, k)$ are such that a complex solution $k(\omega)$ to (2.20) exists, which is very close to the value $k_B(\omega)$ in (2.36). The frequency dependent effective density $\rho(\omega, k(\omega)) = \rho(\omega)$, and effective bulk modulus $\chi^{-1}(\omega, k(\omega)) = \chi^{-1}(\omega)$, are then obtained by putting $k = k(\omega)$ in the aforementioned excitation terms.

Solving the equation (2.20) by the Newton-Raphson method, we varied frequency step by step, taking as initial value for $k(\omega)$ at a given frequency, the solution value obtained at the preceding frequency. Only for the starting frequency ω_0 in the range of interest, we have chosen the value $k_B(\omega_0)$ with a 10% discrepancy. In order to ascertain the validity of the modeling, we have also performed direct FEM simulations to solve the action-response problems, giving, subsequently, FEM evaluations of the functions $\rho(\omega, k)$ and $\chi^{-1}(\omega, k)$. Based on these functions, the computation of the wavenumber of the least attenuated wave was performed in the same way as just seen, with the only difference that (due to computation time) the initial $k(\omega)$ value at a given frequency was systematically taken to be $k_B(\omega)$ with 10% discrepancy. Finally, FEM evaluations of the frequency dependent effective density $\rho(\omega, k(\omega)) = \rho(\omega)$, and effective bulk modulus $\chi^{-1}(\omega, k(\omega)) = \chi^{-1}(\omega)$, were obtained by putting $k = k(\omega)$ in the aforementioned excitation terms. The FEM computations have been performed using FreeFem++,[53] an open source tool solving partial differential equations. Adaptive meshing was employed. According to all of the calculations, the effective density remains practically constant and, therefore, does not play an important role in the macroscopic dynamics of this material.

We see in Fig. 2.6(a), that the real and imaginary parts of $k(\omega)$ computed by nonlocal theory via Newton's method converge exactly to the real and imaginary parts of k_B which have been computed by a simple Bloch-wave modeling without any use of nonlocal theory. The horizontal axis is the dimensionless frequency $k_0 L/\pi$, where

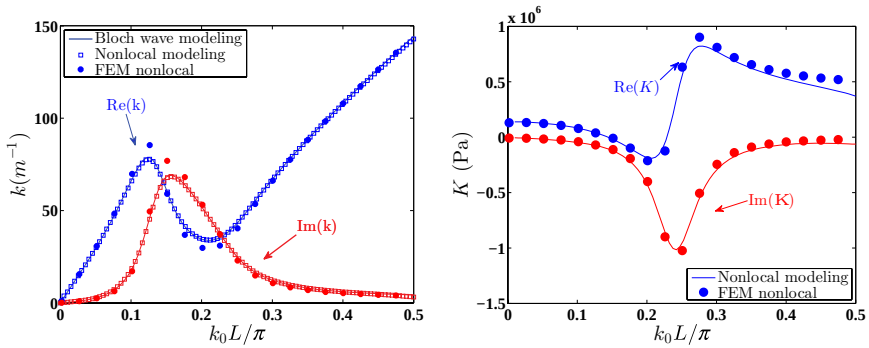


Fig. 2.6. Wavenumber (a) and bulk modulus (b) in terms of a dimensionless frequency, for the 2D structure filled with air. For the wavenumber, results by three calculations are compared: Bloch-wave modeling, nonlocal modeling, and nonlocal theory by FEM.[24]

$k_0 = \omega/c_0$. The results based on the Finite Element Method (FEM) simulations are also in good agreement with those obtained by the Bloch wave modeling and nonlocal modeling. The frequency range has been chosen so that it covers the resonance regime. In the same frequency range, Fig. 2.6(b), shows the real and imaginary parts of $K(\omega) = \phi\chi^{-1}(\omega)$, representing the effective bulk modulus, computed by nonlocal FEM simulations and nonlocal modeling. Here also, we see excellent agreement between the two calculations. We notice the metamaterial behavior demonstrated in the real part of effective bulk modulus which becomes negative in a frequency range within the resonance regime. It is clear that the results by FEM computations based on the exact microscopic equations, can be considered more precise compared with our two modeling results in which we have applied simplifying approximations. As such, the good agreement between FEM results and others, validate the modeling framework. The discrepancies between the results based on the models and FEM simulations can be due in particular, to the fact that the model describes the admittance of the resonator Y_r, without considering the length correction of the neck; what might generate errors in the calculation of the wavenumber.

We observe here the same kind of behavior for the wavenumber and bulk modulus as it has been demonstrated experimentally in Ref. 3 (see Figs. 1 and 2 in that reference) for the case of the 3D material embedded in water. We have observed that removing the thermal effects by decreasing the coefficient of thermal conductivity κ to a value close to zero, would have a negligible effect on the wavenumber and the effective bulk modulus. That is the case also for the second viscosity ζ, associated with losses in the compressional/dilatational motions in the bulk fluid. On the contrary, the material dynamics in terms of the macroscopic wavenumber and bulk modulus is quite sensitive to the values of the shear viscosity η. In a frequency range, for instance, between $k_0 L/\pi = 0.1$ and 0.4, a maximum and minimum appear for the real part of the wavenumber. By decreasing the value of the shear viscosity, the maximum becomes sharper and finally diverges as the viscosity tends to zero at the resonance frequency of the ideal fluid $\omega_H = c_0[\sigma/l(L-2l)(L-\Sigma-2l)]^{\frac{1}{2}}$, namely $k_0 L/\pi = 0.15$ here; the minimum flattens and a bandgap is created. As a matter of fact, the important feature, here, is the resonant behaviour which induces important values of the velocity in the neck, and thus also important viscous dissipation. Furthermore, at small enough η, at frequencies close but smaller than resonance frequency, the corresponding neck flow becomes predominant and the effective wavelength is drastically reduced, leading to a so-called *slow speed*. However, when the shear viscosity increases, the neck flow adjusts to a smaller value, eventually leading to the disappearance of the *slow speed*. The viscous losses also smooth out the extrema of the real and imaginary parts of the modulus in Fig. 2.6(b). Consequently, a wider frequency range of the negative real part of the bulk modulus is achieved by increasing the viscous losses. The thermal boundary layers close to the cavity walls, where the fluid bulk modulus passes from adiabatic to isothermal value, mainly bring a small correction to the cavity spring constant (the cavity dimension is much larger than the boundary layer thickness $\delta_t = (2\kappa/\rho_0 c_p \omega)^{\frac{1}{2}}$). Therefore, their presence do not affect much the effective bulk modulus.

As explained before, the dynamics of the material will be very sensitive to the width of the neck, where a considerable part of the viscous losses take place. In our case (see values of the parameters in Fig. 2.5), between the frequencies $k_0L/\pi = 0.1$ and 0.4, the ratio of the viscous boundary layer thickness $\delta_v = (2\eta/\rho_0\omega)^{\frac{1}{2}}$ to the width of the neck, insensibly changes from 0.35 to 0.39. We observed that, in general, to maintain the similar behavior of the wavenumber and modulus, this ratio should remain in the same order, regardless of changing the scale of the material or the saturating fluid. Here, well above the resonance, at $k_0L/\pi = 0.5$, we can check that the effective wavelength in the material λ_{eff} is comparable to that in air λ_0: $\lambda_{eff}/L \sim 5$, and $\lambda_0/L = 4$. At the resonance frequency $k_0L/\pi = 0.15$, we find that $\lambda_{eff}/L \sim 8$. Roughly, this is a reduction by a factor of two of the wavelength in air ($\lambda_0/L \sim 13.33$), and an illustration of the mentioned trend of a *slow speed* close to the resonance. Although this structure represents a subwavelength material, and therefore, can be regarded in the large wavelength limit $\lambda_{eff} \gg L$, the local theory based on the two-scale homogenization at order zero does not predict correctly the acoustics, ignoring the resonance behavior. The origin of the failure is the presence of widely different length scales, allowing for resonances.

The same modeling framework has been used to study the case of 3D materials.[24] It has been noted that,[24] if the structure with the same geometrical parameters is embedded in water, there would be less loss as the the viscous boundrary layer thickness is smaller compared with that of air. To keep the same dynamic behavior with water as with air, it would be necessary to very significantly decrease the width of the neck; at this point it should be in mind that the complicated effect of nonlinearities would certainly have to be taken into account. Furthermore the thermal effects in water are not important. The general thermodynamic identity $\gamma - 1 = \beta_0^2 T_0/\rho_0 c_p$, shows that the deviation of $\gamma \equiv c_p/c_v$ from unity, is a second order effect on the thermal expansion coefficient β_0. For a liquid, like water, β_0 is very small; what implies that γ is practically 1. In this case, adiabatic bulk modulus $\chi_{0(adiab)}^{-1}$ and isothermal bulk modulus $\chi_{0(isoth)}^{-1}$

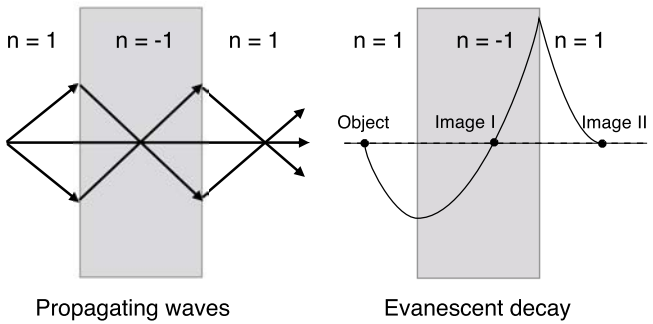
are very close, since in general, $\chi_{0(adiab)}^{-1} = \gamma\chi_{0(isoth)}^{-1}$. Therefore, thermal exchanges have practically no effects.

2.2.4. *Sound focusing and negative index acoustic metamaterials*

The first experimental demonstration of focusing ultrasound waves through a flat acoustic metamaterial lens composed of a planar network of subwavelength Helmholtz resonators has been presented in Ref. 13. A tight focus of half-wavelength in width at 60.5 kHz by imaging a point source has been observed. This result was in excellent agreement with the numerical simulation by transmission line model in which the effective mass density and compressibility has been derived.

Negative Refractive Index Lens

The refractive index is a fundamental parameter describing the interaction between waves and material. In late 1960s, Veselago[54] first considered the theoretical possibility for a medium having simultaneous negative permittivity and negative permeability. When $\varepsilon(\omega) < 0$, $\mu(\omega) < 0$ at certain frequency, the refractive index is $n = -\sqrt{\varepsilon\mu}$ and the negative sign is taken to satisfy causality. When light passes from a positive ($n > 0$) to a negative ($n < 0$) medium, Snell's law implies that the angle of refraction is negative, showing the refracted ray emerges on the same side of the normal as the incident ray. In 2000, Pendry[55] proposed that a thin slab of metamaterial with negative refractive index could make a *perfect lens* without any loss of details. A conventional lens cannot focus light onto an area smaller than a square wavelength due to the diffraction limit. The reason for this limit is that the evanescent waves which carry the subwavelength details of the object are exponentially decaying, leading to the loss of those fine features in the image. However, utilizing negative refractive index material, a *perfect lens* can focus propagation waves and also amplify evanescent waves to generate subwavelength imaging.



(a) (b)

Fig. 2.7. A flat lens brings all the diverging rays from an object into two focused images (a). The NI medium can enhance the evanescent waves across the lens, so the amplitude of the evanescent waves are identical at the object and the image planes (b).

As shown in Fig. 2.7, negative refraction allows a flat slab lens of negative index (NI) to focus all the diverging light rays from an object into two images: one inside the slab and the other one outside the slab. The evanescent waves have been enhanced across the lens and decay again after emerging from the negative index lens. Therefore the amplitude at the two image planes reaches their original level. At the same time, the propagating waves pass through the negative index lens with a reversed phase front, leading to zero phase change at the image planes. By completely recovering both propagating and evanescent waves in phase and amplitude, a perfect image is obtained.

An approximation of the *perfect lens* called as *super lens* was built in optical frequency range under near-field condition. In the near field, since the electric and magnetic components are decoupled, the *super lens* only needs negative dielectric permittivity for one polarization light. The optical superlensing effect on the scale of 60 nm ($\lambda/6$) was observed by excitation of surface plasmons[56] through the metal/dielectric layer structure. The sub-diffraction-limited image was recorded by optical lithography at 365 nm wavelength. In microwave frequency, subwavelength focusing was realized by different groups. An example of subwavelength focusing in

microwave frequency has been realized using a planar transmission-line structure.[57] The negative index lens is a planar slab consisting of a grid of printed metallic strips over a ground plane, loaded with series capacitors and shunt inductors. In the experiment, the loaded grid is sandwiched between two unloaded printed grids that act as effective homogeneous media with a positive refractive index.

Sound Focusing by Acoustic Transmission Line Network

Figure 2.8 shows the experimental setup to study the focusing phenomena of the acoustic metamaterial. To prepare the sample, we machined a 2D array of periodically connected subwavelength Helmholtz resonators in an aluminum plate and the resonators are filled with water. As shown in previous work,[58–60] a main transmission channel with recurrent side branches, which are closed at the

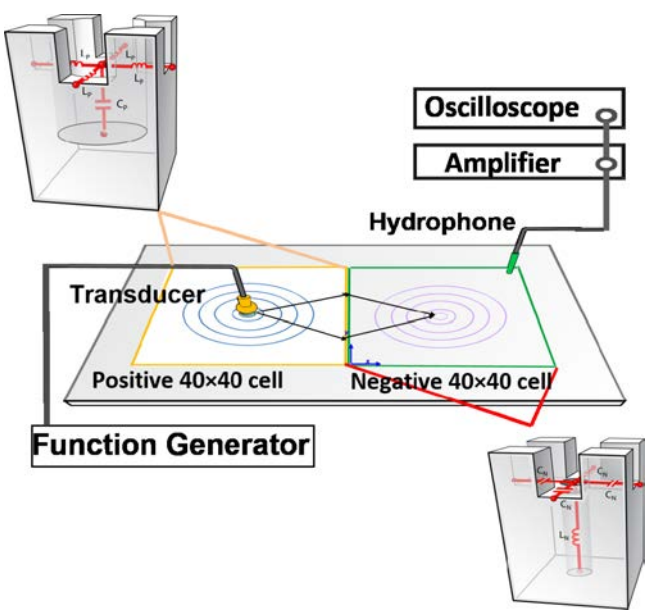


Fig. 2.8. Schematic showing the experimental setup. The sample with PI-NI interface is composed of an array of different designed Helmholtz resonators machined from an aluminum plate. Unit cells of each half part and the corresponding inductor-capacitor circuit analogy are shown in the insets.[13]

outer end, is analogous to a circuit of a series of inductors with shunt capacitors. On the other hand, when the side tubes inserted in the main channel are open on the outer end, the acoustic system can be described by a lumped network of a series of capacitors with shunt inductors. The left and right half parts in the sample are 2D periodic versions of those different types of topology, respectively. One unit cell from each half part is enlarged and shown in the two insets, respectively.

The left half part is composed of a 2D array (40 × 40) of large cavities connected with main channels. The volume of the cavity is around 10 times that of one section of the channels. Such an acoustic system is analogous to an inductor-capacitor circuit as shown in the inset of Fig. 2.8 with the channels acting as a series of inductors (L_P) and the cavity providing the stiffness element as capacitors (C_P). The periodicity (3.175 mm) of the sample is one-eighth of the wavelength at around 60 kHz frequency range. Given this value, the lumped circuit model is a valid approximation for the distributed acoustic system with only 10% error.[62] Following the approach of EM circuit analysis (see also Section 2.2.1),[27,57,61] the effective density and compressibility of this network can be expressed in the form as $\rho_P = L_P S_P/d_P$, $\chi_P = C_P/S_P d_P$, where d_P is the periodicity and S_P is the cross section area of the channels. Both effective density and compressibility are positive. Effective relative acoustic refractive index n_P can be determined by $n_P = c_0\sqrt{L_P C_P}/d_P$ where c_0 is the speed of sound in water. We call this half part the effective positive index (PI) medium. Such an acoustic system is described as a lumped network with a series of capacitors (C_N) for the main channel part and a shunt inductor (L_N) due to the orifice. The periodicity is the same as that in the left part, so the effective mass density and compressibility can be similarly estimated as $\rho_N = -S_N/(\omega^2 C_N d_N)$, $\chi_N = -1/(\omega^2 L_N d_N S_N)$ where d_N is periodicity and S_N is the cross section area of connecting channels. Both parameters are negative. The refractive index $n_N = c_0/c(\omega) = -c_0/(\omega^2 d_N\sqrt{L_N C_N})$ is negative. Therefore, this material acts as a medium exhibiting NI of refraction. The two half parts are designed with effective indices of

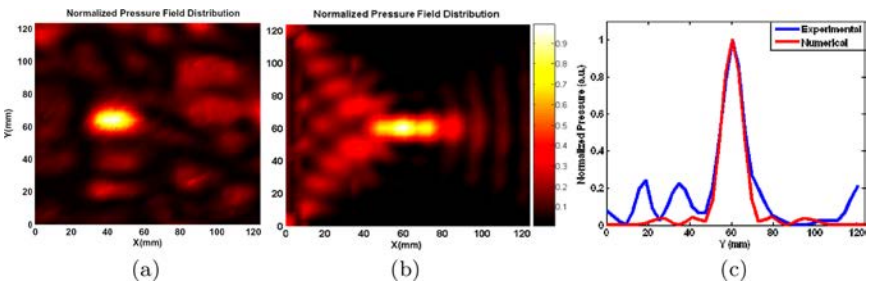


Fig. 2.9. Pseudocolor map of the normalized pressure field distribution at 60.5 kHz. Measured (a) and simulated (b) field map of the acoustic NI metamaterial and line plot of pressure field across the focal plane parallel to interface (c).[13]

equal and opposite value and matched impedance $\sqrt{\rho/\chi}$ at the design frequency 60.5 kHz.

In the experiments, the pressure field through this PI-NI interface has been measured to confirm the focusing in this material. Comparison of Figs. 2.9(a) and 2.9(b) shows that the field plots found through simulation, by circuit simulator SPICE, are in remarkable agreement with the experimental results.

In Fig. 2.9(c), the measured data shown by the blue line is shifted to the left by 3.175 mm for comparison purposes. The comparison demonstrates a very good match in the focus width between the measurement and the numerical simulation. This analysis predicts that the negative refractive index approaches -1 relative to the PI part at 60.5 kHz. In order to achieve high-quality focus imaging, the ratio of the refractive index should be -1 at the PI-NI interface. Only when the index is matched, based on ray acoustics, the angle of refraction equals the angle of incidence for each ray such that all rays can be brought to the same focal spot in the NI part. However, the refractive index is not exactly matched in the experiment. This discrepancy is related to the loss as well as variation in the inductors and capacitors from their designed values due to machining tolerance. It has been noted that single PI-NI interface does not allow enough growth of evanescent fields to achieve subdiffraction focusing[61] while sandwich structure (two PI-NI interfaces) offers a better chance to overcome the diffraction limit.[57]

The emission of a point source at kilohertz frequency which was brought to a focus through the PI-NI interface because of the negative refraction in this ultrasonic metamaterial, has been expected to be a step toward a novel acoustic imaging lens. The resolution of 0.5 wavelength was recorded by focusing the acoustic field of a point source. This is not subdiffraction imaging, but among the best achievable passive acoustic imaging elements. The unit cell of the acoustic network is only one-eighth of the operating wavelength, making the lens a compact size. Compared with conventional lenses, the flat thin slab lens has advantages in that there is no need to manufacture the shapes of spherical curvatures and the focus position is insensitive to the offset of source along the axis. Also, this negative-index lens offers tunable focal length at different frequencies.[13]

2.2.5. *Extraordinary focusing of sound above Helmholtz resonator array*

Recently, Lemoult *et al.*[63] used time reversal to focus sound above an array of soda cans into a spot much smaller than the acoustic wave length in air. However, the time reversal may not be necessary to achieve the extraordinary focusing in such system. In this section, we will experimentally demonstrate the extraordinary focusing above an soda can array using monochromatic sound excitation. We will also explain the experimental observations based on an analytical effective-medium model as well as finite-element calculations.[25]

2.2.5.1. *Experimental configuration and results*

As shown in Fig 2.10(a), we arranged 37 empty soda cans in a hexagonal array. Six commercial speakers, which were continuously driven at a given frequency, were placed symmetrically around the array. In the experiment, the Coca-Cola cans have a volume of 350 cm^3 and an opening area of 4 cm^2 similar to those used by Lemoult *et al.*,[63] with the fundamental Helmholtz resonance at 420 Hz. A microphone mounted on a translation stage was suspended at a height of 12 ± 2 mm above the top of the cans to collect acoustic intensity

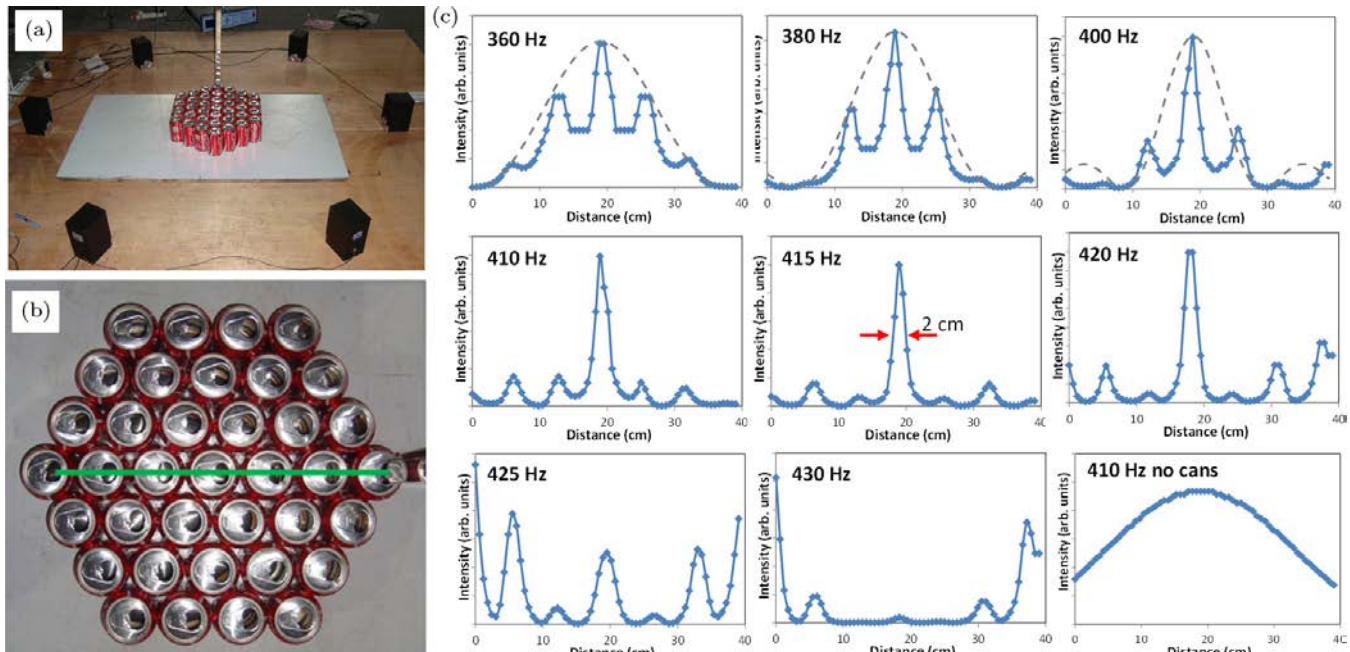


Fig. 2.10. (a) Experimental arrangement inside the anechoic room; (b) top view of the array with the scan line shown; (c) Acoustic intensity profiles along the diameter of the array at different frequencies. The bottom right panel shows an intensity profile measured without soda cans. Symbols are experimental points, connecting lines are guides to the eye. Dashed lines are Bessel function envelopes.[25]

distribution. The experimental results of the acoustic intensity profiles at different frequencies as well as a reference profile measured without cans are plotted in Fig. 2.10(b). Without the soda can array, a focal spot with a FWHM (full width at half maximum) of 31 cm at 410 Hz is observed, which is around 0.37 of the wavelength. With the soda can array, the focal spot at the center of the array gets progressively smaller as the resonant frequency is approached from below, becoming as narrow as 2 cm, or about $\lambda/40$, at 415 Hz. Above the Helmholtz resonance frequency, the intensity profile changes dramatically with maxima at the edges of the array and attenuation towards the center.

2.2.5.2. *Theoretical modeling*

In the effective medium approach, the acoustic wavelength is assumed to be much greater than the average distance between the resonators. The resonators can be modeled as mass-on-a-spring harmonic oscillators with pistons of mass M attached to springs with spring constant K as shown in Fig. 2.11(a). The resonators are regularly distributed in two dimensional plane with the average fractional piston area F.

For a Helmholtz resonator with a zero neck length, the effective mass is estimated as $(16/3)\pi^{-3/2}\rho_0 A^{3/2}$, where A is the opening area and ρ_0 is the density of air, whereas the spring constant K is given by $\rho_0 c^2 A^2/V$, where c is the speed of sound and V is the volume of the resonator.[64] For our soda cans, this model yields an effective mass of 9.23 mg and a spring constant of 64.8 N/m, resulting in a resonance frequency $\omega_0/2\pi = 422$ Hz.

The vertical position of a piston Z is obeying the equation of motion:

$$\ddot{Z} = -\omega_0^2 Z - \frac{pA}{M}, \tag{2.37}$$

where p is a deviation of the pressure above the piston from the equilibrium value and A is the area of the piston. In the case of long wavelength assumption, the average displacement of the boundary can be written as $u_z = FZ$. This leads to an effective boundary condition relating the average displacement and pressure at the boundary

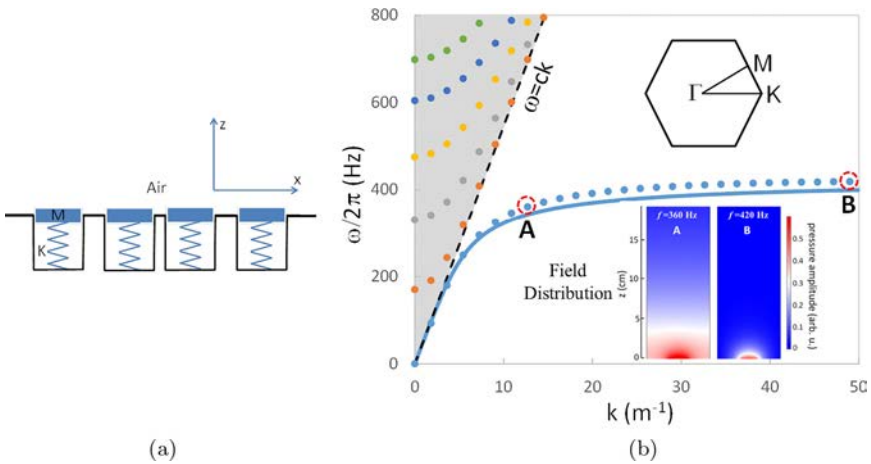


Fig. 2.11. (a) The model system. (b) Dispersion of acoustic waves propagating above the hexagonal array of soda cans calculated by FE (symbols) along the ΓK direction of the reciprocal lattice vs. the effective medium calculation (solid line). Shaded area represents the continuum of bulk modes in the semi-infinite space whereas FE calculations yield discrete modes due to the finite height of the simulation domain. The inserts are distributions of the sound pressure amplitude in the guided mode above a soda can for points A and B of the dispersion curve.[25]

$z = 0$,

$$\ddot{u_z} + \omega_0^2 u_z = -\frac{pFA}{M}, \tag{2.38}$$

which replaces the boundary condition of zero displacement at the rigid boundary in the absence of resonators. By assuming a harmonic wave propagating along the x direction, $\phi = \tilde{\phi}(z) \exp(i\omega t - ikx)$, we can achieve the following equation of motion:

$$\frac{\partial^2 \tilde{\phi}}{\partial z^2} = \left(k^2 - \frac{\omega^2}{c^2}\right) \tilde{\phi}, \tag{2.39}$$

and finally get a dispersion relation for ω and k,

$$\left(k^2 - \frac{\omega^2}{c^2}\right)^{1/2} (\omega_0^2 - \omega^2) = \frac{\omega^2 FA}{M}. \tag{2.40}$$

The dispersion relation is plotted in Fig. 2.11(b) (solid line) for $\rho_0 = 1.23$ kg/m^3, $c = 343$ m/s, $F = 0.106$ (calculated for dense

hexagonal packing of the cans). In the limit of small k the dispersion approaches that of the bulk wave in air, $\omega = ck$, whereas in the opposite limit of large k the frequency asymptotically approaches the resonance frequency ω_0. For a more accurate description of the wave propagation above a hexagonal lattice of soda cans, we use finite element (FE) calculations.

The acoustic module in COMSOL Multiphysics was applied to calculate the dispersion relation of soda cans arranged in a hexagonal lattice. Soda cans were modeled as cylinders with rigid walls of 11.5 cm in height and 6.6 cm in diameter. The opening in the model was circular and centered at the axis of the cylinder, with the same area of 4 cm^2 as the opening of a real can. Due to the computational resource limitation, the height of the simulation domain was set as 1 m, with rigid wall boundary conditions at the "ceiling". Floquet periodic boundary conditions were applied in order to find acoustic eigenmodes of an infinite 2D hexagonal lattice. The calculated dispersion relation along the ΓK direction of the reciprocal lattice is plotted in Fig. 2.11 (symbols). Discrete modes in the shaded area above the sound line $\omega = ck$ are due to a finite height of the simulation domain. For a semi-infinite half space, the shaded area should be filled by a continuum of bulk waves propagating at oblique angles to the floor. The mode below the sound line is guided by the can array, and its dispersion is close to what the simple effective medium theory has predicted.

In conclusion, we have demonstrated that focusing of sound in a metamaterial formed by a 2D array of soda cans results in an increasingly narrow intensity peak as the acoustic frequency approaches the Helmholtz resonance from below. The observed phenomenon results from the small acoustic wavelength in the metamaterial in combination with a near-field effect, i.e., the localization of the acoustic intensity at the opening of a can at frequencies close to the resonance. Furthermore, we found that the acoustic wave propagating along the Helmholtz resonator array is a guided mode becoming increasingly confined to the array as its frequency approaches the Helmholtz resonance from below.

2.3. Acoustic Transformation and Its Application

As we could see earlier, acoustics and electromagnetism have been brought together through different analogies derived from high similarities in their formalisms, despite deep differences in their physical nature. How coordinate transformation was used in both fields, does not make an exception to that rule. In 2006, Pendry *et al.*[16] showed that one could manipulate at will[a] the wave propagation in the framework of a relatively easy formalism -transformation optics, the coordinate transformation in optics- in comparison with the offered possibilities. Soon, the concept was transposed to acoustics.

Among the foreseen applications, the invisibility cloak has been the subject of first experimental studies in late 2006.[20] It was followed later by the acoustic ("inaudible") cloak, first realized in 2011 for acoustic waves[50] and in 2012 for elastic waves.[66] In this section, we start by a brief summary of the recent rebirth of coordinate transformation in optics and acoustics, with an emphasis on the "inaudible" acoustic cloak.[50] Then, we show how coordinate transformation can be broadly applied to different physics and lead for example to surface wave cloaking. Finally, we show that the concept of coordinate transformation can be applied more generally, allowing for example the design of a complementary material meant to remove the aberrations resulting from the transmission of an acoustic wave through any arbitrary heterogeneous medium.

2.3.1. *Transformation acoustics*

2.3.1.1. *Coordinate transformations and invisibilty*

The work by Pendry *et al.* that was published in 2006, is in fact based on the conclusions of an article published 10 years earlier. In 1996, Ward and Pendry showed that Maxwell's equations are invariant under any arbitrary transformation of coordinates.[65] In other words, after the coordinate transformation, their form is the same as the one

[a]The control of wave propagation relies by spatially controlling the material parameters ϵ and μ.

they have in the Cartesian mesh. The only difference is that if we start with an isotropic material for which ϵ and μ are scalar, after the coordinate transformation, we generally end up with anisotropic and heterogeneous material for which ϵ and μ are represented by tensors.

This form invariance can be used in a smart way. Indeed, chosen in a suitable manner, the coordinate transformation makes it possible to exclude the field from any arbitrary surface (2D) or volume (3D). Furthermore, it can be chosen so that the field lines are reformed on each side of the prohibited area, thus making this space invisible to any observer or listener.

In view of the invisibility cloak, the first coordinate transformation scheme that was historically provided, was to exclude waves from a sphere.[16] As we discuss later the two dimensional case, we present here one coordinate transformation which allows us to exclude the field from a disk surface:

$$\begin{aligned} r' &= R_1 + r(R_2 - R_1)/R_2 \\ \theta' &= \theta \\ z' &= z \end{aligned} \tag{2.41}$$

This coordinate transformation makes the region of space that is comprised in the circle of radius R_1 inaccessible, as it moves the points of that region to the region defined by $R_1 < r < 2R_1 - {R_1}^2/R_2$. Meantime, the space that was originally defined by $R_1 < r < R_2$, is squished in the region defined by $2R_1 - {R_1}^2/R_2 < r < R_2$. Space beyond R_2 remains unchanged. For the rays, whether they are of light or sound, the trajectories that initially borrow straight lines within the region $r < R_2$ become more or less severely curved in order to avoid the forbidden region (see Fig. 2.12).

This apparent simplicity has a cost, which is reflected in the way the parameters of the medium are changed by the coordinate transformation. Their calculation involves the Jacobian matrix of the latter, and the parameters of the non-deformed/homogeneous medium. As we have mentioned earlier, the new material parameters generally depend on the position but also on the direction.

Finally, we note that in the invisibility cloak, the trajectories are curved, and therefore they are longer than the straight trajectories

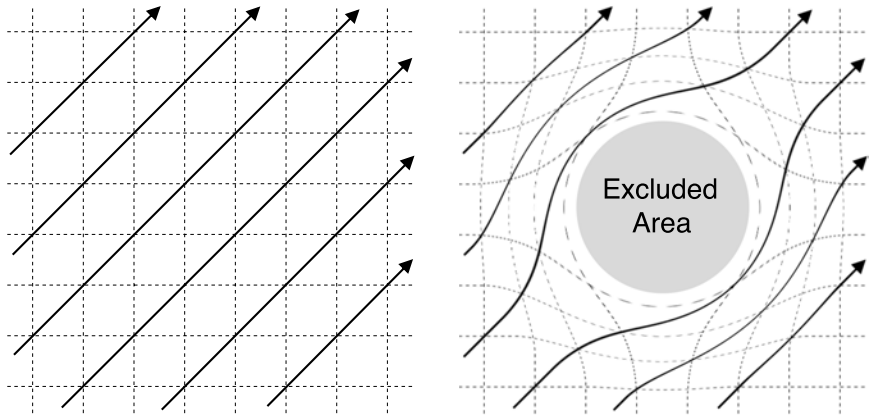


Fig. 2.12. Starting with a homogeneous configuration (a), the medium is deformed to exclude rays — of light or sound — (thick arrow lines) from the desired area (b). The coordinate transformation is *recorded* by comparing the ray arrows on both side, and applied to the equations to find the value of the new material parameters (or metamaterial).

in the medium surrounding the cape. For the wavefront to emerge non-deformed after it went through the cape, it is necessary that the speed of the wave is higher in the cloaking device than in the surrounding material.

This raises some issues in optics, where it is hard to fabricate a medium in which light propagates faster than in a medium of reference such as air.[b] On the contrary, in acoustics, there exists many media in which sound travels faster than in reference media such as water or air.

2.3.1.2. *From transformation optics to transformation acoustics*

In acoustics, the first study on coordinate transformation coincides with an article from Milton *et al.* which shows that the equation of

[b]In the cloaking device, it is the phase speed c_ϕ that is higher than the speed of wave outside the cloaking device. The group velocity c_g is given by the relation $c_g = {c_0}^2/c_\phi$ which ensures that it is always lower than c_0.

elasto-dynamic does not generally remain unchanged by coordinate transformation.[67] From this observation, attempts to apply coordinate transformation to the equation of acoustic, mainly focused on finding forms of the equation of acoustic that are similar to the form of Maxwell's equations.

Steven Cummer and David Schurig[68] consider the propagation of acoustic wave in a z-invariant 2D inviscid homogeneous fluid with anisotropic density $\rho = [\rho_r, \rho_\phi]$ and bulk modulus K_0.[c] Assuming time harmonic wave ($\exp(+j\omega t)$ convention), they write the equations for the conservation of momentum and mass in cylindrical coordinates as follows:

$$j\omega\rho_\phi v_\phi = -\frac{1}{r}\frac{\partial p}{\partial \phi}, \tag{2.42}$$

$$j\omega\rho_r v_r = -\frac{\partial p}{\partial r}, \tag{2.43}$$

$$j\omega K^{-1} p = -\frac{1}{r}\frac{\partial (r v_r)}{\partial r} - \frac{1}{r}\frac{\partial v_\phi}{\partial \phi}. \tag{2.44}$$

where p is the pressure and $\mathbf{v} = (v_r, v_\phi)$ is the particle velocity. At the same time, Cummer and Schurig notice that the z-invariant 2D Maxwell equations for transverse electric (TE) polarization, can be written in the following form:

$$j\omega\mu_r(-H_r) = -\frac{1}{r}\frac{\partial(-E_z)}{\partial \phi}, \tag{2.45}$$

$$j\omega\mu_\phi H_\phi = -\frac{\partial(-E_z)}{\partial r}, \tag{2.46}$$

$$j\omega\epsilon_z(-E_z) = -\frac{1}{r}\frac{\partial(r H_\phi)}{\partial r} - \frac{1}{r}\frac{\partial(-H_r)}{\partial \phi}. \tag{2.47}$$

Comparing these three equations to the equations of acoustics (2.42–2.44), they find that these two sets of equations are similar under the

[c]This medium is a transversely isotropic fluid.

following substitution:

$$[p, v_r, v_\phi, \rho_r, \rho_\phi, K^{-1}] \leftrightarrow [-E_z, H_\phi, -H_r, \mu_\phi, \mu_r, \epsilon_z] \tag{2.48}$$

This analogy between the time harmonic acoustic equations and Maxwell equations also preserves boundary conditions. Indeed, the pressure and the normal velocity component are continuous at a fluid-fluid interface, as the vertical component of the electric field and the tangential component of the magnetic field are at an electromagnetic material interface.

Following analogy (2.48) and applying the coordinate transformation corresponding to the equations (2.41) imposes the value of the material parameters K and ρ, inside the cloaking shell ($R_1 < r < R_2$), as follow:

$$\frac{\rho_r}{\rho_0} = \frac{r}{r - R_1} \tag{2.49}$$

$$\frac{\rho_\phi}{\rho_0} = \frac{r - R_1}{r} \tag{2.50}$$

$$\frac{K^{-1}}{{K_0}^{-1}} = \left(\frac{R_2}{R_2 - R_1}\right)^2 \frac{r - R_1}{r}. \tag{2.51}$$

Within the cloaking device ($R_1 < r < R_2$), the effective material parameters continuously change. In the limit $r \to R_1$, $\rho_r \to +\infty$, $\rho_\phi \to 0$ and $K \to +\infty$. Note here that there is no restriction on the values of K and ρ in the cloaked area, i.e., in the region $r < R_1$.

Other studies have shown such analogy between electromagnetism and acoustic equations. In particular Chen and Chan have shown that, in 3D, the acoustic equation in a heterogeneous fluid and the electrical conductivity equation have equivalent forms, therefore allowing to use coordinate transformation in order to guide acoustic wave in 3D-fluid.[69]

2.3.2. *Broadband acoustic cloak for ultrasound*

First numerical simulations are reported by Cummer and Schurig in Ref. 68 where a cylindrical cloaking shell is simulated based on

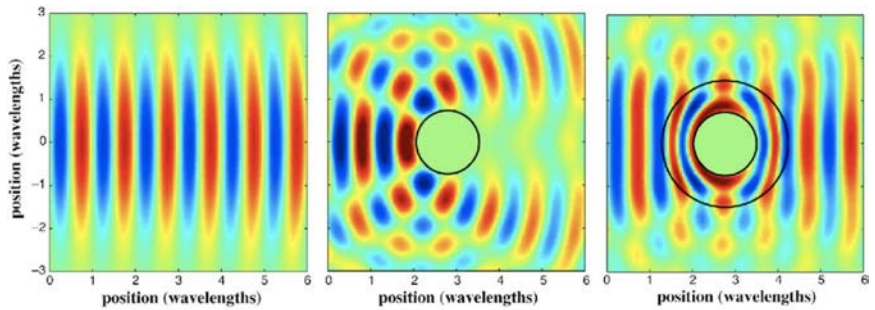


Fig. 2.13. Numerical simulation of an acoustic cloak from Ref. 68. (a) An acoustic beam is propagating in a homogeneous medium. (b) The acoustic beam is scattered by a cylindrical scatterer ($\varnothing = 0.7\lambda$). (c) The scatterer is surrounded by the acoustic cloaking shell.

analogy (2.48). Three configurations are compared: (a) an acoustic beam is sent in a homogeneous medium, (b) the beam encounters an incompressible cylindrical scatterer and (c) the beam encounters the cylindrical cloaking shell (0.75 wavelength thick) surrounding the scatterers. As expected, in configuration (b), the incompressible cylinder is responsible for large scattering and a shadow behind the cylinder. The latter disappears in configuration (c). Indeed, in the forward direction, the wavefront is reformed and retrieves with fairly good agreement the shape of the wavefront in configuration (a), as if there was no scatterer. As in Ref. 67, the difficulties going from the numerical study to the experimental proof are raised. Indeed, in nature, no material has the features of the cloak, namely, a continuous variation of the parameters of the medium along the radius of the cloaking shell, and an anisotropic density.

This challenge was met by Zhang et al. to whom we owe the first experimental realization of the acoustic inaudible cloak[50] for 2D acoustic wave in water. As shown in Fig. 2.14, the cloaking shell is divided in several concentric rings, each one being characterized by a particular set of acoustic material parameters.[d] The rings themselves

[d]As a reminder, ρ_r, ρ_θ and K all vary along r.

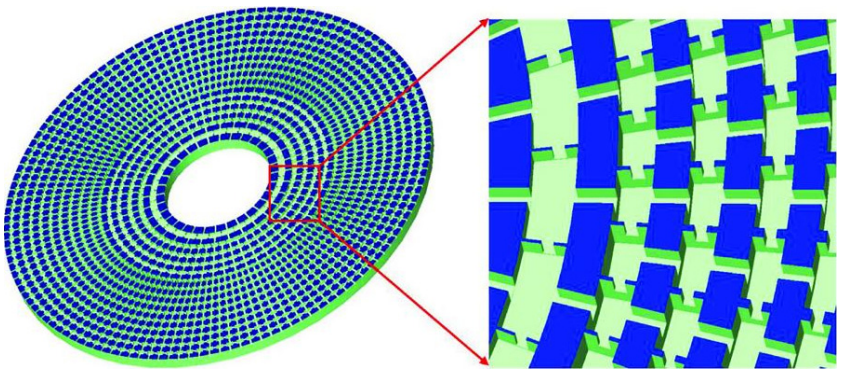


Fig. 2.14. Schematic of cloaking shell for 2D acoustic wave in water, from.[50] The aluminum cloak is devided into 16 concentric rings, with inner and the outer radii of the cloak equal to 13.5 mm and 54.1 mm.

are divided in small cells arranged in a circular array. Each cell is made of a large cavity surrounded by four narrow channels, communicating with the four neighboring cells (two on the same ring and two on the two surrounding rings). The cavity behaves as acoustic capacitor whereas the channels act as serial inductors as seen in Section 2.2.1. Around the central working frequency (60 kHz), each cell is only of the order of one tenth of the wavelength.

Figures 2.15(a–c) shows experimental pressure field mappings resulting from the scattering of an acoustic wave by a bare steel cylinder in water at three different frequencies (60 kHz, 52 kHz and 60 kHz). It is compared to the pressure field mappings when the steel cylinder is surrounded by the acoustic cloak (Figs. 2.15(d–f)). Although the acoustic cloak central frequency is 60 kHz (Figs. 2.15(a) and (d)), the acoustic cloak shows also great results at 52 kHz (Figs. 2.15(b) and (e)) and 64 kHz (Figs. 2.15(c) and (f)). Indeed, the steel cylinder shade is almost completely removed once it is surrounded by the acoustic cloak.

2.3.3. *Molding water, acoustic and electromagnetic waves with a single cloak*

In the most of invisibility cloaking demonstrations, the carefully designed artificial structure can be valid only for one physical

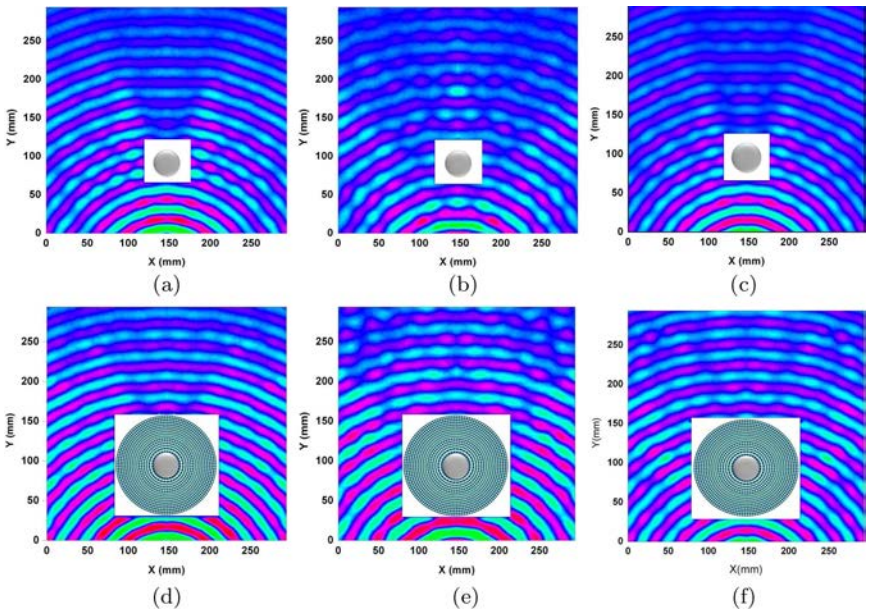


Fig. 2.15. Measured pressure field mappings of the bare cylinder and the cloaked steel cylinder illuminated with a point ultrasound source. The cloak lies in the center of a water tank and surrounds the steel cylinder. The scattering field patterns of the bare cylinder at (a) 60 kHz (b) 52 kHz, and (c) 64 kHz. The pseudo-color map in the immediate environment of the cloaked steel cylinder at (d) 60 kHz (e) 52 kHz, and (f) 64 kHz.[50]

variable, e.g., electromagnetic wave,[20,70–73] acoustic wave,[50,68,69,74,75] elastic wave,[66,76,77] and heat flux,[78–80] etc. In this section, we will introduce a cylindrical cloak which can work equally for linear surface liquid wave, acoustic wave, and electromagnetic wave.[74,81] This structured cloak behaves like a surface liquid wave cloak with an effective anisotropic shear viscosity, an acoustic cloak with an effective anisotropic density, and an electromagnetic cloak with an effective anisotropic permittivity, respectively. The effective anisotropic effect parameters are proceeded with mathematical approach of homogenization,[82,83] which amounts to replacing a structured material by an effective medium that captures the essential wave phenomena for wave with wavelength large compared to the typical heterogeneity

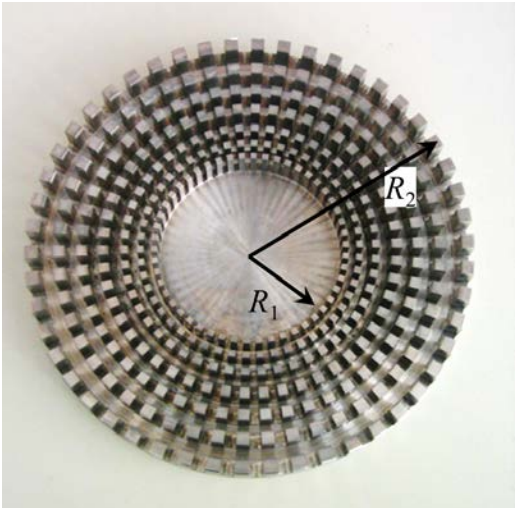


Fig. 2.16. Photo of the structured cloak. 100 rigid sectors are evenly machined in a metallic ring of inner radius $R_1 = 41$ mm and outer radius $R_2 = 100$ mm.[74]

size. We also numerically and experimentally demonstrate the performance of the cylindrical cloak in these three kinds of waves.

2.3.3.1. *Sample configuration*

The invisibility cloak as shown in Fig. 2.16 is manufactured in aluminum using classical numerically controlled machine tools. The outer and inner radii are $R_2 = 100$ mm and $R_1 = 41$ mm, respectively. The cloak is first divided into 14 layers with the layer thickness as the function of r as $(\frac{R_2(r-R_1)}{r(R_2-R_1)})^2$. One layer in two is further divided into 100 identical angular sectors along the azimuthal direction as shown in the figure, and there are seven rows of rods along the radius.

2.3.3.2. *Theoretical and numerical approaches*

For the the surface of a linear fluid medium, the conservation of momentum leads to the Navier-Stokes equations:

$$\rho_0 \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right) \mathbf{v} - \eta \nabla^2 \mathbf{v} = -\nabla p + \rho \mathbf{g} \tag{2.52}$$

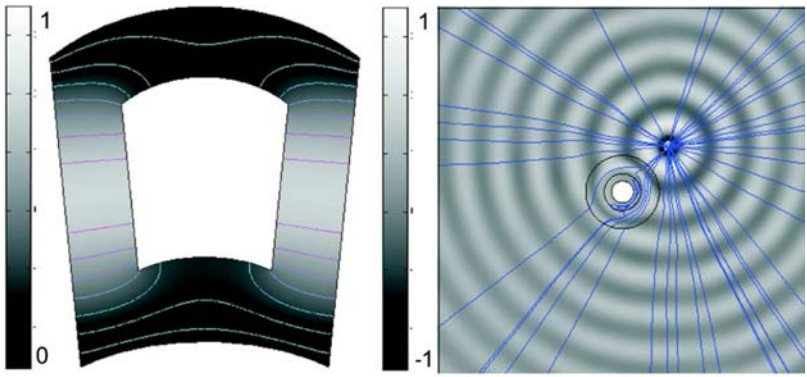
where $\mathbf{v}$ is the velocity field, $\eta\nabla^2\mathbf{v}$ accounts for the fluid's viscosity, p the fluid pressure, ρ_0 is density and $\mathbf{g}$ the vector of gravity force. Note that $\mathbf{g} = -g\mathbf{e}_3$, where g denotes the acceleration caused by gravity and $\mathbf{e}_3$ a vertical unit vector. It is noted that $\eta\nabla^2\mathbf{v}$ can be neglected outside the cloak region due to the fairly low viscosity of the surrounding fluid. In addition, the fluid should be also incompressible (divergence free), irrotational (curl free), and undergo only small fluctuations around a mean vertical position. Therefore, the deduced vertical displacement of the fluid for the harmonic oscillation can satisfy with Helmholtz's equation.[84] By solving the Eq. (2.52) for the cloak region, we can achieve an anisotropic matrix of viscosity whose nontrivial part (transverse shear) is:

$$[\eta_{hom}] = \frac{1}{A(Y^\star)} \begin{pmatrix} A(Y^\star) - \psi_{rr} & \psi_{r\theta} \\ \psi_{\theta r} & A(Y^\star) - \psi_{\theta\theta} \end{pmatrix} \tag{2.53}$$

Here, $A(Y^\star)$ denotes the area of the region $Y^\star$ surrounding a rigid inclusion (subject to Neumann boundary conditions) in an elementary cell Y of the periodic array, and ψ_{ij} represent corrective terms, which is related to periodic hydrostatic fields, as shown in Fig. 2.17(a). For simplification, in our sample, we can introduce some variation in the radial length of sectors for which it seems reasonable to assume that the improved cloak is characterized by an effective anisotropic fluid whose shear viscosity (a diagonal matrix in polar basis) is

$$\eta'_{rr} = \left(\frac{R_2(r - R_1)}{r(R_2 - R_1)}\right)^2, \quad \eta'_{\theta\theta} = \left(\frac{R_2}{R_2 - R_1}\right)^2 \tag{2.54}$$

where R_1 and R_2 are the inner and outer radii of the ring, respectively. Importantly, the effective fluid's density is the same as the fluid density, which does not play any prominent role. These parameters were actually first proposed in Ref. 85 for the case of electromagnetic waves. The snapshot of the surface waves for the homogenized coating is shown in Fig. 2.17(b), which is calculated by COMSOL Multiphysics. The stream lines indicating the direction of the fluid flow, clearly demonstrate the bending effect in the cloaking area.



(a) (b)

Fig. 2.17. (a) Calculated periodic hydrostatic field. (b) Pattern of the concentric surface wave and associated stream lines (indicating the direction of fluid flow) generated by a forced term with frequency 15.84 Hz. A rigid cylinder of radius 38 mm is placed in the center of a homogenized cloak.[74]

Similar approaches are applied for acoustic wave and electromagnetic wave as well. The homogenized wave equations (in polar coordinates) for acoustic waves can be written as:

$$-\nabla \cdot \left(\begin{bmatrix} \rho_{rr}^{-1} & 0 \\ 0 & \rho_{\theta\theta}^{-1} \end{bmatrix} (r) \nabla p(r,\theta) \right) = \frac{\omega^2}{K_{eff}} p(r,\theta) \tag{2.55}$$

where the tensor of effective density is given by

$$\begin{bmatrix} \rho_{rr}^{-1} & 0 \\ 0 & \rho_{\theta\theta}^{-1} \end{bmatrix} = \frac{1}{\text{area}(Y)} \int_0^{2\pi} \int_0^1 \rho(r,r',\theta') \times \begin{bmatrix} \text{area}(Y) - \psi_{rr} & \psi_{r\theta} \\ \psi_{\theta r} & \text{area}(Y) - \psi_{\theta\theta} \end{bmatrix} r' dr' d\theta' \tag{2.56}$$

Here, ψ_{ij} are periodic potentials.

And for electromagnetic wave:

$$-\nabla \cdot \left(\begin{bmatrix} \epsilon_{rr}^{-1} & 0 \\ 0 & \epsilon_{\theta\theta}^{-1} \end{bmatrix} (r) \nabla H_z(r,\theta) \right) = \omega^2 \mu_{eff} H_z(r,\theta) \tag{2.57}$$

where ϵ_{rr}^{-1} and $\epsilon_{\theta\theta}^{-1}$ are the coefficients of the tensor of effective permittivity which are also given by Eq. (2.56) (replacing ρ by μ in the equation) and μ_{eff} is the effective permeability which can be written as:

$$\mu_{eff} = \frac{1}{\text{area}(Y)} \int_0^{2\pi} \int_0^1 \mu(r, r', \theta') r' dr' d\theta' \tag{2.58}$$

It is noted that this equation is simply derived from the vector Maxwell equation by considering a magnetic field $\mathbf{H} = (0, 0, H_z)$.

It can be shown that the reduced effective parameters are good approximations for well-known transformed density and permittivity in their reduced forms:

$$\rho'_{rr} = \left(\frac{R_2(r - R_1)}{r(R_2 - R_1)}\right)^2, \quad \rho'_{\theta\theta} = \left(\frac{R_2}{R_2 - R_1}\right)^2 \tag{2.59}$$

$$\epsilon'_{rr} = \left(\frac{R_2(r - R_1)}{r(R_2 - R_1)}\right)^2, \quad \epsilon'_{\theta\theta} = \left(\frac{R_2}{R_2 - R_1}\right)^2 \tag{2.60}$$

The numerical simulation for all these three cases are conducted by using the commercial finite elements package COMSOL Multiphysics. The geometric configuration in the simulation is same as the real structured cloaking device. Here, we select electromagnetic wave as an example to demonstrate the invisible cloak behavior as shown in 2.18. In order to quantitatively assess the cloaking efficiency of the metamaterial, we numerically compute its total radar cross section (RCS),[86,87] and plot them in Fig. 2.18(d). We can observe that a small infinite conducting (resp. rigid for pressure waves) obstacle surrounded by the cloak is slightly above RCS of the cloak on its own, but lower than RCS of small infinite conducting (resp. rigid) obstacle, which is itself smaller than RCS of the large infinite conducting (resp. rigid) obstacle of same diameter as the cloak up to 7 GHz (resp. 8 KHz for pressure waves). The theoretical and numerical results confirm that these two-dimensional finite element simulations are valid for all water wave, acoustic wave, and transverse electromagnetic wave.

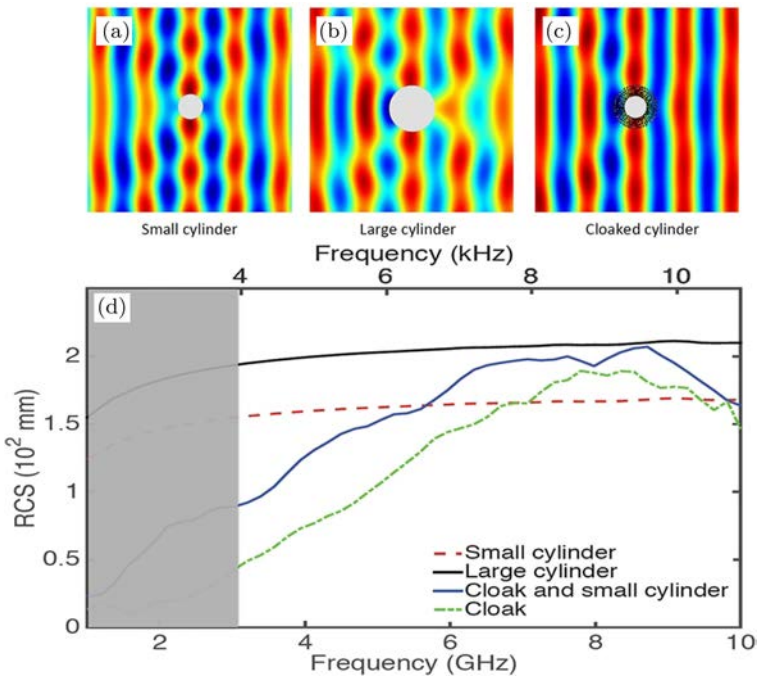


Fig. 2.18. The simulated real part of the transverse magnetic (resp. pressure) field for a plane wave incident at 4.5 GHz (resp. 5 kHz) from the left on the small obstacle surrounded by the cloak on the small obstacle on its own (a), on the large obstacle (b), and on the cloaked obstacle (c). (d) Numerical simulation of radar cross section (RCS) for cloak and cylindrical rigid obstacles.[81]

2.3.3.3. *Experimental characterization*

The performance of the invisible cloak has been characterized for water, acoustic, and electromagnetic waves, respectively. In Fig. 2.19 (a), we present the schematic experimental setup for observing back scattering water wave from the cloak. In the system, a halogen lamp modulated by a perforated rotating disc illuminates a transparent vessel containing the liquid. We choose a small tube to excite a localized pressure with the same frequency as the modulation of the light. The surface waves create local curvatures of the liquid and the light is refracted when crossing the surface. Therefore, on the screen the dark and light zones allow visualizing the liquid surface waves. We

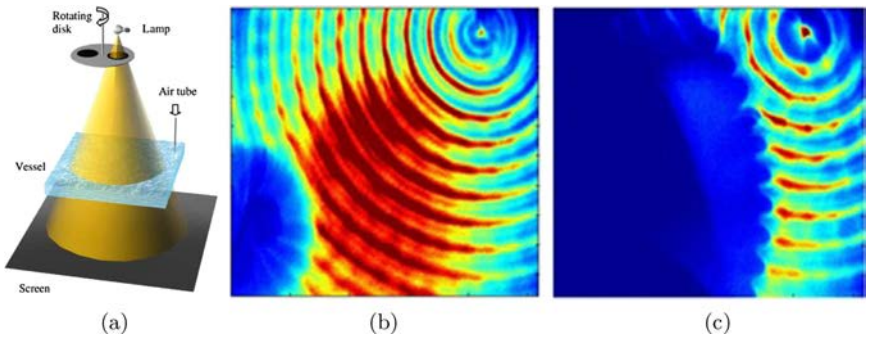


Fig. 2.19. Schematic view of the experimental setup for linear surface water waves. (b–c) Experimental results of reduced backscattering adapted from[74] for a concentric liquid surface wave of frequency 10 Hz interacting with a rigid cylinder (7.6 cm in diameter) on its own (b) and the structured waterwave cloak (20 cm in diameter) (c).[81]

take snapshots of the liquid surface waves when a metallic cylinder is placed in the vessel alone and with the invisibility cloak, which are shown in Fig. 2.19(b) and (c). The dramatic reduced back scattering field can be observed if the metallic cylinder is placed in the cloak.

The schematic view of acoustic wave experimental setup is shown in Fig. 2.20(c). A commercial loudspeaker is placed at 20 cm from the structure cloak. The loudspeaker is driven by a programmable functional generator and a power amplifier. A high sensitive microphone mounted on an x-y translation stage is used to record the two-dimensional acoustic pressure field distribution. The loudspeaker, acoustic cloak, and microphone are placed at the same level for characterizing how the pressure field is affected by the cloak. In the experiment, a five-period sinusoidal wave is launched by the loudspeaker. The time-dependent pre-amplified signal collected by the microphone is recorded by an oscilloscope and downloaded to a computer for further analysis. The signal generation and acquisition are synchronized by the computer program. Therefore, we can map the 2D pressure field distribution with both amplitude and phase information. The scanning area is 25×25 cm^2. The measured forward scattering fields for a cylindrical scatterer with or without the cloak are plotted in Fig. 2.20(a). We can observe that the acoustic wavefront is neatly

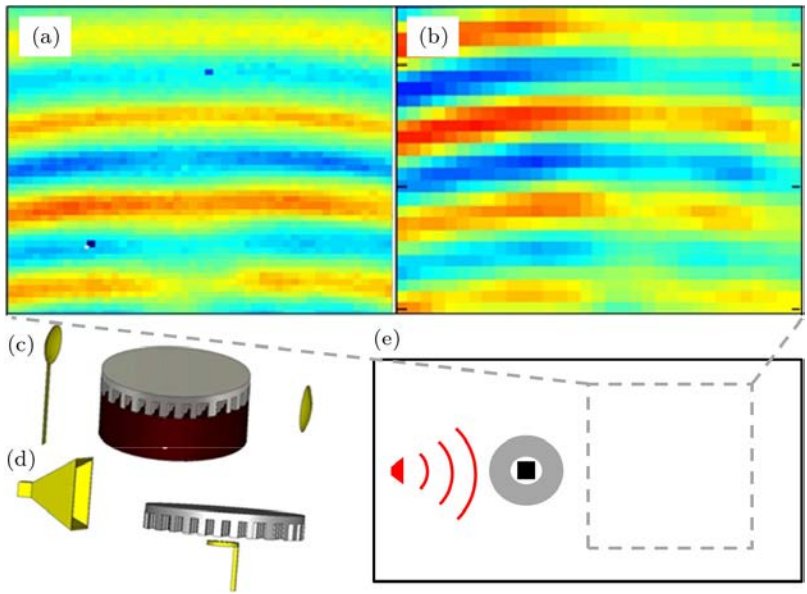


Fig. 2.20. (a, b) Experimental results in forward scattering for pressure waves (a) and transverse electric microwaves (b) for the same cloak containing a large glass bottle obstacle (a) and a small square metallic obstacle (b). (c, d) Schematic view of the experimental setup for acoustic wave (c) and microwave (d). (e) Top view of 2D configuration for forward scattering wave mapping. The measured area are indicated by the dashed box.

reconstructed behind the cloak, with a good agreement with the numerical simulation results.

The schematic view of microwaves experimental setup which is positioned in an anechoic chamber is shown in Fig. 2.20(d). A ridged horn antenna is positioned at 24 cm from the structure cloak. A rectangular metallic obstacle is inside the cloak structure. The magnetic probe is positioned 5 mm above the structure cloak. The probe is a homemade magnetic loop with diameter 5 mm and the loop is positioned perpendicular to the magnetic field of the emitter antenna. In this configuration, the magnetic probe measures a single component (vertical component) of the three Cartesian components of the magnetic field. The magnetic loop scanned a 33×71 cm^2 surface thanks to a 3D axis positioning system.[86] The antenna and the probe

are connected to a vector network analyzer that generates and measures the electromagnetic field. The electromagnetic field is emitted by the antenna and the probe only receives the field. The vector network analyzer measures the transmission coefficient between the emitter antenna and the magnetic probe. The complex transmission coefficient between the antenna and the probe is displayed in terms of magnitude in dB to visualize the intensity of the field. The real part of the magnetic field is also displayed since this is related to the phase of the field. The measured forward scattering longitudinal magnetic field for a square metallic obstacle placed in the cloak is plotted in Fig. 2.20(b). Similar to the results of the acoustic wave, the electromagnetic wavefront is reconstructed.

In conclusion, we present three experiments demonstrating that a cylindrical cloak works equally well for linear surface liquid wave, acoustic wave, and electromagnetic wave. Measured forward scattering for pressure and magnetic fields are in good agreement and confirm broadband cloaking with a central frequency of 5 kHz for sound and 4.3 GHz for microwaves. Microwave experiments further confirm the much reduced forward and backscattering when a rectangular metallic obstacle is surrounded by the structured cloak range of cloaking frequencies between 2.6 and 6.0 GHz. This suggests, as supported by numerical simulations, sound waves are cloaked between 3 and 7 kHz and linear surface liquid waves are cloaked between 8 to 14 Hz. Moreover, microwave experiments confirm the field is reduced by 10 to 30 dB inside the invisibility region, which suggests the multiphysics cloak could be used as a protection against water, sonic or microwaves.

2.3.4. *Anisotropic complementary acoustic complementary metamaterials*

In the previous sections, we have introduced conventional cloaking strategies[50,74,88] by compressing the space and hide the object inside an enclosure in which there is no interaction with the outside world. However, this solution is not suited to the problems for

the application of medical ultrasound or non-destructive evaluation (NDE). In these problems, ultrasound needs to be transmitted through an aberrating layer,[89–95] where either the transmission is desired to be maximized or the reflection needs to be minimized. One of the most representative examples is transcranial ultrasound beam focusing, which could find usage in both brain imaging and treatment.[91,92] However, transcranial beam focusing is extremely challenging due to the presence of the skull. A common approach to achieve transcranial beam focusing is based on the time-reversal/phase conjugate technique and ultrasound phased arrays.[96,97] Although the focal position can be corrected, one significant shortcoming of this strategy is that it does not compensate for the large acoustic energy loss due to the impedance mismatch between the skull and the background medium (water). Recently, Lai *et al.* demonstrated that cloaking/illusion based on electromagnetic wave (EM) complementary metamaterials (CMM)[98] can open up a virtual hole in a wall without distortion.[99,100] In addition, this type of approach does not require the cloaked object to be inside an enclosure/cloaking shell and is valid in free space.[101] Due to the similarity between acoustic and EM wave equations in 2D, CMMs have been also proposed for acoustic cloaking.[102,103]

In this section, we will introduce a type of anisotropic, acoustic complementary metamaterials (CMM) and their application in restoring acoustic fields distorted by aberrating layers.[104] The proposed quasi 2D, non-resonant CMM consists of unit cells formed by membranes and side branches with open ends. Simultaneously anisotropic and negative density is achieved by assigning membranes facing each direction (x- and y-direction) with different thicknesses while the compressibility is tuned by the side branches. Numerical examples demonstrate that, the CMM, when placed adjacent to a strongly aberrating layer, could acoustically cancel out that aberrating layer. This leads to dramatically reduced acoustic field distortion and enhanced sound transmission, therefore virtually removing the layer in a noninvasive manner.

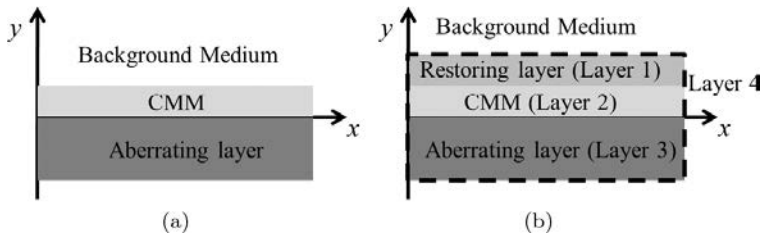


Fig. 2.21. (a) Schematic of the CMM, the aberrating layer, and the background medium. (b) Schematic of acoustic cloaking using CMM, the dashed lines indicate the boundaries of layer 4.

2.3.4.1. *Theoretical approach for CMM*

The CMM is placed on top of the aberrating layer, as illustrated in Fig. 2.21(a). The aberrating layer is assumed to be sufficiently long so that the edges do not significantly affect the acoustic field. The CMM compresses and cancels the information of the selected aberrating layer. Let $\rho^{(c)}(x^{(c)}, y^{(c)}, z^{(c)})$, $\chi^{(c)}(x^{(c)}, y^{(c)}, z^{(c)})$ and $\rho^{(a)}(x^{(a)}, y^{(a)}, z^{(a)})$, $\chi^{(a)}(x^{(a)}, y^{(a)}, z^{(a)})$ be the effective density and compressibility tensors of the CMM and the aberrating layer, respectively. $x^{(c)}, y^{(c)}, z^{(c)}$ and $x^{(a)}, y^{(a)}, z^{(a)}$ are generalized curved coordinates. Based on the acoustic coordinate transformation, we have[69]:

$$\left[\rho^{(c)}\right]^{-1} = \mathbf{A}\left[\rho^{(a)}\right]^{-1}\mathbf{A}^{T}/\det\mathbf{A}, \tag{2.61}$$

$$\chi^{(c)} = \chi^{(a)}/\det\mathbf{A}, \tag{2.62}$$

where $\mathbf{A}$ is the Jacobian transformation tensor of compressing transformation given by:

$$\mathbf{A} = \begin{bmatrix} \frac{\partial x^{(c)}}{\partial x^{(a)}} & \frac{\partial x^{(c)}}{\partial y^{(a)}} & \frac{\partial x^{(c)}}{\partial z^{(a)}} \\ \frac{\partial y^{(c)}}{\partial x^{(a)}} & \frac{\partial y^{(c)}}{\partial y^{(a)}} & \frac{\partial y^{(c)}}{\partial z^{(a)}} \\ \frac{\partial z^{(c)}}{\partial x^{(a)}} & \frac{\partial z^{(c)}}{\partial y^{(a)}} & \frac{\partial z^{(c)}}{\partial z^{(a)}} \end{bmatrix} \tag{2.63}$$

In this study, we focus on 2D problems and only wave propagation in the x-y plane is of interest. The z component in the Jacobian

matrix is therefore dropped. Without loss of generality, the thickness of the CMM is assumed to be half of the aberrating layer, which leads to $\frac{\partial x^{(c)}}{\partial x^{(a)}} = 1$ and $\frac{\partial y^{(c)}}{\partial y^{(a)}} = -0.5$, whose ratio is negative since the acoustic information is folded in the CMM and would cancel out that of the aberrating layer. Other components $\frac{\partial x^{(c)}}{\partial y^{(a)}}$ and $\frac{\partial y^{(c)}}{\partial x^{(a)}}$ in the tensor are equal to zero, as the transformation in each direction (i.e., x- and y-directions) are independent. For more complicated geometries, off-diagonal components may appear. They could, however, be eliminated by coordinate rotations.[105]

Finally, the tensor $\mathbf{A} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -0.5 \end{smallmatrix}\right)$, and consequently, the density and compressibility tensors of the CMM are: $\rho^{(c)} = \left(\begin{smallmatrix} -0.5 & 0 \\ 0 & -2 \end{smallmatrix}\right) \times \rho^{(a)}$ and $\chi^{(c)} = -2 \times \chi^{(a)}$. It is noted that a generalized CMM requires strongly anisotropic density as well as negativity for both density and compressibility. In addition, the density can be isotropic if the thickness of the CMM is chosen to be the same as the aberrating layer. However, in this case, the refractive index is -1 and the k vector along the interface goes to infinity.[106] In other words, such a CMM will be very sensitive to the unit cell size and can be difficult to demonstrate. Furthermore, a generalized CMM is preferred in practice as its thickness can be arbitrarily chosen, i.e., it does not depend on the thickness of the aberrating layer, providing a great flexibility.

2.3.4.2. *Unit cell design for CMM*

Periodic cubic blocks with clamped elastic membranes and side branches as shown in Fig. 2.22 are chosen to achieve the double negative and anisotropic properties for CMM. The membrane is introduced here to tune the effective density. The side branches are open ended and are introduced to tune the effective compressibility. In the demonstration, we choose the operating frequency of the CMM to be 50 kHz, at which the wavelength in water (background medium) is 15 times larger than the size of a unit cell. By adjusting the thicknesses of the membranes facing each direction (x- and y-), the effective density can be tuned therein in order to achieve anisotropy. Assuming the interaction and coupling between membranes in the

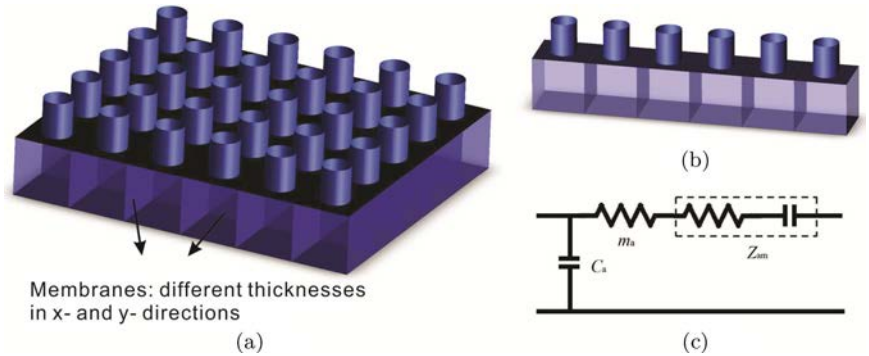


Fig. 2.22. (a) Schematic of a portion of the quasi 2D CMM. (b) Schematic of a 1D side brand and membrane-based metamaterial. (c) Equivalent acoustic circuit of the 1D membrane-based metamaterial.

x- and y- directions is negligible, the effective density and compressibility in either x- or y- direction can be estimated separately by one dimensional (1D) studies.[107] To this end, 1D models are first studied in order to determine the appropriate thicknesses of the membranes and the dimensions of the side branches by both theoretical analysis and numerical simulations (Fig. 2.22(b)).

Theoretically, the effective compressibility with open ended side branches in theory can be written as[108]: $\chi_e = \chi_0(1 - \frac{S}{Ad\rho_0\beta_0 l'\omega^2})$, where β_0 is compressibility of the background medium, S, A, d, l' are, respectively, cross section area of the branch, cross section area of the waveguide (cubic block), length of unit cell, the effective length of the branch, and is the angular frequency. On the other hand, the effective density of the unit cell depends on the properties of the membranes.[108–111] The side branches are assumed to have negligible effect on the effective density.[108] The effective density with clamped membranes can be derived by using the lumped model, with the equivalent acoustic circuit of the 1D membrane-based metamaterial shown in Fig. 2.22(c). Here, $m_a = \frac{\rho_0}{A(d-h)}$ is the effective acoustic mass of the tube, $C_a = A\chi_0(d-h)$ is the acoustic capacitance of the waveguide, where h is the membrane thickness. Z_{am} is the acoustic impedance of the membrane and can be approximated by an inductor

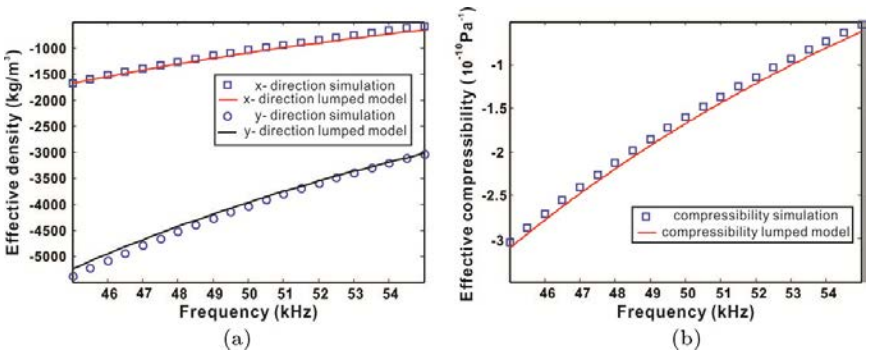


Fig. 2.23. Effective densities and compressibility extracted from full wave simulations and predicted by the lumped model.[104] (a) Effective densities in x- and y-directions. (b) Effective compressibility. At 50 kHz, the desired density and compressibility are achieved.

and capacitor in series in the low-frequency region,[109] and defined as:

$$Z_{am} = \frac{Z_m}{A^2} = \frac{\iint \Delta p dA}{j\omega\xi A^2}, \tag{2.64}$$

where $\Delta p = p_1 - p_2$ is the pressure difference across the membrane, Z_m is the mechanical impedance of the membrane and ξ represents the average displacement of the membrane. Therefore, the expression for the total acoustic impedance of the tube is: $Z_{as} = j\omega m_a + Z_{am}$, with the effective density of a unit cell written as: $\rho_e = \frac{Z_{as}}{j\omega}\frac{1}{dA}$.

The effective acoustic parameters calculated by theoretical prediction are compared with the numerical simulation by COMSOL Multiphysics, which are shown in Fig. 2.23. We can observe a good agreement. We need to address that negative properties of these structures are not relying on the resonance, since the negative properties appear in a broad-band frequency range. The material properties and geometric parameters in the simulation are listed: the membrane made by aluminum film is 2 mm × 2 mm; the Young's modulus is 70 GPa; the Poisson's ratio is 0.33; the density is 2700 kg/m^3; the thicknesses of the membranes are 0.083 mm and 0.11 mm in x and y directions; the radius and length of the side branches are 0.25 mm and 1.25 mm. The density and sound speed of background medium (water) are 1000 kg/m^3 and 1500 m/s.

2.3.4.3. *Numerical simulation*

We conduct two sets of full wave simulation to validate the designed CMM by using COMSOL Multiphysics. The acoustic-solid interaction module is chosen for numerical simulations. The entire CMM consists of 120×10 unit cells: there are 120 units along x-direction and 10 units along y-direction. Perfectly matched layers (PML) are used to minimize reflections from the boundary. The thickness of the aberrating layer is 40 mm, with the material density 2000 kg/m^3 and sound speed 2500 m/s, for mimicking human skulls. The acoustic impedance of aberrating layer is therefore over 3.3 times larger than the background medium, providing a sufficient amount of mismatch. According to the effective material property calculation, the effective density of the unit cells are -1000 kg/m^3 (along x-axis) and -4000 kg/m^3 (along y-axis), as well as the effective compressibility is -1.6×10^{-10} m^2/N, which meet the requirement of the coordinate transformation.

In the two simulations, we configure a curved/focused array and a linear array to generate different acoustic fields. In the first case, the curved acoustic source array which can generate a focal point in the free space (shown in Fig. 2.24(b) as a control case) is placed in front of the aberrating layer. It is straightforward that the aberrating layer can block most of the transmission acoustic energy due to the large impedance mismatch, as shown in Fig. 2.24(a). By placing the CMM layer on the same side with the array, the CMM can effectively cancel out the aberrating layer, so a strong focal point which is 60 mm (total thickness of the CMM and the aberrating layer) behind the original one can be observed. The curved source array is moved 60 mm away from the aberrating layer when the CMM is inserted, so that the focal point remains at the desired location, as shown in Fig. 2.24(c). Quantitative analysis shows that in the curved array case, the intensity amplitude at the focal point is 88% of that of the control case, while it is 28% without the CMM. A significant improvement in terms of the sound transmission ($> 300\%$) is achieved. As indicated in the white dash line in Fig. 2.24(c), the focal point of the CMM case results in an accurate position, which is only 3 mm (1/10 of

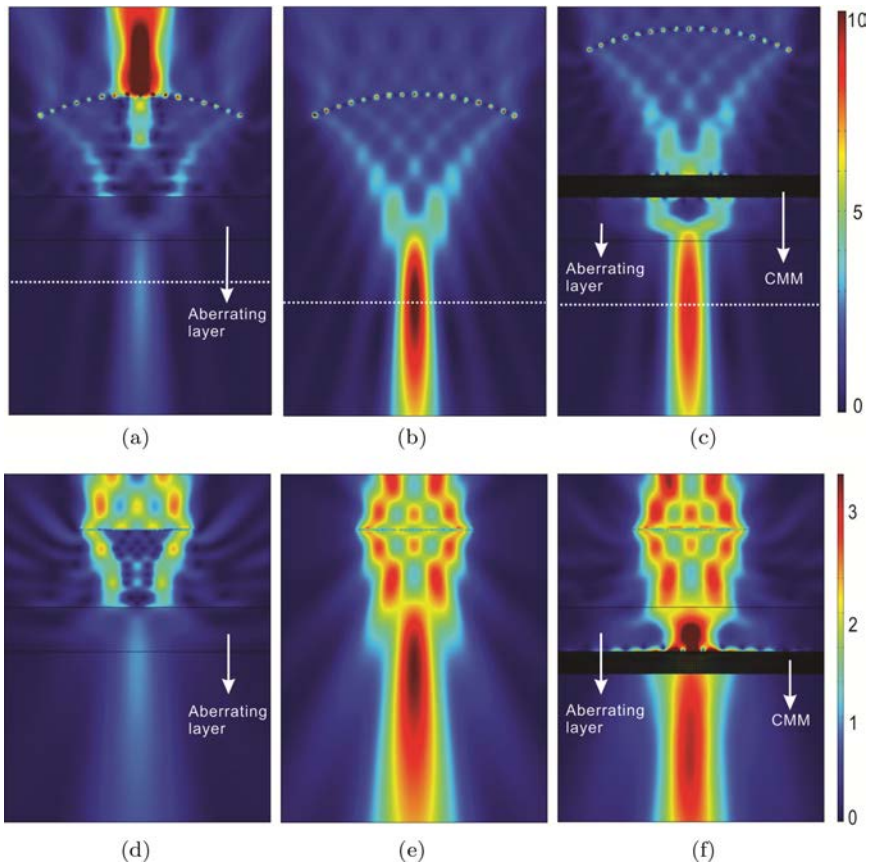


Fig. 2.24. (a–c) Acoustic intensity field for a curved array (focused beam). The CMM is placed in front of the aberrating layer. White dotted lines indicate the position of the focal plane. Three cases are presented: (a) Skull only. (b) Homogeneous medium. (c) With CMM and skull. (d–f) Acoustic intensity field for a linear array (unfocused beam). The CMM is placed behind the aberrating layer. Three cases are presented: (d) Skull only. (e) Homogeneous medium. (f) With CMM and skull.[104]

the wavelength) off the desired location as in the control case, while it is 14 mm (about 1/2 of the wavelength) off without the CMM. This example demonstrates one potential application of CMMs to focus ultrasound behind aberrating layers, which could be extremely useful for improving ultrasound imaging or therapy. In the linear

array case (unfocused acoustic field), which we place the CMM layer behind the aberrating layer, the total acoustic intensity is 97% of that of the control case, whereas the transmission is only 31% without the CMM. The numerical simulation results are plotted in Fig. 2.24(d–f). This example demonstrates another potential application of CMMs to detect passive acoustic source emissions or reflections from an object to be imaged/detected behind aberrating layers. Also, interestingly, this example indicates that the CMM can be placed behind a reflective layer and yet achieve anti-reflection, thus has the advantage of being virtually concealed.

In both cases, the effective density and compressibility are extracted from the simulation results and are very close to the ones predicted by the earlier 1D model. It is noted that in ideal case, the CMM should exactly restore the sound field without considering the energy loss inside the metamaterial. However, it would require an infinitely small unit cell so that the homogenization is perfect. In the current simulation, the material losses are not taken into account in the simulations, since aluminum (membrane material) is known to have a small loss factor and the attenuation in water is also negligible. In addition, the side branch and membrane-based metamaterials used here are non-resonant metamaterials,[108] therefore the material loss is not expected to be a significant factor.

A single CMM layer can significantly enhance the acoustic energy transmission through the aberrating layer. However, the phase of the acoustic field cannot be recovered, which induces the focal point shift in the curved source array case. Similar to the electromagnetic CMMs, an additional restoring layer added on top of the CMM layer can recover both amplitude and phase information of the acoustic field to achieve an acoustic cloaking. The same mathematics of coordinate transformation is applied to retrieve the effective parameters for layer 1 and 2. The effective density and compressibility tensors for layer 2 are the same as the previous result, since layer 2 plays the same function to cancel out the aberrating layer. Layer 1 will restore the information of layer 4 (combine with layer 1, 2, and 3), which has the acoustic properties of the background medium.

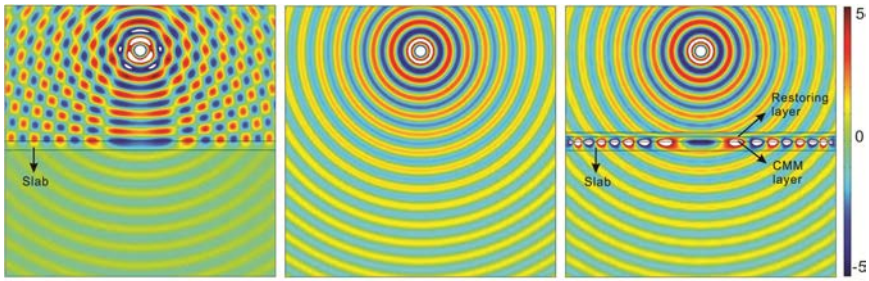


Fig. 2.25. (a–c) Acoustic pressure filed for a point source which is blocked by a slab. An acoustic cloaking device is placed on top of the slab so that it can effectively cloak the slab.[104] (a) Slab only. (b) Homogeneous medium. (c) With cloaking device adjacent to the slab. The wave fields in (b) and (c) are almost identical.

Therefore, the effective density and compressibility tensors for layer 1 are $\rho^{(1)} = \begin{pmatrix} 0.25 & 0 \\ 0 & 4 \end{pmatrix} \times \rho_0$ and $\chi^{(1)} = 4 \times \chi_0$. We conduct the numerical simulation by using the homogenized effective media with desired properties, and plot them in Fig. 2.25. The simulation results show that a slab is effectively cloaked by the cloaking device with a CMM layer and a restoring layer.

In conclusion, we design acoustic CMMs based on the coordinate transformation of acoustic waves to cancel out an object in free space. Numerical examples demonstrate that the CMMs are able to restore the acoustic fields distorted by aberrating layers, which is capable of virtually removing an aberrating layer in a non-invasive manner. Therefore, it could greatly facilitate NDE, transcranial ultrasound imaging and treatment. In addition, the method for designing acoustic CMMs can be readily used to cancel out multiple layers if needed. CMMs are also expected to be useful for the design of acoustic cloaking and all angle anti-reflection materials.[112,113]

Bibliography

[1] Liu, Z., *et al.* (2000). *Science* **289**, p. 1734.
[2] Li, J. and Chan, C. T. (2004). *Phys. Rev. E* **70**, p. 055602.

[3] Fang, N., Xi, D., Xu, J., Ambati, M., Srituravanich, W., Sun, C. and Zhang, X. (2006). *Nature Mater.* **5**, pp. 452–456.
[4] Cheng, Y., *et al.* (2008). *Appl. Phys. Lett.* **92**, p. 051913.
[5] Mei, J., *et al.* (2006). *Phys. Rev. Lett.* **96**, p. 024301.
[6] Ding, Y. Q., *et al.* (2007). *Phys. Rev. Lett.* **99**, p. 093904.
[7] Hu, X. H. and Chan, C. T. (2005). *Phys. Rev. E* **71**, p. 055601.
[8] Sheng, P. (1995). *Introduction to Wave Scattering, Localization and Mesoscopic Phenomena*, Academic, New York.
[9] Ye, Z. and Hsu, H. (2001). *Appl. Phys. Lett.* **79**, p. 1724.
[10] Wang, Z. G., *et al.* (2008). *J. Appl. Phys.* **103**, p. 064907.
[11] Cheng, Y., *et al.* (2008). *Phys. Rev. B* **77**, p. 045134.
[12] Lai, A., *et al.* (2004). *Microwave. Mag. IEEE* **5**, p. 34.
[13] Zhang, S., Yin, L. and Fang, N. X. (2009). *Phys. Rev. Lett.* **102**, p. 194301.
[14] Fokin, V., *et al.* (2007). *Phys. Rev. B* **76**, p. 144302.
[15] Hou, Z., *et al.*, (2005). *Phys. Rev. E* **71**, p. 037604.
[16] Pendry, J. B., Schurig, D. and Smith, D. R. (2006). *Science* **312**, p. 1780.
[17] Leonhardt, U. (2006). *Science* **312**, p. 1777.
[18] Kildishev, A. V. and Shalaev, V. M. (2008). *Opt. Lett.* **33**, p. 43.
[19] Rahm, M., Cummer, S. A., Schurig, D., Pendry, J. B. and Smith, D. R. (2008). *Phys. Rev. Lett.* **100**, p. 063903.
[20] Schurig, D. *et al.* (2006). *Science* **314**, pp. 977–980.
[21] Rahm, M. *et al.* (2008). *Opt. Express* **16**, p. 11555.
[22] Smith, D. R., Padilla, W. J., Vier, D. C., Nemat-Nasser, S. C. and Schultz, S. (2000). *Phys. Rev. Lett.* **84**, p. 4184.
[23] Smith, D. R., Pendry, J. B. and Wiltshire, M. C. K. (2004). *Science* **305**, pp. 788–792.
[24] Nemati, N., Kumar, A., Lafarge, D. and Fang, N. X. (2015). *Comptes Rendus*, doi:10.1016/j.crme.2015.05.001.
[25] Maznev, A., Gu, G., Sun, S., Xu, J., Shen, Y., Fang, N. and Zhang, S. (2015). *New Journal of Physics* **17**(4), p. 042001.
[26] Pendry, J. B., Holden, A. J., Robbins, D. J. and Stewart, W. J. (1999). *IEEE Trans. Microw. Theory Tech.* **47**, pp. 2075–2084.
[27] Iyer, A. K., Kremer, P. C. and Eleftheriades, G. V. (2003). *Opt. Express* **11**, pp. 696–708.
[28] Lafarge, D. and Nemati, N. (2013). *Wave Motion* **50**, pp. 1016–1035.
[29] Zwikker, C. and Kosten, C. W. *Sound Absorbing Materials*, Elsevier Publishing company, Inc., New York (1949). Reprinted 2012 by the NAG (Nederlands Akoestisch Genootschap).
[30] Johnson, D. L., Koplik, J. and Dashen, R. (1987). *J. Fluid Mech.* **176**, pp. 379–402.
[31] Champoux, Y. and Allard, J. F. (1991). *J. Appl. Phys.* **70**, pp. 1975–1979.
[32] Lafarge, D., Lemarinier, P., Allard, J. F. and Tarnow, V. (1997). *J. Acoust. Soc. Am.* **102**, pp. 1995–2006.
[33] Burridge, R. and Keller, J. B. (1981). *J. Acoust. Soc. Am.* **70**, pp. 1140–1146.

[34] Norris, A. N. (1986). *J. Wave Mat. Interact.* **1**, pp. 365–380.
[35] Zhou, M. Y. and Sheng, P. (1989). *Phys. Rev. B* **39**, pp. 12027–12039.
[36] Smeulders, D. M. J., Eggels, R. L. G. M. and van Dongen, M. E. H. (1992). *J. Fluid Mech.* **245**, pp. 211–227.
[37] Auriault, J. L. (1980). *Int. J. Engng Sci.* **18**, pp. 775–785.
[38] Auriault, J. L., Boutin, C. and Geindreau, C., *Homogenization of Coupled Phenomena in Heterogenous Media*, ISTE and Wiley (2009).
[39] Sanchez Palencia, E., *Nonhomogeneous Media and Vibration Theory*, Lectures notes in Physics, 127, Springer, Berlin (1980).
[40] Bensoussan, A., Lions, J. L. and Papanicolaou, G. C., *Asymptotic Analysis for Periodic Structure*, North-Holland, Amsterdam (1978).
[41] Craster, R. V., Kaplunov, J. and Pichugin, A. V. (2010). *Proc. R. Soc. A* **466**, pp. 2341–2362.
[42] Antonakakis, T., Craster, R. V., Guenneau, S. and Skelton, E. A. (2013). *Proc. R. Soc. A* **470**, p. 20130467.
[43] Boutin, C., Rallu, A. and Hans, S. (2012). *J. Acoust. Soc. Am.* **132**, pp. 3622–3636.
[44] Boutin, C. (2013). *J. Acoust. Soc. Am.* **134**, pp. 4717–4729.
[45] Yang, M., Ma, G., Wu, Y., Yang, Z. and Sheng, P. (2014). *Phys. Rev. B* **89**, p. 064309.
[46] Wu, Y., Lai, Y. and Zhang, Z. Q. (2007). *Phys. Rev. B* **76**, p. 205313.
[47] Boutin, C. (2012). *Transp. Porous Med.* **93**, pp. 309–329.
[48] Willis, J. R. (2009). *Mech. Mater.* **41**, pp. 385–393.
[49] Landau, L. D. and Lifshitz, E. M., *Electrodynamics of Continuous Media*, Course of Theoretical Physics Vol. 8 Elsevier, Oxford/Butterworth-Heinemann (2004).
[50] Zhang, S., Xia, C. and Fang, N. (2011). *Phys. Rev. Lett.* **106**, p. 24301.
[51] Nemati, N. and Lafarge, D. (2014). *Wave Motion* **51**, pp. 716–728.
[52] Allard, J. F. and Atalla, N., *Propagation of Sound in Porous Media: Modelling Sound Absorbing Materials*, Second Edition, John Wiley & Sons (2009).
[53] Hecht, F. (2012). New development in FreeFem++, *J. Numer. Math.* **20**, pp. 251–265.
[54] Veselago, V. G. (1968). *Sov. Phys. Usp.* **10**(4), p. 509.
[55] Pendry, J. B. (2000). *Phys. Rev. Lett.* **85**, p. 3966.
[56] Fang, N., Lee, H., Sun, C. and Zhang, X. (2005). *Science* **308**, p. 534.
[57] Grbic, A. and Eleftheriades, G. V. (2004). *Phys. Rev. Lett.* **92**, p. 117403.
[58] Kinsler, L. E., *Fundamentals of Acoustics*, Wiley, New York (1982).
[59] Stewart, G. W. (1922). *Phys. Rev.* **20**, p. 528.
[60] Mason, W. P. (1927). *Bell Syst. Tech. J.* **6**, p. 258.
[61] Eleftheriades, G. V., Iyer, A. K. and Kremer, P. C. (2002). *IEEE Trans. Microwave Theory Tech.* **50**, p. 2702.
[62] Beranek, L. L., *Acoustics*, McGraw-Hill, New York (1954).
[63] Lemoult, F., Fink, M. and Lerosey, G. (2011). *Phys. Rev. Lett.* **107**, p. 064301.

[64] Bruneau, M. (2006). *Fundamentals of Acoustics* (London: ISTE).
[65] Ward, A. J. and Pendry, J. B. (1996). *Journal of Modern Optics* **43**(4), pp. 773–793.
[66] Stenger, N., Wilhelm, M. and Wegener, M. (2012). *Phys. Rev. Lett.* **108**, p. 014301.
[67] Milton, G. W., Briane, M. and Willis, J. R. (2006). *New Journal of Physics* **8**(10), p. 248.
[68] Cummer, S. A. and Schurig, D. (2007). *New Journal of Physics* **9**(3), p. 45.
[69] Chen, H. and Chen, C. T. (2007). *Applied Physics Letter* **91**(8).
[70] Li, J. and Pendry, J. B. (2008). *Phys. Rev. Lett.* **101**, p. 203901.
[71] Ergin, T., Stenger, N., Brenner, P., Pendry, J. B. and Wegener, M. (2010). *Science* **328**, pp. 337–339.
[72] Chen, X, *et al.* (2011). *Nat. Commun.* **2**, p. 176.
[73] Landy, N. and Smith, D. R. (2013). *Nature Mater.* **12**, pp. 25–28.
[74] Farhat, M., Enoch, S., Guenneau, S. and Movchan, A. B. (2008). *Phys. Rev. Lett.* **101**, p. 134501.
[75] Popa, B.-I., Zigoneanu, L. and Cummer, S. A. (2011). *Phys. Rev. Lett.* **106**, p. 253901.
[76] Brun, M., Guenneau, S. and Movchan, A. B. (2009). *Appl. Phys. Lett.* **94**, p. 061903.
[77] Farhat M., Guenneau, S. and Enoch, S. (2009). *Phys. Rev. Lett.* **103**, p. 024301.
[78] Guenneau, S., Amra, C. and Veynante, D. (2012). *Opt. Express* **20**, pp. 8207–8218.
[79] Narayana, S. and Sato, Y. (2012). *Phys. Rev. Lett.* **108**, p. 214303.
[80] Schittny, R., Kadic, M., Guenneau, S. and Wegener, M. (2013). *Phys. Rev. Lett.* **110**, p. 195901.
[81] Xu, J., Jiang, X., Fang, N., Georget, E., Abdeddaim, R., Geffrin, J.-M., Sabouroux, P., Enoch, S. and Guenneau, S. (2015). *Scientific Reports* **5**, p. 10678.
[82] Cherkaev, A. (2000). *Variational Methods for Structural Optimization*, New York: Springer-Verlag.
[83] Milton, G. W., *The Theory of Composites*, Cambridge: Cambridge University Press (2001).
[84] Acheson, D. J., *Elementary Fluid Dynamics*, Oxford University Press, Oxford (1990).
[85] Cai, W., Chettiar, U. K., Kildiev, A. V. and Shalaev, V. M. (2007). *Nat. Photon.* **1**, p. 224.
[86] Georget, E., Abdeddaim, R. and Sabouroux, P. (2014). *IEEE Trans. Antennas Propag.* **56**, pp. 240–248.
[87] Knott, E. F., Shaeffer, J. F. and Tuley, M. T., *Radar Cross Section* (Raleigh: SciTech Publishing Inc, 2nd edition) (2004).
[88] Cummer, S., Popa, B.-I., Schurig, D., Smith, D., Pendry, J., Rahm, M. and Starr, A. (2008). *Phys. Rev. Lett.* **100**, p. 024301.

[89] Smith, S. W., Trahey, G. E. and von Ramm, O. T. (1986). *Ultrasound Med. Biol.* **12**, p. 229.
[90] Clement, G. T., White, P. J. and Hynynen, K. (2004). *Acoust. Soc. Am.* **115**, p. 1356.
[91] Ivancevich, N. M., Pinton, G. F., Nicoletto, H. A., Bennett, E., Laskowitz, D. T. and Smith, S. W. (2008). *Ultrasound Med. Biol.* **34**, p. 1387.
[92] Win, K. K., Wang, J. and Zhang, C. (2010). *5th IEEE Conf. Ind. Electron. Appl.* 1675.
[93] Demirli, R., Amin, M. G., Shen, X. and Zhang, Y. D. (2012). *Adv. Acoust. Vib.* 1.
[94] Jing, Y., Meral, F. C. and Clement, G. T. (2012). *Phys. Med. Biol.* **57**, p. 901.
[95] Wang, T. and Jing, Y. (2013). *Phys. Med. Biol.* **58**, p. 6663.
[96] Clement, G. T. and Hynynen, K. (2002). *Phys. Med. Biol.* **47**, p. 1219.
[97] Pernot, M., Aubry, J.-F., Tanter, M., Boch, A.-L., Marquet, F., Kujas, M., Seilhean, D. and Fink, M. (2007). *J. Neurosurg.* **106**, p. 1061.
[98] Pendry, J. B. and Ramakrishna, S. A. (2003). *J. Phys. Condens. Matter* **15**, p. 6345.
[99] Lai, Y., Chen, H., Zhang, Z.-Q. and Chan, C. (2009). *Phys. Rev. Lett.* **102**, p. 093901.
[100] Lai, Y., Ng, J., Chen, H., Han, D., Xiao, J., Zhang, Z.-Q. and Chan, C. (2009). *Phys. Rev. Lett.* **102**, p. 253902.
[101] Zhu, X., Liang, B., Kan, W., Zou, X. and Cheng, J. (2011). *Phys. Rev. Lett.* **106**, p. 014301.
[102] Liu, B. and Huang, J. P. (2009). *Eur. Phys. J. Appl. Phys.* **48**, p. 20501.
[103] Yang, J., Huang, M., Yang, C., Peng, J. and Chang, J. (2010). *Comput. Mater. Sci.* **49**, p. 9.
[104] Shen, C., Xu, J., Fang, N. and Jing, Y. (2014). *Phys. Rev. X* **4**, p. 041033.
[105] Popa, B.-I. and Cummer, S. A. (2011). *Phys. Rev. B* **83**, p. 224304.
[106] Ambati, M., Fang, N., Sun, C. and Zhang, X. (2007). *Phys. Rev. B* **75**, p. 195447.
[107] Park, C. M., Park, J. J., Lee, S. H., Seo, Y. M., Kim, C. K. and Lee, S. H. (2011). *Phys. Rev. Lett.* **107**, p. 194301.
[108] Lee, S. H., Park, C. M., Seo, Y. M., Wang, Z. G. and Kim, C. K. (2010). *Phys. Rev. Lett.* **104**, p. 054301.
[109] Bongard, F., Lissek, H. and Mosig, J. R. (2010). *Phys. Rev. B* **82**, p. 094306.
[110] Lee, S. H., Park, C. M., Seo, Y. M., Wang, Z. G. and Kim, C. K. (2009). *Phys. Lett. A* **373**, p. 4464.
[111] Jing, Y., Xu, J. and Fang, N. X. (2012). *Phys. Lett. A* **376**, p. 2834.
[112] Torrent, D. and Sánchez-Dehesa, J. (2007). *New J. Phys.* **9**, p. 323.
[113] Wang, Y., Deng, K., Xu, S., Qiu, C., Yang, H. and Liu, Z. (2011). *Phys. Lett. A* **375**, p. 1348.

CHAPTER 3

Flat Lens Focusing of Flexural Waves in Thin Plate

PATRICK SEBBAH*

Institut Langevin, France and University of Bar Ilan, Israel

MARC DUBOIS

University of California Berkeley, USA

3.1. Introduction

This chapter will focus on negative refraction concept and materials with negative index of refraction. Negative index materials for electromagnetic waves were first considered in 1968 by Victor Veselago [Veselago (1968)]. His statement is as follows, if a medium presents simultaneous negative electric permittivity and magnetic permeability then the index of refraction is negative. Several applications of negative index media have been discussed in this seminal paper. Among them, the ability to create a lens from a parallel flat slab of negative index medium will later become the most stimulating perspective. In 2000, John Pendry demonstrate that such a slab of negative index material can beat the diffraction limit and focus waves on a scale smaller than the wavelength [Pendry (2000)].

*Corresponding author: patrick.sebbah@espci.fr

This superfocusing property has triggered an important number of theoretical, numerical and experimental studies. In spite of a potential breakthrough for many lensing application such as imaging or lithography, negative index material have not been observed in nature yet. Two main approaches have been considered in order to achieve negative refraction experimentally: crystalline structures and metamaterials. In this chapter, we will focus on the realization of negative refraction in periodic structures. The first part of the chapter will introduce the concept of negative refraction in periodic structure in a general framework. In a second part, we will see how these notions can be applied to elastodynamics and present the first experimental demonstration of flat lensing for flexural waves in thin plates. In a last part, we will explore the dynamic behavior of such flat lens when excited with broadband source signals.

3.2. Negative Refraction in Periodic Structures

Wave propagation in periodic structures has been widely studied over the last century. Bloch theorem was first used to describe electronic wave function in crystals. It has now been extended to a broad class of waves including electromagnetic, acoustic, elastic or water waves. As an example, the presence of band gap is a universal hallmark of periodic structures. This phenomenon, related to destructive interference, results in a frequency band where waves cannot propagate in a periodic medium. The following will present the theoretical tools and methods used to explore the dispersion properties of periodic structures.

3.2.1. *Waves in periodic structures*

Crystallography and solid state physics methods are used to obtain properties of periodic structures. Periodic structures are described by a direct Bravais lattice. This lattice is an infinite array of points obtained by discrete translations represented by vectors named primitive vectors a_i, $i = [1, 2, 3]$ in three dimensional lattices. A reciprocal lattice is associated to this direct lattice, its primitive vectors b_i, $i = [1, 2, 3]$ are defined by $a_i.b_j = 2\pi\delta_{ij}$. The unit cell of the reciprocal lattice is defined as the first Brillouin zone. This unit cell

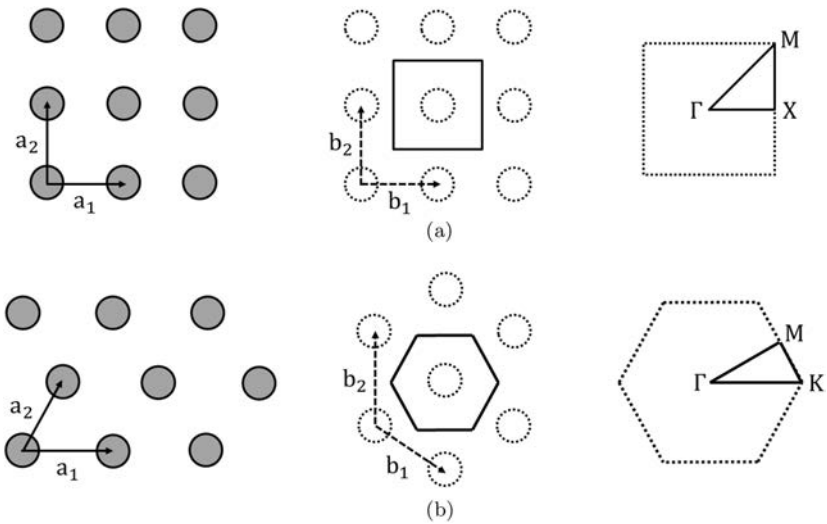


Fig. 3.1. Schematics of direct lattice, reciprocal lattice and first Brillouin zone. Irreducible Brillouin zone and critical points are presented for two dimensional (a) square lattice and (b) triangular lattice.

can be reduced to the irreducible Brillouin zone delimited by symmetry centers known as critical points. Γ represents the origin of the irreducible Brillouin zone in any periodic structure. Figure 3.1 presents the direct lattice, the reciprocal lattice, and the irreducible Brillouin zone for the two dimensional square and triangular lattices.

Bloch's theorem was initially used to describe electronic wave functions in periodic potential [Kittel (2007)]. In a periodic potential with period r_0 wave functions $\Psi(r)$ associated with a wave vector k are Bloch waves

$$\Psi(r) = e^{ik.r} u_k(r)$$

with u_k a periodic function with period r_0. The variable couple (k, u_k) is not unique but we can always write this wave vector as

$$k = k_0 + \sum_i n_i.b_i$$

consequently k_0 is kept inside the irreducible Brillouin zone. Thus, we can calculate the wave function associated with all wave vectors by calculating the response of the crystal only for wave vectors

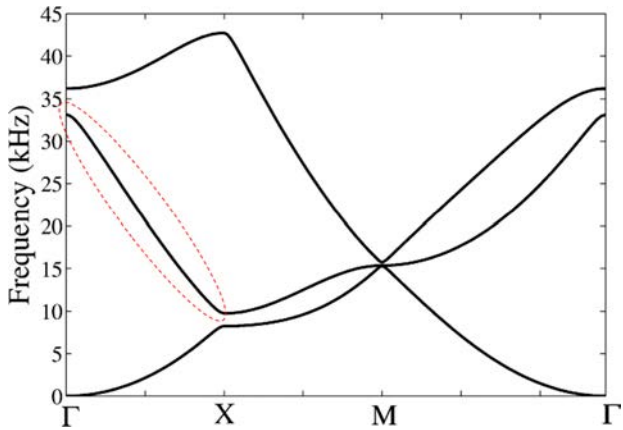


Fig. 3.2. Dispersion diagram of a square lattice phononic crystal for flexural waves in a thin plate. The red ellipse denotes a band with a negative derivative as the frequency increases and wave vector decreases.

bounded in the irreducible Brillouin zone. It is common to represent the relation that links the energy of the wave functions to the wave vectors on a diagram called the band diagram of the crystal.

As mentioned before, Bloch's theorem can be applied to determine the propagation properties of electromagnetic or acoustic wave in periodic structure. Such crystals for light and sound are no longer semiconductor crystals. They result in a periodic modulation of the index of refraction or speed of sound, leading to the so called photonic crystals [Yablonovitch (1987)] and phononic crystals [Kushwaha *et al.* (1993)]. The band diagram now represents the dispersion relation of the structure that links pulsation ω and wave vector k. An example of such a band diagram is presented in Figure 3.2. It is calculated for flexural waves propagating in a square lattice of circular holes drilled in a thin plate. The dispersion curve is plotted on the three segments of the irreducible Brillouin zone, namely ΓX, ΓM and XM for a square lattice.

3.2.2. *Negative refraction at an interface*

Snell's law dictates the refraction of a plane wave at an interface between two materials. Accordingly, the transverse component of the

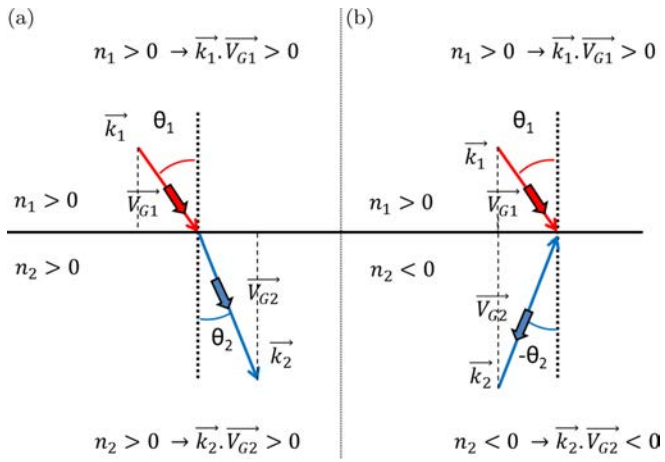


Fig. 3.3. Ray tracing according to Snell's law. Refraction of incident wave vector $\vec{k_1}$ at the interface between two medium with index of refraction n_1 and n_2. (a) Case with both positive index. (b) Case where n_1 is positive and n_2 negative.

incident wave vector is conserved in the refraction process. As shown in Figure 3.3(a), it is possible to calculate the angle of refraction with the following equation $n_1 sin(\theta_1) = n_2 sin(\theta_2)$. It is to be mentioned that the dot product of group and phase velocity is positive $\vec{V_g}.\vec{k} > 0$ in any positive index medium. Snell's law is still valid if one of a medium presents a negative index of refraction. Figure 3.3(b) presents the ray tracing obtained for $n_2 < 0$. Victor Veselago demonstrate that phase velocity and group velocity have to be opposite such has $\vec{V_{g2}}.\vec{k_2} < 0$ [Veselago (1968)]. From causality relations, group velocity $\vec{V_{g2}}$ has to be directed away form the interface. Therefore, the phase velocity must be directed in the opposite direction.

Phase and group velocities in periodic medium can be calculated from the dispersion curves presented in Figure 3.2. Phase velocity V_ϕ is given by

$$V_\phi = \frac{\omega}{k} \tag{3.1}$$

and group velocity V_g is defined as

$$V_g = \frac{\delta\omega}{\delta k} \tag{3.2}$$

In Figure 3.2, the second band of ΓX segment, denoted by the red dotted ellipse, presents a negative slope as wave vector decreases when frequency increases. This negative slope leads the dot product $\vec{V_g}.\vec{k}$ to be negative and as a consequence the effective index of refraction of the crystal can be considered negative in this band. This property arises from band folding that occurs for frequency above the first band gap in photonic or phononic crystal. It has been used to demonstrate negative refraction of a single wave vector in electromagnetics [Cubukcu *et al.* (2003); Berrier *et al.* (2004)], in acoustics [Yang *et al.* (2004)] and for elastic waves in solids [Pierre *et al.* (2010)].

3.3. Flat Lensing with Periodic Structures

The previous section depicted how negative refraction can be observed for a plane wave incident on a periodic structure. Negative index flat lenses are build to image a point source. In order to form the image of a point source, one needs to design a medium where any angle of incidence presents a negative refraction. This property, called All Angle Negative Refraction (AANR), requires specific attention to the crystal design.

3.3.1. *Snell's law and equifrequency surfaces*

Equifrequency surfaces are used in order to represent the dispersion of a medium in function of the direction of propagation. In two dimensional structure, the dispersion relation can be represented in the space (k_x, k_y, ω). Equifrequency surface or equifrequency contour represent the spatial dispersion at a constant pulsation. In the case of homogeneous and isotropic media, equifrequency contours are concentric circles which radius are proportional to the index of refraction. This representation becomes essential to study wave refraction in the presence of anisotropy.

Equifrequency surface allow a simple geometrical relation between wave vector and group velocity, $\vec{V_g} = \vec{\nabla}_{\vec{k}}\,\omega(\vec{k})$. The direction of group velocity is given by the outgoing normal of the equifrequency

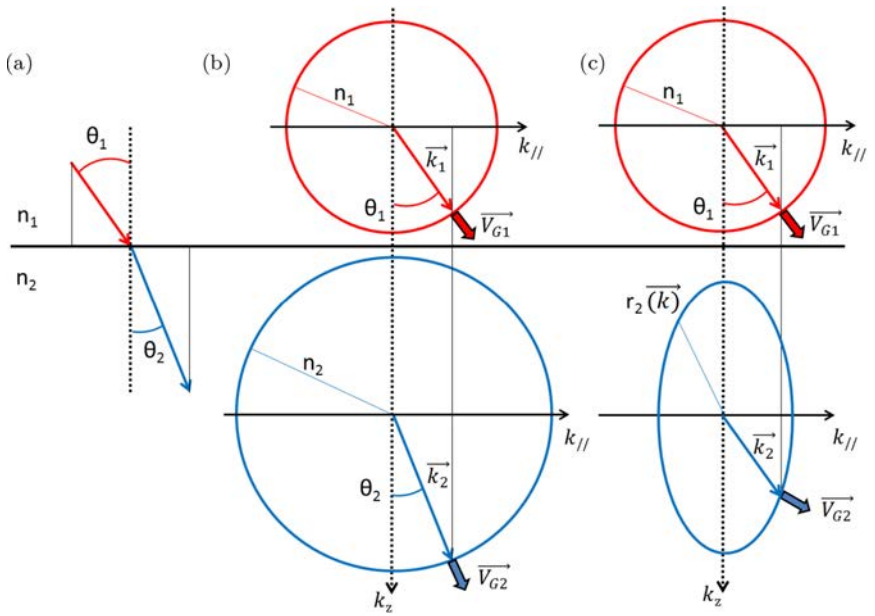


Fig. 3.4. (a) Ray tracing obtained with Snell-Descartes law. (b) Ray tracing obtained from the equifrequency contour method. (c) Refraction obtained in the case where medium 2 is anisotropic (elliptical equifrequency contour).

contour. Figures 3.4(a,b) represent a ray tracing based on Snell's law with the use of equifrequency contours. In this case, both medium are described by an isotropic index of refraction such that equifrequency contours are circle and as a consequence wave vector $\vec{k}$ and group velocity $\vec{V_g}$ are always collinear. In the last case presented in Figure 3.4(c), medium 2 presents a strong anisotropy. One can notice that for different angle of propagation wave vector and group velocity are no longer collinear.

3.3.2. *All angle negative refraction*

Two main strategies have been used in order to achieve AANR in periodic structures. The first one uses the effective negative index of refraction obtained by band folding above the first band gap. Many demonstrations will elect the triangular lattice that reduces the effect of anisotropy and ensure an index of refraction independent from the

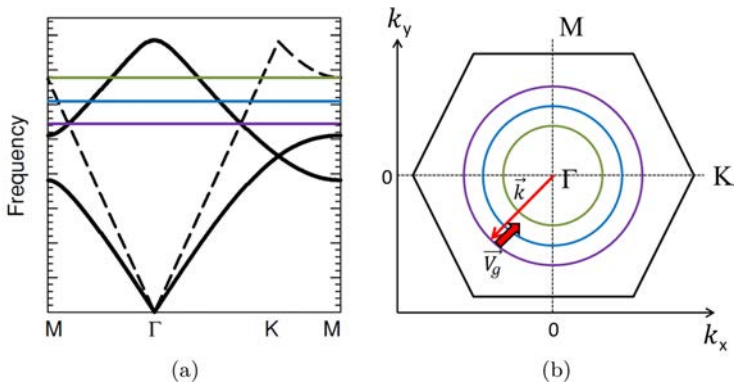


Fig. 3.5. (a) Typical dispersion relation for a two dimensional triangular lattice. The dotted curve represent the dispersion of the homogeneous, isotropic host medium. (b) Equifrequency contour at three different frequencies in the second band, wave vector $\vec{k}$ and group velocity $\vec{V_g}$ are represented.

angle of incidence. Figure 3.5 presents an example of two dimensional triangular lattice dispersion relation and equifrequency contour for three frequencies. As expected, the three equifrequency contours presented in 3.5(b) are perfectly circular. These contours also confirm that the index of refraction can be considered as negative such that the radius of these contours reduces when frequency increase. As a consequence, wave vector and group velocity are aligned but in opposite direction. This confirms the effective index of refraction to be isotropic and negative for these frequencies.

AANR effect has been demonstrated in triangular lattices for photonic crystals at microwave frequency [Lu *et al.* (2005)], infra red frequency [Berrier *et al.* (2004)], and in phononic crystal for ultrasound in water [Sukhovich *et al.* (2008)] and for elastic surface wave [Veres *et al.* (2012)].

3.3.3. *Focusing resolution and index matching*

Veselago's flat lens design presents one more requirement. Magnitude of host medium and negative index slab must be equal. This relation is not always satisfied by the AANR effect in crystals. This index mismatch can be interpreted geometrically with different equifrequency

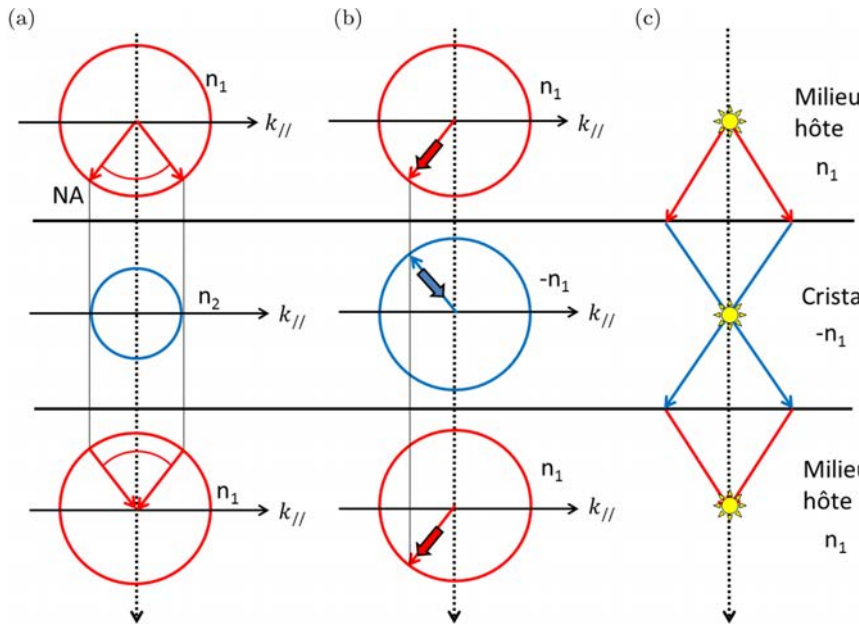


Fig. 3.6. Refraction of a flat slab in different configurations. (a) in the case where $|n_2| < n_1$. (b) if $n_2 = -n_1$. (c) Ray tracing in Veselago's flat lens with the two images inside and outside the lens.

radii between host and crystal medium as depicted in Figure 3.6(a). This schematics shows a slab of negative index n_2 embedded in a host medium with index n_1 such as $|n_2| < n_1$. In that specific case, one can notice that only a portion of incident angles can penetrate the crystal. This effect is perfectly equivalent to a total internal reflection and alters the focusing performance. Figure 3.6(b) presents the case where index are matched $n_2 = -|n_1|$. In this specific condition any incident wave vector is focused and creates two image source, one inside the slab and one on the other side.

The index matching step has been realized experimentally for ultrasonic wave in water [Sukhovich *et al.* (2008)]. A triangular lattice of steal beam was first immersed in water. At this point, the index of the crystal was negative in the AANR frequency range but the magnitude of the index was too small. From the same limitations

discussed in 3.6a, the lateral resolution of the measured focal spot after the slab was 1.15λ. This resolution loss can be interpreted as a reduction of the numerical aperture of the system. In a second experiment, the phononic crystal was again composed of a triangular arrangement of steal beams but immersed in liquid methanol to from the negative index slab [Sukhovich *et al.* (2008)]. Pressure waves are slower in methanol such that the index of refraction of the phononic crystal matches the index of water. The lateral resolution measured improved to 0.55λ demonstrating the importance of index matching.

3.4. Platonic Crystal for Flexural Wave Focusing

3.4.1. *Elastic waves in thin plates*

In three dimensional solids, elastic waves can propagate with two different polarizations, namely longitudinal and transverse. These different polarizations do not couple in homogeneous and isotropic solids. However, the presence of discontinuities in the mechanical parameters will generate interactions between these two modes. Such coupling is at the origin of elastic guided mode or surface mode that are confined at the interface between two media. The most studied surface waves in solids are the Rayleigh waves. This mode is confined at the interface between a solid and a light fluid. Rayleigh waves present an elliptic polarization due to a phase quadrature coupling between longitudinal and vertical transverse waves. The displacement field is confined at the interface and exponentially decays in the solid. A solid medium is considered as a thin plate if one of the dimensions is strongly reduced. In that case, two principal propagation modes exist:

- TH mode corresponds to the in-plane transverse horizontal polarization.
- Lamb modes originate from the coupling of longitudinal and transverse vertical polarizations. The mechanism is similar to Rayleigh wave generation at one interface. However, in the case of thin plates, the two interfaces air-solid and solid-air are separated by a short distance. This configuration induces a coupling between

the two Rayleigh waves of each interface and creates the different Lamb modes.

At low frequency, only two cut-off free Lamb modes can propagate. The S_0 mode or symmetric mode is mostly in-plane and present very small vertical displacement. The A_0 mode or anti-symmetric mode is mainly out of plane and correspond to the flexural waves with important vertical displacement. In the low frequency approximation, Kirchhoff-Love theory of plates can be used to calculate flexural wave propagation [Royer and Dieulesaint (1996)]. Flexural wave propagation is obtained with a fourth order scalar equation on vertical displacement w called bi-harmonic equation (3.3).

$$\frac{D}{\rho h}\Delta^2 w + \frac{\delta^2 w}{\delta t^2} = 0 \tag{3.3}$$

Flexural rigidity D is given by $D = \frac{Eh^3}{12(1-\nu^2)}$ with Young's modulus E, thickness h and Poisson's ratio ν. For harmonic excitation, we can define the following dispersion relation between pulsation ω and wave vector k

$$\omega = \sqrt{\frac{D}{\rho h}}k^2 \tag{3.4}$$

This quadratic behavior at low frequency is typical for the flexural wave in thin plates.

3.4.2. *Negative refraction based on anisotropy*

In the previous sections, we studied how negative effective index of refraction can be obtained in the second band of periodic structures. This paragraph will focus on negative refraction in the first band of highly anisotropic crystals. The anisotropy obtained from circular inclusions in a square lattice can be appreciated in Fig. 3.7. One dimensional arrangement are represented through different angle, normal to the lattice or in diagonal. The filling fraction obtained

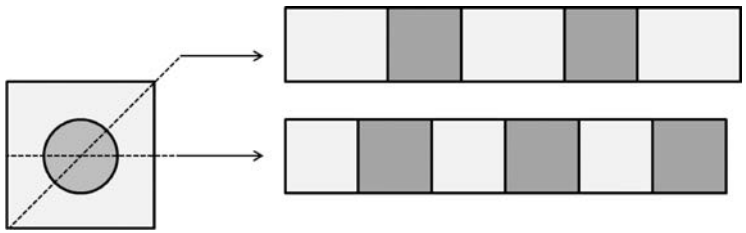


Fig. 3.7. Anisotropy of the square lattice with circular inclusion. Two cuts shows the filling fraction difference.

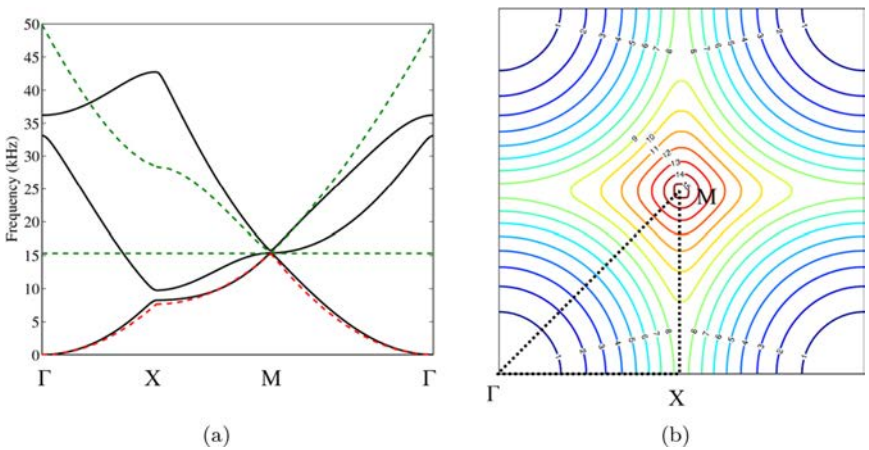


Fig. 3.8. (a) Black solid line, dispersion relation of flexural waves in a square lattice of cylindrical holes in a 1 mm thick duraluminium plate. The lattice constant is 15 mm and hole diameter is 12 mm. This dispersion is obtained by 3D finite element method. Dotted line, beam lattice analytic model valid for the first band calculation (red dotted line). (b) Equifrequency contour in the first band, each line represent a 1 kHz step, the first Brillouin zone is represented in dotted line, from [Dubois *et al.* (2013)].

are 50% for the normal incidence and 35% in the diagonal. As a consequence, the response of the crystal will be anisotropic and band gap frequencies will depend on the direction of propagation.

The Figure 3.8(a) presents the dispersion relation of flexural waves propagating in a square lattice of cylindrical holes in a 1 mm thick duraluminium plate. The lattice constant is 15 mm and hole diameter is 12 mm. Two different methods are used to calculate the

dispersion relation. Finite Element Method is plotted in solid lines, this three dimensional model account for any displacement. The dispersion relation represented in dotted lines is obtained with a beam lattice model where the entire lattice is built from crossing straight beams creating square inclusions. This analytic model is based on Euler-Bernoulli beam theory and accounts for the vertical displacement only. It is valid for the first band of the dispersion relation while wavelength is large compare to the inclusions [Farhat *et al.* (2010); Dubois *et al.* (2013)].

Equifrequency contour are presented in Figure 3.8(b). For frequency below 8 kHz, these contours are circles centered around the point Γ, origin of the Brillouin zone. As the normal to the contour is outgoing, this crystal behaves like a isotropic positive index material. After 8 kHz, there is a bandgap in the direction ΓX. However, because of the lattice anisotropy, the propagation is still allowed along ΓM. Consequently, the equifrequency contour are now wrapped around the M point of the Brillouin zone with a normal pointing toward the M point. This property can be exploited to observe negative refraction. As presented in Figure 3.9, a 45° rotation of the unit cell is required to align the M point with slab normal incidence. For 10 kHz incidence, the group velocity points inside the cell reproducing Veselago's initial design. Negative refraction is obtained but the effective index of refraction remains positive as the relation $\vec{V_g}.\vec{k} > 0$ is satisfied.

This negative refraction effect in the first band of anisotropic crystals has been used in experimental demonstrations for photonic crystals [Cubukcu *et al.* (2003)], phononic crystals [Zhang and Liu (2004)] and for water wave crystals [Hu *et al.* (2004)].

3.4.3. *Index matching with thickness variation*

Flat lens focusing is achievable with this structure, but we can observe from Figure 3.9(b) that the equifrequency contours dimensions are not matched and a small angle portion is collected by the lens. As mentioned in previous section, it is necessary to adapt the index of refraction in order to increase the focusing performance.

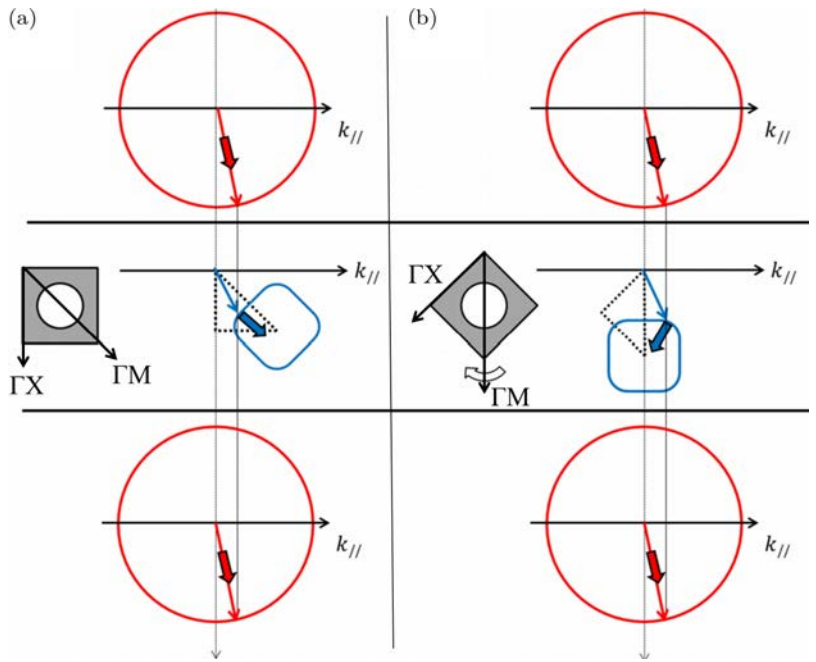


Fig. 3.9. (a) Refraction ray tracing at 10 kHz for normal incidence. Equifrequency contour of the host plain plate and crystal are scaled. (b) Ray tracing obtained at the same frequency once the unit cells of the crystal are rotated by 45°.

The specificity of flexural waves in thin plate allow us to use the thickness of the plate in order to adapt these index of refraction. In the low frequency regime, Kirchhoff-Love theory's provides an analytic expression of the flexural wave dispersion relation (3.5).

$$\omega = \frac{c_P}{\sqrt{12}} h k^2 \tag{3.5}$$

with $c_P = \sqrt{\frac{E}{\rho(1-\nu^2)}}$, E the Young's modulus, ρ the density, h the thickness and ν the Poisson's ratio. This equation can be used in order to define the flexural waves phase velocity (3.6).

$$v_\phi = \frac{\omega}{k} = \sqrt[4]{\frac{E h^2 \omega^2}{12\rho(1-\nu^2)}} \tag{3.6}$$

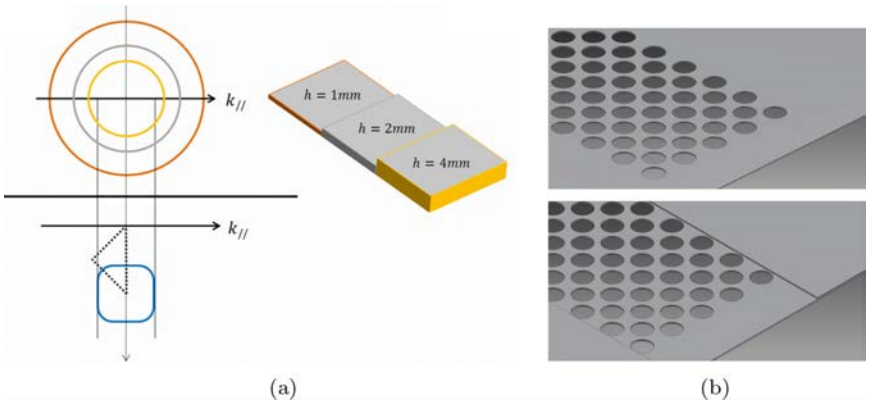


Fig. 3.10. (a) Equifreqncy contour reduction for different thicknesses. Each are taken at 10 kHz frequency. (b) Sketch of the two machined duraluminium plates. *Top*, constant 1 mm thickness. *Bottom*, the thickness profile is 2 mm outside of the crystal and 1 mm inside.

Consequently, increasing the thickness of a plate is equivalent to reduce the size of the equifrequency contour. Figure 3.10(a) presents the evolution of 10 kHz equifrequency contour of the host plate for increasing thickness.

As the equifrequency contour is reduced by increasing thickness, the wavelength of flexural waves is modified accordingly. If the thickness of the host plate is multiply by a factor four, the wavelength is doubled. Large wavelength will be hard to machined as they required larger plates (more than a meter square at 10 kHz). For this reason, we use the plates presented in Figure 3.10, where a 2 mm thick plate is reduced to 1 mm in the crystal area.

3.4.4. *Experimental demonstration of flexural wave focusing*

Duraluminium alloy is chosen as it presents a low loss factor, $\eta = 0.0001$ at kHz frequencies. Two samples are machined in 500 mm $\times$ 300 mm plates using computer aided manufacturing. The phononic crystal represents a surface of 98 mm $\times$ 250 mm. One crystal is drilled in a plate with constant thickness 1 mm. The other

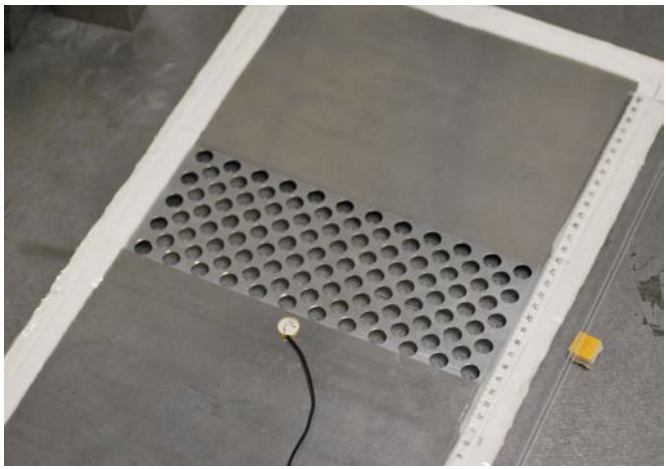


Fig. 3.11. Picture of the backside of the sample with varying thickness. The piezoelectric patch is bound to the surface and used as a point source. Adhesive paste covers the edge of the plates in order to reduce reflections. The sample is supported by felt squares.

sample presented in Figure 3.11 is processed in a 2 mm thick plate that has been milled to 1 mm in the crystal region.

A piezoelectric patch PKS2-4A1 Murata Shock Sensor is used to generate flexural wave in our samples. The patch is bound to the plate with Salol salt. Salol melting point is 41°C, it returns to crystalline phase when cool down. This crystallization is used to create a solid bind between the piezoelectric patch and the plate. The sample edges are covered with 2 cm wide adhesive paste in order to reduce reflection from the boundaries. This reduction is about 3 dB at 10 kHz frequencies. A gaussian pulse $S(t) = sin(2\pi f_0 t)e^{-\frac{(t-t_0)^2}{2\sigma^2}}$ centered at frequency f_0 with standard deviation σ is numerically produced and transmitted to the piezoelectric patch with a digital-to-analog converter Agilent U2542A (500kS/s). The analog input is amplified by a large band amplificator ProSystem US300 and sent to the piezoelectric element. Flexural waves are imaged with a Laser Doppler vibrometer OFV505 Polytec. A continuous monochromatic Laser is shined onto the sample. When the sample presents vertical motion, the reflected light frequency will shift due to Doppler effect.

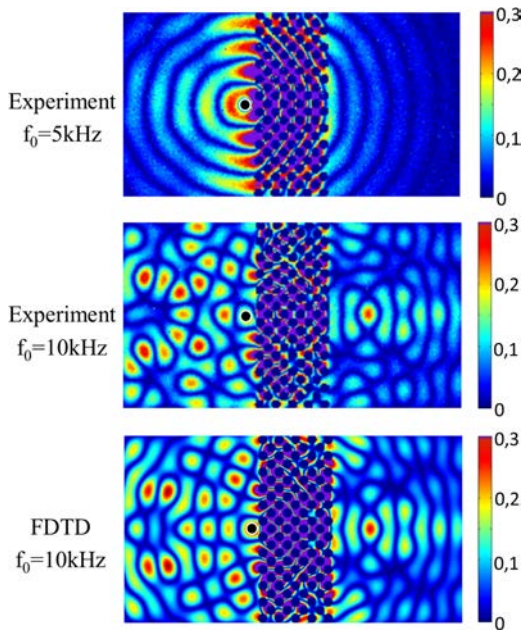


Fig. 3.12. Snapshots of amplitude of normal displacement velocity at t = 780 μs. Each map is measured or simulated with the thickness modified plate. The first map is measured for a pulse centered at 5 kHz and the two lasts are for pulse centered at 10 kHz. Black dots depict the the position of the point source.

A Mach-Zender interferometer is used to measure the Doppler shift and thus extract the vertical displacement at the observation point. This device uses a single photodiode allowing measurement of time dependent vertical displacement upto several hundred kHz. A point by point scan of the vertical displacement velocity is realized over the sample. The scan covers an area of 440 mm $\times$ 260 mm with 2 mm spatial sampling. Ten measurements are averaged on each point in order to increase the signal to noise ratio.

Figure 3.12 presents snapshots of the amplitude of normal displacement velocity at t = 780 μs for the sample where the thickness of the plate has been modified. The top map gives experimental results for a gaussian pulse centered at $f_0 = 5$ kHz and $\sigma = 1.5$ kHz. The map in the center presents the same map for a pulse centered at at $f_0 = 10$ kHz and $\sigma = 2.1$ kHz. The map in the bottom shows FDTD

simulation results for the 10 kHz pulse. The black dots denote the position of the piezoelectric point source.

The top map does not show any focusing effect as expected from the equifrequency contour study. Below 8 kHz, equifrequency contour are circles centered around Γ such that the index of refraction remains positive. The experiment confirms this theory as we can observe cylindrical wavefronts going through the crystal without complex interaction with the structure. However this behavior is drastically modified for frequencies above 8 kHz. For a pulse centered at 10 kHz a bright focus point is observed on the other side of the phononic crystal. This experiment is the first demonstration of flexural waves focusing with negative refraction [Dubois *et al.* (2013)].

The Figure 3.13 presents the transverse cut made from the focus at t = 830 μs. This representation is used to measure the transverse dimension of the focal point and extract the transverse resolution

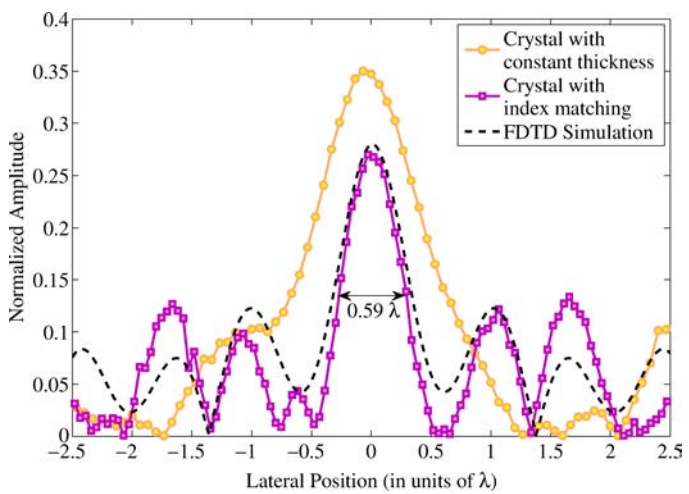


Fig. 3.13. Lateral cut of amplitude of normal displacement velocity at t = 830 μs. The horizontal axis is normalized by the wavelength. The experimental results from the constant thickness plate are presented in orange circle line. The experimental results from the varying thickness sample are plotted in purple square line. The FDTD simulation of the varying thickness plate are shown in black dotted line. The lateral resolution is calculated from the full width at half maximum of the focus point. From [Dubois *et al.* (2013)]

achieved by the lens. The horizontal axis is normalized by the wavelength at 10 kHz, $\lambda(2\text{ mm}) = 44.6$ mm and $\lambda(1\text{ mm}) = 31.6$ mm in Duraluminium plate. The figure presents results obtained for the phononic crystal with index matching (purple square line) and for the phononic crystal with a 1 mm constant thickness (orange circle line). The dotted line show results from FDTD simulation in the index matching case. One can observe an important reduction of the focal spot lateral size thanks to the thickness adaptation. The only drawback observe is a small reduction of the maximum amplitude focused that reduces from 36% to 27%. Nonetheless, this amplitude decrease is negligible considering the improvement of the lateral resolution from 1.05λ to 0.59λ. This work is a clear experimental demonstration of flexural wave index matching through thickness modification [Dubois *et al.* (2013)].

3.5. Superlens in Time Domain

As mentioned in the introduction, the main benefit of flat lenses is the super resolution effect. Indeed, the flat lens is able to focus the propagative information but also restores the evanescent field emanated from a point source. If the focusing of the propagative part can be represented with simple ray tracing, the amplification of evanescent field is a consequence of the existence of sharp resonances inside the lens [Gómez-Santos (2003); Wee and Pendry (2011); Ambati *et al.* (2007)]. The inevitable presence of loss in any system brings several challenges that need to be addressed in order to experimentally observed the superlensing effect. As a direct consequence, source plane and image plane have to be located in the near field of the structure as the evanescent field cannot be fully restored due to limited amplification in the lens [Fang *et al.* (2005); Sukhovich *et al.* (2009)]. Another restriction appears in the time scale of this super focusing effect. The amplification of evanescent information requires high quality factor resonances. Thus, the field associated with this resonance is extremely well confined within the lens and it will be difficult to excite these resonance with a point source located outside of the lens. This inefficient coupling between the source and

the resonances in the lens will introduce a transient regime which duration scales with the quality factor of the resonances [Gómez-Santos (2003); Wee and Pendry (2011); Archambault *et al.* (2012)]. It has been shown numerically that the super resolution needs a certain transient time before being observe [Archambault *et al.* (2012)] and the fundamental resolution limit can be obtained by comparing the loss rate to the transient dynamics of the lens.

3.5.1. *Time-resolved experiments*

Experimental investigation of this time dependent resolution are challenging at optical frequencies because time scale is on the order of picosecond. On the other hand, time-resolved experiments for acoustics or elastic waves are commonly carried out as ultrasonic frequencies do not required ultrafast electronics. The flat lens for flexural waves presented in the previous sections represents an appropriate opportunity to explore this lensing dynamics. Unlike previous experimental studies restricted to narrow band excitation, our experimental setup enables pulsed excitation and time resolved measurement of the focal spot obtained [Dubois *et al.* (2015)]. The snapshot represented in Figure 3.12 corresponds to the time when the amplitude at the focal spot is maximum. However, it may not correspond to the time when the resolution is maximum.

Figure 3.14(a) presents the lateral dimension of the focal spot in function of time for two gaussian pulses centered at 9 kHz and 10 kHz with equal standard deviation $\sigma = 2.1$ kHz. These curves are extracted from the temporal evolution of the focal spot observed frame by frame. The trend of both curves confirm the reduction of the focal spot size with time. The horizontal black dotted line represents the diffraction limit 0.5λ. For the pulse centered at 9 kHz, the lateral resolution becomes smaller than the diffraction limit with a minimum of 0.35λ at $t = 1150\ \mu s$. This first result experimentally demonstrates that the resolution of a flat lens improves with time for pulse excitation. Figure 3.14(b) shows the equifrequency contours of the crystal and the host plate. Brace brackets represent the width of each pulses. One can observe that the 9 kHz gaussian pulse excites contours with

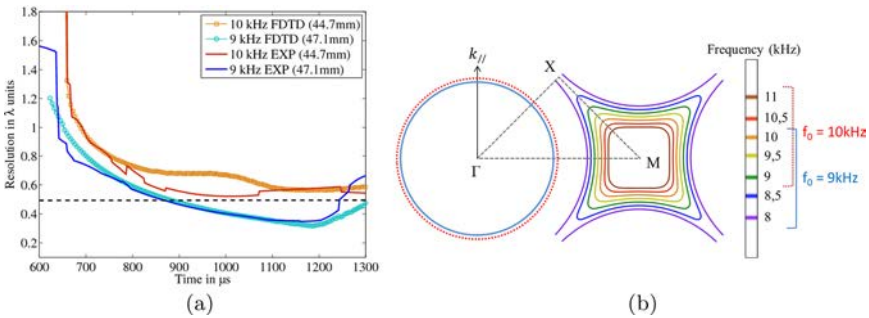


Fig. 3.14. (a) Lateral resolution in function of time. Red curves correspond to 10 kHz centered gaussian pulse, experiments with solid line and FDTD numerical results with open circles. Blue curves correspond to 9 kHz pulse, respectively. The horizontal black dotted line represents the diffraction limit 0.5λ. (b) Equifrequency contour of the square lattice and the host plate. Brace brackets represent the width of each pulses.

larger transverse wavevector, leading to a better resolution in the focal plane. Nonetheless, the equifrequency contour representing the dispersion of an infinite square lattice fails to predict the super resolution observed for this 9 kHz pulse excitation. It is mandatory to observe the field inside the lens and identify the resonance of the finite system.

Figure 3.15 presents measured and simulated intensity snapshots for the gaussian pulse centered at 9 kHz at t = 800 μs when the focal spot amplitude is maximum, and at t = 1150 μs when the lateral resolution is maximum. The ability to scan the displacement field inside the crystal is crucial in order to unravel the different resonances of the finite size lens.

The decay of the intensity inside the lens is used to investigate the resonances present in the lens. Figure 3.16(a) presents the average vertical displacement velocity in the lens in function of time. The red dotted line represents the duration of the pulse used to excite the lens. The persistent linear tail observed after the end of the pulse is a clear indication of resonances existence in the medium. Figure 3.16(b) represents the Fourier transform of the vertical displacement velocity averaged over the lens surface. Several resonances can be spotted

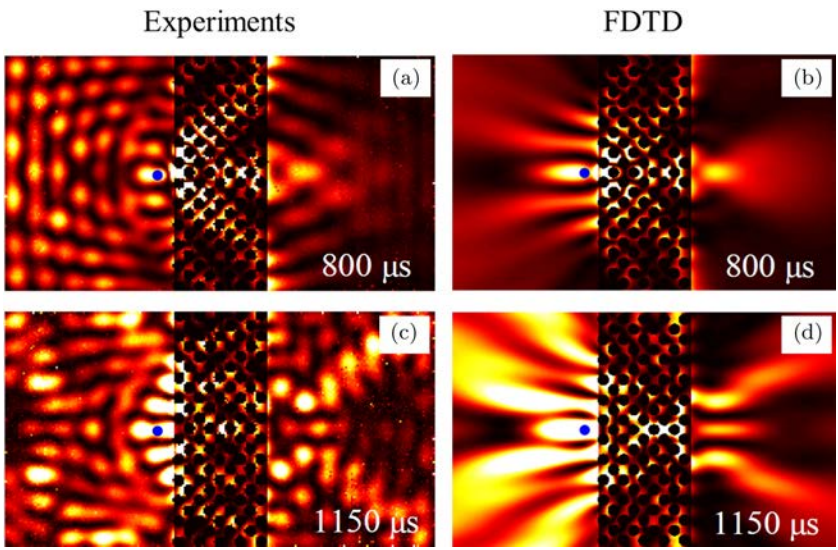


Fig. 3.15. (a) Snapshot of experimental normalized intensity of vertical displacement velocity at t = 800 μs for the gaussian pulse centered at 9 kHz. (b) Experimental intensity at t = 1150 μs. (c) Numerical intensity at t = 800 μs. (d) Numerical intensity at t = 1150 μs. Blue dots are located at the source position. The intensity inside the crystal is divided by a factor 9 to avoid the colorscale saturation. From [Dubois *et al.* (2015)]

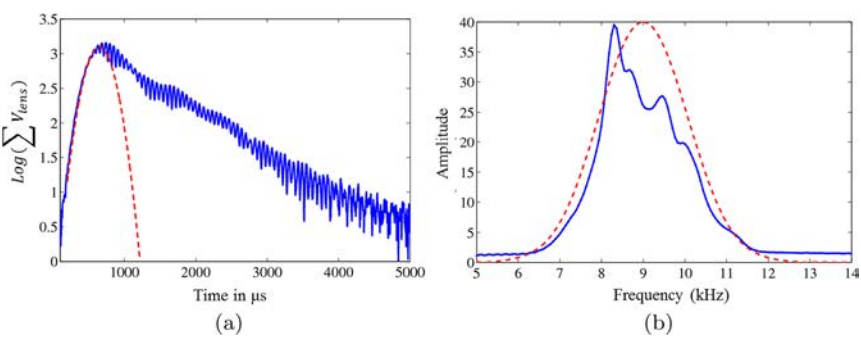


Fig. 3.16. (a) Decay of the intensity inside the lens in logarithmic scale. The red dotted line represents the duration of the pulse used to excite the lens. (b) Spectral amplitude spatially averaged over the lens. The red dotted line represents also the spectral width of the pulse used to excite the lens.

in this spectrum. However, due to the large density of peaks, it is impossible to resolve these resonance and extract information.

3.5.2. *Spatial Fourier transform*

Resonances are described by one eigenfrequency and one eigenvector associated to the spatial distribution of the field. We have seen in Figure 3.16 that it is impossible to resolved the frequency spectrum observed in our lens because of the high peak density. In order to address this issue, we perform the Fourier transform over the spatial variables and resolve the different resonances in the reciprocal space. Such analysis has been carried out previously [Engelen *et al.* (2007); Veres *et al.* (2012); Otsuka *et al.* (2013)] in order to experimentally observed the equifrequency contour of photonic and phononic crystals.

Figure 3.17 presents the reciprocal space obtained at t = 1176 μs after the spatial Fourier transform. The white dotted circle at the center of the map shows the equifrequency contours of the 2 mm thick host plate for 9 kHz frequency. The equifreqency of the phononic crystal for the frequency 8.5 kHz and 10.5 kHz are superposed to the reciprocal space. These two contours are centered around the M point of the first Brillouin zone. Figure 3.17 shows two different reciprocal spaces, one obtained from the entire real space and another one extracted from the lens space only. The entire signal processing based on Fourier transform is a linear process such that by cropping the real space, it is possible to observe the eigenvectors present in the lens. As expected, these resonance are located on top of the corresponding equifrequency contour, each spot corresponding to a specific resonance of the finite lens. Unlike the Fourier transform over the time variable, these different spots obtained by spatial decomposition are well resolved. Therefore, we can filter them out individually to determine the contribution of each of these resonances to the lateral resolution.

Figure 3.18(a) depicts the locations of the different resonances. Due to the four fold symmetry, 7 resonances are identified in the upper right corner of the reciprocal space. Again, the white dotted

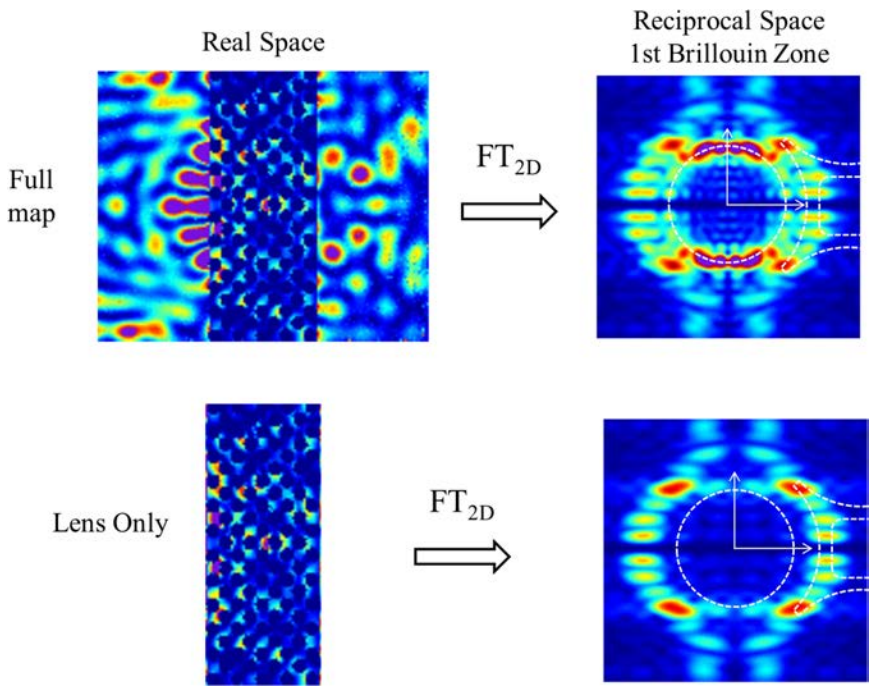
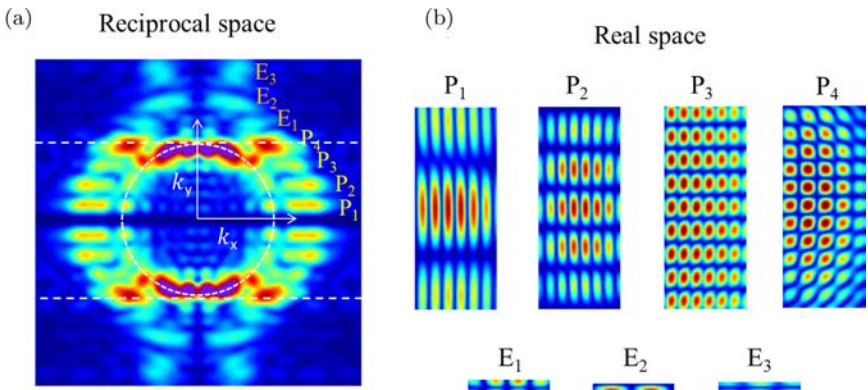


Fig. 3.17. Spatial Fourier transform processing applied to experimental measurement at t = 1176 μs for the 9 kHz pulse excitation. On the left are represented the different signals used in the real space. The right hand side presents the reciprocal space obtained for each map. The white dotted lines represents the equifrequency contours at 9 kHz for the 2 mm host plate and at 8.5 kHz and 10.5 kHz for the phononic crystal.

circle represents the equifrequency contour of 2-mm-thick the host plate at 9 kHz. Depending on the magnitude of the transverse component of the wavevectors, it is possible to sort the different resonances into two categories. If the resonance spot can be horizontally projected to the circle, the field associated with this resonance will be able to propagate in the host medium. These resonances are labeled $P_{1,2,3,4}$ for "propagating". On the other hand, the resonances labeled $E_{1,2,3}$ are located above the circle. Therefore, the field radiated by these modes will be evanescent in the host plate. Figure 3.18(b) shows the real space distribution associated with each resonance. It is



(c)

Batch	Rows	Columns	Frequency	Mode
7,5 kHz	15	4	7,7 kHz	*E1*
8,5 kHz	5	6	8,2 kHz	P2
	11	5	8,1 kHz	P4
	17	2	8,2 kHz	*E2*
9,5 kHz	2	6	9,8 kHz	P1
	23	3	9,8 kHz	*E3*
10,7 kHz	9	6	10,7 kHz	P3

Fig. 3.18. (a) Snapshot of the reciprocal space at t = 1176 μs. The white dotted circle represents the equifrequency contour of 2-mm-thick the host plate at 9 kHz. The two horizontal lines represents the limit between propagating and evanescent resonance. Seven resonances are spotted, 4 in the propagating region $P_{1,2,3,4}$ and three in the evanescent counterpart $E_{1,2,3}$. (b) Real space representation of the normalized amplitude of the field distribution associated with each resonance. (c) Table indicating the eigenfrequencies of each resonance. The batch frequency correspond to the central frequency of the narrow band excitation used in the FDTD simulations.

obtained by isolating the resonance spot and apply the inverse spatial Fourier transform. The filtering act as a low pass filter, such that the field inside the holes is interpolated. The table in Figure 3.18(c) presents the eigenfrequency of each of these resonance. These frequencies are obtained by numerical simulations with narrow band pulses, $\sigma = 0.7$ kHz against $\sigma = 2.1$ kHz in the original experiment. Perfect absorbing boundaries are set in numerical experiment in order to extract these eigenfrequencies.

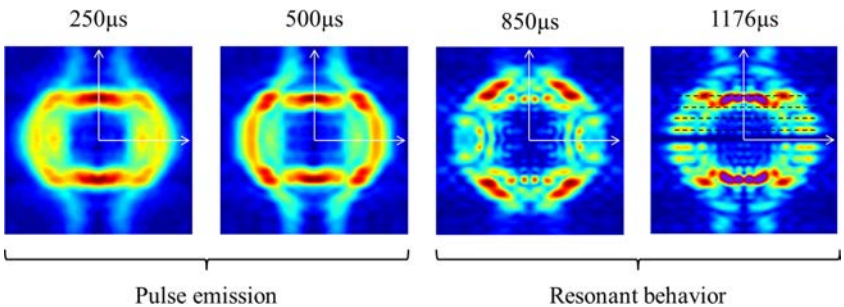


Fig. 3.19. Snapshots of the reciprocal space at times 250 μs, 500 μs, 850 μs, and 1176 μs. The horizontal dotted lines in the snapshot t = 1176 μs represents the conservation of the transverse component of the wavevector when the field goes from the lens to the exterior host medium.

As the different resonances existing in the lens are identified, we can relate the field radiated by the lens to these resonances and understand the time-dependent behavior of the focal spot. Figure 3.19 presents four snapshots of the reciprocal space obtained at four different times. The first two at t = 250 μs and t = 500 μs are taken while the pulse source is emitted. As all the wave vectors are emitted by the point source, the reciprocal space do not present the different spots shown previously. However, once the emission is stopped at t = 850 μs, only the wavevectors corresponding to the resonances are maintained in the structure. Thus, we can apply the Snell's law in order to extract the field radiated by each resonance. When the field radiated by the resonance crosses the interface between the lens and the host medium, we observe the horizontal projection of the wavevectors from the crystal equifrequency contour (centered around the M point) to the equifrequency contour of the host medium (inner circle centered in Γ point). This projection, corresponding to the conservation of the transverse component of the wavevector, is depicted by horizontal dotted line on the snapshot at t = 1176 μs. Therefore, we can isolate the field radiated by the different resonances and observe the spatial distribution in the real space with the inverse transformation method. This filtering can be done frame by frame in order to obtain the amplitude and phase of the radiated fields in function of time.

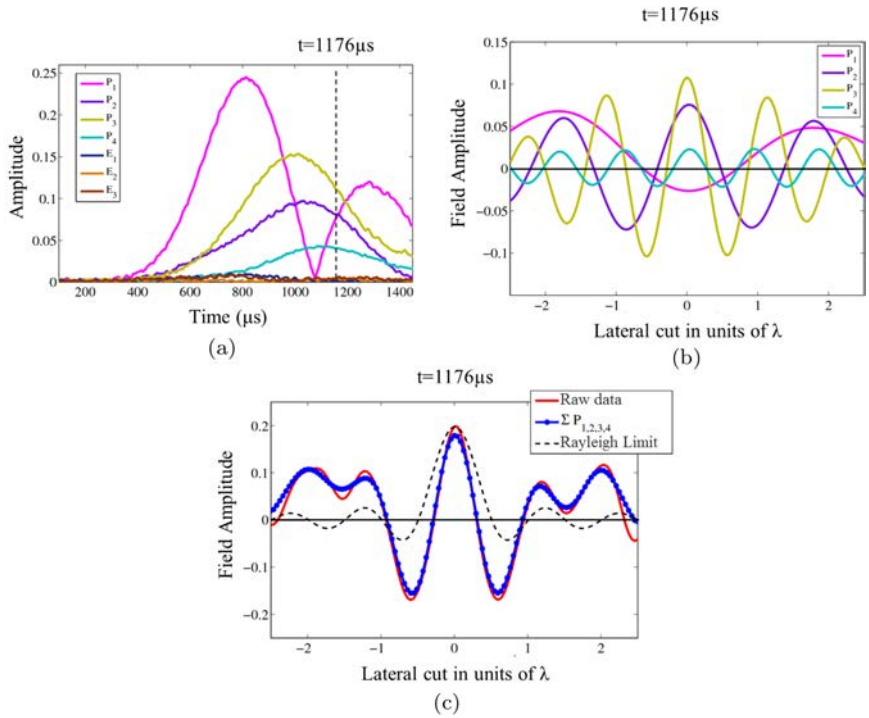


Fig. 3.20. (a) Amplitude of the field radiated by each resonance in function of time. The amplitude is taken at the center of the focal spot. (b) Lateral cut at the focal point of the field radiated by each resonance at t = 1176 μs. Only the "propagating" resonances $P_{1,2,3,4}$ are shown. (c) Lateral cut of the field at the focal point at t = 1176 μs. The original signal without processing is plotted in solid red curve. The sum of the "propagating" contributions is plotted in marked blue line. The focusing Rayleigh limit at 9 kHz in a 2 mm thick plate is presented in black dotted line.

Figure 3.20(a) shows the amplitude of the field radiated by each resonance in function of time. This amplitude is recorded at the focal spot frame by frame. Surprisingly, the contribution form the evanescent resonances $E_{1,2,3}$ is negligible even at time t = 1176 μs when the superfocusing is observed. Figure 3.20(b) presents the lateral cut at the focal point of the fields radiated by the "propagating" resonances $P_{1,2,3,4}$ at the time t = 1176 μs. Contrary to the classic image construction through a classic lens, here the different

contributions are not in phase and have different amplitudes resulting in a complex interference pattern. This interference pattern is presented in Figure 3.20(c) at time $t = 1176\,\mu s$. The solid red line shows the lateral cut of the focal spot without signal processing. The marked blue line presents the profile obtained by coherently adding the four contributions presented in Figure 3.20(b). The cardinal sine function in black dots represents the focusing Rayleigh limit. One can observe that the lateral profile of the focal spot is obtained without any evanescent field radiated by the lens. This result goes against the previous studies where the evanescent contribution is responsible for the superfocusing effect [Luo *et al.* (2003); Sukhovich *et al.* (2009)].

3.5.3. *Superfocusing without evanescent waves*

In 1990, Yakir Aharonov and his colleagues demonstrated that a mathematical function can present features that oscillate faster than their higher Fourier component [Aharonov *et al.* (1990)]. These functions are called *superoscillations*. Michael Berry and Sandu Popescu demonstrated that these functions can be obtained form plane wave interference. A suitable grating might lead to subwavelength field oscillation without the need of evanescent waves [Berry and Popescu (2006)]. Pioneering studies from Giuliano Torraldo di Francia [Francia (1952)] are further explored by Nikolay Zheludev and colleagues and lead to the experimental demonstration of superoscillations of light waves obtained in the far field of a quasi-crystalline lattice of holes [Huang *et al.* (2007)]. In the introduction of the previous article, authors proposed a simple scheme to obtain a superoscillatory function out of the coherent sum of six sinusoidal harmonics: $f(x) = \sum_{n=0}^{5} a_n \cos(2\pi n x)$ with coefficients a_n such as $a_0 = 1$, $a_1 = 13\,295\,000$, $a_2 = -30\,802\,818$, $a_3 = 26\,581\,909$, $a_4 = -10\,836\,909$ and $a_5 = 1\,762\,818$. This simple function is able to create a central spot that oscillates nine times faster than the higher harmonic used. However, such performance is not achievable without a counterpart, here the energy present inside this superoscillatory window is extremely small. The subwavelength spot is surrounded by side lobes which are seven orders of magnitude larger. The same team

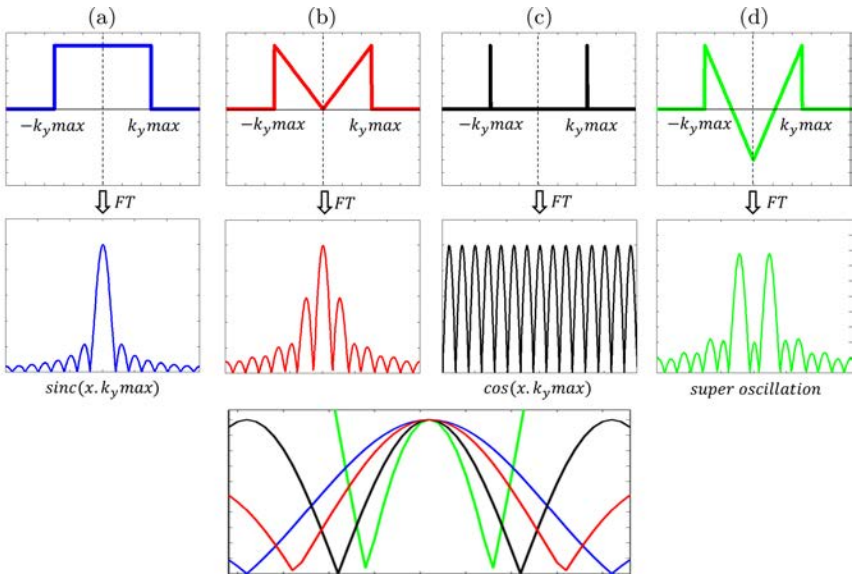


Fig. 3.21. Principle scheme representing the link between the Fourier plane and the lateral cut of the focal spot obtained in real space. The complex amplitude of the different components are presented on the top line. The center line presents the field distribution in the focal plane for the different case. The bottom line present the size of central spot by normalizing the different amplitude at the center.

demonstrated later that for less arbitrary fast features, the contrast with the side lobes can be reduced [Rogers *et al.* (2012)].

The Figure 3.21 presents different case where the amplitude and phase of the different components used in the construction of a focal spot can lead to different focal profile. Figure 3.21(a) presents the classic lens scheme where all the different component have the same phase and the same amplitude. In the Fourier plane, this situation correspond to a low pass filter or a gate function. The focal spot obtained by this a configuration is the well known cardinal sine function that leads to the diffraction limit. The other configurations show how the profile of the focal spot can be affected by the modifications in the Fourier plane. Figure 3.21(b) shows that you can already create a subwavelength spot by weighting carefully

the different components. The Figure 3.21(c) presents the most extreme case where only the largest transverse wavector is filtered. The spatial details are smaller than the cardinal sine function but the contrast is zero. The superoscillation behavior can be observed in the Figure 3.21(d) where the amplitude of the different components can access negative values, meaning that the different components are no longer in phase. In this specific case, the central spot can present oscillations faster than the cosine function presented in Figure 3.21(c). As discussed before, the counterpart of such extreme behavior is a negative contrast where the amplitude of the side lobes is larger than the amplitude at the central spot. This paragraph describes an intuitive approach to the superoscillation behavior in the wave lensing effect. Studies have explored superoscillations in a mathematical framework and unraveled the fundamental concepts associated [Kempf (2000); Calder and Kempf (2005)].

3.5.4. *Time-driven superoscillations*

We demonstrated previously that only the propagating field emanating from our phononic flat lens is involved in the creation of the focal spot. These four components can be described by four sinusoidal functions in the focal plane of the lens, the lateral profile $f(y)$ can be described as

$$f(y) = \sum_{i=1}^{4} a_i \, cos(k_y(P_i).y) \tag{3.7}$$

with a_i the complex amplitude associated with each component. The Figure 3.22 presents these complex coefficients in phasor diagram. This representation gives a simultaneous access to the amplitude, the phase and the frequency of each coefficient. Three phasors are presented in the Figure 3.22, they correspond to three different regime of focusing:

- at t = 718 μs when the spot appears
- at t = 882 μs when the diffraction limit is reached
- at t = 1176 μs when the focal spot is subwavelength.

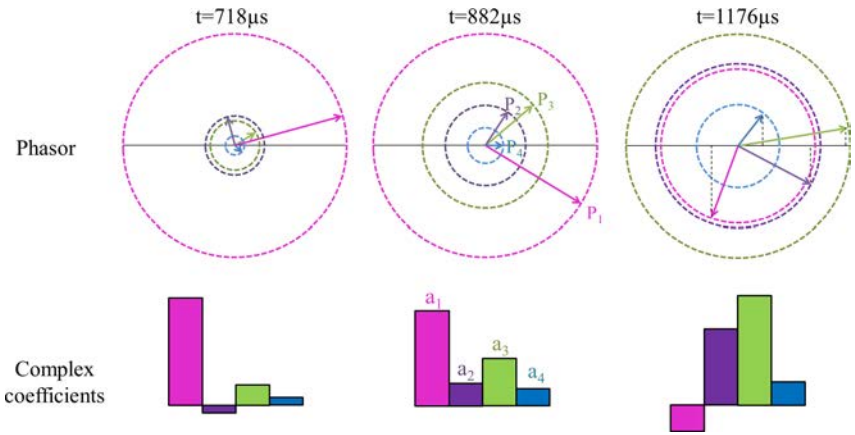


Fig. 3.22. Top, Amplitude and phase of the fields radiated by each propagating resonance P_i for three different times t = 718 μs, t = 882 μs and t = 1176 μs. Bottom, Complex coefficients a_i presented in histograms. Coefficient are obtained by vertical projection of the phasor vectors.

The vertical projection of the phasor vectors leads to the complex coefficients a_i for the three different times selected. As showed in the phasor at time t = 1176 μs, it is possible to understand the subwavelength focusing as the coefficient associated with the low transverse wavevector P_1 becomes out of phase with the other contributions. This negative coefficient results in a superoscillation pattern in the focal plane of the lens explaining the superfocusing without the presence of evanescent information. More than this observation, the value of these coefficients in function of time can explain in details the time dependent behavior of the focal spot.

By measuring the value of the transverse wavevector component of each resonance in the reciprocal space, we can reconstruct analytically the pattern observed in the focal plane frame by frame with the following function

$$f(y) = \sum_{i=1}^{4} a_i \, cos(k_y(P_i).y). \tag{3.8}$$

with a_i the complex coefficients, $k_y(P_1) = 35.1\, rad.m^{-1}$, $k_y(P_2) = 74\, rad.m^{-1}$, $k_y(P_3) = 116\, rad.m^{-1}$ and $k_y(P_4) = 140.1\, rad.m^{-1}$, the

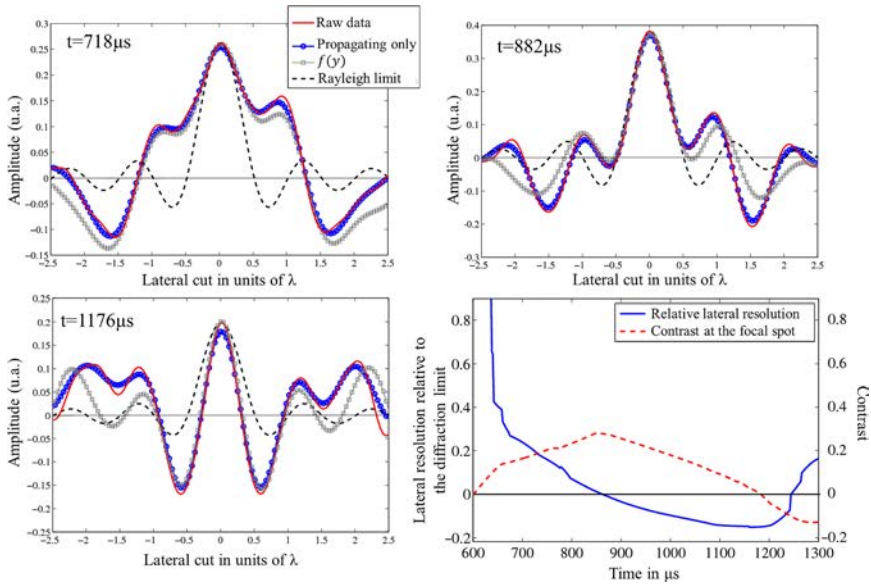


Fig. 3.23. Transverse profile of the field at the focal point at times t = 718 μs, t = 882 μs and t = 1176 μs. The red solid line is obtained from the raw experimental data without processing. The blue line marked with circles shows the contribution of the four propagating resonance excited in the lens. The grey line marked with squares represents the profile obtained analytically from the coefficients extracted previously. The cardinal sine function in dotted line shows the theoretical diffraction limited focal spot. The bottom right quadrant presents the lateral resolution relative to the diffraction limit and the contrast of the focal point in function of time.

transverse wavevectors of the resonances. Figure 3.23 presents the transverse profile of the field at the focal point at times t = 718 μs, t = 882 μs and t = 1176 μs. The red solid line is obtained from the raw experimental data without processing. The blue line marked with circles shows the contribution of the four propagating resonance excited in the lens. The grey line marked with squares represents the profile obtained analytically from the coefficients extracted previously. The cardinal sine function in dotted line shows the theoretical diffraction limited focal spot. The bottom right quadrant presents the lateral resolution and the contrast of the focal point in function of time.

First, the agreement between the calculated profile and the profile observed experimentally confirms and validates the approach used to describes the construction of the focal point through our phononic lens. Secondly, this approach can be used to finally link the temporal evolution of the lateral resolution to the different resonances excited inside the lens.

The time t = 718 μs corresponds to the early manifestation of the focal point. At this specific time, the contribution of the resonance P_1 is predominant as showed in Figure 3.22. However, the resonance P_1 corresponds to the lowest transverse wavevector. As a consequence the focal spot observed at this time presents a low lateral resolution, coarser than the actual diffraction limit.

As the point source fades above t = 850 μs, the amplitudes associated to the different resonances start to even. As a result, the lateral profile obtained at time t = 882 μs match the cardinal sine function denoting the diffraction limit. Consequently, the relative resolution represented in the bottom right corner crosses the zero line and the contrast at the focal spot is maximum.

The extinction of the source signal affects both the amplitude and the relative phase between resonances. The table presented in the Figure 3.18(c) shows that each resonance possess a different eigenfrequency. These eigenfrequencies correspond to different rotation speeds of the vectors presented in Figure 3.22. As soon as the point source vanishes, these vectors will rotate at their own rotation speed. Consequently, the phase between the different resonances is going to evolve with time. At time t = 1176 μs, the contribution of P_1 is negative as it is out of phase in regard to the other resonances. The lateral profile obtained in the focal plane can be described as a superoscillation and present subwavelength features. The negative value of the contrast at this specific time is the direct aftermath to this superoscillation behavior.

This spontaneous organization of the field radiated by the resonances of the lens into a superoscillation behavior is an unexpected result [Dubois *et al.* (2015)] as recent studies insist on the extreme care of the amplitude and phase of the different wavevectors [Wong

and Eleftheriades (2013); Rogers and Zheludev (2013)]. However, the exhaustive study of the finite lens resonance help us to identify their role in the focusing process. This study addresses fundamental issue of dynamics in anisotropic and dispersive system. The next paragraph will show that this self organization due to time-dependent excitation leads to a better resolution independently from the central frequency of the pulse or the source position.

3.5.5. *Time-domain resolution improvement*

In this section, we want to show that even if the temporal evolution of the system does not always lead to superfocusing or superoscillation, a significant improvement of the resolution is observed between monochromatic and pulsed regime. As mentioned previously, the experimental measurement of the monochromatic response is hindered by the presence of partial reflections at the boundaries of the sample. Nonetheless, we have shown that the Finite Difference Time Domain simulations can reproduce accurately the experimental observations and allow the implementation of perfectly absorbing boundaries. The goal of this section is to compare the resolution obtained with our phononic flat lens for monochromatic and pulsed source at different carrier frequencies.

Figure 3.24 presents the two source signals used for this demonstration. Results are presented for five different carrier frequencies, $f_0 = 8$ kHz, 8.5 kHz, 9 kHz, 9.5 kHz and 10 kHz. The gaussian pulses represented in dotted black have the same standard deviation $\sigma = 2.1$ kHz. The monochromatic signal represented in solid grey starts with a half of a gaussian envelope but the amplitude stay constant once the maximum value is reached. These two envelopes are kept identical through all the simulations, we vary the carrier frequency of the sinusoidal part of the signal. For every frequency used, the pulse signal leads to a significant better resolution than the stationary resolution observed with the monochromatic signals. In the next figure we are going to compare the properties of the reciprocal space for monochromatic and pulsed excitation.

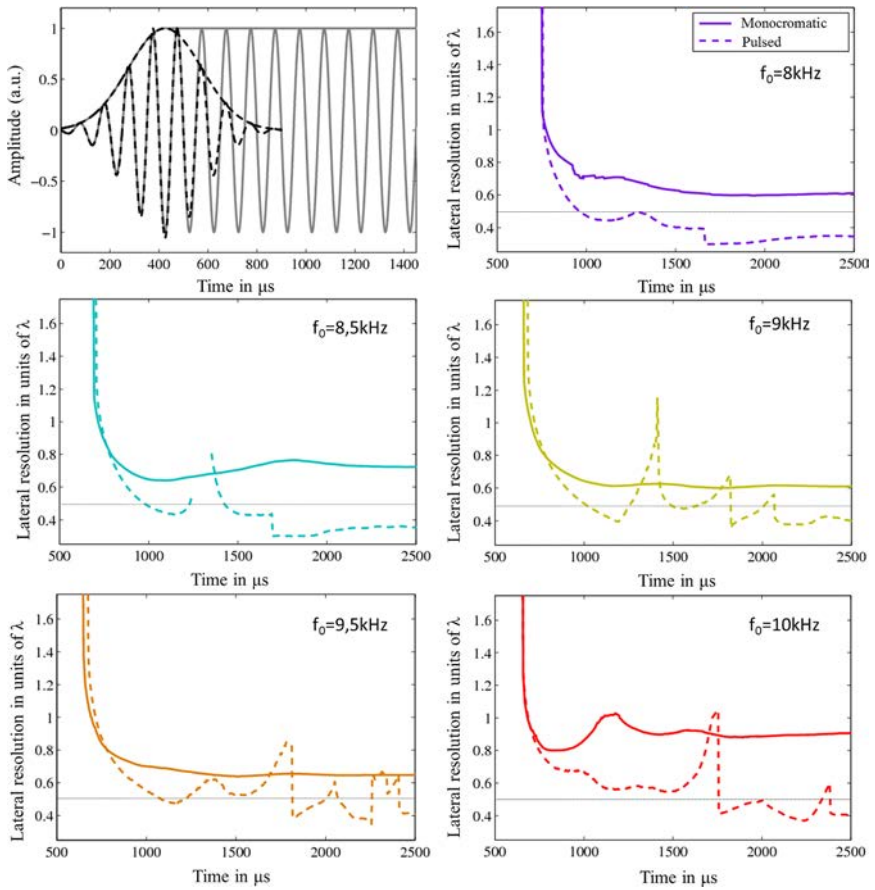


Fig. 3.24. Top left corner, representation of the two different excitation used in the different simulations. The dotted black line represents the gaussian pulse with a standard deviation of $\sigma = 2.1$ kHz. The grey solid line represent the monochromatic source signal. Lateral resolution in function of time is presented for the five different carrier frequency for both monochromatic and pulsed excitation.

Figure 3.25 presents the reciprocal space associated with the two different excitation for the carrier frequency $f_0 = 9$ kHz. The snapshots are taken at time t = 1190 μs for the pulsed excitation (maximum lateral resolution) and at time t = 1660 μs for the

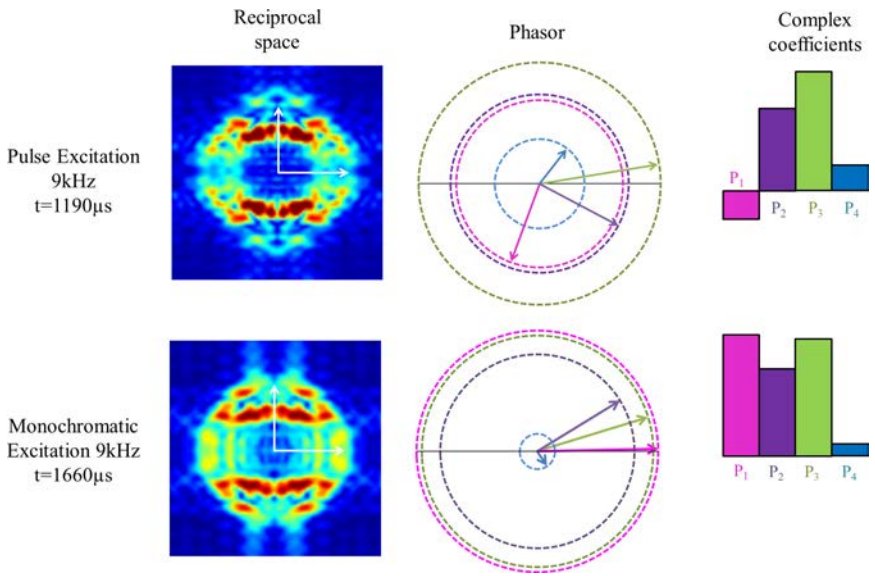


Fig. 3.25. Snapshots of the reciprocal space for the pulse and monochromatic excitation for the frequency $f_0 = 9$ kHz. The phasors and the complex coefficients associated with each propagating resonance are also represented.

monochromatic excitation when the profile is stationary. We first realize that the monochromatic excitation does not imply that a single mode is excited in the lens. The same resonances will be excited inside the lens since their quality factor is low. Secondly, the phasors associated with both excitation present the amplitude and phase of the fields radiated by each propagating resonance. As explained previously, the pulsed excitation allows the amplitude and phase of the different resonances to evolve with time and improves the lateral resolution. In the case of the monochromatic excitation, after a transient regime, these amplitudes and phases remain stationary since the lens is continuously driven at the frequency of the monochromatic source. This property does not depend strongly on the carrier frequency since all the simulations presented in the Figure 3.24 confirm an improvement of the lateral resolution for pulsed excitation.

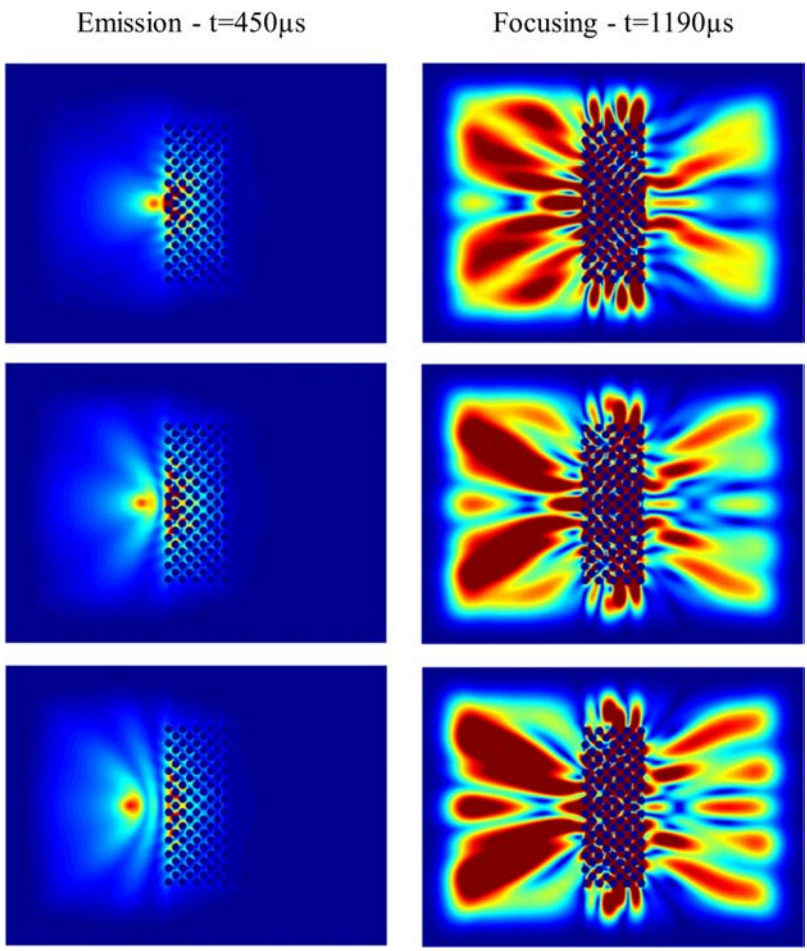


Fig. 3.26. Snapshots of the amplitude of the normal displacement velocity at time t = 450 μs during the emission of the signal source and at time t = 1190 μs during the focusing process. Point source is located at a distance $L/6$, $L/3$ and $L/2$ from the left side of the lens, where $L = 98$ mm represents the width of the lens.

3.5.6. *Influence of the source location*

This last section is dedicated to the influence of the location of the point source. We have seen in the introduction that the superfocusing demonstration in a real experiment requires the point source to be in

the near field of the lens since the evanescent field is involved in the superresolution obtained. As presented previously, the subwavelength focusing observe in our phononic crystal lens is based on superoscillation behavior and does not involve evanescent information emitting from the point source. The Figure 3.26 presents snapshots of the amplitude of the vertical displacement velocity for a point source located at a distance $L/6$, $L/3$ and $L/2$ form the left edge of the lens, where $L = 98$ mm represents the width of the lens. The gaussian pulse centered in $f_0 = 9$ kHz with $\sigma = 2.1$ kHz is used in the three different simulations. Snapshots are taken at time $t = 450$ μs during the emission of the signal source and at time t $= 1190\,\mu$s when the focal point is subwavelength.

The original experiment was using a point source located at a distance $L/6$ form the edge of the lens. One can observe that the focal point remains sharp as the point source is moved away from the near field of the lens ($L/2$ is equivalent to one wavelength at $f_0 = 9$ kHz in a 2-mm-thick duraluminium plate). This numerical experiment confirms that the initial location of the source does not deteriorate the spontaneous construction of the superoscillation pattern in the focal plane of the lens.

3.6. Conclusions

In summary, we have first theoretically explored the negative refraction in periodic structures and review pioneer experimental demonstrations in electromagnetics and in acoustics. We have demonstrated that such notions could also be applied to flexural waves in thin plates. This is an important first step towards the generalization of these concepts to elastodynamics which involves more complex analytic treatment as many polarizations coexist. After demonstrating the flat lens focusing of flexural waves, we studied the dynamic behavior of this flat lens. Along with previous theoretical and numerical efforts, this first experimental study in the time domain confirms that the subwavelength resolution builds up over time due to the interplay of multiple resonance inside the lens. Finally, the analysis

of these discrete resonances in the reciprocal space has shown that the resolution observed was the result of a spontaneous organization of the radiated fields in a superoscillatory function. Although, this effect tends to reduce the contrast at the focal point, it presents several benefits as the evanescent fields are not required to focus waves down to the subwavelength regime. We believe this work provides a new perspective for the design of optimized Veselago flat lenses and tailored superfocusing devices.

Acknowledgments

M.D. acknowledges Ph.D. funding from the Direction Generale de l'Armement (DGA). P.S. is thankful to the Agence Nationale de la Recherche support under grant ANR PLATON (No. 12-BS09-003-01), the LABEX WIFI (Laboratory of Excellence within the French Program Investments for the Future) under reference ANR-10-IDEX-0001-02 PSL* and the PICS-ALAMO. This research was supported in part by The Israel Science Foundation (Grant No. 1781/15 and 2074/15) and the Binational Science Foundation BSF-NSF (Grant No. 204358).

Bibliography

[1] Aharonov, Y., Anandan, J., Popescu, S., and Vaidman, L. (1990). Superpositions of time evolutions of a quantum system and a quantum time-translation machine, *Physical Review Letters* **64**, 25, pp. 2965–2968, doi: 10.1103/PhysRevLett.64.2965.

[2] Ambati, M., Fang, N., Sun, C. and Zhang, X. (2007). Surface resonant states and superlensing in acoustic metamaterials, *Physical Review B* **75**, 19, p. 195447, doi:10.1103/PhysRevB.75.195447.

[3] Archambault, A., Besbes, M. and Greffet, J.-J. (2012). Superlens in the Time Domain, *Physical Review Letters* **109**, 9, p. 097405, doi:10.1103/PhysRevLett.109.097405.

[4] Berrier, A., Mulot, M., Swillo, M., Qiu, M., Thylén, L., Talneau, A. and Anand, S. (2004). Negative Refraction at Infrared Wavelengths in a Two-Dimensional Photonic Crystal, *Physical Review Letters* **93**, 7, p. 073902, doi:10.1103/PhysRevLett.93.073902.

[5] Berry, M. V. and Popescu, S. (2006). Evolution of quantum superoscillations and optical superresolution without evanescent waves, *Journal of Physics A:*

Mathematical and General **39**, 22, pp. 6965–6977, doi:10.1088/0305-4470/39/22/011.

[6] Calder, M. S. and Kempf, A. (2005). Analysis of superoscillatory wave functions, *Journal of Mathematical Physics* **46**, 1, p. 012101, doi:10.1063/1.1825076.

[7] Cubukcu, E., Aydin, K., Ekmel, O., Foteinopoulou, S. and Soukoulis, C. M. (2003). Negative refraction by photonic crystals, *Nature* **423**, JUNE, pp. 604–605.

[8] Dubois, M., Bossy, E., Enoch, S., Guenneau, S., Lerosey, G. and Sebbah, P. (2015). Time-Driven Superoscillations with Negative Refraction, *Physical Review Letters* **114**, 1, p. 013902, doi:10.1103/PhysRevLett.114.013902.

[9] Dubois, M., Farhat, M., Bossy, E., Enoch, S., Guenneau, S. and Sebbah, P. (2013). Flat lens for pulse focusing of elastic waves in thin plates, *Applied Physics Letters* **103**, 7, p. 071915, doi:10.1063/1.4818716.

[10] Engelen, R. J. P., Sugimoto, Y., Gersen, H., Ikeda, N., Asakawa, K. and Kuipers, L. K. (2007). Ultrafast evolution of photonic eigenstates in k-space, *Nature Physics* **3**, 6, pp. 401–405, doi:10.1038/nphys576.

[11] Fang, N., Lee, H., Sun, C. and Zhang, X. (2005). Sub-diffraction-limited optical imaging with a silver superlens. *Science* **308**, 5721, pp. 534–7, doi:10.1126/science.1108759.

[12] Farhat, M., Guenneau, S., Enoch, S., Movchan, A. B. and Petursson, G. G. (2010). Focussing bending waves via negative refraction in perforated thin plates, *Applied Physics Letters* **96**, 8, p. 081909, doi:10.1063/1.3327813.

[13] Francia, G. D. (1952). Super-gain antennas and optical resolving power, *Il Nuovo Cimento* (1943–1954), 1.

[14] Gómez-Santos, G. (2003). Universal features of the time evolution of evanescent modes in a left-handed perfect lens, *Physical Review Letters* **90**, 7, p. 077401, doi:10.1103/PhysRevLett.90.077401.

[15] Hu, X., Shen, Y., Liu, X., Fu, R. and Zi, J. (2004). Superlensing effect in liquid surface waves, *Physical Review E* **69**, 3, p. 030201, doi:10.1103/PhysRevE.69.030201.

[16] Huang, F. M., Chen, Y., Garcia de Abajo, F. J. and Zheludev, N. I. (2007). Optical super-resolution through super-oscillations, *Journal of Optics A: Pure and Applied Optics* **9**, 9, pp. S285–S288, doi:10.1088/1464-4258/9/9/S01.

[17] Kempf, A. (2000). Black holes, bandwidths and Beethoven, *Journal of Mathematical Physics* **41**, 4, p. 2360, doi:10.1063/1.533244.

[18] Kittel, C. (2007). *Physique de l'état solide.*

[19] Kushwaha, M., Halevi, P., Dobrzynski, L. and Djafari-Rouhani, B. (1993). Acoustic band structure of periodic elastic composites, *Physics Review Letters* **71**, 2022, doi.org/10.1103/PhysRev Lett.71.2022.

[20] Lu, Z., Chen, C., Schuetz, C. A., Shi, S., Murakowski, J. A., Schneider, G. J. and Prather, D. W. (2005). Subwavelength imaging by a flat cylindrical lens using optimized negative refraction, *Applied Physics Letters* **87**, 9, p. 091907, doi:10.1063/1.2035317.

[21] Luo, C., Johnson, S., Joannopoulos, J. and Pendry, J. (2003). Subwavelength imaging in photonic crystals, *Physical Review B* **68**, 4, p. 045115, doi:10.1103/PhysRevB.68.045115.

[22] Otsuka, P. H., Nanri, K., Matsuda, O., Tomoda, M., Profunser, D. M., Veres, I. a., Danworaphong, S., Khelif, A., Benchabane, S., Laude, V. and Wright, O. B. (2013). Broadband evolution of phononic-crystal-waveguide eigenstates in real- and k-spaces. *Scientific Reports* **3**, p. 3351, doi:10.1038/srep03351.

[23] Pendry, J. (2000). Negative refraction makes a perfect lens, *Physical Review Letters* **85**, 18, pp. 3966–9.

[24] Pierre, J., Boyko, O., Belliard, L., Vasseur, J. O. and Bonello, B. (2010). Negative refraction of zero order flexural Lamb waves through a two-dimensional phononic crystal, *Applied Physics Letters* **97**, 12, p. 121919, doi:10.1063/1.3491290.

[25] Rogers, E. T. F., Lindberg, J., Roy, T., Savo, S., Chad, J. E., Dennis, M. R. and Zheludev, N. I. (2012). A super-oscillatory lens optical microscope for subwavelength imaging. *Nature Materials* **11**, 5, pp. 432–5, doi:10.1038/nmat3280.

[26] Rogers, E. T. F. and Zheludev, N. I. (2013). Optical super-oscillations: subwavelength light focusing and super-resolution imaging, *Journal of Optics* **15**, 9, p. 094008, doi:10.1088/2040-8978/15/9/094008.

[27] Royer, D. and Dieulesaint, E. (1996). *Ondes élastiques dans les solides Propagation libre et guidée*.

[28] Sukhovich, A., Jing, L. and Page, J. (2008). Negative refraction and focusing of ultrasound in two-dimensional phononic crystals, *Physical Review B* **77**, 1, p. 014301, doi:10.1103/PhysRevB.77.014301.

[29] Sukhovich, A., Merheb, B., Muralidharan, K., Vasseur, J., Pennec, Y., Deymier, P. and Page, J. (2009). Experimental and theoretical evidence for subwavelength imaging in phononic crystals, *Physical Review Letters* **102**, 15, p. 154301, doi:10.1103/PhysRevLett.102.154301.

[30] Veres, I. a., Berer, T., Matsuda, O. and Burgholzer, P. (2012). Focusing and subwavelength imaging of surface acoustic waves in a solid-air phononic crystal, *Journal of Applied Physics* **112**, 5, p. 053504, doi:10.1063/1.4747931.

[31] Veselago, V. G. (1968). The electrodynamics of substances with simulatneously negative values of permittivity and permeability, *Soviet Physics Uspekhi* **10**, 4, pp. 509–514, doi:10.1070/PU1968v010n04ABEH003699.

[32] Wee, W. H. and Pendry, J. B. (2011). Universal evolution of perfect lenses, *Physical Review Letters* **106**, 16, p. 165503, doi:10.1103/PhysRevLett.106.165503.

[33] Wong, A. M. H. and Eleftheriades, G. V. (2013). An optical super-microscope for far-field, real-time imaging beyond the diffraction limit. *Scientific Reports* **3**, p. 1715, doi:10.1038/srep01715.

[34] Yablonovitch, E. (1987). Inhibited spontaneous emission in solid-state physics and electronics, *Physical Review Letters* **58**, 20, pp. 2059–2062, doi:10.1103/PhysRevLett.58.2059.

[35] Yang, S., Page, J., Liu, Z., Cowan, M., Chan, C. and Sheng, P. (2004). Focusing of sound in a 3D phononic crystal, *Physical Review Letters* **93**, 2, p. 024301, doi:10.1103/PhysRevLett.93.024301.

[36] Zhang, X. and Liu, Z. (2004). Negative refraction of acoustic waves in two-dimensional phononic crystals, *Applied Physics Letters* **85**, 2, p. 341, doi: 10.1063/1.1772854.

CHAPTER 4

Space-time Cloaking

MARTIN W. McCALL*

Imperial College London, UK

PAUL KINSLER

Lancaster University, UK

4.1. Introduction: What is a Space-time Cloak?

It all seems like something out of Star Trek. A cloak in space and time. Not just those working in the new science of metamaterials, but a much wider general audience have virtually all heard something about the idea of cloaking and invisibility as popularised by J. K. Rowling in the Harry Potter novels. When wearing his cloak, the aspiring wizard Harry cannot be seen. It is not merely that he is hidden from the light by which we would otherwise see him, such as if he were to hide in a box, for example; light is cleverly refracted around him so that the rays from all the objects behind Harry reach any observer's unsuspecting eyes just as if he wasn't there (Fig. 4.1(a)). Now imagine replacing the y-axis in Fig. 4.1(a) with time as shown in Fig. 4.1(b). Light's uniform speed in either vacuum or a uniform medium mean that the 'rays' are now slanted. Distorting the rays

*Corresponding author: m.mccall@imperial.ac.uk

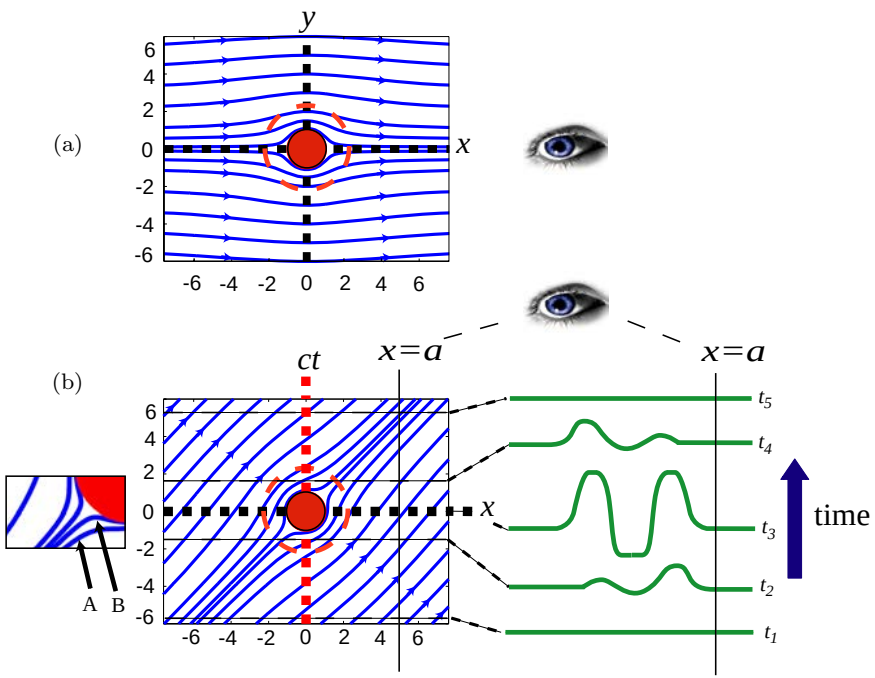


Fig. 4.1. (a) A cloak in space. Optical rays are refracted around an object so that an observer located to the far right does not see the object. (b) A cloak in spacetime that conceals *events* near the spacetime origin. The intensity distribution for various times is shown on the right indicating the formation and subsequent evaporation of the intensity null that is fully developed at $t = 0$. The observer to the right never suspects the occurrence of any non-radiating events near the spacetime origin and sees a uniform intensity for all time. The vertical dashed line represents the world line of an object whose visible history is temporally redacted. Figure 1 from Ref. 1 © IOP Publishing. Reproduced with permission. All rights reserved.

as shown results in a neighbourhood of the spacetime origin being unvisited by *any* of the rays. Since points in a spacetime diagram refer to *events*, the schematic depiction of Fig. 4.1(b) is of a cloak that conceals happenings or events. By reconstituting the rays to their parallel undistorted form, a downstream observer at $x = a$ does not suspect the occurrence of any events that occur at or during the cloak's operation. For such an observer the concealed events

are edited out of their version of observed history. Figuratively, the operation of an event cloak allows a safecracker to enter a scene, open a safe and steal the contents, and exit the scene (being careful to close the safe before he leaves!) whilst a surveying cctv camera all the while records the safe as being closed.

The remarkable thing about the new science of metamaterials is that methods now exist to transform these delightful, if rather fanciful concepts, into reality. The underpinning concept is that of *Transformation Optics*, a method that delivers a theoretical recipe for any cloak design. The spatial cloak, after being proposed in 2006 by Pendry *et al.*[2] and Leonhardt[3,4] was first demonstrated by Schurig *et al.*[5] A 2-D object of about 6 cm diameter was cloaked at microwave frequencies ($\lambda_0 = 4$cm). Since that paper there have been several improvements in the experimental realization of various spatial cloaks.[6–14] The event cloak, or spacetime cloak, was first proposed theoretically by McCall *et al.* in 2011.[1] Its first experimental realization was demonstrated by Fridman *et al.* a few months later.[15]

Experimental demonstration of these concepts brings applications to closer realization. Hiding objects in space has obvious attractions for the military, whilst event cloaking opens up new possibilities, including an important 'interrupt-without-interrupt' functionality, that may find application in optical signal processing.

In this chapter we will overview spacetime cloaking from its conception, through to its experimental realization and beyond. In Section 4.2 we describe the theory of electromagnetic cloaking in general, showing how both spatial cloaking and spacetime cloaking are brought within the same theoretical framework. In particular, we show the close connection between spacetime cloaking and a moving dielectric medium. Then, in Sec. 4.3 we discuss how, remarkably quickly, the abstract notion of a spacetime cloak was turned into an experimental reality. In Sections 4.4.1, 4.4.2 and Sec. 4.4.3, we discuss respectively the applications, illusions and implications of spacetime cloaking. Beyond optics, in Sec. 4.5 we discuss how an acoustic model can embrace spacetime transformations. Finally, in Sec. 4.6 we conclude.

Although the term 'Temporal Cloaking' has come to be associated with the concept introduced in Ref. 1, we, as originators of the concept, still prefer the term 'Space-Time Cloak', as it captures more precisely the nature of the cloaking operation it represents. Cloaking in a single dimension (e.g. time) is topologically impossible, and any concealment of events in time must necessarily be localized in space.

4.2. Transformation Optics

Transformation Optics is the recipe by which any space or space-time distortion can be actualized in an electromagnetic medium. The constitutive parameters of the medium are designed so that the field vectors $\mathbf{E}, \mathbf{B}, \mathbf{D}$ and $\mathbf{H}$ are actually morphed by the prescribed distortion. When discussing general distortions that can include time as well as space, it is convenient to use the indexed covariant form for the electromagnetic constitutive relations:

$$G^{\alpha\beta} = \frac{1}{2}\chi^{\alpha\beta\mu\nu}F_{\mu\nu}, \tag{4.1}$$

where in Cartesian coordinates

$$G^{\alpha\beta} = \begin{pmatrix} 0 & D_x & D_y & D_z \\ -D_x & 0 & H_z & -H_y \\ -D_y & -H_z & 0 & H_x \\ -D_z & H_y & -H_x & 0 \end{pmatrix}, \quad \text{and}$$

$$F_{\alpha\beta} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix}. \tag{4.2}$$

The index $\mu = 0$ refers to the time component, while the remaining three indices $\mu = 1, 2, 3$ refer to the spatial dimensions, and the Einstein summation convention is used. The fourth rank tensor $\chi^{\alpha\beta\mu\nu}$ contains the linear electromagnetic constitutive parameters of the medium. The independent components of $\boldsymbol{\chi}$ can be represented by

the block form

$$\boldsymbol{\chi} = \begin{pmatrix} -\epsilon_0\boldsymbol{\epsilon} & \eta_0^{-1}\boldsymbol{\alpha} \\ \eta_0^{-1}\boldsymbol{\alpha}^\dagger & \mu_0^{-1}\boldsymbol{\mu}^{-1} \end{pmatrix}, \tag{4.3}$$

where $\boldsymbol{\epsilon}, \boldsymbol{\mu}^{-1}$ and $\boldsymbol{\alpha}$ represent respectively the (relative) permittivity tensor, the (inverse) permeability tensor and the magneto-electric tensor. The symbol † in Eq. (4.3) indicates Hermitian conjugate, and $\eta_0 = (\mu_0/\epsilon_0)^{1/2}$ is the vacuum impedance. For subsequent convenience we take the vacuum speed of light $c = (\epsilon_0\mu_0)^{-1/2} = 1$. The indexing scheme for $\chi^{\alpha\beta\mu\nu}$ can be adduced from comparing Eqs. (4.1), (4.2) and (4.3). In an isotropic, non-magneto-electric medium we would have, for example $D_y = -\chi^{0202}E_y$ and $H_z = \chi^{1212}B_z$. The block form of Eq. (4.3) allows the linear constitutive relations between $\mathbf{E}, \mathbf{B}, \mathbf{D}$ and $\mathbf{H}$ to be written as

$$\begin{pmatrix} \mathbf{D} \\ \mathbf{H} \end{pmatrix} = \begin{pmatrix} -\epsilon_0\boldsymbol{\epsilon} & \eta_0^{-1}\boldsymbol{\alpha} \\ \eta_0^{-1}\boldsymbol{\alpha}^\dagger & \mu_0^{-1}\boldsymbol{\mu}^{-1} \end{pmatrix} \begin{pmatrix} -\mathbf{E} \\ \mathbf{B} \end{pmatrix}. \tag{4.4}$$

The apparently odd-looking appearance of the signs in Eq. (4.4) arises from a combination of $\chi^{\alpha\beta\mu\nu}$ effectively raising the covariant indices of $F_{\mu\nu}$ in Eq. (4.1), and the Minkowskii metric signature being taken as $-+++$. The symmetries of the constitutive tensor are[16] $\chi^{\alpha\beta\mu\nu} = \chi^{\mu\nu\alpha\beta} = \chi^{-\beta\alpha\mu\nu} = -\chi^{\alpha\beta\nu\mu}$.

We can now state the Transformation Optics algorithm succinctly and very generally. Suppose we have a medium (which might be vacuum) characterized by $\chi^{\alpha\beta\mu\nu}$ in which there exists the electromagnetic fields $G^{\alpha\beta}, F_{\mu\nu}$, at each spacetime point x^γ. Under the morphism $\varphi : x^\gamma \to \tilde{x}^\rho$ the electromagnetic field is similarly morphed provided the medium is replaced by one described by

$$\tilde{\chi}^{\mu\nu\rho\lambda} = \Delta^{-1}\varphi^\mu{}_\alpha\varphi^\nu{}_\beta\varphi^\rho{}_\gamma\varphi^\lambda{}_\delta\chi^{\alpha\beta\gamma\delta}\,, \tag{4.5}$$

where $\varphi^\mu{}_\alpha = \partial\tilde{x}^\mu/\partial x^\alpha$ and $\Delta = |\det(\phi^\mu{}_\alpha)|$. Incidentally, the above equation can also be interpreted as the prescription for expressing the χ tensor in different local coordinate systems; between Cartesian and polar coordinates for example. Under the morphism interpretation, in which points in space or spacetime are relocated, Eq. (4.5) is the

most general statement of Transformation Optics. For morphisms of space that do not involve time it reproduces the transformation optics that led to the spatial object cloak, for example. This arises from the polar transformation on the disc $r \leq b$

$$r \to \left(1 - \frac{a}{b}\right) r + a, \quad \theta \to \theta, \quad z \to z, \tag{4.6}$$

This morphism bends linear paths around an origin-centred disc of radius a (cf. Fig. 4.1(a)) and is thus the recipe for an electromagnetic cloak. Taking $\chi^{\alpha\beta\gamma\delta}$ to be vacuum it is straightforward to show from Eq. (4.5) that in conventional notation the required recipe is in (r, θ, z) coordinates

$$\boldsymbol{\epsilon} = \boldsymbol{\mu} = \begin{pmatrix} 1 - a/r & 0 & 0 \\ 0 & (1 - a/r)^{-1} & 0 \\ 0 & 0 & (1 - a/r)(1 - a/b)^{-2} \end{pmatrix}, \tag{4.7}$$

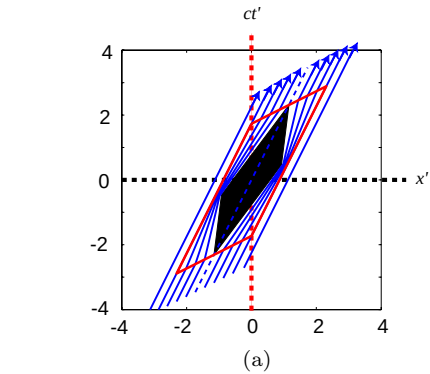
which was the basis for the design of the first experimental electromagnetic cloak made by Schurig *et al.* in 2006.[5] Firstly, the coordinate transformation interpretation of Eq. (4.5) is applied to express the vacuum constitutive parameters in cylindrical polar coordinates. Then, the polar cloaking morphism of (4.6) is applied using the morphism interpretation of Eq. (4.5). The result is the material prescription of Eq. (4.7). In the next section we will apply Eq. (4.5) to show how a spacetime cloak can be designed.

4.2.1. *The space-time cloak as an example of transformation optics*

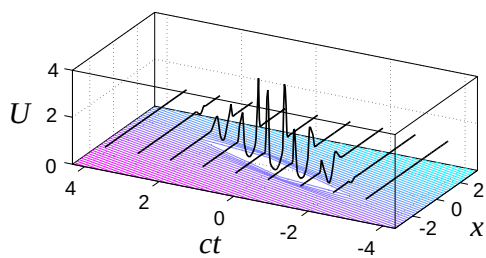
Within the framework of transformation optics, consider the space–time transformation $(x, t) \to (\tilde{x}, \tilde{t})$ and restrict to light propagating forwards along the x-axis. In Fig. 4.1(b) for example, a spacetime transformation is carried out in the (x, t) plane that is analogous to the spatial transformation carried out in the (x, y) plane of Fig. 4.1(a). In contrast to the spatial cloak, where the direction of propagation in the x–y plane is arbitrary, the vacuum light rays must follow the straight lines $x = ct + \text{const}$. These rays are then mapped under the transformation to the curved rays shown so that

the *events* within the disc surrounding the origin are avoided by the new rays. The new light trajectories are actualized when they propagate through a suitable inhomogeneous time-dependent medium, and will then curve around the event occurring at $x = 0, t = 0$. The intensity along the x–axis for different times is shown on the right of Fig. 4.1(b), indicating the formation of the intensity null that is fully developed at $t = 0$. For later times the null closes up restoring a uniform intensity distribution so that the observer at $x = a$ records a constant intensity for all time. All events in the (y, z) plane that occur near $(x = 0, t = 0)$ are concealed, so that effectively a space-time corridor is opened along which non-radiating events, such as the movement of matter, exchange of information etc., can occur undetected.

While Fig. 4.1 highlights the conceptual similarity between spatial and spacetime cloaking, it is schematic. Some rays in Fig. 4.1(b) would be required to have a phase speed exceeding the vacuum speed of light (e.g. ray A in the inset), and some rays very close to the cloaked region (e.g. ray B) propagate backwards in time. Through careful design, however, both of these shortcomings can be circumvented. Consider carrying out the coordinate transformation on a space that is filled with a medium with uniform and identical permittivity $\epsilon = n\epsilon_0$ and permeability $\mu = n\mu_0$, rather than vacuum. The medium then has a refractive index $n = (\epsilon\mu/\epsilon_0\mu_0)^{1/2} \xrightarrow{c=1} (\epsilon\mu)^{1/2}$, and superluminance can be avoided provided $n > 1$ and the dynamic modulation of the phase speed of light around its average value of $1/n$ is not too large. One proposed spacetime transformation to achieve this is shown in Fig. 4.2. Figure 4.2(b) shows a detailed calculation of the resulting electromagnetic energy density in the medium defined by the transformation. Unlike the cloak of Fig. 4.1(b), this spacetime cloak only works for light travelling in the $+x$ direction. An observer to the left of the origin does see the cloaked events, though time separations between the cloaked events are first speeded up, then time shifted, and then finally slowed down, before progression returns to normal. The recipe does not compromise the forward action of the spacetime cloak.



(a)



(b)

Fig. 4.2. (a) Construction of a sub-luminal space–time cloak. As with Fig. 4.1(b) the map $(x,t) \to (x',t')$ creates a void near the spacetime origin. However, the base space is filled with a medium of refractive index n rather than vacuum, so that prior to the transformation light rays are straight lines of gradient n. The transformation is a composition of a Lorentz boost, a 'hole-opening' transformation, followed by an inverse Lorentz transformation (see McCall *et al.*[1] for details). The deformed rays all have positive gradients (i.e. propagate forwards in time) and have speed $\leq c$. (b) Electromagnetic energy density for various times for the map of (a). Figure taken from Ref. 20.

4.2.2. *The relationship between a space-time cloak and a moving electromagnetic medium*

A very simple example of a spacetime transformation is a Lorentz transformation. Suppose the base medium is isotropic with $\epsilon = \mu = n$, i.e.

$$\boldsymbol{\chi} = \begin{pmatrix} -n\epsilon_0 \mathbb{1} & \mathbf{0} \\ \mathbf{0} & (n\mu_0)^{-1}\mathbb{1} \end{pmatrix}, \tag{4.8}$$

where $\mathbb{1}$ is the 3-D identity. Setting such a medium in motion with velocity v along the x-axis is equivalent to subjecting $\boldsymbol{\chi}$ to the Lorentz boost

$$\begin{pmatrix} \tilde{t} \\ \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix} = \begin{pmatrix} \gamma & \gamma v & 0 & 0 \\ \gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}, \tag{4.9}$$

where $\gamma = (1 - v^2)^{-1/2}$. In the notation required for Eq. (4.5) we have $\varphi^0{}_0 = \varphi^1{}_1 = \gamma$, $\varphi^0{}_1 = \varphi^1{}_0 = \gamma v$, $\varphi^2{}_2 = \varphi^3{}_3 = 1$, yielding after straightforward calculation the block components of $\tilde{\boldsymbol{\chi}}$ as

$$\boldsymbol{\epsilon} = \begin{pmatrix} n & 0 & 0 \\ 0 & \epsilon_\perp & 0 \\ 0 & 0 & \epsilon_\perp \end{pmatrix}, \quad \boldsymbol{\mu}^{-1} = \begin{pmatrix} n^{-1} & 0 & 0 \\ 0 & \mu_\perp^{-1} & 0 \\ 0 & 0 & \mu_\perp^{-1} \end{pmatrix},$$

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\alpha_\perp \\ 0 & \alpha_\perp & 0 \end{pmatrix}, \tag{4.10}$$

where $\epsilon_\perp = n\left[\frac{1-v^2/n^2}{1-v^2}\right]$, $\mu_\perp^{-1} = n^{-1}\left[\frac{1-n^2v^2}{1-v^2}\right]$ and $\alpha_\perp = \frac{v}{n}\left[\frac{n^2-1}{1-v^2}\right]$. When Eq. (4.4) is recast as

$$\begin{pmatrix} \mathbf{D} \\ \mathbf{B} \end{pmatrix} = \begin{pmatrix} -\epsilon_0\left(\boldsymbol{\epsilon} + \boldsymbol{\alpha}\boldsymbol{\mu}\boldsymbol{\alpha}^\dagger\right) & \boldsymbol{\alpha}\boldsymbol{\mu} \\ -\boldsymbol{\mu}\boldsymbol{\alpha}^\dagger & \mu_0\boldsymbol{\mu} \end{pmatrix} \begin{pmatrix} -\mathbf{E} \\ \mathbf{H} \end{pmatrix}, \tag{4.11}$$

it is found that for Eq. (4.10) that $\boldsymbol{\epsilon} + \boldsymbol{\alpha}\boldsymbol{\mu}\boldsymbol{\alpha}^\dagger = \boldsymbol{\mu}$ i.e. the equivalent medium has identical dielectric and magnetic properties. It is also seen that the equivalent medium is magneto-electric, which in fact is always the case for any transformation that mixes space and time. A meta-medium with constitutive parameters designed according to Eq. (4.10) would emulate an electromagnetic medium with refractive index n, moving with velocity v. Although such non-reciprocal bi-anisotropic meta-media would be difficult to manufacture, the possibility of such a medium has been studied by Tretyakov[17] who has shown that such media can be constructed from

small magnetized ferrite spheres combined with planar-chiral metallic inclusions.[18]

For a monochromatic plane wave travelling along the x-axis $\mathbf{E} = \mathbf{E}_0 \exp i(kx - \omega t)$, the constitutive relations of Eqs. (4.10) and (4.4) can be substituted into the Maxwell's relations $\nabla \times \mathbf{E} = i\omega\mathbf{B}$ and $\nabla \times \mathbf{H} = -i\omega\mathbf{D}$ to find that the phase velocity $v_p = \omega/k$ of light travelling in the moving medium is given by

$$v_p = \frac{v + n^{-1}}{1 + n^{-1}v}\,, \tag{4.12}$$

which is just the relativistic velocity addition formula. In the absence of dispersion this is also the group velocity of the propagating light. The possibility of a longitudinal light velocity profile $v(x)$ in a metamedium gives rise to the conceptualization of the spacetime cloak illustrated in Fig. 4.3. If the equivalent medium speed is small ($v \ll c$), then the speed of light along the direction of motion is just $v_p \approx c/n + v$ and the medium velocity v can be used to modulate the speed of light. Imagine initially that the medium is moving to the right in the Lab frame with speed v so that the phase speed of light along $+x$ is $c/n + v$. Now consider the situation from the medium's rest frame as shown in Figure 4.3(a) in which the light travels at phase speed c/n. Viewed from the same frame, Figure 4.3(b) shows the situation when the medium acquires a velocity profile in which the front part of the medium moves with velocity v_m, whilst the rear half moves with velocity $-v_m$. The leading part of the light is then sped up to $c/n + v_m$, whilst the trailing part is slowed down to $c/n - v_m$. In the Lab frame these two velocities are $c/n + v + v_m$ and $c/n + v - v_m$ respectively. A dark region then emerges at the mid-point and the cloak is opened. When the velocity profile is reversed (Figure 4.3(c)), the leading part is slowed down, whilst the trailing part sped up, and the cloak is closed. From the Lab perspective, the mid-point, where the cloak is opened and closed, moves at the medium's average speed v. Provided $v < v_p$, then the light never travels 'backwards' in the Lab frame. The reformed light contains no information about what happened at the origin when the cloak was open. Any non-radiating

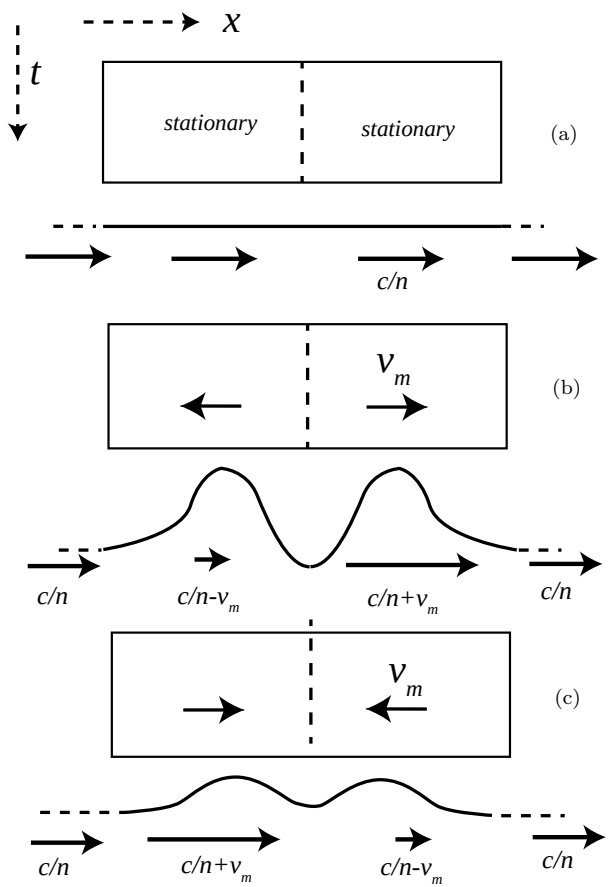


Fig. 4.3. Relating the formation of a spacetime cloak to a moving medium. In the Lab frame the medium moves to the right with speed $v \ll c$, so that the speed of light in the $+x$ direction is $c/n + v_0$. The three sketches are drawn from the viewpoint of an observer in the rest frame of the medium, where the nominal phase speed of light is c/n. (a) medium rest frame (b) medium with outward velocity profile (opening the cloak) (c) medium with inward velocity profile (closing the cloak).

events, such as the movement of matter, exchange of information etc. can occur undetected. It must be emphasised that the event that occurs within the dark region opened by the cloak must be *non-radiating* if it is to be hidden from the view of a distant observer.

An ideal spacetime cloak therefore requires the construction of a metamaterial that mimics propagation in a medium with a velocity $v(x,t)$. Near $t=0$, the effective medium is arranged so that $v(x,t)$ is negative for $x<0$ and positive for $x>0$. The trailing and leading parts of the light are thus respectively slowed down and sped up to produce the required intermediate null.

Moreover, each of the parameters $\epsilon_\perp, \mu_\perp$ and $\alpha_\perp$ is required to depend on both space (x) and time (t), the details depending on the transformation $t'(x,t), x'(x,t)$.

The magneto-electric parameter, $\alpha(x,t)$, in Eq. (4.11) arises invariably whenever the transformation $(x,t)\to(x',t')$ mixes space and time. This occurs, for example, when the base medium $\epsilon=\mu=n$ moves with velocity v. Then, by applying the Lorentz transformation $L_0^{0'}=L_1^{1'}=(1-v^2)^{-1/2}, L_1^{0'}=L_0^{1'}=v=(1-v^2)^{-1/2}$, the constitutive parameters of Eq. (4.10) are found to be

$$\epsilon_{||}=\epsilon,\ \mu_{||}^{-1}=\mu,\quad \epsilon_\perp=\epsilon\left[\frac{1-v^2/n^2}{1-v^2}\right],$$
$$\mu_\perp^{-1}=n^{-1}\left[\frac{1-n^2v^2}{1-v^2}\right],\quad \alpha=n^{-1}\left[\frac{1-n^2}{1-v^2}\right]v. \tag{4.13}$$

An ideal spacetime cloak therefore requires the construction of a metamaterial that mimics propagation in a medium with a velocity $v(x,t)$. The occurrence of the null can then be understood via the relativistic velocity addition formula $v'(x,t)=[1+nv(x,t)]/[n+v(x,t)]$ where $v'(x,t)$ is the phase velocity of light in the moving medium. Near $t=0$, the effective medium is arranged so that $v(x,t)$ is negative for $x<0$ and positive for $x>0$.

4.3. Experimental Realizations of Spacetime Cloaking

It will be noticed that throughout the above, dispersion was neglected. This is an important omission that has only very recently started to be addressed in the theory of spacetime cloaking.[19] Ironically, dispersion turned out to be the key to realizing the first experimental spacetime cloak by Fridman *et al.* in 2012.[15] Rather than utilize the acceleration and deceleration of different parts of

a ray by manipulating the refractive index, their experiment, illustrated in Fig. 4.4 achieved the temporal gap by utilizing a 'split time lens'. This technique relies on translating, via four-wave mixing, a chirp imposed on an intense pump wave, on to a weaker cw probe (1569 nm) over a time window in which the gap will be produced (cf. Fig. 4.4(b)). The initial and trailing parts of the probe beam then carry linearly chirped regions of opposite sign ($\sim$1 nm ps^{-1}). On propagating through a prescribed length of normal dispersing fiber, the longer wavelengths travel relatively slowly, whilst the shorter wavelengths relatively quicker. After a prescribed propagation distance through a single mode fiber, the wavelengths are brought into temporal alignment at the 'focus' of the temporal lens (cf. Fig. 4.4(c)). The method was able to open a gap of about 15 ps, during which time a further pulsed four-wave-mixing event took place in the central region. Generation of new frequencies relied on the presence of the probe, although the 'event' of the pump pulses (5 ps) interacting in the central region occurred independently of the presence of the probe. After passing through a length of complementary negatively dispersing fiber, the chirp profile is restored to its original shape, to be finally 'de-chirped' and brought back to a single frequency cw beam (cf. Fig. 4.4(d)). Without the cloak, an additional frequency (1539 nm) is generated every 24 μs in the central region (cf. Fig. 4.4(II)). With the cloak on, these pulses are not seen, demonstrating the hiding of the mixing event during the 50 ps for which the cloak operates. The experiment thus succeeded in hiding for the first time, and however briefly, an event from an observer's view, vindicating the concept of the spacetime cloak.

Impressive though the first demonstration was, the fraction of time cloaked was very small, amounting to about 10^{-4} of the available temporal window. An improvement on Fridman *et al.*'s proof-concept experiment was made by Lukens *et al.* at Purdue as reported in Ref. 20, another Nature paper. Similar to the Cornell experiment, the Purdue scheme also exploited a temporal analogue of a spatial diffractive effect, this time the time-analogue of the Talbot effect.[21] The spatial Talbot effect occurs when light passes through

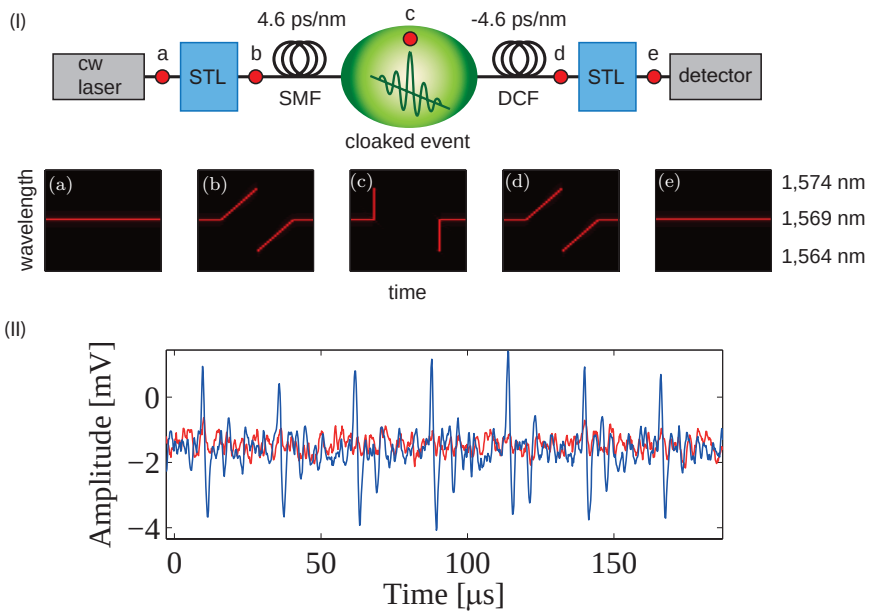


Fig. 4.4. (I) Experimental configuration and (II) Experimental results for cloaking an event in spacetime. The temporal gap is opened via a 'split time lens' in which a double chirp is imposed on a probe (1569 nm) over a time window in which the gap will be produced (Ib). On propagating through a normal dispersing fibre, the longer wavelengths travel relatively slowly, whilst the shorter wavelengths relatively quickly. After a prescribed propagation distance, the wavelengths are brought into temporal alignment at the 'focus' of the lens (Ic). A gap of ∼50 ps is opened, during which a further pulsed four-wave-mixing event takes place in the central region. Generation of new frequencies relied on the presence of the probe, although the 'event' of the pump pulses (5 ps) interacting in the central region occurs independently of the presence of the probe. After passing through a length of complementary negatively dispersing fiber, the chirp profile is restored to its original shape, to be finally 'de-chirped' and brought back to a single frequency cw beam (Id). Without the cloak, an additional frequency (1539 nm) is generated every 24 μs in the central region (II), whereas with the cloak on, these pulses are not seen. Taken from Ref. 15. Reprinted by permission from Macmillan Publishers Ltd: *Nature*, **481**, 62–65, copyright 2012.

an aperture whose transmittance is modulated periodically as in a diffraction grating. The optical distribution just beyond the aperture is regenerated periodically at multiples of the so-called Talbot distance $z_T = d^2/\lambda$, where d is the grating period. If the transmittance

is a sinusoidal phase variation (i.e. leaving the intensity distribution uniform), then at $z_T/4$ the intensity distribution becomes spatially periodic. The temporal Talbot effect works analogously. A signal acquiring a temporal sinusoidal phase variation via a phase modulator at $z = 0$, becomes periodically intensity modulated after propagating a distance $z_T/4$ through a medium whose dispersion characteristics analogize free-space diffraction — see Fig. 4.5. This already contains the genesis of spacetime cloaking, because intensity modulating/absorbing 'events' that occur at the periodic intensity nulls have no effect, and hence leave no evidence when the signal is reconstituted to a constant intensity. Moreover, the technique provides a *periodic series* of spacetime cloaks, occurring at a repetition rate roughly equal to the period of the signal. Rapid, repetitive spacetime cloaking at telecoms frequencies thus becomes possible. As a detail, the essentially sinusoidal peaks passing through $z = z_T/4$ were 'sharpened up' through applying an additional phase-modulation and propagation step — essentially the temporal equivalent of passing a sinusoidal spatial distribution through an array of lenses that focus each sinusoidal peak. The result in the experiment was a cloaking window of about 20 ps at a repetition rate of about 40 GHz. Attempts to intensity modulate at the null points of each spacetime cloak will not result in any modulation being passed to the signal beam, which will be restored to its uniform state after passing through appropriate inverting dispersion delay/phase modulation steps. Figure 4.6 shows the result of applying a sequence of ones and zeros in the event-cloaking plane both with and without the cloak switched on.

4.4. Implications, Applications, and Illusions

The idea of spacetime cloaking generalizes in a rather similar way as did the original spatial cloak to various types of distorting, illusion generating, or wave guiding devices. In its simplest form, however, as we have seen, spacetime transformation design can be seen in terms of speeding or slowing signals, and not solely or necessarily just steering them. This speed control then affects the pace at which events will

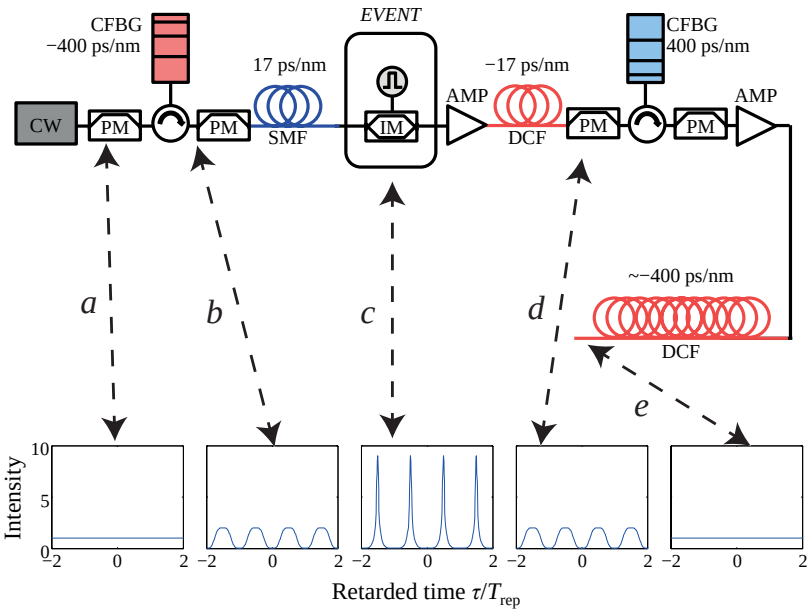


Fig. 4.5. Experimental set-up to achieve spacetime cloaking at telecom rates. The CW signal (a) is first phase modulated (PM) before passing through a Chirped Fiber Bragg Grating (CFBG). Together these steps are the temporal analogue of imposing a periodic phase distribution on a plane wave and then propagating through a distance $z_T/4$, at which point the distribution becomes periodically modulated in intensity (b). A further phase modulation/dispersion delay serves to focus the sinusoidal modulation to the sharp peaks seen at (c). This temporal distribution constitutes a rapid repetitive series of spacetime cloaks. Any data imposed on the signal via the Intensity Modulator (IM) during the intensity nulls between peaks will not be recorded by the signal. The subsequent dispersion delay/phase modulation steps serve first to defocus the sharp peaks to the sinusoidal distribution at (d) and then restore the constant intensity at (e). Taken from Ref. 20. Reprinted by permission from Macmillan Publishers Ltd: *Nature*, **498**, 205–208, copyright 2013.

finally be perceived; and opening up a gap in the illumination is the core concept in building a spacetime cloak. A range of possibilities for spacetime transformation devices has been discussed in more detail elsewhere.[22,23]

To understand spacetime transformation design we need to think of observers as those attempting to deduce cause and effect from light

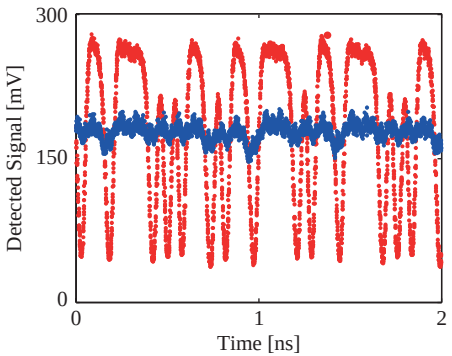


Fig. 4.6. A defined bit sequence of ones and zeros observed when the spacetime cloak is concealed when the cloak sequence is switched on. Taken from Ref. 20. Reprinted by permission from Macmillan Publishers Ltd: *Nature*, **498**, 205–208, copyright 2013.

or sound signals providing information about the environment. And since most of us are so used to the rapid and accurate view of the world offered to us by our eyes, it is often more instructive, when considering transformation design, to think of how bats or dolphins investigate their environment using sonar.

4.4.1. *Applications*

It has already been suggested[23] that not only continuous signals but also pulse trains are candidates for the velocity-modulation approach to spacetime cloaking. This would require time-dependent (i.e. dynamic) group velocity control[24,25] that can change the spacing between pulses; first opening up a wider time interval between a selected pair of pulses, then closing it again.

To elucidate this point, we will first discuss how a signal processing application can be altered to incorporate spacetime cloaking. In the realm of digital computation the pulse train is not "illumination" in the strobe-light or sonar-ping sense, but being used as a clock signal to control the behaviour of some signal processing unit (SPU). This is depicted in Fig. 4.7, where a kind of interrupt-without-interrupt functionality[1] can be achieved using a spacetime

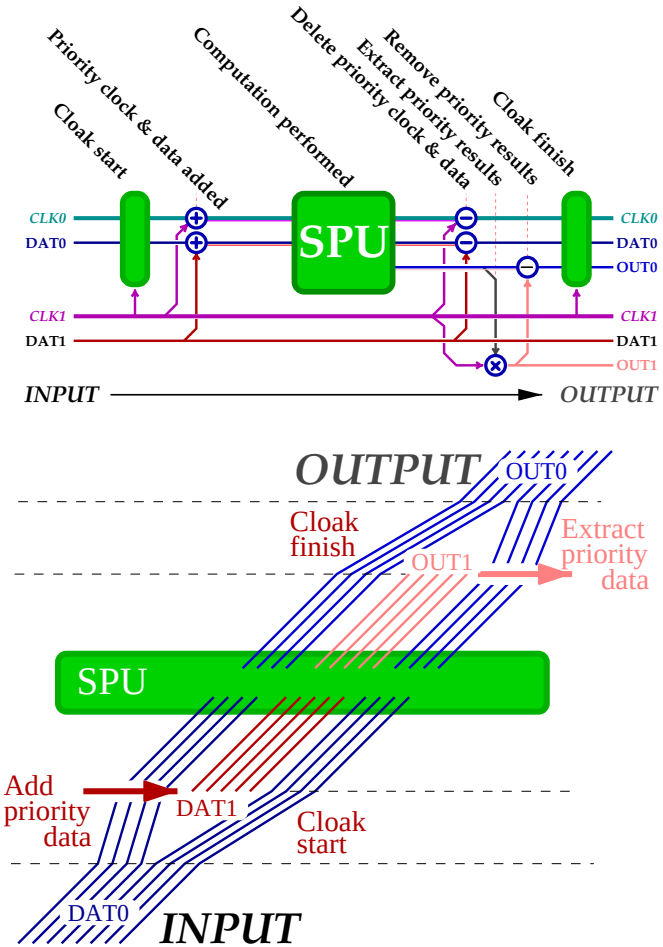


Fig. 4.7. A digital signal processing circuit modified to incorporate a space-time cloak. The normal ("background") operation of the SPU is controlled by clock CLK0 and processing data DAT0. This is modified by the insertion of a high-priority ("cloaked") computation on lines CLK1 & DAT1. After the signal processing unit (SPU) has processed the incoming data, the output has the extraneous CLK1 & DAT1 signals deleted, and the priority results diverted to OUT1. Lastly, the CLK0, DAT0, and OUT0 lines have their original timings restored and apparently pristine; however the priority results on OUT1 are available for use. The shifts in the signal timings are indicated in the lower panel, showing how the interior of the cloak is used to perform the priority computations. Taken from Ref. [23, Fig. 4] and used in accordance with its Creative Commons license.

cloaking method. The advantage of this over a straightforward temporary hijacking is that it can hide the disruption by a clever and reversible timing control that merely smoothly and fractionally overclocks the processor– perhaps only demanding e.g. only one extra clock pulse in ten. Such a pulse-based scheme could also be adapted to any telegraph-like electrical or electronic signal, to particle-like 'illumination' such as cars on a road, or to pulsed wave models of almost any type.

4.4.2. *Illusions*

A typical spacetime cloaking diagram such as Fig. 4.2 can have its different regions considered in isolation, so that each piece can be seen to perform elementary spacetime transformations. One part slows the illumination, another speeds it up; these provide the observer with the impression that events before the manipulation are occurring more or less rapidly. These are the spacetime analogs of spatial T-devices that do not cloak but shrink or expand the apparent size of an object; such time lens like effects — albeit not transformation-based ones — have a well-established track record in optics. Other parts divide the illumination into separate spacetime paths, another joins them together.

It has already been noted that the spatially magnifying T-device changes signals so that they not only imply a changed size for objects, but also of a space bigger on the inside than on the outside, in a sort of 'tardis' illusion. Likewise, a spacetime tardis[26] would be a T-device that seems to allow more time to pass than would be expected – unfortunately not a time machine, but only the illusion of one. Naturally, the opposite 'smaller on the inside' effect is also possible.

We can also consider manipulating an observer's view with the aim of confusing cause and effect. Figure 4.8 shows how this might be achieved, by first separating and distinguishing between the cause and effect segments of the sound stream forming the observer's view. Various methods of distinguishing between segments are possible, such as frequency conversion, polarization switching, or physical separation by routing into different waveguides. At the end of the

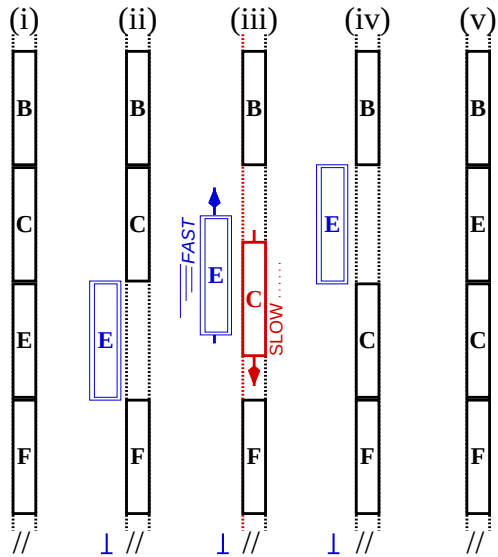


Fig. 4.8. Causality editing is where the a sequence of events BCEF is observed as reordered. With all signals initially in one 'parallel' (//) channel (i), those signals in interval E (the effect) are converted (ii) into a distinct 'perpendicular' channel (⊥). This allows (iii) the cause signals C to be slowed down, whilst the E signals are speeded up, so that they can switch places in the temporal sequence (iv). Finally, a continuous (but rearranged) history is constructed (v) by mapping the perpendicular channel back into the original parallel one.

process, the observer will then detect a view of history containing all of the expected information, but in a misleading sequence. Further, in the spacetime cloak experiment of Fridman *et al.*,[15] four wave mixing was used to imprint a *continuous* modulation of the illumination. Although this has not yet been attempted, it should allow us to not just re-order but actually *reverse* a segment of the historical record, as depicted in Fig. 4.9.

4.4.3. *Implications*

One implication of the discussion here is the change in emphasis of transformation design from implementation to the possibilities of signal manipulation. Just as the re-pitching of notions of superpositions

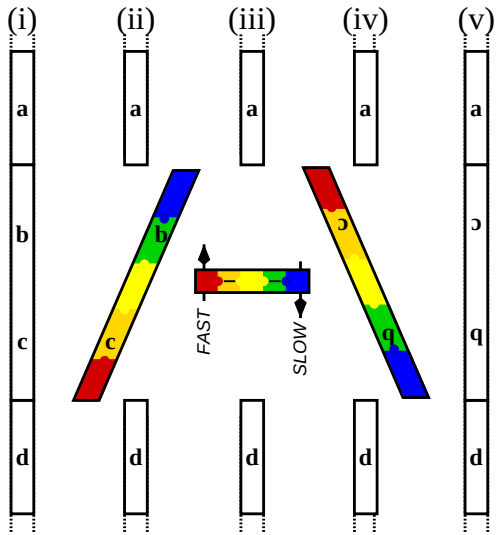


Fig. 4.9. Causality reversal, where the middle portion of a single frequency view of a sequence of events (i) is mapped continuously onto a chirped version (ii), as suggested by the rainbow coloured bar. Next, dispersion (i.e. frequency dependent wave speed) is then used in (iii) to slow early events (here lower frequencies, or red) and speed up later events (higher frequencies, or blue). Once the colour sequence is reversed (iv), the frequency chirp can be removed (v). This produces a continuous historical record, albeit one where one interval has been time reversed.

and coherence in quantum mechanics into the modern discussion of entanglement as a *resource* has unleashed a wave of new thinking and exciting results in quantum information and quantum cryptography, here we can see a way in which transformation design gives rise to a new operational way of thinking about signal manipulation.

4.5. Beyond Optics — Acoustic Space-time Transformations

In this subsection we will use a deliberately simplified acoustic model in order to clarify the features of spacetime transformations in an acoustic system.

4.5.1. *Acoustic model*

The version of p-acoustics presented here is generalised version of that used in the investigation of spacetime carpets[26]; it has also been used in an investigation of a general theory of transformation design.[27] Rather like macroscopic electromagnetism, it consists of two first order differential equations and one constitutive equation relating pairs of dynamical system properties. The system properties are a velocity field v^i and momentum density V^i, in combination with population P and a generalised pressure/stress p^{ij}; the indices span the three spatial dimensions x, y, z.

The p-acoustic equations are

$$\partial_t P = -\partial_i v^i, \qquad \partial_t V^m = -\partial_n p^{mn}, \tag{4.14}$$

$$p^{jk} = -\kappa^{jk} P \qquad V^m = \rho^m{}_i v^i, \tag{4.15}$$

where we also need to know that $\kappa^{rs}[\bar{\kappa}]_{rs} = 1$, and $[\bar{\rho}]_m{}^{i'} \rho^m{}_i = \delta^{i'}{}_i$. Here the first differential equation relates a velocity field to local population, and so can be interpreted as a conservation (of number) law; the second differential equation relates momentum to the generalised pressure/stress, and can be interpreted as a force law. As would be expected, the momentum density field is related to the velocity field by a matrix of mass-density parameters $\rho^m{}_i$. We also assume that in the rest frame of the acoustic material, there are no exotic cross-couplings between v^i and p^{jk} or between P and V^l; however as we will see they can be induced by spacetime transformations.

For ordinary p-acoustics, P is a scalar field representing the local population, and $\kappa^{ij} = \kappa_o \delta^{ij}$ repesents the bulk modulus; as a result $p^{kl} = p_o \delta^{kl}$ so that p_o is a pressure field. There is also a version of p-acoustics that mimics pentamode materials,[28] where the modulus κ^{ij} is a symmetric matrix but $\rho^{ij} = \rho_o \delta^{ij}$. Most generally, p-acoustics allows the case where κ^{ij} and p^{jk} can be (at least in principle) any symmetric matrix; in this case P represents the amplitude of an oscillating stress field whose orientation is determined by κ^{jk}, and where the restoring stress is p^{jk} is proportional to P.

The usual process for generating a second order wave equation then leads straightforwardly to

$$\partial_t^2 P = \partial_i C^{ij} \partial_j P \tag{4.16}$$

with a speed-squared matrix C^{is} that depends on the bulk modulus κ and the mass density matrix ρ, i.e.

$$C^{ij} = \bar{\rho}_n{}^i \kappa^{nj}, \tag{4.17}$$

A complementary wave equation for the momentum density V^m can also be obtained.

In a spacetime picture, as is most convenient for spacetime transformations, it is best to convert the above formalism into a tensor representation, just as we can convert the vector (Heaviside) formulation of Maxwell's equations into a tensor form. As part of this process we need to rescale the components so that all elements of each individual tensor have identical units; however from a practical perspective we can avoid doing this explicitly, since the relative changes induced by the transformation design process are the same irrespective of the units system.

We therefore pack a two index F with the scalar F^{00} component P, the velocity field v^i in F^{0i} and F^{i0}, and an auxiliary symmetric matrix quantity q^{ij} in F^{ij}. Likewise, we pack another two index G with the momentum density field V^i in G^{0i} and G^{i0}, the stress/pressure matrix p^{ij} in G^{ij}, and an auxiliary scalar Q in G^{00}. The whole p-acoustic theory then can be written in a way that closely matches an EM tensor form[a] but for symmetric tensors rather than antisymmetric ones. We have, in the source free case

$$\partial_\alpha F^{\alpha\beta} = 0, \quad \partial_\nu G^{\mu\nu} = 0, \tag{4.18}$$

$$G^{\alpha\beta} = \chi^{\mu\nu}{}_{\alpha\beta} F^{\alpha\beta}. \tag{4.19}$$

[a]Although the standard EM formulation uses a lower-index object $F_{\mu\nu}$, here we refer to the upper index 2-form version[26,27] that maps more directly onto a vector formulation.

We can now determine that the standard constitutive parameters ρ and κ are slices of χ such that

$$\rho_n{}^i = \chi_{n0}{}^{i0}, \quad \kappa^{nj} = c^2\,\chi^{nj}{}_{00}, \tag{4.20}$$

where c is the the p-acoustic equivalent of the vacuum speed of light. This c is most naturally set to unity in the tensor form, but usually has both a value and units in the vector calculus picture.

In addition to the usual constitutive parameters, two exotic couplings can also be induced, namely $\alpha^{jk}{}_i = c\chi^{jk}{}_{i0}$ between F^{i0} and G^{jk} (i.e. from velocity v^i to stress p^{jk}), and another $\beta^i = c\chi^{0i}{}_{00}$ between F^{00} and G^{0i} (i.e. from occupation P to momentum V^i). These are the p-acoustic analogues of EM's magneto-electric couplings.

The auxiliary quantities q^{ij}, Q also have a role to play in the tensor reformulation. However, these (and their respective constitutive parameters and relations) must be matched exactly to the standard quantities described above; this assumed, they do not need to be treated independently or explicitly.

In what follows, we follow the approximation used in Ref. 27, namely that we only transform the effective p-acoustic metric by means of transforming the constitutive parameters. This means we do not consider or allow for the effects of scattering induced by impedance changes induced by the transformation (see e.g. Ref. 29), nor do we allow for the effect of the transformation on the differential equations eqn. (4.18) (see e.g. Ref. 30). However, for sufficiently gradual transformations, we would expect that both these omitted effects would be negligible; so that the new effective metric induced by the transformation will at least get the steering of p-acoustic waves correct.

As a final note, the upper index nature of the field tensors $F^{\mu\nu}$ and $G^{\alpha\beta}$ means that they transform with a factor according to the determinant of the transformation, whereas the mixed index nature of the constitutive $\chi^{\mu\nu}{}_{\alpha\beta}$ is independent of the determinant.[27]

4.5.2. *Space-time transformations*

It is worth investigating the simplest transformations to start with. There are two cases, either (a) a purely spatial transformation, with

a compression or expansion (scaling) along one spatial axis (see e.g. Refs. 31 and 27) or (b) a spacetime transformation, with a time-dependent spatial offset or shear. Further, we will consider the effect of such transformations when starting from an isotropic material characterized by a single density parameter ρ_o and a single modulus parameter κ_o, with no exotic couplings.

Spatial case: here we choose an x-axis scaling such that the only change between primed and unprimed coordinates is $x' = \gamma x$. Note that we cannot scale two axes and still maintain continuity across the plane perpendicular to the x scaling. An alternative transfomation to consider would be a shear.

The spatial transformation matrix, and its inverse, for this is

$$T^{i'}{}_{i} = \left[\frac{\partial x^{i'}}{\partial x^{i}}\right] = \begin{bmatrix} \gamma & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad T_{i'}{}^{i} = \left[\frac{\partial x^{i}}{\partial x^{i'}}\right] = \begin{bmatrix} \gamma^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{4.21}$$

The transformed density, perhaps unexpectedly, does not change because the transformation-induced factors of γ cancel out. This happens as a result of it having mixed indices (one upper and one lower). The transformed (but nevertheless unchanged) density is

$$\rho' = \left[\rho'^{i}_{j}\right] = \rho_o \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{4.22}$$

The modulus has two upper indices and so transforms differently, so that the zz component accumulates a factor of γ^2. It becomes

$$\kappa' = \left[\kappa'^{ij}\right] = \kappa_o \begin{bmatrix} \gamma^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{4.23}$$

Spacetime case: here we choose an x-axis offset such that the only change between primed and unprimed coordinates is with $x' = \gamma x + \xi t$. The retention of a x scaling factor in this transformation allows us to use it as part of the construction of a 1D carpet cloak.[26]

The transformation matrix (and inverse) for this is

$$T^{\alpha'}{}_{\alpha} = \left[\frac{\partial x^{\alpha'}}{\partial x^{\alpha}}\right] = \begin{bmatrix} \gamma & \xi & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$T_{\alpha'}{}^{\alpha} = \left[\frac{\partial x^{\alpha}}{\partial x^{\alpha'}}\right] = \begin{bmatrix} \gamma^{-1} & -\xi\gamma^{-1} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.24}$$

The full transformation for the p-acoustic constitutive tensor χ is

$$\left[\chi'^{\mu'\nu'}{}_{\alpha'\beta'}\right] = \left[T^{\mu'}{}_{\mu}\right]\left[T^{\nu'}{}_{\nu}\right]\left[\chi^{\mu\nu}{}_{\alpha\beta}\right][T_{\alpha'}{}^{\alpha}]\left[T_{\beta'}{}^{\beta}\right]. \tag{4.25}$$

The altered constitutive parameters can now be calculated; but we need to be careful since the spacetime transformation means that the transformed constitutive parameters can depend on *any* of the original ones. We therefore have to allow for all of the summations in Eq. (4.25), which might cause κ_o to influence ρ', and vice versa. Further, we also check for the presence of new elements in χ – namely, the exotic couplings α and β, which we had started off without.

For the transformation chosen, which depends on only t and x, we only need to consider a subset of elements, i.e. those where spatial indices i become (only) x. For clarity, we also write the temporal index as 't' rather than the more common '0'.

For the density $[\rho'^{i}_{j}] = [\chi'^{it}{}_{jt}]$ we have

$$\chi'^{x't'}{}_{x't'} = T^{x'}_{x}T^{t'}_{t}\ \chi^{xt}{}_{xt}\ T^{x}_{x'}T^{t}_{t'} + T^{x'}_{x}T^{t'}_{x}\ \chi^{xx}{}_{tt}\ T^{t}_{x'}T^{t}_{t'} \tag{4.26}$$

$$\rho'^{x}_{x} = \gamma.1.\ \rho_o\ .\gamma^{-1}.1 + \gamma.0.\ \frac{\kappa_o}{c^2}\ .(-\xi\gamma^{-1}).1 = \rho_o. \tag{4.27}$$

And for the modulus $[\kappa'^{ij}] = c^2[\chi'^{ij}{}_{tt}]$ we have

$$\chi'^{x'x'}{}_{t't'} = T^{x'}_{x}T^{x'}_{t}\ \chi^{xt}{}_{xt}\ T^{x}_{t'}T^{t}_{t'} + T^{x'}_{x}T^{x'}_{x}\ \chi^{xx}{}_{tt}\ T^{t}_{t'}T^{t}_{t'} \tag{4.28}$$

$$\kappa'^{xx} = c^2\gamma.\xi.\ \rho_0\ .0.1 + \gamma.\gamma.\ \kappa_o\ .1.1 = \gamma^2\kappa_o. \tag{4.29}$$

We also need to check for exotic couplings that might be induced by the transformation. For $[\alpha'^{ij}{}_{k}] = c[\chi'^{ij}{}_{kt}]$ we find that this space-time transformation induces a coupling between the velocity and stress fields, i.e.

$$\chi'^{x'x'}{}_{x't'} = T^{x'}_{x} T^{x'}_{t} \chi^{xt}{}_{xt} T^{x}_{x'} T^{t}_{t'} + T^{x'}_{x} T^{x'}_{x} \chi^{xx}{}_{tt} T^{t}_{x'} T^{t}_{t'} \tag{4.30}$$

$$= \gamma.\xi.\ \rho_o\ .1.1 + \gamma.\gamma.\ \frac{\kappa_o}{c^2}\ .(-\xi\gamma^{-1}).1 \tag{4.31}$$

$$\alpha'^{xx}_{x} = c\xi\rho_o - \gamma\xi\frac{\kappa_o}{c}. \tag{4.32}$$

In contrast to the situation with α', for $[\beta'^{ti}] = [\chi'^{ti}{}_{tt}]$ we find no induced coupling β' between occupation and momentum. This is because the transformed time t' has no dependence on position x. In detail, the calculation runs as follows

$$\chi'^{x't'}{}_{t't'} = T^{x'}_{x} T^{t'}_{t} \chi^{xt}{}_{xt} T^{x}_{t'} T^{t}_{t'} + T^{x'}_{x} T^{t'}_{x} \chi^{xx}{}_{tt} T^{t}_{t'} T^{t}_{t'} \tag{4.33}$$

$$\beta'^{x} = c.\gamma.1.\ \rho_o\ .0.1 + c.\gamma.0.\ \kappa_o\ .1.1 \tag{4.34}$$

$$= 0. \tag{4.35}$$

4.5.3. *Time-dependent spatial expansion*

Cloaks in expanding spacetime has been a topic of recent interest.[32] As a partial investigation of how this kind of transformation device might be implemented in p-acoustics, we now consider the expanding (or contracting) space transformorphism where $\vec{r}' = \vec{r}\exp(\zeta t)$. Because the scalings are independent of orientation, we can just consider the t, z coordinate axis, and copy the consequences over to x and y as necessary. Fortunately, this case maps closely onto the previous one, but we now write $\zeta z \exp(\zeta t)$ in place of ξ, and $\exp(\zeta t)$ in place of γ. Thus we find that we need

$$\rho' = \rho_o, \tag{4.36}$$

$$\kappa' = e^{2\zeta t}\kappa_o, \tag{4.37}$$

$$\alpha' = \zeta z'[c\rho_o - e^{\zeta t}\kappa_o/c] = \zeta z' c^2 \rho_o[1 - n_o^{-2} e^{\zeta t}], \tag{4.38}$$

where $n_o = c/c_o$ is the original speed ("refractive") index.

Although it is not all that well justified, it is instructive to consider the approximation that the induced velocity-stress coupling terms α' can be ignored. With such a viewpoint, then the z-direction wave speed would be predicted to be

$$c'_{guess} = e^{\zeta t}\sqrt{\kappa_o/\rho_o} = e^{\zeta t}c_o. \tag{4.39}$$

We can now see that the retained terms account correctly for the change in scaling induced by the transformation. We can then interpret the α' coupling as the rest of the physical modification caused by the transformation — i.e. α' represents how the acoustic medium is moved relative to itself as the space expands/contracts.

An alternative transformation might be to instead use $\vec{r}' = \vec{r}\cosh(\zeta t)$. This provides the useful property that there is no α' at $t = 0$, regardless of the c_o/c ratio. As a result, we can always have an initial condition corresponding to our preferred "ordinary" case. This choice requires time-dependent constititive parameters which are given by

$$\rho' = \rho_o, \tag{4.40}$$

$$\kappa' = \cosh^2(2\zeta t)\kappa_o, \tag{4.41}$$

$$\alpha' = \xi z c^2 \rho_o \sinh(\zeta t)\left[1 - n_o^{-2}\cosh(\zeta t)\right] \tag{4.42}$$

$$= \xi z' c^2 \rho_o \tanh(\zeta t)\left[1 - n_o^{-2}\cosh(\zeta t)\right]. \tag{4.43}$$

From this, we have seen how one useful spacetime transformations is implemented under our model, and described the meaning of the two induced transformation effects. This and the other transformation used earlier can now be combined into more complicated transformation devices such as event cloaks.

4.6. Conclusion

In conclusion, in this chapter we have briefly summarized the field of spacetime transformation optics, firstly from its theoretical inception, through to experimental realization. We then showed how these concepts can be converted over to an acoustic realization, using a quite

general p-acoustic model. This model has the advantage of a simple construction, avoiding many of the complicating details involved in more realistic (microscopic) formulations of acoustics. This simple and straightforward presentation enables us to clearly show how the basic principles of spacetime transformation design can be applied generally. We see this as an enabler for workers intending to investigate spacetime transformation design — cloaking — in their own fields of interest.

Acknowledgements

PK and MWM acknowledge support from EPSRC grant number EP/K00305/1. PK acknowledges support from the STFC (The Cockcroft Institute, ST/G008248/1).

Bibliography

[1] McCall, M. W., Favaro, A., Kinsler, P. and Boardman, A. (2011). A spacetime cloak, or a history editor, *Journal of Optics* **13**, p. 024003, doi: 10.1088/2040-8978/13/2/024003, http://dx.doi.org/10.1088/2040-8978/13/2/024003.

[2] Pendry, J. B., Schurig, D. and Smith, D. R. (2006). Controlling electromagnetic fields, *Science* **312**, 5781, pp. 1780–1782, doi:10.1126/science.1125907, http://dx.doi.org/10.1126/science.1125907.

[3] Leonhardt, U. and Philbin, T. G. (2006). General relativity in electrical engineering, *New Journal of Physics* **8**, p. 247, doi:10.1088/1367-2630/8/10/247, http://iopscience.iop.org/1367-2630/8/10/247/.

[4] Leonhardt, U. (2006). Optical conformal mapping, *Science* **312**, 5781, pp. 1777–1780, doi:10.1126/science.1126493, http://dx.doi.org/10.1126/science.1126493.

[5] Schurig, D., Mock, J. J., Justice, B. J., Cummer, S. A., Pendry, J. B., Starr, A. F. and Smith, D. R. (2006). Metamaterial electromagnetic cloak at microwave frequencies, *Science* **314**, pp. 977–980, doi:10.1126/science.1133628, http://dx.doi.org/10.1126/science.1133628.

[6] Cai, W., Chettiar, U. K., Kildishev, A. V. and Shalaev, V. M. (2007). Optical cloaking with metamaterials, *Nature Photonics* **1**, 4, pp. 224–227.

[7] Li, J. and Pendry, J. B. (2008). Hiding under the carpet: A new strategy for cloaking, *Phys. Rev. Lett.* **101**, p. 203901.

[8] Liu, R., Ji, C., Mock, J., Chin, J., Cui, T. and Smith, D. (2009). Broadband ground-plane cloak, *Science* **323**, 5912, pp. 366–369.

[9] Ergin, T., Stenger, N., Brenner, P., Pendry, J. B. and Wegener, M. (2010). Three-dimensional invisibility cloak at optical wavelengths, *Science* **328**, 5976, pp. 337–339.

[10] Gharghi, M., Gladden, C., Zentgraf, T., Liu, Y., Yin, X., Valentine, J. and Zhang, X. (2011). A carpet cloak for visible light, *Nano Letters* **11**, 7, pp. 2825–2828.

[11] Chen, X., Luo, Y., Zhang, J., Jiang, K., Pendry, J. B. and Zhang, S. (2011). Macroscopic invisibility cloaking of visible light, *Nature Communications* **2**, p. 176.

[12] Zhang, B., Luo, Y., Liu, X. and Barbastathis, G. (2012). Macroscopic invisibility cloak for visible light, *Physical Review Letters* **106**, p. 033901.

[13] Chen, H., Zheng, B., Shen, L., Wang, H., Zhang, X., Zheludev, N. I. and Zhang, B. (2013). Ray-optics cloaking devices for large objects in incoherent natural light, *Nature Communications* **4**.

[14] Feng, Z., Wu, B.-H., Zhao, Y.-X., Gao, J., Qiao, L.-F., Yang, A.-L., Lin, X.-F. and Jin, X.-M. (2016). Invisibility cloak printed on a photonic chip, *Scientific Reports* **6**.

[15] Fridman, M., Farsi, A., Okawachi, Y. and Gaeta, A. L. (2012). Demonstration of temporal cloaking, *Nature* **481**, pp. 62–65, doi:10.1038/nature10695, http://dx.doi.org/10.1038/nature10695.

[16] Post, E. J. (1997). *Formal Structure of Electromagnetics*. Dover, Mineola, N.Y.

[17] Tretyakov, S. A., Nefedov, I. S. and Alitalo, P. (2008). Generalized field-transforming metamaterials, *New Journal of Physics* **10**, p. 115028, doi:10.1088/1367-2630/10/11/115028, http://dx.doi.org/10.1088/1367-2630/10/11/115028.

[18] Tretyakov, S. A. (2009). On a possibility to imitate media moving with superluminal velocity, in *3rd Int. Conf. on Advanced Electromagnetic Materials in Microwaves and Optics*. London.

[19] Gratus, J., Kinsler, P., McCall, M. W. and Thompson, R. T. (2016). On spacetime transformation optics: Temporal and spatial dispersion, *New Journal of Physics* vol. 18 (2016): 123010.

[20] Lukens, J. M., Leaird, D. E. and Weiner, A. M. (2013). A temporal cloak at telecommunication data rate, *Nature* **498**, 7453, pp. 205–208, doi:10.1038/nature12224, http://dx.doi.org/10.1038/nature12224.

[21] Wen, J., Zhang, Y. and Xiao, M. (2013). The Talbot effect: Recent advances in classical optics, nonlinear optics and quantum optics, *Advances in Optics and Photonics* **5**, 1, pp. 83–130, doi:10.1364/AOP.5.000083, http://dx.doi.org/10.1364/AOP.5.000083.

[22] Kinsler, P. and McCall, M. W. (2013). Cloaks, editors and bubbles: Applications of spacetime transformation theory, *Annalen der Physik* **526**, pp. 51–62, doi:10.1002/andp.201300164, http://dx.doi.org/10.1002/andp.201300164.

[23] Kinsler, P. and McCall, M. W. (2015). The futures of transformations and metamaterials, *Photonics and Nanostructures* **15**, pp. 10–23, doi:10.1016/j.photonics.2015.04.005, http://dx.doi.org/10.1016/j.photonics.2015.04.005.

[24] Tian, K., Arora, W., Takahashi, S., Hong, J. and Barbastathis, G. (2009). Dynamic group velocity control in a mechanically tunable photonic-crystal coupled-resonator optical waveguide, *Phys. Rev. B* **80**, 13, p. 134305, doi:10.1103/PhysRevB.80.134305, http://link.aps.org/doi/10.1103/PhysRevB.80.134305.

[25] Kumar, P. and Dasgupta, S. (2014). Sharply tunable group velocity in alkali vapors using a single low-power control field, *J. Phys. B* **47**, 17, p. 175501, doi:10.1088/0953-4075/47/17/175501, 1309.3581, http://dx.doi.org/10.1088/0953-4075/47/17/175501.

[26] Kinsler, P. and McCall, M. W. (2014). Transformation devices: Event carpets in space and space-time, *Phys. Rev. A.* **89**, p. 063818, doi:10.1103/PhysRevA.89.063818, http://dx.doi.org/10.1103/PhysRevA.89.063818.

[27] Kinsler, P. and McCall, M. W. (2015). Transformation design: metrics and speeds, 1510.06890. https://arxiv.org/pdf/1510.06890.pdf

[28] Norris, A. N. (2008). Acoustic cloaking theory, *Proc. Royal Soc. A* **464**, 2097, pp. 2411–2434, doi:10.1098/rspa.2008.0076, http://dx.doi.org/10.1098/rspa.2008.0076.

[29] Kinsler, P. (2016a). Cloak imperfect: Determinants, preprint.

[30] Kinsler, P. (2016b). Cloak imperfect: Impedance, https://arxiv.org/abs/1708.01071.

[31] McCall, M. W., Kinsler, P. and Topf, R., D. M. (2016). The refractive index of reciprocal electromagnetic media, *J. Opt.* **18**, 4, p. 044017, doi:10.1088/20408978/18/4/044017, http://dx.doi.org/10.1088/2040-8978/18/4/044017.

[32] Thompson, R. T. and Fathi, M. (2015). Shrinking cloaks in expanding spacetimes: The role of coordinates and the meaning of transformations in transformation optics, *Phys. Rev. A* **92**, 1, p. 013834, doi:10.1103/PhysRevA.92.013834, 1506.08507, http://dx.doi.org/10.1103/PhysRevA.92.013834.

CHAPTER 5

Soda Cans Metamaterial: Homogenization and Beyond

FABRICE LEMOULT*, **GEOFFROY LEROSEY, NADÈGE KAÏNA and MATHIAS FINK**

Institut Langevin, France

Phononic crystals and metamaterials are both man made media that allow us to tune the propagation of waves. Due to their very different typical spatial scales — wavelength and deep subwavelength — and underlying physical mechanisms — Bragg interferences or local resonances — they are often considered to be very different. As such, while the former are commonly used to manipulate and control waves at the scale of the wavelength, the latter are usually considered for their effective properties. Yet we have shown in the last few years that under some approximations, metamaterials can be used as phononic crystals, with the great advantage that they are much more compact. In this chapter, we review our results obtained on an acoustic metamaterial whose the unit cell is very simple and consists on an everyday object: a mere soda can, that is, a Helmholtz resonator of deep subwavelength dimension.

*Corresponding author: f.lemoult@gmail.com

We will first show that their properties can be understood, likewise phononic crystals, as resulting from interferences only, through multiple scattering effects and Fano interferences. Then, we will demonstrate that below the resonance frequency of its unit cell, a soda can metamaterial supports a band of subwavelength varying modes, which can be excited coherently using time reversal, in order to beat the diffraction limit from the far field. Above this frequency, the metamaterial supports a band gap, which will be used to demonstrate cavities and waveguides, very similar to those obtained in phononic crystals, albeit of deep subwavelength dimensions. We finally show that multiple scattering can be taken advantage of in these metamaterials, by correctly structuring them. This allows to turn a metamaterial with a single negative effective property into a negative index metamaterial, which refracts waves negatively, hence acting as a superlens.

5.1. Introduction

Waves, whatever their nature, acoustic or electromagnetic for instance, are subject, while propagating, to phenomena such as reflection, refraction and diffraction. Refraction and reflection, in particular, result from the mismatch between the propagating properties of two different media at a given interface. The propagating properties that we refer to are the typical parameters that permit to describe the way a wave propagates in a bulk material. In the case of acoustics, these are the sound velocity c and the bulk density ρ. Obviously, those parameters can be related to other physical parameters characterizing the medium, and for instance the wave velocity can be written in terms of the adiabatic fluid compressibility χ as $c = \sqrt{1/\rho\chi}$. In electromagnetic materials, the macroscopic parameters that are commonly used are the permittivity ϵ, the permeability μ and the resulting index of refraction $n = \sqrt{\epsilon\mu}$. Such macroscopic parameters are very useful because they allow us to predict most of the waves phenomena, but one has to keep in mind that those parameters do not necessarily describe the physics happening at the "microscopic" scale (where microscopic refers here to a scale much smaller than the typical scale of the waves, that is the wavelength).

In the late 60s of the last century, Veselago [Veselago (1968)] demonstrated in the specific case of electromagnetic waves that a medium having simultaneously negative permittivity and permeability is theoretically not prohibited, even if it had not yet been observed, and should lead to very unusual wave phenomena. In the late 90s, this theoretical prediction became real with the proposal of heterostructures that can be described respectively with a negative permittivity [Pendry *et al.* (1996)] and a negative permeability [Pendry *et al.* (1999)]. A mixing of those two media exhibiting both the properties to be negative were eventually proposed [Smith *et al.* (2000)]. The revolution in those proposals comes from the fact that the propagating media are heterostructures organized on a scale much smaller than the wavelength of operation. Thus, they can be homogenized to an effective bulk material described by effective propagating properties: the concept of a metamaterial was born. This idea did not take long to reach the acoustic community, and no later than in 2000 the first mention of an acoustic metamaterial that can be described with a negative density appeared [Liu *et al.* (2000)]. Though metamaterials were historically introduced to reach the negative effective property, it is not the only interesting one of a subwavelength resonant medium. In this chapter we will for example see that even the positive effective properties can lead to interesting phenomena.

The research presented in this Chapter results from a very similar approach. Everything started after initial experiments obtained in electromagnetics in the microwave range where focussing waves below the diffraction limit from the farfield inside a random collection of conducting wires was demonstrated [Lerosey *et al.* (2007)]. At this time, it was not thought that such a complex and heterogeneous medium could be described as a metamaterial, and a multiple scattering point of view was adopted. To better understand those results, the medium was simplified to a periodic arrangement of identical wires and this permitted to highlight the importance of resonance coupling when packing individual resonant wires on a subwavelength scale [Lemoult *et al.* (2010, 2011c,b)].

Because of our multiwave background, we decided to transpose those results from electromagnetics to acoustics and replaced the half-wavelength microwave resonator by an acoustic resonator that also presents a dimension smaller than the wavelength. In airborne acoustics, there is a typical object that fulfills this condition: the Helmholtz resonator [von Helmholtz (1885)]. It consists of a rigid container enclosing a volume of air, coupled to the outside through a small open-ended neck in order to admit sound. The resonance frequency of such an object can be tuned by changing its volume, its aperture or the length of the neck, and can thus completely behave as a subwavelength object at its resonance frequency. We have to mention that this resonance can be easily probed by anybody in the daily life since blowing in a bottle-neck generates a sound that is a signature of this effect. Starting from this simple observation, instead of manufacturing our own Helmholtz resonator, we decided to work with a mass-produced object exhibiting this resonance. After few tests and for some reasons that we will discuss later on, we ended up on using a simple soda can as our Helmholtz resonator. Experiments performed on a medium composed of several soda cans gave similar results [Lemoult *et al.* (2011a)] as the microwaves ones and will be discussed in this chapter.

The commonly used homogenization procedure in the metamaterial community to describe the propagation requires the wavelength to be large compared to the material scale, which is the case of the soda cans medium, so it happened that our research completely fell in the field of metamaterials even if at this moment we did not really mention it. Therefore, among the huge variety of acoustic metamaterials that have been proposed within the last 15 years [Liu *et al.* (2000); Fang *et al.* (2006); Guenneau *et al.* (2007); Ding *et al.* (2007); Zhang *et al.* (2009); Lee *et al.* (2010); Christensen *et al.* (2013); Yang *et al.* (2013); García-Chocano *et al.* (2014); Ma *et al.* (2014); Brunet *et al.* (2015); Ma and Sheng (2016)], we will now focus on a specific class: the locally resonant metamaterials. They consist on an arrangement of unit cells that resonantly interact with waves. So, in the first section of this chapter, we will show that arranging several soda cans

on a subwavelength scale builds a propagating medium that typically behaves as a locally resonant metamaterial. Notably, it exhibits a polariton-like dispersion relation, that results from the coupling of a continuum of propagating waves and a local resonance. Equivalently, the propagation in this medium can be described by effective parameters, one of them being typical of a resonant behavior.

In the next two sections, we will exploit this specific dispersion relation for applications in different frequency ranges. Below the intrinsic resonance of one resonator, the wave propagation exhibits an effective wavelength that is of the order of the spacing between two cans, meaning deeply subwavelength compared to the free space wavelength. We will show that one can exploit this property in order to beat the free-space diffraction limit from the farfield. Slightly higher in frequency, the medium can be described with a negative compressibility which overall results in an inhibited propagation of waves, a phenomenon which we will refer to as hybridization bandgap. We will show how one can exploit this property in order to trap and guide waves on a very small scale by introducing small defects in the structure. Those defects cannot be caught by the homogenization approach which averages the response over the free space wavelength scale, and it will eventually lead us to the last section of this chapter where we study the possibility to build a double negative medium just by using one type of resonator that should expectedly only bring one type of negativeness, using a clever structuration of the metamaterial.

5.2. The Soda Can as a Building Block of an Acoustic Metamaterial

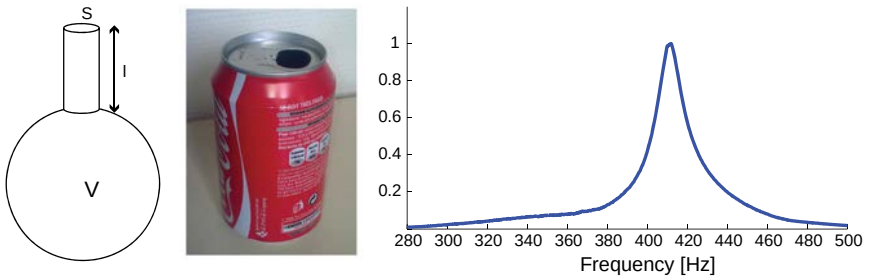
5.2.1. *Why the soda can?*

As already mentioned in the introduction, our research started in electromagnetics in the microwave range where we studied a wire medium [Lemoult *et al.* (2010, 2011c)]. The latter consists in a subwavelength periodic arrangement of identical conducting wires. It is well known that each of these wires exhibits a resonance resulting

from the stationary current in the wire when its length is a multiple of half the freespace wavelength. Therefore, such a wire can be seen from the top as a physically tiny object compared to the wavelength. Packing such resonant wires on a small scale we end up on a composite medium that falls within the class of locally resonant metamaterials.

The first idea that came out to mimic this electromagnetic example in acoustics was therefore to reproduce the behavior of the half-wavelength-long conducting wire. In acoustics, the complete equivalent of this resonator is used in the fabrication of many wind instruments: an open-ended half-wavelength-long pipe [Kinsler *et al.* (2000); Kergomard and Chaigne (2008)]. Indeed, there is always a propagating mode in a rigid-walled pipe whatever its cross section. Especially, at low frequency when the wavelength becomes large compared to the typical scale of the cross section, the pipe behaves as a single mode waveguide. This propagating mode presents a uniform profile of pressure along the cross section and propagates at the speed of sound. When cutting the pipe with two open-ended terminations, the impedance mismatch at both extremities builds a stationary resonant mode inside the pipe. The resonance frequency occurs when the frequency $f = n\frac{c}{2L}$ (where L is the pipe's length and n an integer). Such a pipe is therefore the acoustic equivalent of the conducting wire in our microwave experiment.

Therefore the acoustic equivalent of the wire medium was built, and it was unfortunately impossible to obtain similar results. This comes from a strong difference between the two types of resonators: while the electromagnetic resonator can be considered as almost lossless because copper is a good conductor, the pipe resonance suffers from losses. Indeed, the attenuation in the pipe originates from the viscous damping along its walls induced by the sliding layer near the interfaces. The ratio between the damping volume, which is located on a skin layer along the pipes walls, and the overall volume in the pipe increases while reducing the cross-section of the pipe. As a consequence, a pipe with a small cross-section (remember that we want to have a subwavelength resonator at the end) is described by an



(a) (b) (c)

Fig. 5.1. (a) A basic Helmholtz resonator with its important geometric parameters. (b) The chosen experimental resonator: a soda can. (c) Measure of the resonance frequency with a microphone placed inside the can.

attenuation length which can be much smaller than the attenuation length in air [Kosten and Zwikker (1949)].

After this unsuccessful attempt, we had to look for another acoustic resonator that fulfills the requirements of being small compared to the wavelength and of presenting relatively low losses due to viscous damping. This drove us to consider the case of the well known acoustic resonator introduced by Herman von Helmholtz [von Helmholtz (1885)] more than one century ago. The latter consists of a rigid container enclosing a volume V, terminated by an open-ended neck of length l and cross section S (Fig. 5.1(a)). The resonant behavior of such an object stems from the oscillation of the mass of the gaz column located in the neck, while the cavity behaves as the restoring force of the harmonic oscillator. The resonance frequency of the Helmholtz resonator therefore is simply:

$$f = \frac{c}{2\pi}\sqrt{\frac{S}{Vl}} \tag{5.1}$$

The resonance frequency can occur in the low frequency range where the operating wavelength is much larger than the resonator dimension. In our case, we want a subwavelength unit cell so we prefer to choose a relatively high volume V. We also know from the pipe experiment that most of the damping effects occur in the small

cross section, so we would rather prefer a short neck. As a consequence, in order to keep the resonance in the low frequency range it imposes to decrease S. Then started a real experimental strategy during which we tested dozens of mass-produced objects: glass bottles, glasses of different geometries, Christmas baubles or even chemical flasks... Out of those experiments stood out the "ideal" resonator for which both the neck was short enough to diminish the viscous damping and the geometrical dimensions appropriate to ensure the subwavelength nature of the resonator: we ended up on the use of a 33 cL soda can as a Helmholtz resonator (Fig. 5.1(b)). Experimentally, a resonance frequency of 420 Hz is measured (Fig. 5.1(c)), meaning that the lateral dimension of the can (i.e. 6.6 cm) is roughly $\lambda/12$ at resonance.[a]

5.2.2. *Wave propagation in a soda can medium*

Now that the unit cell is found, we want to study the propagation of waves in a locally resonant metamaterial based on it. To do so, we move from the single soda can to a medium made of several ones. Since the resonance occurs in the low frequency range, namely in the large wavelength regime, this medium in first approximation can be seen as a homegenous medium as sketched in Fig. 5.2(a). This homogenization procedure which is the scope of most of the metamaterials research is the exact analogous of dielectrics for light. Indeed, in this case the resonant unit cells are the atoms, which are deep subwavelength resonators. They are excited by the incoming light waves and their relaxation participate to the total transmitted optical field. This gives rise to variations of the optical index of refraction, which can present values larger or lower than that of vacuum. Stated otherwise, the interaction of light with microscopic resonant scatterers creates, at the macroscopic scale, an effective index of refraction. Our locally resonant metamaterials behave just the same way: the interaction

[a]Note that increasing the volume and keeping the lateral scale constant was doable by using a 50 cL can, which would have resulted on a more subwavelength resonator, but in France the 33 cL container is more common.

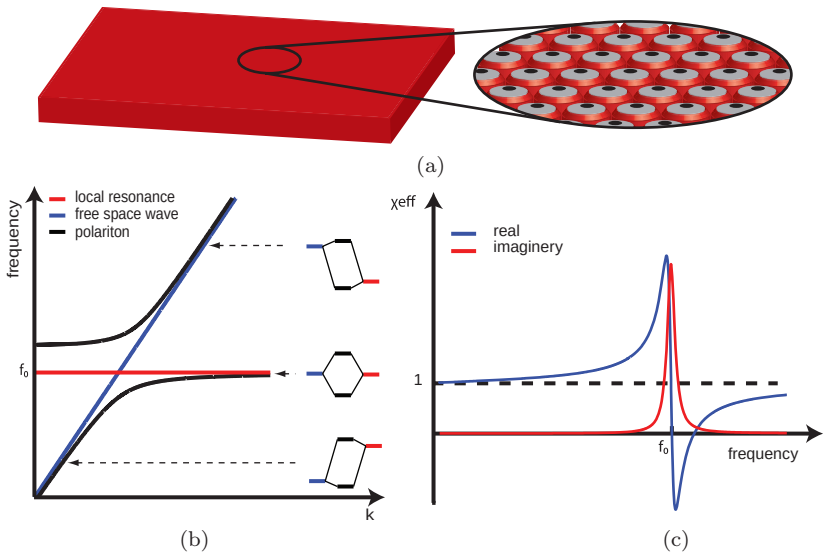


Fig. 5.2. (a) The metamaterial made of soda can can be seen as a homogeneous medium in the long wavelength regime. (b) The dispersion relation of waves inside the metamaterial is a so-called polartion which results from the avoiding crossing of the free space dispersion relation and the local resonance. (c) Description of the dispersion relation in terms of an effective property which in the case of the soda can is the effective compressibility χ_{eff}.

of acoustic waves with the soda cans gives birth to effective properties that relates to the resonant behaviour of each of them. There is, however, a big difference between atoms in dielectrics and soda cans. While the former present an albedo close to zero, meaning that their scattering cross-section is much smaller than their absorption cross-section, the latter present an albedo closed to unity [Lagendijk (1993)]. In other words, while most of the incoming interacting light is absorbed by atoms through inelastic scattering, the soda can is able to re-radiates most of the stored energy at resonance, as a result of our quest for the least lossy acoustic resonator.

Nevertheless, the physics of dielectrics is very similar to that of locally resonant metamaterials if we make the approximation that the metamaterial unit cells are not strongly coupled by any near field interaction. Though this approximation is not valid for any unit cell,

in the present case, the soda can walls are rigid enough to guarantee that the pressure field inside a given one does not influence the pressure field on its nearest neigtbours. Under this approximation, the interaction of acoustic waves and subwavelength resonators creates a polariton, just as light interacts with atoms in a dielectric. Namely, there is an avoided crossing between the local resonance of the soda can and the plane wave dispersion line (Fig. 5.2(b)). This gives rise to a binding branch of subwavelength modes below the resonant frequency f_0 of a single can, a band gap above it, and eventually an anti-binding branch of supra-wavelength modes. We have interpreted recently this behaviour in terms of Fano interferences [Fano (1961)] between the continuum of plane waves propagating in the matrix of the metamaterial and the local resonance [Lemoult *et al.* (2013)].

A more common way of describing the physics of locally resonant metamaterials consists in using the idea of effective properties. In acoustics, a given unit cell will act on macroscopic properties which are typically the effective mass density ρ_{eff} and the effective compressibility χ_{eff} depending on which type of excitation it is sensitive to [Brunet *et al.* (2013)]. In the case of the soda can, the volume V of the soda can behaves as an extra compressive volume for the incident wave encountering it and will therefore locally affect the effective compressibility seen by the pressure field. As shown in Fig. 5.2(c) the soda can creates below its resonance frequency a band of very high effective compressibility — equivalent to the branch of subwavelength modes in the polariton description — then a band of negative effective properties — the band gap — and finally a band of low effective property — the supra-wavelength modes branch.

5.3. Exploiting the Propagating Band for Subwavelength Focussing from the Far-Field

5.3.1. *Metamaterials for subwavelength focussing*

In the metamaterial community, much attention has been paid in the past 10 years to using the negative effective property band for focusing waves below the diffraction limit. Indeed, if one can realize a

metamaterial presenting two almost co-localized subwavelength unit cells which are resonant at the same frequency and which act on both properties of the medium, one can obtain a metamaterial that has both its effective properties negative. As was pointed out more than 40 years ago by [Veselago (1968)], this results in a medium whose effective index is negative. John Pendry proposed in a seminal paper in 2000 [Pendry (2000)] that a slab of such a medium should behave as a perfect lens for imaging and focussing purposes since it amplifies infinitely the evanescent waves coming from a source or object, making them measurable in the far field with conventional optical components. This approach has been shown to be largely hampered by losses of materials, especially for applications in optics where materials are relatively dissipative.

Subwavelength control of acoustic waves has not been studied as much as in electromagnetics, but there have been a few proposals in order to realize super-resolution imaging based on canalization [Zhu *et al.* (2011)] or hyperlens [Li *et al.* (2009)]. Concerning focussing under the diffraction limit there have been propositions based on the analogue of the optical "Bulls eye" [Christensen *et al.* (2007)], or by using negative index material [Zhang *et al.* (2009)]. Nevertheless, none of those proposals clearly demonstrated subwavelength control of the acoustic waves below the wavelength scale. The only experimental proofs that clearly showed super-focussing comes from the use of an acoustic sink [de Rosny and Fink (2002)] which requires an active source at the focal point, or a proposal which uses a phononic crystal where both the source and the image are in the near field of it [Sukhovich *et al.* (2009)].

While most of groups were focused on the negative index metamaterial, we realized that the high effective property band offered by the locally resonant metamaterial could be used for such a purpose. Indeed, this band is inherently composed of evanescent waves since at a fixed frequency the wavenumber is higher than the freespace one. In the following, using our soda cans metamaterial, we will explain how and under which conditions they can indeed be used for subwavelength focussing. The reader has to keep in mind that

the soda can medium is a very good airborne acoustic example which has been published in [Lemoult *et al.* (2011a)], but similar results have been obtained in underwater acoustics in the ultrasonic range with a bubble as a unit cell [Lanoy *et al.* (2015)], with Lamb waves in a thin plate [Rupin *et al.* (2015)], in electromagnetics in the microwave regime [Lemoult *et al.* (2010, 2011b)], and in the optical range [Lemoult *et al.* (2012); Pierrat *et al.* (2013)], making the approach both robust and general.

5.3.2. *Eigenmodes of a two dimensional array of soda cans*

The infinite medium described in the previous section is not realizable and we will see in this section that the finite dimension of a real medium plays a crucial role for the subwavelength focussing application. So, let us first start by the experimental measurments that we have been able to perform. We conducted experiments on two-dimensional array of 7×7 closed-packed square lattice of soda cans as shown in Fig. 5.3(a). The medium is surrounded by a set of 8 computer controlled speakers, while a motorized microphone is placed on top of the array of Helmholtz resonators. One speaker emits a short pulse (actually we used chirped emission and applied matched filtering to recover a short pulse) and the microphone placed 1 cm on top of the aperture of one can records the temporal signal. The typical temporal signal obtained is shown in Fig. 5.3(b). It extends over hundreds of milliseconds compared to the initial pulse duration of 20 ms which is a clear signature of resonant effects. The signal's spectrum (Fig. 5.3(c)) reveals that they actually correspond to the existence of many resonant peaks ranging from 250 Hz up to the resonance frequency of a single resonator, that is $f_0 = 420$ Hz.

We then repeat the same experiment when each speaker emits a short pulse and we move the microphone on top of the array. Knowing the entire set of temporal Green's functions relating the pressure field at a given position in the medium related to the emission from one of the speakers, we can then mimic different monochromatic experiments. We show in Fig. 5.3(d) the monochromatic field

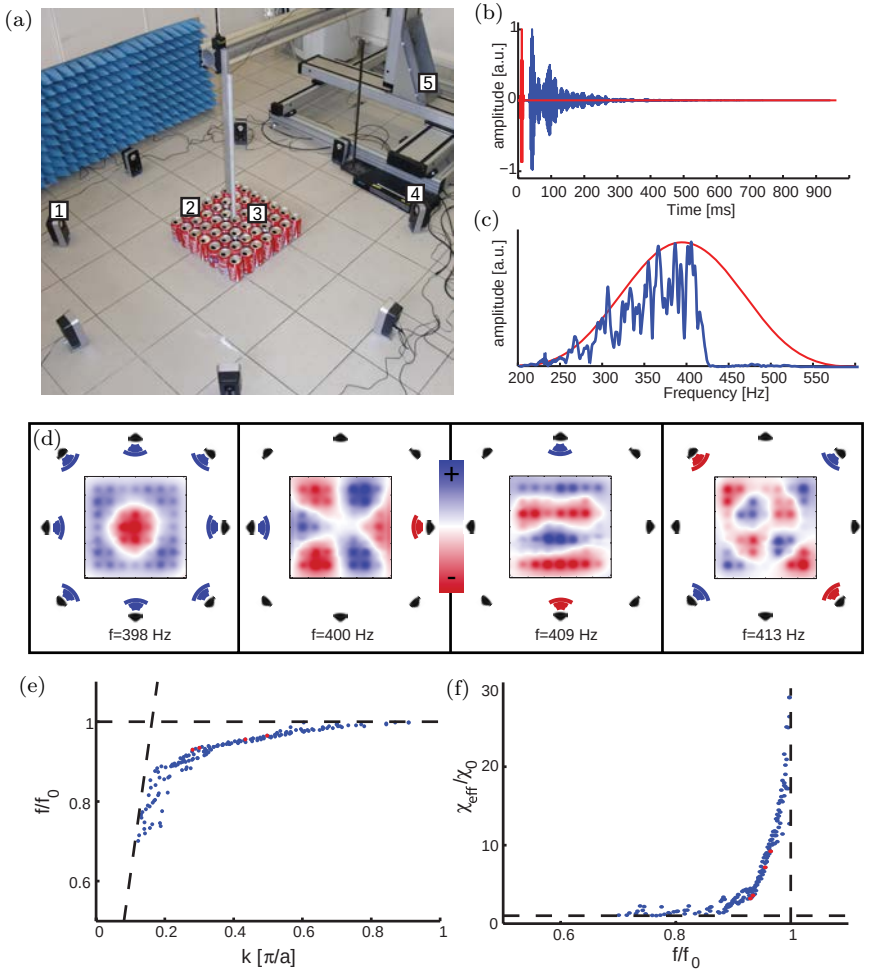


Fig. 5.3. (a) Experimental setup: 8 commercial computer speakers (1) are controlled using a multi-channel soundcard (4) and create sounds exciting the soda cans metamaterial (2). Mounted on a 3D moving stage (5), a microphone records the pressure field. (b) Typical emitted pulse (red) and the measured signal (blue) on top of one can and their corresponding spectra (c). (d) Measured monochromatic patterns at distinct frequencies and with different emitting patterns. (e) Experimental dispersion relation and its equivalent in terms of effective compressibility (f).

maps at various frequencies and with different emission patterns (a monopolar, a dipolar along the x direction, a dipolar along the y direction and a quadrupolar). All of those maps clearly show the subwavelength nature of the modes supported by the medium. For example the first mode shows two nodes of the field while the entire dimension of the medium is roughly $\lambda/2$.

The measured bank of data does not limit to those 4 modes but covers the whole frequency range so that we can extract the dispersion relation out of it. To do so, the wave field for each frequency is spatially Fourier transformed to retrieve an effective wavenumber for the corresponding mode. The result of such a treatment is summarized in Fig. 5.3(e). The experimental dispersion relation exhibits the expected polariton behavior that has been introduced in the first section. One can note that we measured wave fields oscillating on a scale as small as the medium's lattice. From this dispersion curve, we can then extract the effective compressibility as a function of frequency (Fig. 5.3(f)) which is in very good agreement with the general behavior of locally resonant metamaterials: we see the high effective compressibility below the resonance frequency f_0, and we note the absence of modes above it. The latter has two origins: in the bandgap regime there is obviously no propagating wave so no mode actually exists. For higher frequencies however we should have observed supra-wavelength modes corresponding to the anti-binding branch of the polaritonic dispersion relation, but the latter actually falls within the propagating cone and correspond to leaky modes that radiate acoustic waves out of plane.

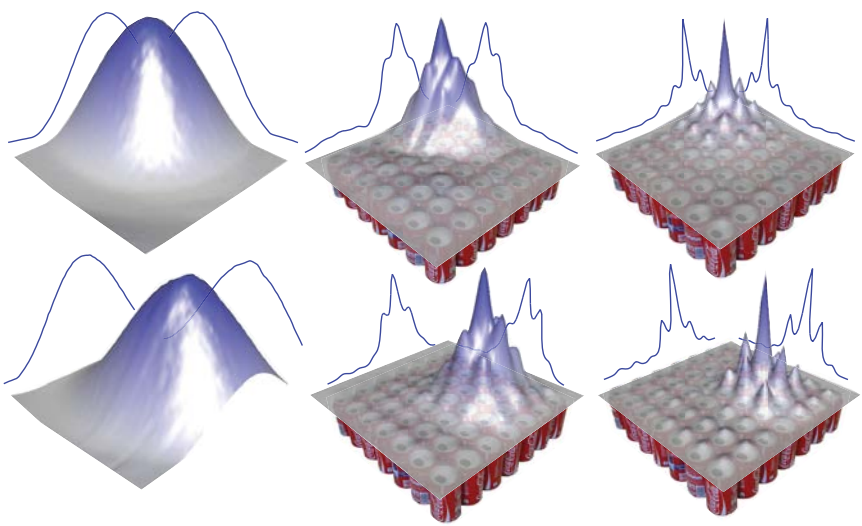
Interestingly and quite surprisingly, the homogenization procedure that applies for an infinite medium can be probed on a physical two-dimensional metamaterial which typical dimension is only half a free space wavelength. Because of the finiteness of the medium, however, other phenomena have to be discussed. Indeed, as the typical spectrum shown in Fig. 5.3(c) the finite size medium only supports a discrete set of resonant eigenmodes. This has several consequences that are the key features for subwavelength focussing. First, because of the finiteness, the supported eigenmodes are stationary modes

trapped inside the medium and therefore they experience a resonance phenomenon. Second, such collective eigenmodes can leak some energy through radiative damping to the farfield area, and this is the reason why we have been able to probe them when doing the reciprocal operation, namely exciting from the far-field and probing in the near-field of the medium. As a consequence, the high effective compressibility which results in the existence of subwavelength (with respect to the freespace wavelength) modes cannot been used as it in order to build a focal point: exciting the medium with a single monochromatic source placed in the far-field one will only excite the combination of eigenmodes that can radiate energy in the direction of the source, and for our small medium (i.e. $\lambda/2 \times \lambda/2$) it reduces to 4 modes (monopolar, dipolar and quadrupolar radiation patterns). Furthermore, one has no control over the relative phases between the 4 trapped monochromatic eigenmodes by using only one source. Increasing the number of sources allows the manipulation of the relative phases between the different types of radiation patterns but actually it is not sufficient enough to be able to focus everywhere on the sample. This strategy has been adopted in [Maznev *et al.* (2015)] where the authors demonstrated the ability to build a focal hot spot in the middle of the sample that is thinner than the free space wavelength. If we want to focus everywhere in the sample we need to have access to all the different spatial scales which is only realizable by taking advantage of the dispersion of the medium: at one frequency the spatial variation of the field is given by the dispersion relation. Consequently, as we will see in the next section we proposed to use a polychromatic approach, namely time reversal [Fink (1997)], in order to treat coherently the spatio-temporal degrees of freedom [Derode *et al.* (2001); Lemoult *et al.* (2009); Mosk *et al.* (2012)].

5.3.3. *Subwavelength focussing from the farfield*

Before entering the case of the soda can medium, let us start by a control experiment while removing the Helmholtz resonators' array. The latter consists in recording with the microphone placed at a given position the set of 8 Green's functions when each of the speaker emits

a short pulse. We then time reversed each of those signals and simultaneously reemit them by their corresponding speaker, meaning that each speaker emits first what arrives later in time and reciprocally. Note that in this experiment, we also take advantage of the spatial reciprocity of the wave equation since we did not place a source at the position where we want to focus, but we learn the Green's functions from the speakers to the microphone. We then map the wave field generated around the initial microphone's position by moving it (and repeating the simultaneous emission each time the microphone has moved). As a result of such a procedure, we plot the square of the maximum in time of each received signal. This is equivalent to show the maximum power received on each position. The result of such an operation is displayed in Fig. 5.4(a). The obtained field map show that 8-channel time reversal in a typical laboratory room permits to focus wave on isotropic focal spots whose dimension is



(a) Control (without cans):$\lambda/2$ (b) Time reversal: $\lambda/8$ (c) Iterative time reversal: $\lambda/25$

Fig. 5.4. Sub-diffraction focusing of sound. (a) Diffraction limited focal spots obtained using time reversal without the array of cans. (b) The foci obtained using time reversal onto the same locations with the array of Helmholtz resonators. (c) Foci obtained with inverse filter signals demonstrating focal spots as thin as $\lambda/25$.

half-wavelength, only limited by the diffraction limit [Lord Rayleigh (1879)].

We then perform the same procedure with the cans array. This time, the set of emitting signals span a longer time range as a signature of the modes resonances. The time reversal permits to synchronize at the desired location and at the desired time all of the eigenmodes that are non-zero. This results in a spatio-temporal focusing of the acoustic waves. The measure of the focal width is as thin as $\lambda/8$ (Fig. 5.4(b)), far beyond the diffraction limit. This is a consequence of the subwavelength nature of the modes inside the metamaterial. This experiment clearly shows that thanks to the resonant nature of the eigenmodes trapped inside the metamaterial, the conversion from free space wavelength to subwavelength varying fields is guaranteed. This permits us to beat the diffraction limit from the far field by focussing waves on a thinner scale than the free space wavelength.

But, even if this is a breakthrough, we have not reached the limit of the device since we initially probed eigenmodes that oscillate on scales as thin as the distance between two cans. This limitation comes from the fact that time reversal does not compensate for losses during the propagation. It only recombines the different frequency components by coherently adding them — they all add in phase at the focus — but does not play any role on their relative amplitudes. And we know that the modes that suffer most from the losses are the highest Q ones, or equivalently the ones with the smallest group velocity, which exactly corresponds to the most subwavelength ones. So, in order to circumvent this issue, we need to compensate for the losses by increasing the relative weight of these modes at the focus. To do so, we proposed to build signals that are the equivalent of an inverse filter [Tanter *et al.* (2001)]. This procedure first requires the knowledge of the set of all impulse responses between the 8 speakers and each of the desired focal positions on top of the resonators array, which we limited to 49 as the number of cans. We then numerically computed a bank of 8×49 signals based on an iterative scheme of time reversal [Montaldo *et al.* (2004)] that supposedly focus on each

position with the lowest possible level of spatio-temporal side lobes level. We then use 8 of those signals and simultaneously emit them with the speakers. Eventually, we map the wavefield on top of the cans while emitting those signals and we end up on the result shown in Fig. 5.4(c) with focal spots as thin as $\lambda/25$. Of course, because we cannot focus waves in between cans, the focussing resolution is actually limited by the period of the medium. Overall, we prove that we can beat the diffraction limit by a factor of 12.5 with a positioning accuracy of $\lambda/12$. We only present two maps but this focussing can be performed on any position on top of a can.

Apart from its evident fundamental interest, this experiment opens up many avenues in terms of applications for sound and ultrasound. We believe that our approach is very promising for the design of arrays of actuators, micro-mechanical actuators in general and by reciprocity, of sensors. Indeed, using subwavelength coupled resonators offers three tremendous advantages. First, it introduces the possibility to engineer matrix of actuators or sensors that are arranged on a subwavelength scale. Second, because our approach takes advantage of dispersion, it allows to address independently many sensors using their temporal signature. Finally, as we will prove it next, it also enhances the intensity deposited onto one location, because of the subwavelength dimensions of the focal spots.

As a principle proof of concept, we performed a visual experiment: we suspended a 20-μm-thin sheet of metalized Mylar on top of the array of cans, that is supposedly transparent to acoustic waves because it follows the displacement of the surrounding air (Fig.5.5(a)). We deposited few glass beads (diameter around 120 μm) on top of this sheet. While emitting the time reversal signals (or inverse filters ones) we saw the glass beads moving on the desired subwavelength area on top of the soda cans array. Anyone who visited the laboratory at that time could enjoyed this visual experiment, but translating it into publishable results revealed was a bit challenging.

We used a high frame rate camera in order to catch the beads displacements. A white light projector illuminates the sheet of Mylar,

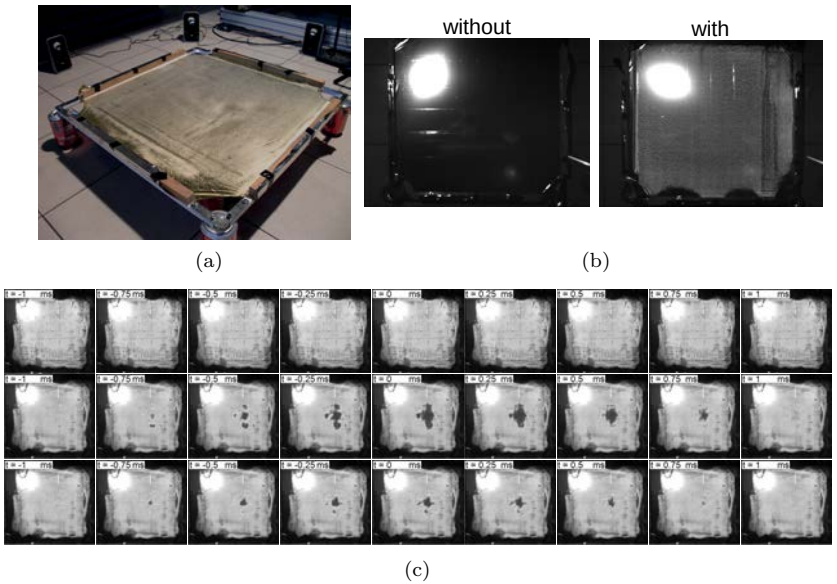


Fig. 5.5. Subwavelength actuators. (a) We added a Mylar sheet on top of the soda can array and deposited some glass beads. (b) Due to a photo-reflective effect the sheet appears dark without beads and very bright when the beads are stuck to the surface. (c) We film the slow motion of glass beads on top of a Mylar sheet for the control experiment without the cans (top) or when creating subwavelength focal spots using time reversal (middle) and our iterative scheme (bottom). Clearly, the beads are in movement solely on subwavelength areas and at a given time, demonstrating the possibilities offered by our approach.

oriented a few degrees from its normal, while the camera is placed exactly at normal incidence. Because of the small angle between the projector and the camera, the Mylar sheet appeared dark except for the direct image of the projectors bulb in the absence of beads. On the contrary, after depositing the beads the Mylar sheet appeared very shiny. This effects occurs thanks to the retro-reflection of the light on the glass beads that are placed on top of a metalized surface. This means that when beads stick to the Mylar sheet the image on the camera appears white, while it reveals darker when the beads are no more touching the sheet (Fig. 5.5(b)).

We have utilized this experimental procedure in order to make high frame rates movies of the field created when we focus onto

various locations on top of the array of cans. Figure 5.5(c) shows few frames from films obtained for the three different types of emissions we previously performed. While nothing happens on the control experiment without cans, one can notice that we have actually been able to darken the image on a small area by using time reversal (and inverse filters signals) during roughly 1 ms. This means that the beads located on a $\lambda/12$ squared area jumped on top of the Mylar sheet and this has been controlled from speakers placed in the far field of this medium. This entails that our approach can be exploited for subwavelength size actuators and microelectromechanical systems [Lani *et al.* (2016)].

5.4. Hybridization Bandgap: Subwavelength Trapping and Waveguiding

5.4.1. *Existence of the hybridization bandgap*

Let us now move to a slightly higher frequency range, just above the resonant frequency of a single soda can. In the previous experiment, we were not able to measure any wavenumber because of the absence of resonant eigenmodes, and therefore we were not able to extract the effective compressibility this way. In order to show what happens in this spectral domain, we represent in Fig. 5.6(a) the sketch of the new experiment that is conducted. A speaker placed in the far-field emits a short pulse that spans the frequencies ranging from 200 Hz to 600 Hz and a microphone measures the signal received on top of one can. The spectrum represented in a logarithmic scale (Fig. 5.6(b)) clearly shows two distinct regimes. Below the resonant frequency of one can (blue shaded area) the transmission is high and shows the existence of the resonant peaks that we discussed in the previous section; Above the resonance, the transmission drops down to -60 dB and almost no energy is transmitted from the speaker to the microphone. This frequency region is therefore associated to a so-called band gap (red shaded area). One has to keep in mind that the overall dimension of the soda can remains smaller than the free space wavelength meaning that the attenuation length is really short. In order

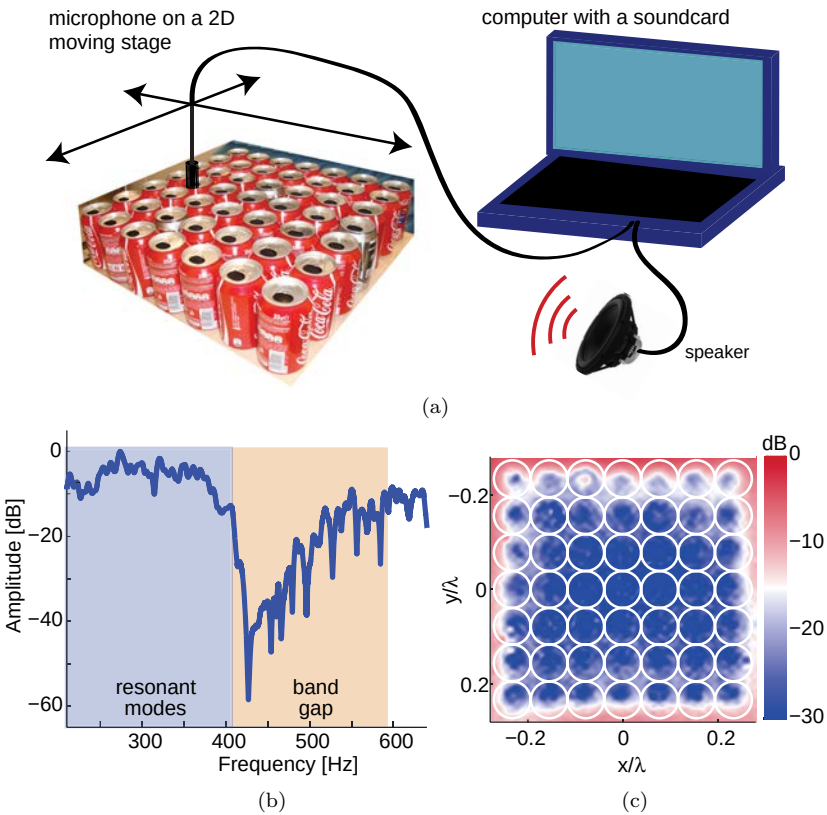


Fig. 5.6. (a) Single channel experimental setup: the sound-card of a computer is connected to a single speaker placed in front of the metamaterial and to a microphone that is mounted on a 2D moving stage and positioned above the sample in its near field. (b) Transmission measured between the speaker and the microphone when the latter is placed in the middle of the acoustic metamaterial. The transmission is normalized to unity. Below 420 Hz, resonant deep subwavelength modes can be observed while above a dip asymmetric band gap can be identified. (c) Map of the absolute value of the pressure field at 450 Hz in logarithmic scale. No field can penetrate inside the metamaterial and after only one layer an attenuation of 20 dB is observed.

to further investigate this attenuation effect, we mapped the field at a frequency within the bandgap (450 Hz) while the loudspeaker emits. We represent in Fig. 5.6.c the absolute value of the measured monochromatic field in decibel scale again. This map clearly shows

that the field cannot penetrate inside the soda can metamaterial at this frequency and that the attenuation length is really short since after only one layer of cans the transmission has been reduced by one order of magnitude.

This bandgap was actually already introduced initially when we mentioned the polariton dispersion relation. In terms of effective parameter, the existence of the bandgap is associated to the negativeness of the effective compressibility which results in a complex propagating velocity, that is, evanescent waves. We have to mention here that in the acoustic community this type of bandgap has been observed before the emergence of metamaterials and in the early 1990s researches referred to it as "hybridization bandgap" [Sigalas and Economou (1993); Psarobas *et al.* (2002); Penciu *et al.* (2003); Leroy *et al.* (2009b,a); Cowan *et al.* (2011)]. It results from the interaction of the propagating waves and the waves re-radiated from the resonant unit cell: it is the result of destructive interferences between those two waves. We have therefore reinterpreted this bandgap in terms of Fano interferences [Fano (1961); Miroshnichenko *et al.* (2010)] since it is clearly the interaction between a continuum of modes (the free space waves) and a local resonator [Lemoult *et al.* (2013)].

5.4.2. *Creating a defect within the hybridization bandgap*

This kind of bandgap that occurs in the low frequency regime, i.e. when the wavelength is large compared to the typical distance between resonators, presents an important difference compared to the well documented ones in the context of phononic crystals [Kushwaha *et al.* (1993); Sigalas and Economou (1993); Martinezsala *et al.* (1995); De Espinosa *et al.* (1998); Yang *et al.* (2002); Page *et al.* (2004); Sukhovich *et al.* (2008)]. In the phononic crystals, the forbidden propagating band is due to Bragg interferences which are inherently the consequence of the periodic nature of the medium: the destructive interferences occur thanks to the periodic patterning of the medium. As a consequence, the proposed applications

exploiting the forbidden propagation consists in locally breaking the translational symmetry of the medium. Indeed, when doing so, solely evanescent waves can penetrate from this location on small distances inside the crystal. For example, a point defect can be created in a phononic crystal by locally removing a scatterer. This results in a small cavity because a resonant mode is created by the defect within the band gap. Following this concept, many components have been demonstrated based on periodic media such as waveguides using line defects [Joannopoulos *et al.* (2011); Kafesaki *et al.* (2000); Khelif *et al.* (2003); Miyashita (2005)] and the envisioned applications span a large amount of domains from optronics [Chutinan and Noda (1999)] to light matter interactions [Painter *et al.* (1999)]. However, because of their wavelength scale period, phononic crystals result in relatively large devices. This seriously restrains the range of applications, specifically in the low frequency regimes where the wavelength is large. Contrary to Bragg interferences based bandgap, the hybridization one, which originates from the destructive interferences between the resonant response and the incident wave, is robust to a spatial disorder: breaking the translational periodicity of the medium does not close the bandgap [Penciu *et al.* (2003); Cowan *et al.* (2011); Kaina *et al.* (2013)]. As a consequence we cannot use the trick of breaking the periodicity of the medium as performed in phononic crystals in order to create a defect state.

In order to illustrate the strong difference between those Bragg bandgaps and the hybridization one observed in the soda cans medium, we ran a set of simulations where we created a defect by removing one can inside the array. We show in Fig. 5.7 (middle) the spectrum of the transmission between an incident plane wave and the location where we removed the soda can. This one has to be compared to the one of the case without the defect in Fig. 5.7 (left). No significant difference between the two spectra in the bandgap is observed, a result that is further confirmed by the similitude of the pressure field maps recorded at 440 Hz. This makes a big difference compared to Bragg based bandgaps: the defect that we created this way is actually too small to support a resonant mode, while in a

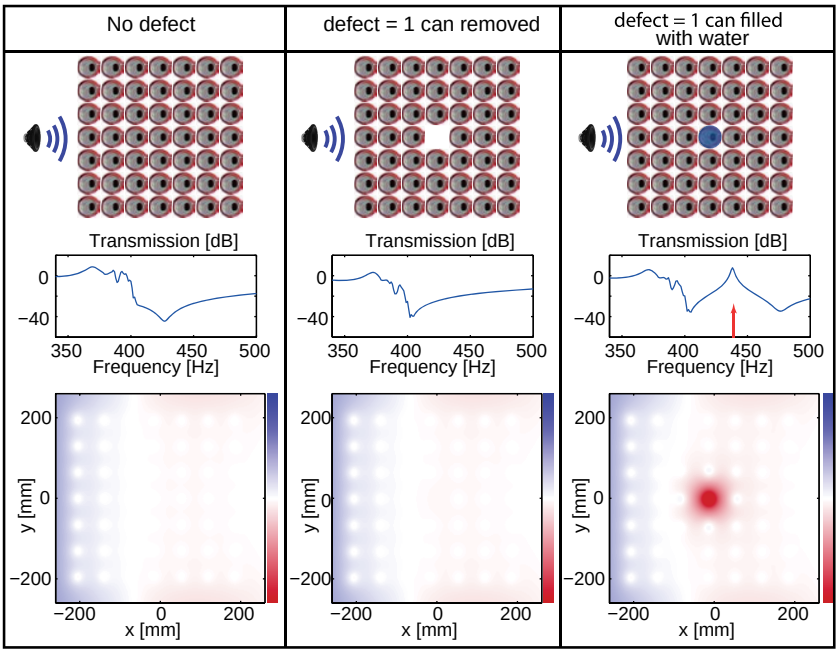


Fig. 5.7. Simualtions corresponding to 3 different configurations: the regular 7 soda can array (left), an array in which we removed the central can (middle) and an array where volume of the central can is reduced by 6 cL. For each, we show the spectrum of the transmission 1 cm above the array and the pressure field map at the frequency of 440 Hz. While removing a can cannot create a resonant cavity, introducing a detuned resonator permits to create a $(\lambda/15)^2$ area cavity.

phononic crystal removing one scatterer allows the existence of a stationary defect mode since the typical scale of such a defect is the wavelength. In order to tackle this issue, we actually have to physically add a resonant defect inside the soda can medium, or in other words to introduce a detuned resonator compared to the rest of the medium. In the case of the soda can medium, this is fairly easy to build this defect resonator: the Helmholtz resonance of a single can is parameterized by the volume of the air cavity. So, we can create a detuned resonator just by changing this volume and we do so by filling the can with a few centiliters of water. It results in a resonance frequency that is upward shifted, thus falling within the forbidden

band. The simulation corresponding to a water volume of 6 cL clearly shows that this creates a defect: the spectrum of the transmission shows a resonant peak near 440 Hz and the corresponding field map displays that it corresponds to a resonant mode confined at the defect position (Fig. 5.7 (right)).

The existence of this defect mode deserves few comments. First, given the very small spatial scale of the soda can medium, one would expect very strong near field interactions between the resonators. Yet the polariton origin of the bandgap, that is the coupling of a freespace plane wave and a local resonance, actually proves that the dispersion in this deep subwavelength scaled medium is solely due to interference effects which therefore guarantees the existence of this defect mode. Second, because the unit cell is very small compared to the freespace wavelength the modal volume of the confined mode in this medium is deeply subwavelength. Third, while the quality factor of a single can is around 10 it reaches 80 for the defect mode, solely limited by the viscous losses. This comes from the fact that the can filled with 6 cL of water cannot radiate waves toward far-field, thus canceling any radiative damping of its intrinsic resonance. Those two effects, namely a high quality factor and a low mode volume, are typical quantities that opticians are researching to enhance the emission rate of an emitter and it is known under the name of Purcell factor [Purcell (1946)]. Here, we have shown with an acoustic example that high Purcell factors are attainable in metamaterial defect cavities. Similar results in microwaves [Lemoult *et al.* (2013); Kaina *et al.* (2013)] have been obtained that permit to be confident on the fact that the optical community can certainly reach this goal.

5.4.3. *Molding experimentally the flow of acoustic waves at a subwavelength scale*

We now come back to the experimental setup described in Fig. 5.6(a) and we use it to experimentally prove that it is possible to control the waves at the subwavelength scale by introducing defects similarly as we numerically performed. Obviously, the first experiment consists in reproducing the point defect. So, we locally introduce a

defect state by introducing 6 cL of water in the central can. While emitting with the loudspeaker, we use our microphone mounted on the 2D moving stage to map the field on top of the array. The real part of the monochromatic map of the field at the defect resonant frequency is given in Fig. 5.8(a) as well as the absolute value profile along the dashed line. Clearly, we observe that the pressure field is localized within one can of the array, meaning that waves have tunneled through the metamaterial, and filled the cavity. The effective cavity size of $\lambda/15$ is much smaller than in any realized phononic crystal. And again, the measured quality factor for the cavity is around 80, in very good agreement with the numerical results.

We now concentrate on a second type of defect that probably can lead to more applications in acoustics: a line defect. To that aim, using a 5×12 cans array, a subwavelength line defect is obtained by filling the soda cans of the central line with 6 cL of water. Now, instead of having a resonant cavity we actually create a subwavelength waveguide in the soda can array. Again, we measure the acoustic field distributions on top of this array. The map of the measured acoustic field for a frequency of 447 Hz is presented in Fig. 5.8(b), as well as the transverse profile of the waveguide mode. This demonstrates a $\lambda/14$ wide waveguide and we measure an attenuation transversely of 25 dB after the first row of cans. This kind of waveguide is therefore extremely confined on the defect line at the operating frequency. We cannot present all of the results here but similar propagating behavior has been obtained for frequencies around the presented one, except that the effective wavelength within the waveguide is changing while changing the frequency. The bandwidth is centered on the defect's resonance frequency, which is typical of a tight binding type of coupling. Indeed, the filled cans creating the waveguide are embedded in a medium that does not support any propagating waves, therefore the only channel for the coupling is a tunneling from can to can. Note that we stated initially that the physics of the soda can medium can be described by neglecting any near field coupling because the picture of the polariton catches all of

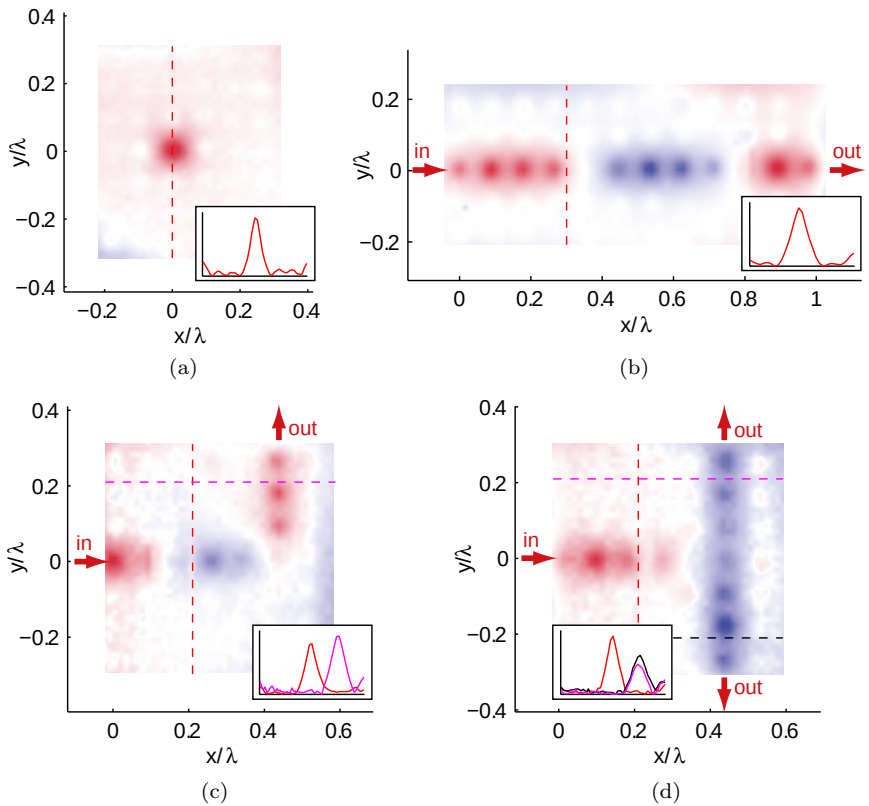


Fig. 5.8. Molding the flow of acoustic waves at the deep subwavelength scale. Spatial distributions of the amplitude of the acoustic near field of the sample when (a) the can positioned at (0,0) is filled with 6 cL of water, resulting in a $(\lambda/15)^2$ area cavity at 465 Hz, (b) a line of partly filled Helmholtz resonators at $y = 0$ constitutes a $\lambda/14$ wide waveguide for acoustic waves at 447 Hz, (c) a 90° curved similar waveguide permits to bend with unity efficiency acoustic waves on a $\lambda/14$ distance at 451 Hz and (d) a T-shaped waveguide splits acoustic waves into two identical arms with equal amplitudes at 450 Hz. The insets show the profiles of the magnitude of the fields along dashed colored lines.

the physics. Though it might seem contradictory with this previous statement, note that the propagation mechanism is here constrained by the presence of the hybridization bandgap. As no wave is allowed to freely propagates the only mechanism for coupling two neighbours is a tight-binding model.

Using the same experimental protocol, we now move on to more complex components to manipulate waves. We designed a corner waveguide and a splitter by inserting those partly filled Helmholtz resonators in a 7×7 array of empty cans. Maps of the spatial distributions of the fields in those samples are presented in Fig. 5.8(c) and (d) alongside the profiles of their modulus in the directions of interest. From those new experiments, we make a series of conclusions. First, one can force the guided waves in any given direction, with no specific engineering of the material structure, owing to the resonant nature of the unit cell. In our case, we bend the waves 90° with unity efficiency and within a $\lambda/14$ long unit cell. Actually, because a soda can has an isotropic radiation pattern, there is no difference between the 90° bend and the linear waveguide. This is in great contrast with designs based on designers plasmons for instance [Martin-Cano *et al.* (2010)], where scattering impairs seriously the ability of those structures to bend waves within small propagation distances. Here, since the waveguide lies in a band gap material, no scattering occurs. Second, we can use our approach to split waves into two arms, which ensures the possibility to realize deep subwavelength interconnections and routing of acoustic energy.

As a conclusion, the presented approach gives unprecedented solutions to manipulate acoustic waves, especially those for which the wavelength seems large compared to the envisioned applications. It then paves the way to the design of ultra compact components. We insist here that this chapter is dedicated to the soda can medium but this approach is very general and applies to any locally resonant medium with neglectable near field coupling. Indeed, in our initial paper [Lemoult *et al.* (2013)], we demonstrated very similar results in the context of electromagnetic waves where the unit cell was a half-wavelength-long metallic wire. We also performed experiments in microwaves where the subwavelength resonators were placed in a disordered way: locally introducing a detuned resonator in such a medium still creates a very confined cavity [Kaina *et al.* (2013)]. Eventually, parametric experiments again performed in the microwave range demonstrate very interesting applications for this

kind of waveguides: thanks to the S-shaped dispersion relation the waves can travel very slowly with unprecedented bandwidth-group index product [Kaina *et al.* (2016)].

In the same time, those results raise interesting questions regarding the physics of metamaterials. Indeed, with the conventional homogenization approach one cannot predict what happens inside metamaterials containing one or several defects because this procedure averages the response of all the resonators. For example, a straight waveguide or a 90° bending one present the same number of defects but give completely different results. Our microscopic approach nevertheless permits to understand that the interferences phenomena occurring at the deep subwavelength scale of the metamaterial are actually very similar to the ones occuring at the wavelength scale in phononic crystals.

5.5. Spatial Structuration and Multiple Scattering Lead to Negative Refraction

5.5.1. *Breaking locally the symmetry*

In this section, instead of introducing localized defects we decide to engineer locally the material with local defects that respect a given translational symmetry of the metamaterial. Indeed, the previous experiments have highlighted the fact that local modifications of the metamaterial gives rise to interferences effect at the scale of the material no matter what is the free space wavelength. Therefore we now propose to slightly break the periodic nature of the initial metamaterial either by changing the resonance frequency of one resonator out of two, or by slightly off-centering one resonator (Fig. 5.9). Consequently, the unit cell of the medium becomes a dimer made of two resonant elements. This type of new cells, because of their dimeric natures, should support a dipolar resonance overlapping with the monopolar one that gives the metamaterial its effective macroscopic property.

For the sake of simplicity, we start with the simplest example of locally resonant metamaterial which consists in a one

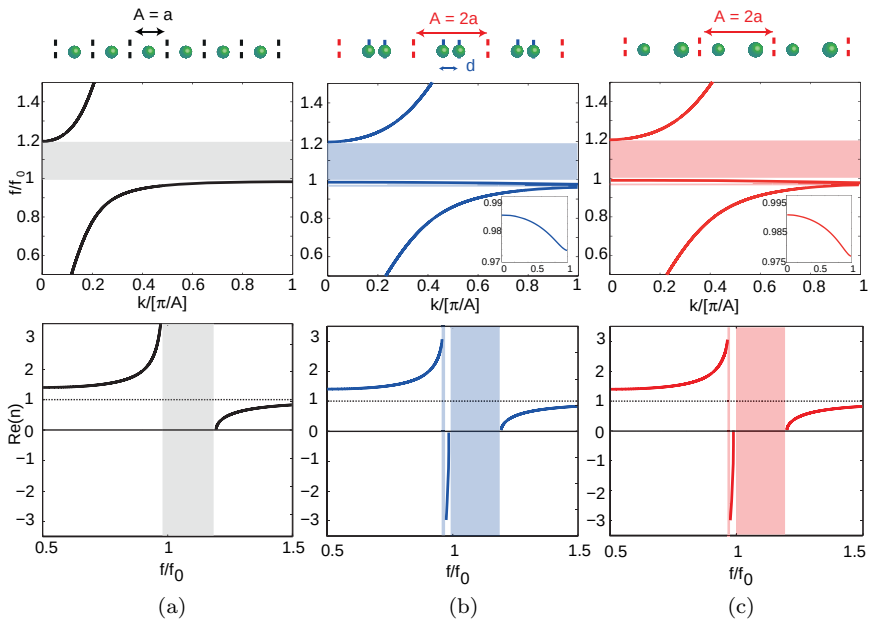


Fig. 5.9. (a) Typical polaritonic disperion relation for a periodic arrangement of resonant point scatterers (mathematical equivalent of the soda can) or its representation in terms of an effective index of refraction. By locally breaking the periodicity of the chain either by off-centering one resonator out of two (b) or by changing its resonant frequency (c) we ends up on dispersion relations showing a negative band (the effective index of refraction is negative).

dimensional chain of point scatterers organized periodically on a deep subwavelength scale. The dispersion relation of such a chain of resonant unit cells, with $a = \lambda_0/12$, is calculated analytically using a combination of a Green's function formalism and a transfer matrix approach [Kaina *et al.* (2015); de Vries *et al.* (1998)]. As seen before, the obtained dispersion is typical of the soda can array already studied and exhibits the polaritonic behavior (Figure 5.9(a)). Or again, equivalently, the medium can be described by a set of two effective parameters with only one of them being negative. It then can be named as a single negative metamaterial. This one dimensional chain physically behaves as the soda can medium which can be characterized with a negative compressibility.

From this single negative medium, two new configurations are created just by breaking the symmetry in two different ways: either a bi-periodic chain is built by off-centering one resonator out of two, or a bi-disperse one is generated by slightly shifting the resonance frequency of half of the resonators. In both cases, we again analytically calculate the dispersion relation and extract the corresponding effective refractive index (Figure 5.9(b) and (c)). Both of the new symmetry broken metamaterials now exhibit, in the bandgap of the single negative medium, a new propagating band that is characterized by a negative phase velocity.

5.5.2. *Physical origin of the negative index branch*

To understand the origin of the negativeness of the band, we parametrically study those two symmetry-broken formal metamaterials. In the case of the bi-periodic chain, the parameter is the shift in position, while for the bi-disperse one the frequency detuning between the resonators is modified. For each parameter, we analytically calculate the new dispersion relation with the same approach that takes into account multiple scattering and extract an effective index out of it. The results are shown in a colour-coded maps (Figure 5.10). For the bi-periodic chain, we note that the index of the new band is negative whatever the shift in position. In the bi-disperse case, however, this negative band only appears on a narrow frequency detuning range. This suggests that multiple scattering may be involved since it seems that the resonances due to each type of resonators have to overlap. To confirm this intuition, we extract for the same set of parameters, the effective index only by considering the independent scattering approximation at the scale of the new unit cell made of two resonators. Namely, we do not consider the Fabry-Perot resonator that is built on a deep subwavelength scale between the two resonators. From the maps on Fig. 5.10, it is clear that this approximation does not permit to retrieve the existence of a negative index, and positive index bands, typical of a double polariton, appear. This shows that the negative index arises from multiple scattering between the resonators of the unit cell, even if the distance is far below the

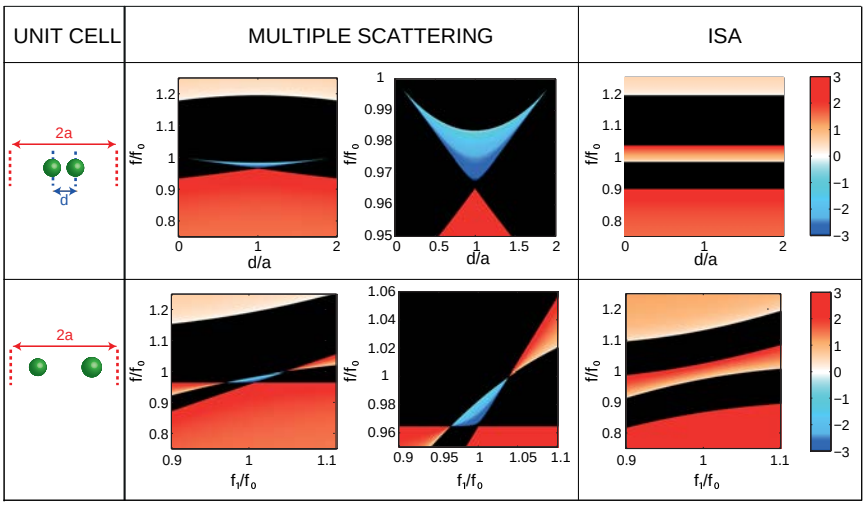


Fig. 5.10. Either for the biperiodic (a) or the bidisperse (b) chains, we extracted the effective index of refraction for various strength of the symmetry breaking with a full multiple scattering approach or with the independent scattering approximation (ISA). While the maps with the full multiple scattering calculation show the existence of the negative band (blue color in the map) the ISA does not retrieve it.

wavelength. This, in turn, explains why, for the bi-disperse chain, the existence of the negative index depends strongly on the chosen detuning: for too large a resonance frequency mismatch, the two resonators cannot couple any longer owing to multiple scattering.

To grasp the physics of the approach, we carefully studied the fields created by a dimer (the new unit cell of the symmetry-broken media) and made the following observation: multiple scattering creates a dipolar resonance (the two resonators are out-of-phase) overlapping with a monopolar resonance of the dimer. This dipolar resonance is responsible for the opening of a narrow transparency window within the large out-of-phase response of the monopolar one. This is analogous to electromagnetic induced transparency in quantum physics [Fleischhauer *et al.* (2005)], or more precisely to its metamaterial equivalents [Papasimakis *et al.* (2008)]. This dipolar resonance results from multiple scattering occurring between the two adjacent resonators. Moving from the unit cell to the infinite

medium, this dipolar mode gives rise to a band of propagating waves within the bandgap of the single negative medium, the latter being a consequence of the monopolar resonance of the unit cell. This band has a negative slope, or equivalently, the metamaterial now presents a negative index. This originates physically from the fact that owing to the symmetry breaking, the lower polaritonic band folds in the first Brillouin zone, analogous to optical branches in diatomic crystals or band folding in phononic crystals. Here, though, the band folding, owing to the change of sign of the Bloch mode between the two edges of the unit cell has a different origin. Indeed, while in phononic crystals, it arises from the fact that the host medium wavelength becomes smaller than twice the lattice constant, in the former the change of sign results from the dipolar nature of the resonant mode within the unit cell. This implies that, contrary to negative refraction in crystals, this new phenomenon exists even if the scale of the metamaterial is deeply subwavelength, and happens at the same frequency as the resonance of the original building block, which we can qualify as the low frequency regime. Furthermore, since the original single negative effective property does not rely on spatial order [Kaina *et al.* (2013)], this negative index band should be robust even in a metamaterial constituted of randomly placed dimers: both the monopolar and the multiple-scattering-induced dipolar resonances should remain, hence leading to a negative index medium [Li and Chan (2004)].

5.5.3. *From 1D to 2D media*

This idea of breaking the symmetry is rather simple in one dimension and before starting experiments on the quasi-bidimensional medium based on soda cans we need to move on to the two dimensional world. Finding a lattice that paves the space and presents a bi-periodicity in all directions is far much complicated than in one dimension. In order to stay as isotropic as possible, we choose the Bravais lattice that presents the first Brillouin zone that is the closest to the circle which is the hexagonal lattice. Then, as we mentioned in the one dimensional case the unit cell needs two resonant objects in order to have the dipolar resonance creating the transparency window within

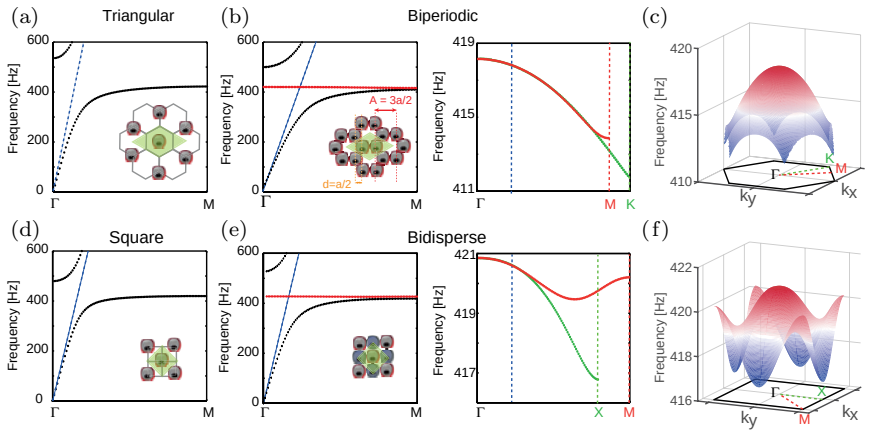


Fig. 5.11. Dispersion relations from numerical simulations with Bloch boundary conditions. (a) a triangular lattice of soda cans displays the polaritonic dispersion behaviour while (b) the honeycomb lattice which consists of the superposition of two triangular ones exhibit the presence of a negative band in the two principal directions of the crystal. (c) The surface plot of this band shows the almost isotropic nature of this band. Similarly, (d) the square lattice of soda cans displays the polaritonic dispersion while (d) a square lattice made of two resonators with two different resonant frequencies shows a negative band. (f) The surface plot for this bidisperse crystal shows less isotropy but near its upper edge it seems isotropic.

the bandgap created by the monopolar one. Therefore, we naturally end up on the well-known honeycomb lattice. Carefully looking at the positions of the Bragg planes one can see that it exhibits this double periodicity in any of the M directions. This "crystal" is made of a diamond-shaped unit cell comprising two resonators, and is compared to the triangular lattice which has the same unit cell but with only one resonator. Numerical simulations using Comsol Multiphysics give the dispersion of both the regular and the symmetry-broken lattices (Figure 5.11(a) and (b)). The triangular lattice medium presents a polaritonic dispersion relation, while the honeycomb lattice (which actually consists of the superposition of two identical triangular lattice crystals) displays a negative band. The dispersion depends slightly on the propagation direction but remains rather isotropic as shown in the surface plot (Figure 5.11(c)) and can thus be described

with an isotropic negative effective index of refraction for almost all frequencies.

For the bi-disperse two-dimensional lattice, we have not found ideal solution to deal with the isotropy issue. We chose to favor simplicity by mixing two square lattices of slightly detuned Helmholtz resonators in order to build a new square lattice whose unit cell contains two resonators (Figure 5.11(d) and (e)). Again, the detuning is really easy to achieve since it is experimentally realizable by pouring some water in the can in order to reduce its volume. This bi-disperse resonant crystal exhibits a negative branch although it is simply the superposition of two almost identical single negative media. In this case, the propagation is less isotropic, since the geometry of a square unit cell tends to deform the isofrequency contours near the corners of the first Brillouin zone as shown from the surface plot describing the dispersion relation of the negative band within the entire first Brillouin zone (Figure 5.11(f)). There is, however, no doubt that one can find a more isotropic medium.

Then, in order to show that a band having a negative slope is actually the signature of a medium exhibiting negative refraction, we now focus on the honeycomb crystal and we perform simulations on a slab geometry. We did not implement it experimentally because it would have required a much larger sample than our rooms dimensions. To avoid a too high number of mesh cells, those simulations only consist in simulating a slab layer of height $\sqrt{3}a$ (where a is the distance between two adjacent soda cans) which corresponds to the vertical unit cell of the honey comb arrangement of soda cans. Periodic boundary conditions are applied on the vertical boundaries to simulate the infinitely extended slab. This slab unit cell is excited by an incident plane wave with an incident angle θ with respect to the normal of the slab, and the phase shift applied on the periodic boundary conditions matches this value. We perform a set of 61 simulations with θ ranging from -90° to 90°. From those simulations, we are then able to build an incident wave field impinging on the slab that corresponds to a Gaussian beam with an incident angle θ_0. To do so, we perform the complex summation of the wave fields

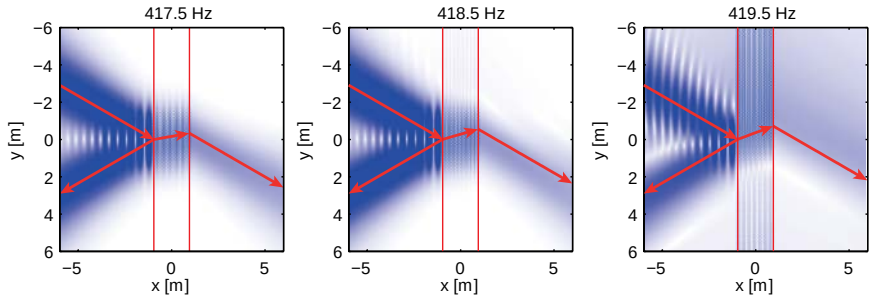


Fig. 5.12. Intensity maps of a simulated Gaussian beam impinging on an infinite slab made of cans without losses with an incident angle of 30°. The superimposed ray tracing corresponding respectively to a refractive index of (a) −3 at 417.5 Hz, (b) −2.4 at 418.5 Hz and (c) −1.5 at 419.5 Hz clearly shows the negative refraction effect.

extracted from each simulation and multiplied by a Gaussian coefficient $\exp\left(-\left(\frac{\theta-\theta_0}{\sqrt{2}\sigma_\theta}\right)^2\right)$, where σ_θ corresponds to the angular aperture of the beam and is chosen to be equal to 9°. The complete field map on positions that are not within the simulation area is built by the use of the Bloch theorem for each incident angle.

Such a superposition of simulations is realized for an incident angle of $\theta_0 = 30°$ and for distinct frequencies in the negative band: 417.5 Hz, 418.5 Hz and 419.5 Hz (Figure 5.12). From the isofrequency contour calculated previously we extract an effective index of refraction (respectively −3, −2.4 and −1.5) and we superimpose on the figures arrows corresponding to the refraction at both interfaces with respect to the Snell laws. Such a ray tracing is in very good agreement with the propagation of the beam through the slab. This simulation first confirms that we have negative refraction since the outgoing beam seems to exit the slab from a y-coordinate that is higher than the coordinate of the incident beam. Second, similar results for different angles, but the same norm of the index of refraction, confirm that the refraction law can be described by an effective index of refraction instead of anisotropic parameters. Third, those results do not show any diffraction order at the exit of the slab as one would expect from periodic media. This is a direct consequence of

the subwavelength organization of the medium: all of the diffraction orders are evanescent at the exit interface. This again legitimates the description of the medium in terms of a negative index of refraction.

5.5.4. *Experimental demonstration: acoustic superlensing*

Such an experiment with a Gaussian beam impinging on a meters-scale slab is not realizable experimentally is not realizable given the required dimensions. So, we decided to prove the index of refraction through an experiment which consists on a point source illumination where the slab should behave as a flat lens.

We again work with the bi-periodic medium, i.e. the honeycomb arrangement of soda cans. We build a slab with 124 cans, surrounded by acoustic absorbers to avoid reflections of sound off the boundaries of the room as shown in the photography of the experiment (Figure 5.13(a)). An 8-cm-wide loudspeaker located approximately 5 cm away from the input interface of the medium is used as the source of sound while 2 microphones mounted on a two-dimensional translational stage measure the acoustic field above the medium. The loudspeaker emits a long chirp ranging from 100 Hz to 800 Hz and the data are treated by Fourier transform in order to obtain the field maps at the desired frequency. Here, we work at the frequency of 417.5 Hz, which corresponds to the frequency close to the lower edge of the negative band, that is, where the effective negative index norm is the highest.

The real part of the field at this frequency is shown Figure 5.13(b) and is hard to interpret due to the dissipation. We thus show the intensity within the slab by numerically compensating for the losses that occur during the propagation in the lens (Figure 5.13(c)). We can clearly distinguish the path for sound refraction, with a focal spot inside the lens, in very good agreement with the Snell law for a metamaterial with an effective index of 3, consistent with the numerical results. On the other side of the slab, in the vicinity of the surface, we record the image of the source with a $\lambda_0/15$ full-width at half-maximum (Figure 5.13(d)). This is much smaller than

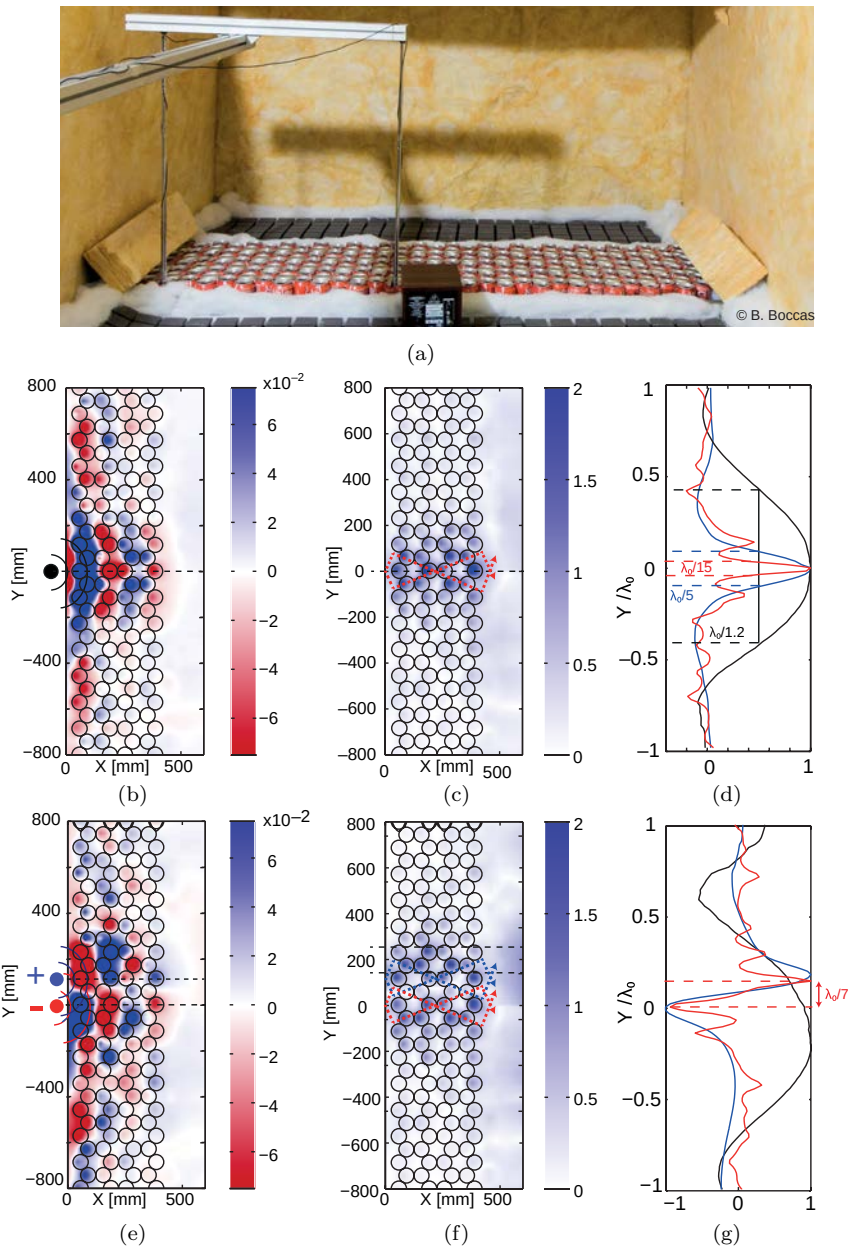


Fig. 5.13.

←

Fig. 5.13. (figure on facing page). Experimental demonstration of subwavelength focusing and imaging using a flat acoustic lens. (a) A photography of the experimental setup: the flat lens, composed of a compact honeycomb arrangement of soda cans, is insonified by a loud speaker placed close to the surface of the medium. Two microphones mounted on a 2D moving stage record the acoustic pressure field less than 1 cm above from the top of the cans. Absorbers surround the lens to prevent from undesired reflections. (b) The real part of the pressure field at 417.5 Hz and (c) its absolute value while compensating for the losses due to the propagation within the lens. The direction of the refracted beams is highlighted with the dashed arrows displaying the features of the negative refraction. (d) The normalized amplitude of the field in the close vicinity of the output surface proves a focusing area of $\lambda_0/15$ (red) while the source (blue) is $\lambda_0/5$ wide and the control experiment (black), that is without the lens, $\lambda_0/1.2$ wide. The same experiment is conducted with two sources playing sounds out of phase to demonstrate super-resolution (e-g). It clearly proves the same negative refraction results with a resolution of $\lambda_0/7$

the diffraction limited focus obtained without the lens (black curve), and even smaller than the width of the source, $\lambda_0/5$, owing to a hotspot created by the aperture of a single soda can: not only this demonstrates the negative refraction property of the medium but it also proves a superlensing effect. We stress here that this superlensing effect owes to the high norm of the effective index of refraction that allows the propagation of waves within the slab that are evanescent in air as well as the negative refraction occurring at both interfaces.

Super-resolution can also be demonstrated by being able to discriminate two sources separated by less than half a wavelength. Two loudspeakers, emitting out-of-phase, are placed near the input interface and separated by 13 cm ($\lambda_0/7$). The measured pressure field, as well as the loss-compensated intensity maps (Figure 5.13(e) and (f)), show that the slab produces two distinguishable foci inside the superlens. In the focal plane, the two images are efficiently separated (Figure 5.13(g)), thereby demonstrating a $\lambda_0/7$ imaging resolution, far beyond the diffraction limit. On the contrary, the control experiment realized without the superlens (black curve) only displays diffraction limited spots, again proving the focusing properties stem from the negative index flat medium. We have further verified

that the two sources can be distinguished whatever the phase shift between them [Kaina *et al.* (2015)].

As a conclusion of this part, we have demonstrated that it is fairly easy to build double negative media starting from a single negative one. Breaking the symmetry of the unit cell of a single negative medium (either by changing the spacing or by adding a frequency detuning), multiple scattering of waves guarantees the existence of an overlap between a dipolar resonance and a monopolar one. This ends on a negative effective index of refraction when considering the infinite medium. This approach is very general and brings a new paradigm to the physics of metamaterials since multiple scattering is often neglected owing to the subwavelength spatial scale of those media. This proves that not only the nature of the scatterers is important but also their spacial ordering: this therefore opens the realm of metamaterial crystals.

5.6. Conclusion

In this chapter we have exemplified, through the acoustic example of the soda cans medium, that the physics of locally resonant metamaterials is very rich thanks to interferences effects occurring at a deeply subwavelength scale. Because the soda can interacts resonantly with the continuum of freespace plane waves, the coupling results in a polaritonic behaviour. This is well explained by the quantum picture of the polariton and in a wave physics like description it is easily understandable by Fano like interferences. Indeed, as for any resonance, the waves scattered by a soda can experience π phase shift near the resonant frequency. This results in a dispersive response that adds additional phase to the transmitted wave. Depending on the frequency range hypersonic modes, a bandgap or leaky modes are then observed.

By using the modes that are below the cone shape of the freespace dispersion relation, i.e. the lower band of the polaritonic dispersion relation, we have demonstrated that one can focus acoustic waves below the freespace diffraction limit. Indeed at a given frequency those modes oscillate at a scale much smaller than the freespace

wavelength and by using a finite-sized medium, those modes are trapped and radiate some energy towards farfield. Reciprocally, by smartly exciting simultaneously those modes we can build interferences that are constructive only at a desired position and at a given time, thus resulting on focal spots that have dimensions smaller than the freespace wavelength.

Exploiting the band gap that occurs at low frequency, i.e. for freespace wavelength much bigger than the average distance between unit cells, one can mold the flow of acoustic waves at a scale that is completely independent of the wavelength. This is implemented by locally introducing resonant defects which have a resonant frequency falling into the bandgap created by the others. In the context of the soda can medium it is really easy to realize since diminishing the volume of the can by adding few centiliters of water at the bottom increases its resonant frequency. We therefore demonstrate trapping, waveguiding, bending and splitting abilities at the subwavelength scale without much engineering.

Eventually, we showed that observations made in the context of the phononic crystals which are wavelength scale media also occur at the deep subwavelength scale of the soda can medium. Indeed, we demonstrated that multiple scattering also occurs on scales that are much smaller than the freespace wavelength. The accumulated phase on a wavelength scale in phononic crystals is replaced by the dispersive response of resonant objects. As a consequence, breaking the translational symmetry of the medium, highlighted with a biperiodic medium and a bidisperse one, results in new propagation properties such as the negative index of refraction. This in turn creates many opportunities in the context of metamaterials because we can now take advantage of the spatial structure in order to create new macroscopic properties: this paves the way to a new class of media that are the "metamaterial crystals".

There is no doubt that we are far from having exploited all of the properties offered from such a locally resonant medium. We have exploited separately multiple scattering, the bandgap or the subwavelength modes, but by combining any of those properties one

can easily find new applications. For example by adding resonators creating a bandgap between the adjacent cans in the honey-comb crystal we turn the coupling between them from a polariton-like to a tight binding one, and we can therefore build a macroscopic analogue of graphene for the propagation of electrons. By adding some disorder in the structure but keeping the pair-correlations in positions we should also keep the negative effective property. All of this will be scope of future works and the soda can medium is a fairly good platform to demonstrate experimentally all of these fascinating phenomena.

Bibliography

[1] Brunet, T., Leng, J. and Mondain-Monval, O. (2013). Materials science. soft acoustic metamaterials. *Science* **342**, 6156, pp. 323–324.

[2] Brunet, T., Merlin, A., Mascaro, B., Zimny, K., Leng, J., Poncelet, O., Aristégui, C. and Mondain-Monval, O. (2015). Soft 3d acoustic metamaterial with negative index, *Nature Materials* **14**, 4, pp. 384–388.

[3] Christensen, J., Fernandez-Dominguez, A. I., de Leon-Perez, F., Martin-Moreno, L. and Garcia-Vidal, F. J. (2007). Collimation of sound assisted by acoustic surface waves, *Nature Physics* **3**, 12, pp. 851–852.

[4] Christensen, J., Liang, Z. and Willatzen, M. (2013). Metadevices for the confinement of sound and broadband double-negativity behavior, *Physical Review B* **88**, 10, p. 100301.

[5] Chutinan, A. and Noda, S. (1999). Highly confined waveguides and waveguide bends in three-dimensional photonic crystal, *Appl. Phys. Lett.* **75**, 24, pp. 3739–3741.

[6] Cowan, M. L., Page, J. H. and Sheng, P. (2011). Ultrasonic wave transport in a system of disordered resonant scatterers: Propagating resonant modes and hybridization gaps, *Phys. Rev. B* **84**, p. 094305.

[7] De Espinosa, F. M., Jimenez, E. and Torres, M. (1998). Ultrasonic band gap in a periodic two-dimensional composite, *Physical Review Letters* **80**, 6, p. 1208.

[8] de Rosny, J. and Fink, M. (2002). Overcoming the diffraction limit in wave physics using a time-reversal mirror and a novel acoustic sink, *Phys. Rev. Lett.* **89**, p. 124301.

[9] de Vries, P., van Coevorden, D. V. and Lagendijk, A. (1998). Point scatterers for classical waves, *Rev. Mod. Phys.* **70**, pp. 447–466.

[10] Derode, A., Tourin, A. and Fink, M. (2001). Random multiple scattering of ultrasound. II. Is time reversal a self-averaging process? *Phys. Rev. E* **64**, 3, pp. 036606–+.

[11] Ding, Y., Liu, Z., Qiu, C. and Shi, J. (2007). Metamaterial with simultaneously negative bulk modulus and mass density, *Phys. Rev. Lett.* **99**, p. 093904.
[12] Fang, N., Xi, D., Xu, J., Ambati, M., Srituravanich, W., Sun, C. and Zhang, X. (2006). Ultrasonic metamaterials with negative modulus, *Nature materials* **5**, 6, pp. 452–456.
[13] Fano, U. (1961). Effects of configuration interaction on intensities and phase shifts, *Phys. Rev.* **124**, pp. 1866–1878.
[14] Fink, M. (1997). Time reversed acoustics, *Physics Today* **50**, 3, pp. 34–40.
[15] Fleischhauer, M., Imamoglu, A. and Marangos, J. P. (2005). Electromagnetically induced transparency: Optics in coherent media, *Rev. Mod. Phys.* **77**, pp. 633–673.
[16] García-Chocano, V. M., Christensen, J. and Sánchez-Dehesa, J. (2014). Negative refraction and energy funneling by hyperbolic materials: An experimental demonstration in acoustics, *Phys. Rev. Lett.* **112**, 14, p. 144301.
[17] Guenneau, S., Movchan, A., Pétursson, G. and Ramakrishna, S. A. (2007). Acoustic metamaterials for sound focusing and confinement, *New Journal of Physics* **9**, 11, p. 399.
[18] Joannopoulos, J. D., Johnson, S. G., Winn, J. N. and Meade, R. D. (2011). *Photonic Crystals: Molding the Flow of Light* (Princeton University Press).
[19] Kafesaki, M., Sigalas, M. M. and García, N. (2000). Frequency modulation in the transmittivity of wave guides in elastic-wave band-gap materials, *Phys. Rev. Lett.* **85**, pp. 4044–4047.
[20] Kaina, N., Causier, A., Fink, M., Berthelot, T. and Lerosey, G. (2016). ?? *Science Advances*.
[21] Kaina, N., Lemoult, F., Fink, M. and Lerosey, G. (2013). Ultra small mode volume defect cavities in spatially ordered and disordered metamaterials, *Appl. Phys. Lett.* **102**, 14, p. 144104.
[22] Kaina, N., Lemoult, F., Fink, M. and Lerosey, G. (2015). Negative refractive index and acoustic superlens from multiple scattering in single negative metamaterials, *Nature* **525**, pp. 77–81.
[23] Kergomard, J. and Chaigne, A. (2008). *Acoustique des instruments de musique* (Paris, Belin).
[24] Khelif, A., Choujaa, A., Djafari-Rouhani, B., Wilm, M., Ballandras, S. and Laude, V. (2003). Trapping and guiding of acoustic waves by defect modes in a full-band-gap ultrasonic crystal, *Phys. Rev. B* **68**, p. 214301.
[25] Kinsler, L. E., Frey, A. R., Coppens, A. B. and Sanders, J. V. (2000). *Fundamentals of Acoustics, 4th Edition* (New York, Wiley).
[26] Kosten, C. W. and Zwikker, C. S. (1949). *Sound Absorbing Materials* (Amsterdam, Elsevier).
[27] Kushwaha, M. S., Halevi, P., Dobrzynski, L. and Djafari-Rouhani, B. (1993). Acoustic band structure of periodic elastic composites, *Phys. Rev. Lett.* **71**, 13, p. 2022.

[28] Lagendijk, A. (1993). Vibrational relaxation studied with light, *Ultrashort Processes in Condensed Matter* **314**, p. 197.
[29] Lani, S., Sabra, K. G. and Degertekin, F. L. (2016). Super-resolution ultrasonic imaging of stiffness variations on a microscale active metasurface, *Applied Physics Letters* **108**, 8.
[30] Lanoy, M., Pierrat, R., Lemoult, F., Fink, M., Leroy, V. and Tourin, A. (2015). Subwavelength focusing in bubbly media using broadband time reversal, *Phys. Rev. B* **91**, p. 224202.
[31] Lee, S. H., Park, C. M., Seo, Y. M., Wang, Z. G. and Kim, C. K. (2010). Composite acoustic medium with simultaneously negative density and modulus, *Phys. Rev. Lett.* **104**, 5, p. 054301.
[32] Lemoult, F., Fink, M. and Lerosey, G. (2011a). Acoustic resonators for far-field control of sound on a subwavelength scale, *Phys. Rev. Lett.* **107**, 6, p. 064301.
[33] Lemoult, F., Fink, M. and Lerosey, G. (2011b). Far-field sub-wavelength imaging and focusing using a wire medium based resonant metalens, *Waves in Random and Complex Media* **21**, 4, pp. 614–627.
[34] Lemoult, F., Fink, M. and Lerosey, G. (2011c). Revisiting the wire medium: An ideal resonant metalens, *Waves in Random and Complex Media* **21**, 4, pp. 591–613.
[35] Lemoult, F., Fink, M. and Lerosey, G. (2012). A polychromatic approach to far-field superlensing at visible wavelengths, *Nature Communications* **3**, p. 889.
[36] Lemoult, F., Kaina, N., Fink, M. and Lerosey, G. (2013). Wave propagation control at the deep subwavelength scale in metamaterials, *Nature Physics* **9**, 1, pp. 55–60.
[37] Lemoult, F., Lerosey, G., de Rosny, J. and Fink, M. (2009). Manipulating spatiotemporal degrees of freedom of waves in random media, *Phys. Rev. Lett.* **103**, 17, p. 173902.
[38] Lemoult, F., Lerosey, G., de Rosny, J. and Fink, M. (2010). Resonant metalenses for breaking the diffraction barrier, *Phys. Rev. Lett.* **104**, p. 203901.
[39] Lerosey, G., de Rosny, J., Tourin, A. and Fink, M. (2007). Focusing beyond the diffraction limit with far-field time reversal, *Science* **315**, p. 1120.
[40] Leroy, V., Bretagne, A., Fink, M., Willaime, H., Tabeling, P. and Tourin, A. (2009a). Design and characterization of bubble phononic crystals, *Appl. Phys. Lett.* **95**, 17, 171904.
[41] Leroy, V., Strybulevych, A., Scanlon, M. G. and Page, J. H. (2009b). Transmission of ultrasound through a single layer of bubbles, *The European Physical Journal E* **29**, 1, pp. 123–130.
[42] Li, J. and Chan, C. T. (2004). Double-negative acoustic metamaterial, *Phys. Rev. E* **70**, p. 055602.
[43] Li, J., Fok, L., Yin, X., Bartal, G. and Zhang, X. (2009). Experimental demonstration of an acoustic magnifying hyperlens, *Nature Materials* **8**, 12, pp. 931–934.

[44] Liu, Z., Zhang, X., Mao, Y., Zhu, Y. Y., Yang, Z., Chan, C. T. and Sheng, P. (2000). Locally resonant sonic materials, *Science* **289**, 5485, pp. 1734–1736.

[45] Lord Rayleigh, F. R. S. (1879). Investigations in optics, with special reference to the spectroscope, *Philosophical Magazine Series 5* **8**, 49, pp. 261–274.

[46] Ma, G. and Sheng, P. (2016). Acoustic metamaterials: From local resonances to broad horizons, *Science Advances* **2**, 2.

[47] Ma, G., Yang, M., Xiao, S., Yang, Z. and Sheng, P. (2014). Acoustic metasurface with hybrid resonances, *Nature Materials* **13**, 9, pp. 873–878.

[48] Martin-Cano, D., Nesterov, M. L., Fernandez-Dominguez, A. I., Garcia-Vidal, F. J., Martin-Moreno, L. and Moreno, E. (2010). plasmons for subwavelength terahertz circuitry, *Opt. Express* **18**, 2, pp. 754–764.

[49] Martinezsala, R., Sancho, J., Sánchez, J., Gómez, V., Llinares, J., Meseguer, F., *et al.* (1995). Sound-attenuation by sculpture, *nature* **378**, 6554, pp. 241–241.

[50] Maznev, A. A., Gu, G., yuan Sun, S., Xu, J., Shen, Y., Fang, N. and yi Zhang, S. (2015). Extraordinary focusing of sound above a soda can array without time reversal, *New Journal of Physics* **17**, 4, p. 042001.

[51] Miroshnichenko, A. E., Flach, S. and Kivshar, Y. S. (2010). Fano resonances in nanoscale structures, *Rev. Mod. Phys.* **82**, pp. 2257–2298.

[52] Miyashita, T. (2005). Sonic crystals and sonic wave-guides, *Measurement Science and Technology* **16**, 5, p. R47.

[53] Montaldo, G., Tanter, M. and Fink, M. (2004). Real time inverse filter focusing through iterative time reversal, *Journal of the Acoustical Society of America* **115**, 2, pp. 768–775.

[54] Mosk, A. P., Lagendijk, A., Lerosey, G. and Fink, M. (2012). Controlling waves in space and time for imaging and focusing in complex media, *Nature Photonics* **6**, 5, pp. 283–292.

[55] Page, J., Sukhovich, A., Yang, S., Cowan, M., Van Der Biest, F., Tourin, A., Fink, M., Liu, Z., Chan, C. and Sheng, P. (2004). Phononic crystals, *Physica Status Solidi* (*B*) **241**, 15, pp. 3454–3462.

[56] Painter, O., Lee, R. K., Scherer, A., Yariv, A., O'Brien, J. D., Dapkus, P. D. and Kim, I. (1999). Two-dimensional photonic band-gap defect mode laser, *Science* **284**, 5421, pp. 1819–1821.

[57] Papasimakis, N., Fedotov, V. A., Zheludev, N. I. and Prosvirnin, S. L. (2008). Metamaterial analog of electromagnetically induced transparency, *Phys. Rev. Lett.* **101**, p. 253903.

[58] Penciu, R. S., Kriegs, H., Petekidis, G., Fytas, G. and Economou, E. N. (2003). Phonons in colloidal systems, *The Journal of Chemical Physics* **118**, 11, pp. 5224–5240.

[59] Pendry, J. B. (2000). Negative refraction makes a perfect lens, *Phys. Rev. Lett.* **85**, 18, pp. 3966–3969.

[60] Pendry, J. B., Holden, A. J., Robbins, D. J. and Stewart, W. J. (1999). Magnetism from conductors and enhanced nonlinear phenomena, *IEEE Transactions on Microwave Theory and Techniques* **47**, 11, pp. 2075–2084.

[61] Pendry, J. B., Holden, A. J., Stewart, W. J. and Youngs, I. (1996). Extremely low frequency plasmons in metallic mesostructures, *Phys. Rev. Lett.* **76**, 25, pp. 4773–4776.

[62] Pierrat, R., Vandenbem, C., Fink, M. and Carminati, R. (2013). Subwavelength focusing inside an open disordered medium by time reversal at a single point antenna, *Phys. Rev. A* **87**, p. 041801.

[63] Psarobas, I. E., Modinos, A., Sainidou, R. and Stefanou, N. (2002). Acoustic properties of colloidal crystals, *Phys. Rev. B* **65**, p. 064307.

[64] Purcell, E. M. (1946). Spontaneous emission probabilities at radio frequencies, *Phys. Rev.* **69**, p. 674.

[65] Rupin, M., Catheline, S. and Roux, P. (2015). Super-resolution experiments on lamb waves using a single emitter, *Appl. Phys. Lett.* **106**, 2, p. 024103.

[66] Sigalas, M. and Economou, E. N. (1993). Band structure of elastic waves in two dimensional systems, *Solid State Communications* **86**, 3, pp. 141–143.

[67] Smith, D. R., Padilla, W. J., Vier, D. C., Nemat-Nasser, S. C. and Schultz, S. (2000). Composite medium with simultaneously negative permeability and permittivity, *Phys. Rev. Lett.* **84**, pp. 4184–4187.

[68] Sukhovich, A., Jing, L. and Page, J. H. (2008). Negative refraction and focusing of ultrasound in two-dimensional phononic crystals, *Phys. Rev. B* **77**, 1, p. 014301.

[69] Sukhovich, A., Merheb, B., Muralidharan, K., Vasseur, J. O., Pennec, Y., Deymier, P. A. and Page, J. H. (2009). Experimental and theoretical evidence for subwavelength imaging in phononic crystals, *Phys. Rev. Lett.* **102**, p. 154301.

[70] Tanter, M., Aubry, J. F., Gerber, J., Thomas, J. L. and Fink, M. (2001). Optimal focusing by spatio-temporal inverse filter. I. Basic principles, *Journal of the Acoustical Society of America* **110**, pp. 37–47.

[71] Veselago, V. G. (1968). The electrodynamics of substances with simultaneous negative values of ϵ and μ, *Physics-Uspekhi* **10**, 4, 509.

[72] von Helmholtz, H. (1885). *On the Sensations of Tone as a Physiological Basis for the Theory of Music* (London, Longmans, Green and Co.).

[73] Yang, M., Ma, G., Yang, Z. and Sheng, P. (2013). Coupled membranes with doubly negative mass density and bulk modulus, *Physical Review Letters* **110**, 13, p. 134301.

[74] Yang, S., Page, J. H., Liu, Z., Cowan, M. L., Chan, C. T. and Sheng, P. (2002). Ultrasound tunneling through 3d phononic crystals, *Phys. Rev. Lett.* **88**, p. 104301.

[75] Zhang, S., Yin, L. and Fang, N. (2009). Focusing ultrasound with an acoustic metamaterial network, *Phys. Rev. Lett.* **102**, p. 194301.

[76] Zhu, J., Christensen, J., Jung, J., Martin-Moreno, L., Yin, X., Fok, L., Zhang, X. and Garcia-Vidal, F. J. (2011). A holey-structured metamaterial for acoustic deep-subwavelength imaging, *Nature Physics* **7**, 1, pp. 52–55.

CHAPTER 6

New Trends Toward Locally-Resonant Metamaterials at the Mesoscopic Scale

PHILIPPE ROUX*

Université Grenoble Alpes, France

MATTHIEU RUPIN, FABRICE LEMOULT and GEOFFROY LEROSEY

Institut Langevin, France

ANDREA COLOMBI and RICHARD CRASTER

Imperial College London, UK

SÉBASTIEN GUENNEAU

Institut Fresnel and Aix-Marseille Université, France

WILLIAM A. KUPERMAN

University of California San Diego, USA

*Corresponding author: philippe.roux@ujf-grenoble.fr

EARL. G. WILLIAMS

Naval Research Laboratory, USA

6.1. Introduction

Man-made composite materials have been the subject of many investigations in the wave physics community over the last decade, as they show wavefield characteristics that cannot be obtained with natural materials. Theoretical and experimental work at the turn of the 21st century [Pendry, 1999; Pendry, 2000; Shelby *et al.*, 2001] has shown that only ad-hoc subwavelength arrangements of resonant elements, later on coined *metamaterials*, could provide the extraordinary negative refraction property first advocated by the visionary study by Veselago in the late 60's [Veselago, 1968]. Metamaterials nowadays provide an opportunity to control and manipulate waves at different wavelength scales [Engheta and Ziolkowski, 2006; Craster and Guenneau, 2012; Pendry *et al.*, 2006; Lemoult *et al.*, 2013; Kadic *et al.*, 2013].

In general, the wave propagation properties of structured manmade materials find their roots in two distinct origins: first, the ordered or disordered spatial organization of their components; second, the resonant or non-resonant nature of their unitary element [Deymier, 2013; Sigalas *et al.*, 2009]. When waves propagate in a complex medium with structural order, they can undertake multiple scattering, which leads to frequency bands of permitted and forbidden propagation, also called band gaps. These forbidden frequency bands are analogous to electronic band gaps in natural crystals [Kittel, 1996]. These materials have a typical spatial scale comparable to the average wavelength of the wave under investigation. They are termed photonic crystals in electromagnetics [John, 1987; Yablonovitch, 1987; Fan *et al.*, 1998; Notomi, 2000; Joannopoulos *et al.*, 2008; Shen *et al.*, 2014] and phononic crystals in acoustics [Martinez-Sala *et al.*, 1995; Vasseur *et al.*, 2008; Sigalas and Economou, 1993; Poulton, 2003; Benchabane *et al.*, 2006; Page *et al.*, 2004; Sukhovich *et al.*, 2009; Farhat *et al.*, 2010; Maznev *et al.*, 2011; Khelif *et al.*, 2003; Pierre *et al.*, 2010; Dubois *et al.*, 2013, Lagarrigue

et al., 2013, Page *et al.*, 2013], and they constitute powerful tools to shape the wavefield in various manners [Joannopoulos *et al.*, 2008].

Other types of composite materials are made of a collection of unitary resonant components (sub-wavelength resonators) from which they extract their dispersive properties at the macroscopic scale [Liu *et al.*, 2000a, 2000b; Fang *et al.*, 2006; Guenneau *et al.*, 2007; Lemoult *et al.*, 2011(a), 2013; Leroy *et al.*, 2009; Psarobas *et al.*, 2002; Cowan *et al.*, 2011; Antonakakis and Craster (2012); Christiansen and de Abajo, 2012; Achaoui *et al.*, 2013; Deymier, 2013]. A key parameter of these so-called locally resonant metamaterials lies in the arrangement of many sub-wavelength resonators within a wavelength which also means that their collective behavior cannot be explained by Bragg scattering. In fact, the proximity of sub-wavelength resonators permits the coupling between a large set of individual resonators [Pendry *et al.*, 1999; Smith *et al.*, 2000]. In this case, the metamaterial physics can be interpreted through interferences between the incident field and the scattered waves generated by each resonant unit [Lemoult *et al.*, 2013], which result in a Fano interference due to the resonant nature of the unit cell [Fano, 1961; Miroshnichenko *et al.*, 2010]. At the macroscopic scale, local resonances lead to the hybridization phenomenon [Sigalas and Economou, 1993], which is analogous to the concept of polariton in solid state physics [Lagendijk, 1993] that deeply modify the dispersion relation in the host medium, generating two hybridized modes resulting from the anti-crossing effect between a continuum of states with a localized one. By tuning the unit-cell resonance, it is possible to obtain media with effective parameters (such as compressibility and density for acoustic waves, for example) that can be positive [Shen *et al.*, 2005; Choi *et al.*, 2011], null [Silveirinha and Engheta, 2006] or negative [Fang *et al.*, 2006; Guenneau *et al.*, 2007; Gracia-Salgado *et al.*, 2012; Kaina, 2015]. When both of the effective parameters are simultaneously negative [Veselago, 1968; Smith *et al.*, 2000; Smith *et al.*, 2004; Li and Chan, 2004], exotic behavior occurs such as negative refraction [Pendry, 2000; Shelby, 2001; Sukhovich *et al.*, 2008; Ramakrishna and Grzegorcyk, 2008; Farhat,

2010; Pierre *et al.*, 2010; Craster and Guenneau, 2012]. Contrary to photonic or phononic crystals, which exhibit band gaps due to the Bragg diffraction effect, the spatial organization of locally resonant metamaterial is no longer relevant which means that band gaps can be observed independently of its ordered or disordered nature [Cowan *et al.*, 2011; Kaina *et al.*, 2013b, Kaina *et al.*, 2015].

As a matter of fact, the width and efficiency of band gaps in locally resonant metamaterials depend upon both the spatial density of the resonators and the radiative quality factor of the resonances. In the limit of small resonators compared to the wavelength, the smaller the resonator, the higher its quality factor. This justifies why locally resonant metamaterials classically support band gaps limited to narrow bandwidths. To overcome this fundamental limitation, metamaterials can be created with resonator units that are small in two dimensions that support wave propagation, but elongated in the third one, where each resonator can expand [Lemoult *et al.*, 2013]. Thus, one dimension is sacrificed to define the resonant property of the resonators, while in the two-dimensional plane of propagation, the resonators are much smaller than the wavelength [Lemoult *et al.*, 2010, 2011(b); Belov *et al.*, 2006]. Such a uni-axial metamaterial was created with a so-called wire medium in electromagnetism, which has been used to control and/or focus waves below the diffraction limit. When curved, such wire media can be used to design invisibility cloaks with deep subwavelength features, as experimentally shown for microwaves [Ktorza *et al.*, 2015].

Most of the metamaterial research has been done experimentally on small scales that are typically fractions of a millimeter. They are typically either electromagnetic or phononic/ultrasonic regimes. For larger elastic and acoustic propagation scales, the subwavelength resonators can be identifiable elastic components such as finite length rods. For example, in acoustics, a collection of narrow, but long pipes was recently used to form a super lens [Zhu *et al.*, 2011]. This approach is also perfectly suited to thin 2D plate supporting Lamb waves that interact with a collection of resonators (also known as meta-surfaces [Khelif *et al.*, 2010; Oudich *et al.*, 2011; Xiao *et al.*,

2012; Rupin *et al.*, 2014(a)]), as will be shown in the present review paper. The purpose of this chapter is to explore examples of this larger scale metamaterical regime from which we envision extending mechanical metamaterials at a geophysical scale.

We start with an example of a flexurally vibrating plate to which we add a collection of vertical rods that act as resonators analogous to the smaller scale structures most commonly studied in the literature.

We present an experimental and theoretical analysis of the resulting metamaterial and we relate these results to the multimodal scattering required for optimal propagation control. For example, we show that high resolution focusing can be achieved in the metamaterial region at subwavelength scales associated with the Lamb waves of the plate. Further, it is demonstrated from the single mode experimental data that cloaking can be achieved at the level expected by theory. Finally, the results suggest that analogous composite structures exist in nature at the geophysical scale. We provide an example in which a forest of trees makes adequate resonant substructures for surface waves propagating at the interface of a semi-infinite elastic medium.

6.2. Part I: Experimental Results

The investigation of the uniaxial metamaterial for Lamb waves is made through experiments at a larger mesoscopic scale (and hence at a lower frequency) than usually reported in the literature [Pennec *et al.*, 2009; Wu *et al.*, 2009; Khelif *et al.*, 2010; Oudich *et al.*, 2011; Assouar *et al.*, 2012; Xiao *et al.*, 2012; Hsu, 2013]. A 6-mm-thick and 1.5-m $\times$ 2-m-long aluminum plate is locally excited by a dynamic shaker with a 1.8-s-long broadband chirp that ranges from 500 Hz to 11 kHz (Fig. 6.1(a)). At low frequency, two Lamb modes [Lamb, 1904] propagate in the plate, which are defined as the anti-symmetric $A0$ and the symmetric $S0$ modes. The $A0$ mode mostly corresponds to out-of-plane displacement of the plate while the $S0$ mode corresponds to in-plane longitudinal displacement. For symmetry reasons, the shaker only excites the $A0$ mode and only the vertical displacement of the plate is locally probed with the laser interferometer. The shape of the plate (a Bunimovich stadium which has chaotic

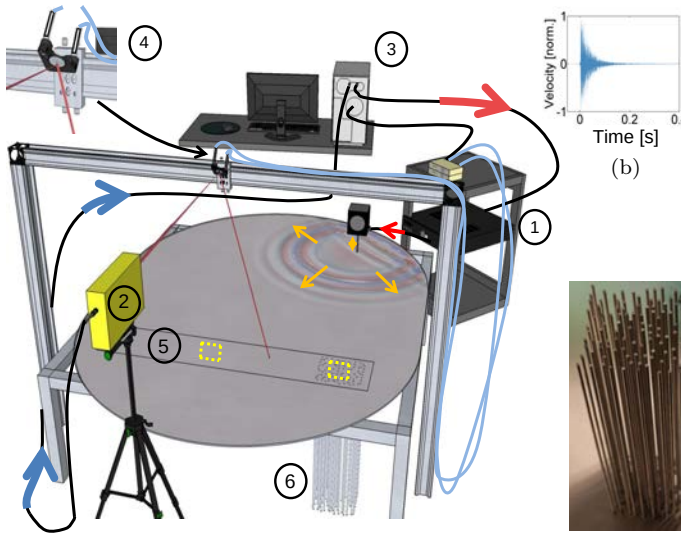


Fig. 6.1. (a) Experimental set-up. A shaker (1) generates $A0$ Lamb waves in the aluminum plate. The wavefield is measured with a Doppler velocimeter (2) and a PC-controlled (3) motorized mirror (4) on the rectangular area (5) on the upper side of the plate. The metamaterial (6) is attached below the plate on one side of the recording map. (b) The typical temporal dispersion exceeds 0.2 s, which corresponds to more than 20 round-trips inside the plate. (c) The metamaterial is made of 100 vertical aluminium rods that are arranged on a random pattern, with an average inter-rod distance of the order of 2 cm. (Modified from Rupin *et al.*, 2014).

dynamical ray trajectories [Bunimovitch, 1979]) ensures that, whatever the source position, the waves propagate at long time in all directions with equal probabilities [Gutzwiller, 1990; Bunimovitch, 1979]. Note that the $A0$ Lamb waves have a quadratic dispersion relation [Royer and Dieulesaint, 2000] with typical wave speed of 340 m/s at 2 kHz (wavelength $\lambda = 17.5$ cm). The multiply-reverberated waves are measured on the upper side of the plate and, after cross-correlation with the emitted chirp, the plate response is spread over more than 200 ms (Fig. 6.1(b)), to be compared to the 1-ms-long autocorrelation signal. This corresponds to more than 20 round trips of the propagating waves across the plate. Using computer-controlled

motorized mirrors, the vertical displacement of the plate surface can be scanned on a 0.20 m × 1.15 m surface S (Fig. 6.1(a), rectangle numbered 5) with a resolution of 3 mm. This leads to the accurate estimation of the spatio-temporal wavefield that describes the interference pattern created by the shaker in the plate.

Within this scanned area, a metamaterial is built with a set of 100 cylindrical 61-cm-long, 6.35-mm-diameter aluminum rods glued on a 400 cm^2 area on the lower side of the plate (Fig. 6.1(c)). The spatial distribution of the vertical rods on the plate is random with an average distance between resonators of the order of 2 cm (i.e. $\lambda/9$ and $\lambda/4$ for the $A0$ mode at 2 kHz and 10 kHz) and a 5-mm minimum inter-rod distance. For Lamb waves, this collection of rods is equivalent to a set of sub-wavelength resonant scatterers exhibiting several resonances of different nature (longitudinal and flexural resonances of rods) in the frequency range under study.

When the plate vertical displacement is probed outside the metamaterial, the spatially averaged Fourier transform shows a maximum energy density below 2 kHz, followed by a plateau up to 11 kHz (Fig. 6.2). The sudden 10-dB drop at 2 kHz is due to the radiation leakage of the $A0$ Lamb mode in air. The Fourier spectrum

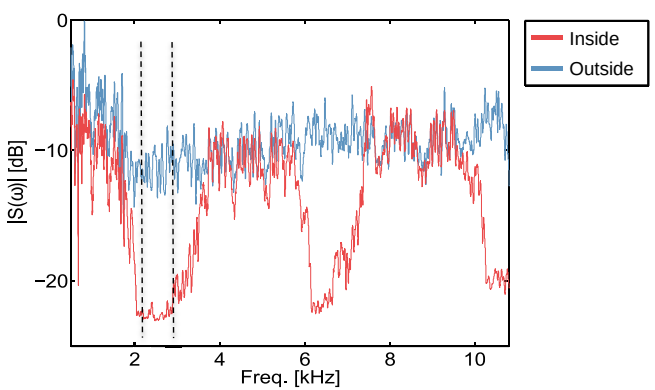


Fig. 6.2. Average Fourier spectra for the signals measured inside (red) and outside (blue) the metamaterial. The averaging is performed over two surfaces (yellow dotted squares) within the rectangle measurement area (number 5) in Fig. 6.1. The two vertical dashed lines correspond to the frequency band in which the wavefield is filtered in Fig. 6.3 (Modified from Rupin *et al.*, 2014).

measured above the multi-resonator disordered metamaterials reveals three wide band gaps starting at 2, 6, and 10 kHz. The shape and intensity of the band gaps are independent of the random organization of the rods [Rupin *et al.*, 2014(a)]. This demonstrates that spatial disorder does not affect the properties of the metamaterial whose macroscopic behavior cannot be explained with constructive or destructive interferences associated with Bragg scattering.

In Figs. 6.3(a)–(d), the wavefield was filtered in the first band gap (dashed vertical lines in Fig. 6.2), and then mapped after a propagation time of 3.3 ms (corresponding to the ballistic field), 20 ms,

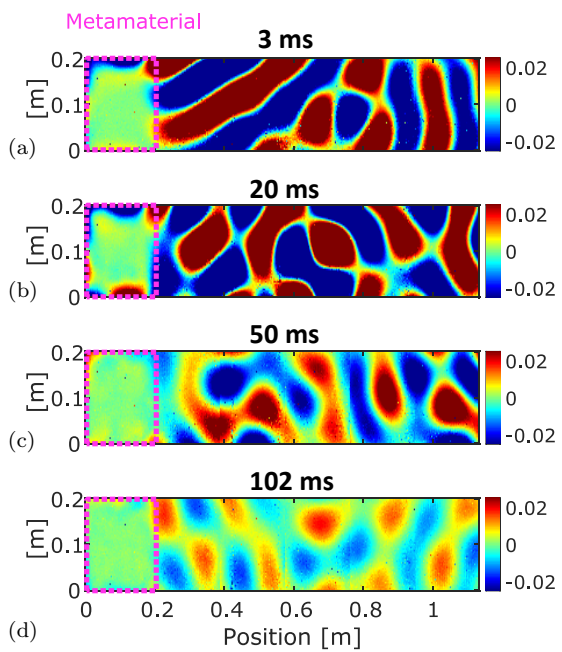


Fig. 6.3. Field maps measured with a Doppler velocimeter on the rectangular area [see Fig. 6.1(a)] inside and outside of the metamaterial, (dashed purple area). Snapshots of the temporal field at (a) 3 ms, (b) 20 ms, (c) 50 ms and (d) 102 ms for signals filtered in the first band gap [2100 Hz – 2800 Hz]. After 100 ms, the field has propagated more than 40 m inside the reverberating plate which results in an omni-directional incident field onto the metamaterial. The colorbar is the same for each panel with a field amplitude maximum equal to 1 for the incident wave in the plate, leading to a strong saturation in panels (a) and (b) that further confirm the very low field amplitude inside the metamaterial.

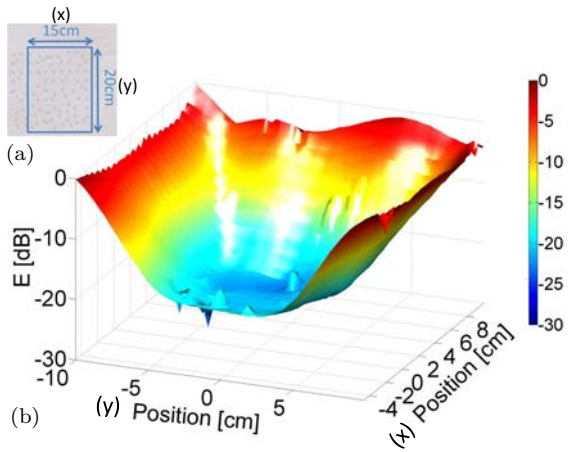


Fig. 6.4. (a) Size of the (x)–(y) area scanned with the Doppler velocimeter inside the metamaterial. (b) Spatial distribution (in dB scale) of the field intensity recorded inside the first band gap [2100 Hz − 2800 Hz].

50 ms and 100 ms, corresponding to multi-reverberated Lamb waves with travel distances of ∼7 m, ∼20 m and 40 m inside the 1.5-m × 2-m plate, respectively. These 'snapshots' reveal two effects. First, the band gap is efficient, as waves have been clearly attenuated within a small fraction of the $A0$ mode wavelength. Secondly, the band gap is valid for all incident angles, as can be seen from the last snapshots with a superposition of incident plane waves in all directions.

When filtered in the bandgap, the field amplitude decreases exponentially as one penetrates inside the metamaterial along one direction perpendicular to its interface (Fig. 6.4). This penetration length can be interpreted as a skin effect and depends on frequency. It varies from 3 cm at 2.5 kHz in the first bandgap to 1.7 cm at 6.4 kHz in the second one (Fig. 6.5). The formulation for the frequency-dependent attenuation coefficient is proposed later in the chapter.

6.2.1. *Dispersion curve within the metamaterial*

Outside of the bandgaps, the wavefield propagates deep inside the metamaterial. The billiard shape of the plate makes the field spatially

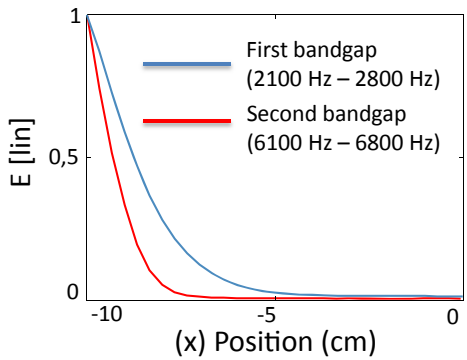


Fig. 6.5. Normalized averaged intensity measured along the y-direction perpendicular to the metamaterial rectangular area (see Fig. 6.4a) in both the first and second bandgaps.

random after a few reverberations from the plate boundaries. Considering a time window associated with long-time reverberations, the spatial Fourier transform of the field recorded inside or outside of the metamaterial (Fig. 6.3) reveals a circle when filtered in a small bandwidth that confirm the equi-distribution of the wave components on all possible azimuthal directions (Fig. 6.6). At each frequency, the circle radius is an accurate measure of the effective velocity inside or outside of the metamaterial, from which the dispersion curve can be plotted.

The dispersion curve of the metamaterial is thus obtained from the collection of grid points measured solely inside the purple rectangles in Fig. 6.3. Both propagative wavenumber and attenuation can be extracted over the frequency band of the source excitation. For each measured point, the temporal signal is Fourier transformed and a 2D spatial Fourier transform of the wave field is then calculated [Capon, 1973]. At each frequency, this 2D Fourier spectrum is transformed into a frequency–wavenumber dispersion curve by averaging the spectrum over a circle [see Figs. 6.6(e), (f)]. The wave physics of the metamaterial is then encapsulated in a single dispersion relation that shows the real part of the wave vector and the corresponding attenuation length within the band gaps (Fig. 6.7). The dispersion

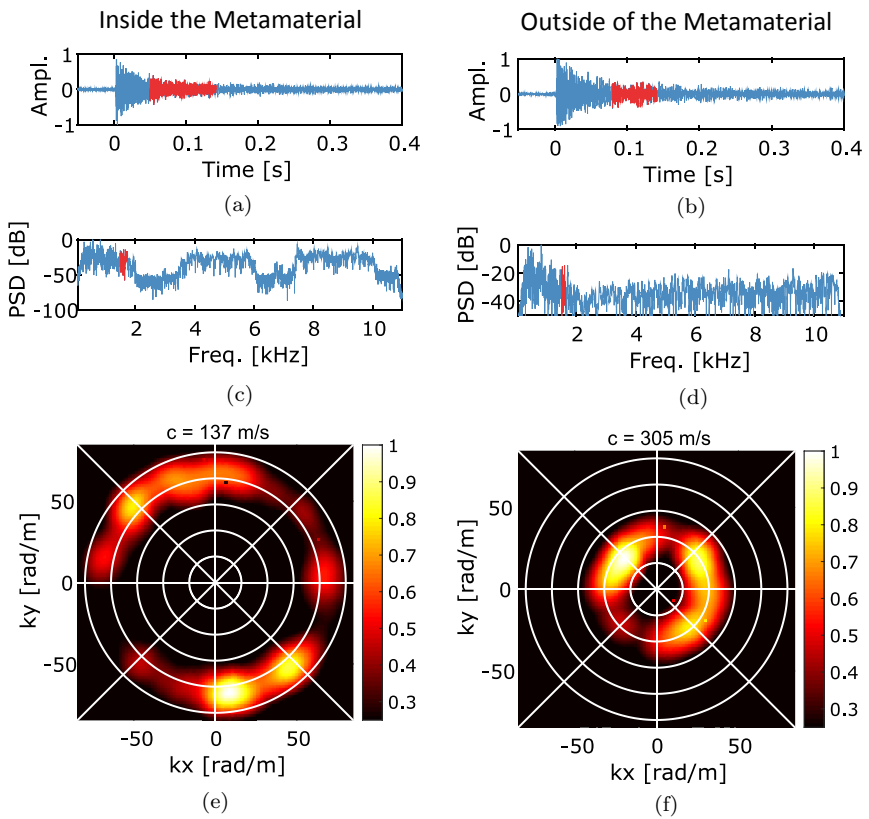


Fig. 6.6. Frequency-wavenumber representation of the field inside and outside of the metamaterial. (a)–(b) A time window is selected (in red) that corresponds to the multi-reverberated part of the field on the plate. (c)–(d) After Fourier transform, a narrow frequency band is selected (in red : 1700 Hz − 1800 Hz). (e)–(f) The k_x–k_y wavenumber representation is obtained both inside and outside the metamaterial area for the time and frequency windows shown above. A radial averaging provides the maximum phase velocity: (e) wave speed $c = 137$ m/s and (f) $c = 305$ m/s, respectively inside and outside of the metamaterial.

relations of the $A0$ Lamb modes in the homogeneous plate are superimposed.

As written above, thin plates support two types of waves at low frequencies: the symmetric and antisymmetric modes, $S0$ and $A0$. In practice, $A0$ waves are mostly vertically polarized and can be

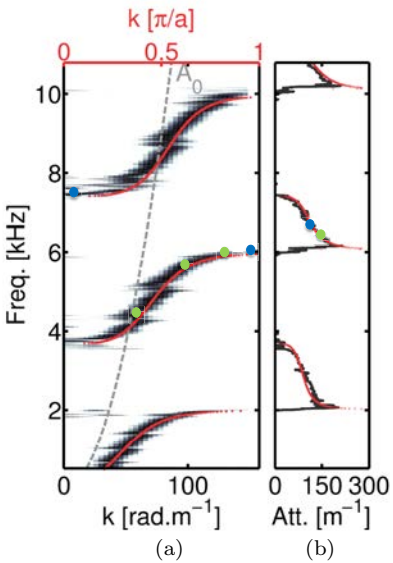


Fig. 6.7. (a) Frequency — wavenumber representation of the wavefield measured inside the locally-resonant metamaterial. Theoretical dispersion of the $A0$ curve for the homogeneous plate is superimposed (dashed gray). The red curves correspond to the theoretical dispersion curve for a periodic distribution of resonators inside the metamaterial using Bloch theorem. The horizontal axis at the top of the panel corresponds to the wavenumber scaled by the average distance between the resonators. The three blue dots and the four green dots refer to the frequency components for the field maps displayed in Figs. 6.8 and 6.9, respectively. (b) Imaginary part of the k-vector measured inside the metamaterial with the same color code as in (a). (Modified from Williams *et al.*, 2015).

characterized by out-of-plane (vertical) displacement when in-plane (horizontal) displacements in the plate are described by $S0$ waves. Numerical simulations performed with 3D elastic finite-element code in the plate together with the metamaterial confirm that nearly no energy is present in the $S0$ mode. This means that the potential conversion of the excited $A0$ waves to $S0$ modes upon scattering of the resonating beams can be neglected in the present configuration. Indeed, given the plate thickness (6 mm), the flexural resonances of the beams are weakly excited by the $A0$ incident wave, and we can limit the analysis to the interaction between the $A0$ mode and

the compressional resonances inside the beams. As will be shown later, we expect things to be very different with a thinner plate (i.e., more flexible), where both $A0$ and $S0$ modes may be excited at the local coupling between the plate and the resonators. In this case, both flexural and compressional resonances may modify the band gap structure observed in Fig. 6.7.

In the "rigid" 6-mm thickness plate configuration, the vertical $A0$ mode of the plate solely hybridizes with the compressional resonances of the resonators, which gives rise to a polaritonic behavior [Lagendijk, 1993] at each crossing of the $A0$ dispersion curve with these low-quality-factor resonances (Fig. 6.7). Actually the dissipation part of the quality factor of the rods is negligible and the wide frequency spreading of the compressional resonances conveys the optimal coupling with the $A0$ Lamb modes. Indeed there is a symmetry matching between the polarization of the $A0$ mode and the local motion of the plate imposed by the resonators. The latter being placed on a single side of the plate, the compressional resonances of the rods induce an asymmetric motion of the plate. The energy trapped in the resonators is then fully converted into $A0$ Lamb waves in the plate. As the metamaterial behavior results from Fano interference between the incoming waves and the scattered waves [Lemoult *et al.*, 2013], the wide compressional resonances of the rods result in wide band gaps that start at the resonance and end at the anti-resonance of one single rod. These band gaps show strong attenuation due to the high spatial density of resonators and are insensitive to ordered or disordered configurations. Outside the band gaps, the hybridizations between the $A0$ mode and the rod compressional resonances give rise to two branches, namely, the anti-binding branches and the binding ones. At frequencies above the band gaps, supra-wavelength modes are observed. On the contrary, the binding branches at frequencies below the band gaps are composed of subwavelength modes with spatial scales comparable to the average distance between rods.

Some of us developed an analytical wave approach to describe the physics properties of multi-resonant metamaterials for Lamb waves

propagating in plates [Williams *et al.*, 2015]. The metamaterial that we characterize consists of a periodic array of long rods attached to the surface of the plate that forms the substrate in which antisymmetric $A0$ Lamb waves are excited. We show that the $A0$ Lamb wave propagation through the metamaterial can be accurately modeled using a simplified theory that replaces the two-dimensional array with a one-dimensional horizontal beam on which is attached a linear array of 10 vertical rods. This unidimensional approach is legitimate because of the isotropic response of the metamaterial. The wave propagation problem is solved rigorously for this one-dimensional system using the scattering matrix formalism for a single rod [Williams *et al.*, 2015] in a manner related to the propagation matrix approach used to derive the realization of an anisotropic medium [Gilbert, 1983; Schoenberg and Sen, 1983] but now with resonance substructures. The exact eigenvalues of the rod and beam system are approximated in a long wavelength expansion to determine a simple expression for the effective velocity c_{eff} and dispersion of the metamaterial:

$$c_{eff}/c_p = \left[\frac{M_b}{M}\frac{\tan\left(k_b L_b\right)}{k_b L_b} + 1\right]^{-1/4}. \tag{6.1}$$

In Eq. 6.1, c_p refers to the speed of flexural waves in plate defined as $c_p = \sqrt{\omega}(\frac{EI}{\rho A})^{1/4}$ with, for the beam parameters, density $\rho = 2700$ kg.m^{-3}, area $A = hb$ with thickness $h = 6$ mm and width $b = 2$ cm, Young's modulus $E = 69 \times 10^9$ m^{-1}.kg.s^{-2}, and inertial moment $I = h^3 b/12$; M_b is the total mass of a rod and M the mass of the beam segment of length $L = 2$ cm corresponding to the inter-rods distance, with $M_b/M = 8.02$ in the present configuration. Finally, the rod length is $L_b = 61$ cm and we have $k_b = \omega/c_b$ with c_b defined as the non-dispersive wave velocity in the rod with Young Modulus $E_b = E$ and density $\rho_b = \rho$ such as $c_b = \sqrt{E_b/\rho_b} = 5055$ m.s^{-1}.

The modeled dispersion is compared with the experimental dispersion curve inside the metamaterial and we obtain excellent agreement (Fig. 6.7(a)). The multi-resonant rods, restricted to longitudinal vibration consistent with $A0$ Lamb waves excited in the

plate, produce two wide stop bands in the frequency domain from 0 to 10 kHz where the stop or passband boundaries align with the minima and maxima of the rod's impedance response.

The imaginary part of the eigenvalue problem solved using the scattering matrix also provides an analytical expression for the attenuation inside the bandgaps:

$$\alpha = \frac{k}{\sqrt{2}} \left[\frac{M_b}{M} \frac{\tan(k_b L_b)}{k_b L_b} + 1 \right]^{1/4} \tag{6.2}$$

with

$$k = \omega / c_p.$$

Analysis of the attenuation (Fig. 6.7(b)) indicates that the attenuation is infinite at the anti-resonances of the rod (given by $\tan \to \pm\infty$ or $k_b L_b = \pi/2,\ 3\pi/2$) and that these frequencies mark the onset of the stop bands, shown in Fig. 6.7(a). Thus, at the onset of the stop bands, the Lamb wave attempting to pass through a segment experiences an extremely high attenuation due to the longitudinal anti-resonances of the attached rod. These anti-resonances clamp the motion of the beam. As the frequency increases, Eq. (6.2) indicates that the attenuation coefficient decreases to zero at the end of the stop band since $tan(k_b L_b) \to 0$. Of course, the decay through a forest of N rods is given by attenuation of $exp(-N\alpha L)$ in the distance NL.

Finally, Eq. (6.3) for the effective density ρ_{eff} inside the metamaterial shows that a negative effective density is obtained in the stop band:

$$\rho_{eff}/\rho = -\left| \frac{M_b}{M} \frac{\tan(k_b L_b)}{k_b L_b} + 1 \right| \tag{6.3}$$

In the characterization of metamaterials there is a large volume of literature describing them as having negative mass density and/or negative compressibility in the stop bands [Shelby *et al.*, 2001; Smith *et al.*, 2004; Li and Chan, 2004; Koschny *et al.*, 2004; Guenneau *et al.*, 2007; Wu *et al.*, 2007; Park *et al.*, 2012]. Generally this relationship is made using monopole and dipole expansions of the metamaterial element (our rod), which is assumed elastic through its cross section.

The monopole (breathing mode of the cross section) relates to negative effective compressibility and the dipole (lateral displacement of the cross section) to negative effective density [Wu *et al.*, 2011]. In our case, however, the rod is modeled as a *point* attachment, and the force it applies to the plate is omnidirectional and thus monopole-like in its excitation of $A0$ waves. Since in-plane, longitudinal $S0$ waves in the plate are not excited, there is no dipole-like excitation, and the result is a mono-modal system. Because of this, we can characterize the metamaterial as having either negative effective density or compressibility, but not both in the stop band.

In conclusion, theoretical results superimposed on experimental results in Fig. 6.7 confirm that a 1D theory has been extremely successful at predicting the 2D scattering from a rectangular patch of 10×10 rods attached to the surface of a plate with a reverberating boundary shape.

These theoretical results confirm the hybridization phenomenon between the dominant flexural $A0$ mode in the plate and the compressional resonances inside the rods and demonstrated the presence of sub-wavelength and supra-wavelength modes within the metamaterial before and after each band gap (Fig. 6.7). The spatial measurements described in Fig. 6.8 at three different frequencies show the diversity of results that can be obtained within the metamaterial (the purple dashed squares on the left between 0 and 0.2 m). In particular, one can see clearly that the spatial wavelengths may be very different within and outside of the metamaterial.

6.2.2. *Subwavelength focusing*

In the following analysis, we targeted four particular frequencies, which correspond to four different spatial intensity distributions (Fig. 6.9). Starting from the lowest frequency, it can be seen in Fig. 6.7(a) (lowest green dot at 4.5 kHz) that it roughly coincides with the wave-number of the $A0$ plate mode. In Fig. 6.9(a), the wavefield amplitude distribution is similar inside (purple dashed square) and outside of the metamaterial. The next two frequencies are closer to the band gap (Fig. 6.7(a), green dots at 5.8 kHz and 6.0 kHz). The

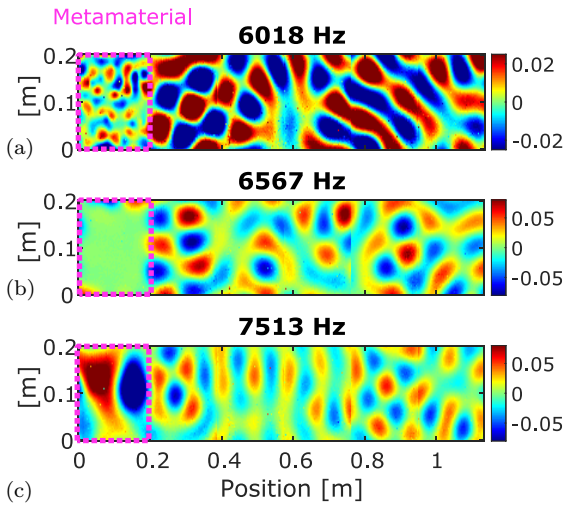


Fig. 6.8. Field maps measured for different frequency components (a) 6018 Hz, (b) 6567 Hz, (c) 7513 Hz with a Doppler velocimeter on the rectangular area [see Fig. 6.1(a)] inside and outside of the metamaterial, (dashed purple area). The colorbar is chosen with a maximum frequency field amplitude of 1 for each panel in order to magnify (a) the small-wavelength or (c) the large-wavelength components inside the metamaterial at frequencies close to the boundaries of each band gap.

corresponding amplitude distribution (see Figs. 6.9(b) and 6.9(c)) shows a smaller correlation length of the wavefield, compared to that in the plate. In parallel, the wavefield becomes less and less energetic. This attenuation is not due to an increase in dissipation since dissipation is null outside of the bandgap, but rather to the fact that the phase velocity varies quickly as the band gap is approached, which makes the group velocity decrease towards zero. Finally, for the frequency inside the band gap [green dot at 6.2 kHz in Fig. 6.7(b)], the wavenumber is purely imaginary and no propagation can be observed in the metamaterial as shown in Fig. 6.9(d).

These experimental results are connected with sub-wavelength resolution [Fang *et al.*, 2005; Lerosey *et al.*, 2007; Lemoult *et al.*, 2010, 2011, 2013] from a single emission channel located in the far-field of an elastic metamaterial for Lamb waves. The experiment is performed in the kHz regime, with centimetric waves that are

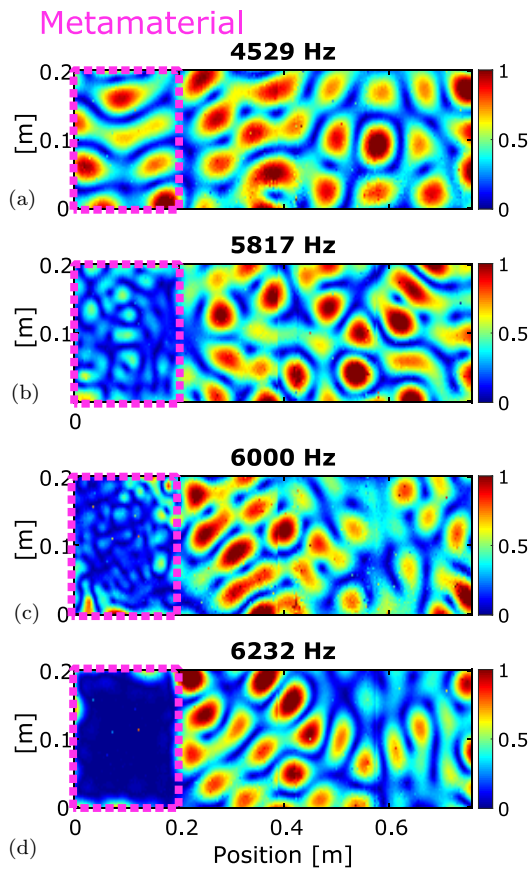


Fig. 6.9. Normalized spatial representation of the energy distribution within the rectangular surface represented in Fig. 6.1 at four frequencies ranging from the propagative regime (a) toward the full bandgap (d). Compared to Fig. 6.8, the colorbar is such that no saturation is observed in each panel. The purple dashed square indicates the location of the metamaterial. (Modified from Rupin *et al.*, 2015(a)).

refocused "in the computer" through cross-correlation computations (Fig. 6.10). The basis of this approach lies in the direct link established between sub-wavelength resolution results and the wavefield dispersion induced by a metamaterial made of long rods that are attached to a plate, and that basically transform the field to shorter wavelength propagation.

To realize sub-wavelength focusing from one channel in the far-field, it is necessary to optimize the use of sub-wavelength modes inside the metamaterial that correspond to modes located at the right-hand side of the $A0$ dispersion curve in Fig. 6.7(a). But, due to their strong attenuation inside of the metamaterial, a time reversal based technique (One-Channel Time-Reversal, OCTR, [Fink, 1997; Draeger and Fink, 1997]) is not the best suited. Indeed, even though time-reversal, or equivalently cross-correlation, is well known to produce focal spots on the order of one wavelength from one single channel in the presence of strong reverberation, it cannot go well beyond the diffraction limit which requires near-field components or, equivalently, locally-damped sub-wavelength modes. To overcome this limitation, we use a more elaborated focusing technique– so-called One-Channel Inverse Filter (OCIF) — based on the inverse filter approach [Tanter *et al.*, 2001, Gallot *et al.*, 2012]. This technique was recently reported to allow optimal use of all of the temporal degrees of freedom available in a reverberated wave-field for focusing purposes. When using the OCIF focusing technique [Rupin *et al.*, 2014b, 2015a] within the metamaterial, the -3 dB width of the focal spots is no longer wavelength-dependent (Fig. 6.10, black dashed lines). This is a consequence of the successive hybridizations that generate binding branches that give rise to extreme wavenumbers of π/a, where a is the average distance between resonators, as shown in Fig. 6.7(a).

In conclusion, the key parameter of the successive hybridizations is the spacing between the rods, which gives the fundamental upper limit of the reachable wave numbers (π/a). A decrease in the intensity in the metamaterial goes along with the hybridization effect, which we showed through spatial distribution maps of the wave-field at different frequencies. Finally, we have demonstrated the direct connections between the highest wave-numbers and the optimal focal spot, through the use of an advanced single-channel focusing technique. The One-Channel Inverse Filter allows us to reach optimal super-resolution of $\lambda/6.2$ without involving any near-field effects.

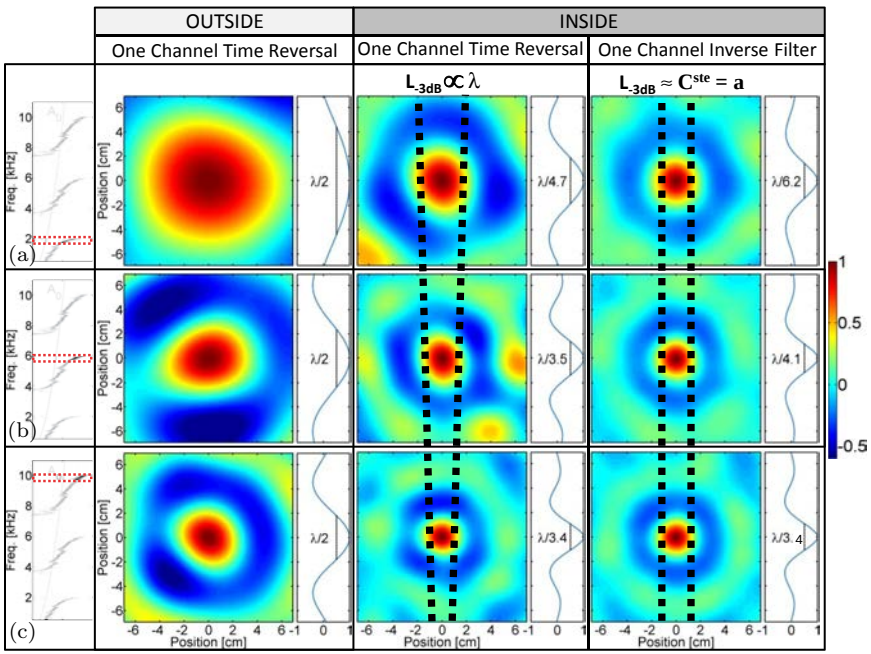


Fig. 6.10. Spatial representation of the focal spots obtained for the three frequency ranges that correspond to the binding branches due to the hybridization with the three compressional resonances of the rods: (a) $f_0 \sim 2$ kHz; (b) $f_1 = 3f_0$; and (c) $f_2 = 5f_0$ (see red dotted rectangles in the dispersion relation in the left column). For every frequency range, the reference -3 dB width (at $\lambda/2$) is defined as the focusing obtained outside the metamaterial with the cross-correlation method (it gives a -3 dB width, in agreement with the wavelength of the corresponding central frequency). Inside the metamaterial, the single-channel focusing techniques are compared: one-channel time reversal and the one channel inverse filter. (Modified from Rupin *et al.*, 2015(a)).

6.3. Part II: Numerical Results

Spectral element simulations were chosen to numerically compute the plate motion using the SPECFEM3D software (Komatitsch and Vilotte, 1998; Peter *et al.*, 2011). The discretization of three-dimensional models that embed several sub-wavelength elements is notoriously challenging with respect to meshing and the numerical cost required for the time integration. The mesh was designed using an adaptive scheme that confines a finer mesh region around the

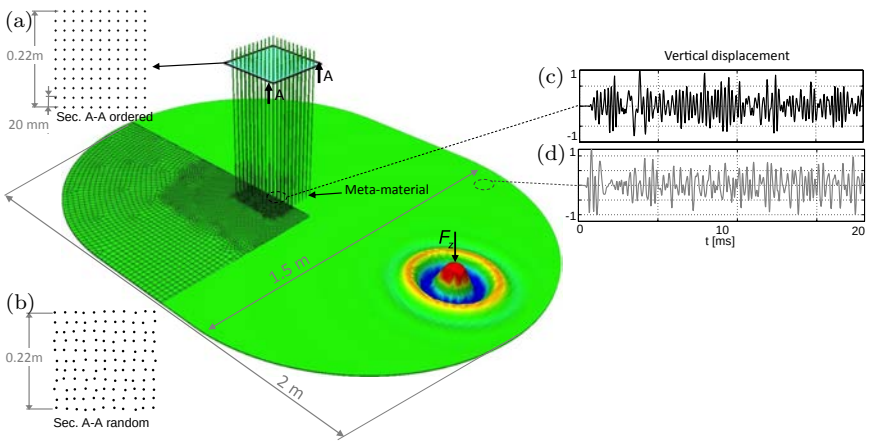


Fig. 6.11. Side view of the plate geometry. A cluster of 11×11 sub-wavelength resonators constituting the meta-material are vertically attached to a 6 mm plate. The result of the adaptive meshing, characterized by three different coarsening levels, is superimposed to part of the model. A Ricker point source (F_z) with a central frequency of 4 kHz excites flexural waves (A_0 mode) in the plate. The cross-sections (a)–(b) show the relevant dimensions of the cluster for a regular and a random distribution of resonators. Each beam has a 4×4 mm square section and a length of 60 cm. Note that the wavelength of the A_0 mode is much longer than the section of the resonators. (c)–(d) The signals depict the displacement recorded for a point outside (grey line) and inside the cluster (black line) at the surface of the plate. Amplitudes are normalized. (Modified from Colombi *et al.*, 2014).

metamaterial, with a coarser mesh elsewhere (Fig. 6.11). The very small time step (10^{-8} s) that results from the stability condition (Peter *et al.*, 2011) is mainly responsible for the computational cost. After the application of a domain decomposition algorithm, the simulations that are run in parallel require approximately 100 CPU hours for a 20-ms long wave-field. The source excitation is a Ricker pulse that is centered at 4 kHz, and is intentionally misaligned with the center of the metamaterial to avoid symmetries, while it is kept at a finite distance to avoid nearfield effects. The plate is assumed to be isotropic, homogeneous, and non-dissipative. The three-component displacement vector is recorded at any point of the mesh, with a sampling frequency of 100 kHz [e.g., Figs. 6.11(c)-(d)]. A bandpass filter

between 1 and 8 kHz is applied to the signal, to eliminate numerical noise. The vertical point source excites flexural waves ($A0$ mode) in the plate (Colombi *et al.*, 2014(a), 2014(b)). Lamb wave dispersion in the 1 to 8 kHz frequency band leads to wavelengths that range from about 5 to 15 cm, which confirms the sub-wavelength characteristic of the cluster of beams. As expected from experimental results, reverberation on the plate boundaries in this ergodic cavity induces long time-dispersed signals that spread over more than 20 ms.

To limit the influence of the ballistic wave [Fig. 6.12(a)], only signals that are recorded after 10 ms, with mainly reverberated waves [Figs. 6.12(b)–(d)] are considered in the following study [Colombi *et al.*, 2014]. The motion of $A0$ Lamb waves in the plate is mostly polarized along the vertical z axis, although a minor part of the field is converted on the basis of the resonators along the horizontal

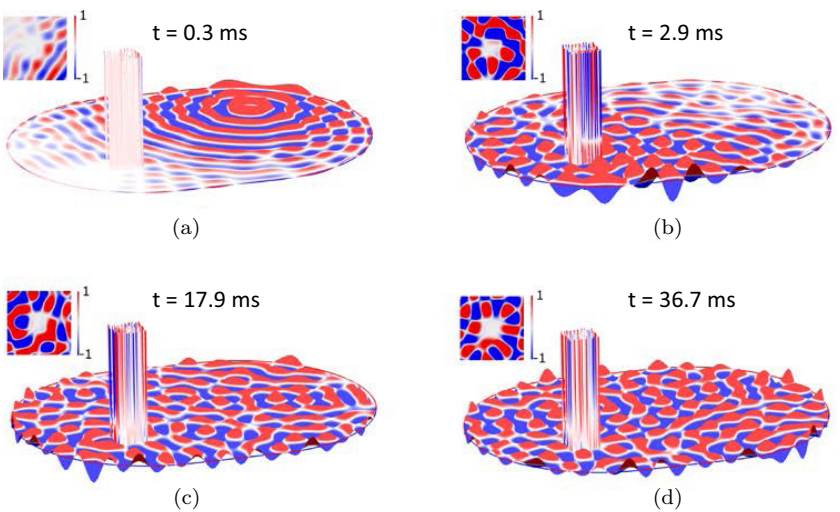


Fig. 6.12. Spatial display of wave-field snapshots for the vertical component at different times. Records have been filtered in the first band-gap [2100 Hz − 2800 Hz] to highlight the wave-field cancellation inside the meta-material. Vertical resonators show the position of the metamaterial on the plate surface. In each panel, the inset represents a zoom of the field on a 40 cm × 40 cm area centered on the metamaterial. The full video is available as multimedia material in Colombi *et al.*, 2014a.

x and y directions too, thus creating a complex pattern. Numerical simulations facilitate the analysis, as each component of the elastic wave-field is available at any discrete point of the mesh, both in the horizontal plate and along the vertical dimension of the beams. This numerical study provides a novel approach for the characterization of metamaterials made of closely spaced resonators, which can be considered to be complementary to theoretical effective medium approaches [Torrent and Sanchez-Dehesa, 2006; Yang *et al.*, 2014] or analog experiments at the laboratory scale. In an actual experiment, the measurement of the elastic wave-field accurately in time and space for such a complex structure is impossible, because of the limits in the probing of the whole plate surface (or worse, the vertical dimension of each beam). From a visual inspection of the field in the plate filtered in the first bandgap [Figs. 6.12(c)–(d)], the wavefield cancellation is obvious in the whole region enclosed by the first outer layer of the resonators.

6.3.1. *Presence of one defect*

Although bandgaps are an essential characteristic of the metamaterial considered here, the scope of this section is rather to address another phenomenon, which is defined as the energy trapping in the presence of a defect within the resonators [Kaina *et al.*, 2013b; Khelif *et al.*, 2003, 2004]. The reference models [presented in Figs. 6.11(a)–(b)] are characterized by either a regular or a random spatial distribution of beams of the same length. The insertion of a defect, e.g., a shorter beam at the center of the structure, modifies the wave properties inside the cluster by trapping energy at the defect. Attention is now drawn to the propagative band from 3 to 6 kHz with the signals treated here assumed to be bandpassed accordingly. To account for the elastic wave-field in the resonator without making a component-by component analysis, it is convenient to calculate the local intensity I in each beam as:

$$I(x,y) = \frac{1}{T}\sum_{i=1}^{3}\sum_{k=1}^{N}\int_{t_0}^{t_0+T} u_i^2(x,y,z_k;t)\,dt, \tag{6.4}$$

where (x, y) are selected pairs on the $x-y$ plane, L is the total length of the beam, z_k is one element of the regularly sampled bar length (with N elements dz between 0 and L), u_i is the ith component of the displacement (x, y, z) in the whole frequency band [Figs. 6.11(c)–(d)], and $[t_0, t_0 + T]$ is the time integration window (10–20 ms). The (x, y) pairs are selected on an 11×11 grid, such that there is a sampling point for each beam in both the regular and the random arrangements. The value of $I(x, y)$ for the reference configuration [Fig. 6.13(b)] where the beams share the same length ($L = 60$ cm)

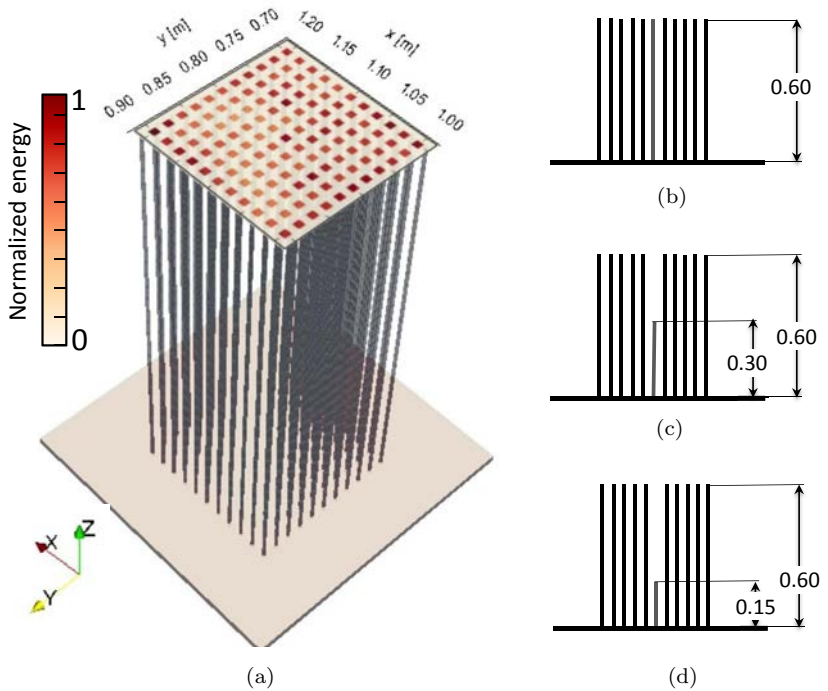


Fig. 6.13. (a) The value of the beam intensity $I(x, y)$ is superimposed on the metamaterial (only the resonator region is depicted) for data filtered in the propagative band [3000 Hz − 6000 Hz]. The color map is normalized. The resonator size corresponds to the reference configuration depicted in (b). (b) Slice view of the central row (or line) of the beam cluster used in the reference configuration. (c) Same as (b), but the central resonator is halved in length. (d) Same as (c), with a shorter central resonator. (From Colombi *et al.*, 2014).

is depicted in Fig. 6.13(a). The resonators accumulate and convert the vertical motion caused by the $A0$ mode excitation. The energy trapped in the cluster of resonators is then released progressively to the plate. Thus, the local energy is much larger in the beam than at any point of the plate surface, and for this reason the plate surface is not involved in the analysis. By taking into account only the reverberated part of the signals, the distribution of energy appears quite homogeneous among the resonators.

Figures 6.13(c) and 6.13(d) show two configurations of the cluster that include a defect that consists of a central beam of different length. The plot of $I(x, y)$ in Fig. 6.14 demonstrates that in this case, the energy is concentrated in the shorter resonator. By normalizing the local energy function by its maximum value in Fig. 6.13, a $\sim$2 factor is found that is attributed to the magnitude of the energy trapping. A further reduction of length [Fig. 6.13(d)] results in a stronger amplification [Fig. 6.14(b)]. The longitudinal resonant frequency f_0 is shifted to a higher frequency, at 4.2 kHz in the case of the 30-cm long beam, and at 8.5 kHz for the 15-cm-long beam. The shift of f_0 inside the propagative band cannot explain the energy trapping for the shortest beam, and further work is required to understand this

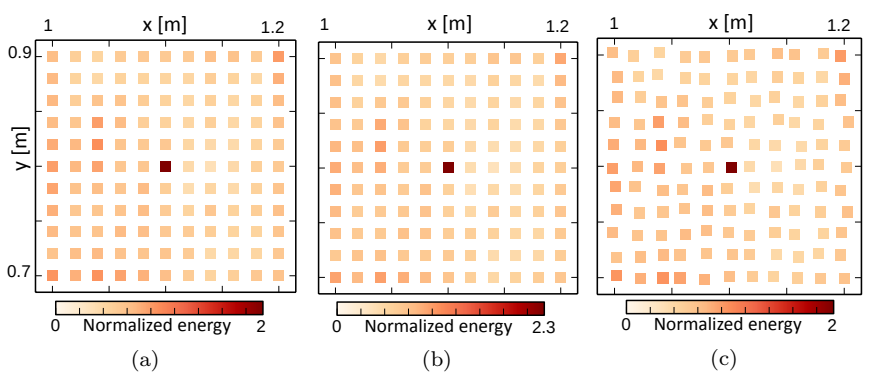


Fig. 6.14. (a) The value of the beam intensity $I(x, y)$ for the configuration depicted in Fig. 6.13(c). The normalization is the same as in Fig. 6.13(a). (b) Same as (a), but for the configuration in Fig. 13(d). (c) Same as (a), but using a spatially random arrangement. The data were filtered in the propagative band [3000 Hz − 6000 Hz] in all cases. (From Colombi *et al.*, 2014).

phenomenon. Finally, note that the energy amplification would have been more spectacular if instead of the total energy, the energy per unit length had been considered, by dividing $I(x, y)$ by the length L of the beam.

As observed in other studies (Kaina *et al.*, 2013(b)), the elimination of one beam does not lead to any focusing effect. This configuration would result in an empty cavity within the cluster that does not support any local resonance effect. To demonstrate that the energy trapping is not caused by collective modes (e.g., Bragg scattering), Fig. 6.14(c) shows an example where the beams are spatially disordered. The energy boost is similar to the ordered case, and hence this confirms that local resonances are at the origin of the phenomenon. It is worth noting that in each of the proposed configurations, the energy trapping occurs from far-field excitation in a frequency band outside the bandgaps, while in the study from Kaina *et al.* (2013(b)), it was only observed within the bandgap from a near-field source. In another simulation (not shown here), the case where the shorter beam is not at the center but is closer to the outer layer of the random cluster [Fig. 6.11(b)] was tested, with the same conclusion.

6.3.2. *Directional cloaking*

We now address the numerical design of a directional invisibility cloak for backward-scattered elastic waves propagating in a thin plate (excited with $A0$ Lamb waves, [Farhat *et al.*, 2009]). The directional cloak is based on a set of resonating beams that are attached perpendicular to the plate and are arranged at a sub-wavelength scale in 10 concentric rings [Fig. 6.15(b)]. The exotic effective properties of this locally resonant metamaterial ensures coexistence of bandgaps and directional cloaking for certain beam configurations over a large frequency band. The best directional cloaking was obtained when the resonator length decreases from the central to the outermost ring.

The metamaterial is realized using $\sim$180 vertical rods (resonators) of varying lengths that are fixed perpendicular to a 6-mm-thick aluminium plate (Fig. 6.15(a)). The beams have an equal, square 4×4 mm^2 section. Depending on the directional cloak

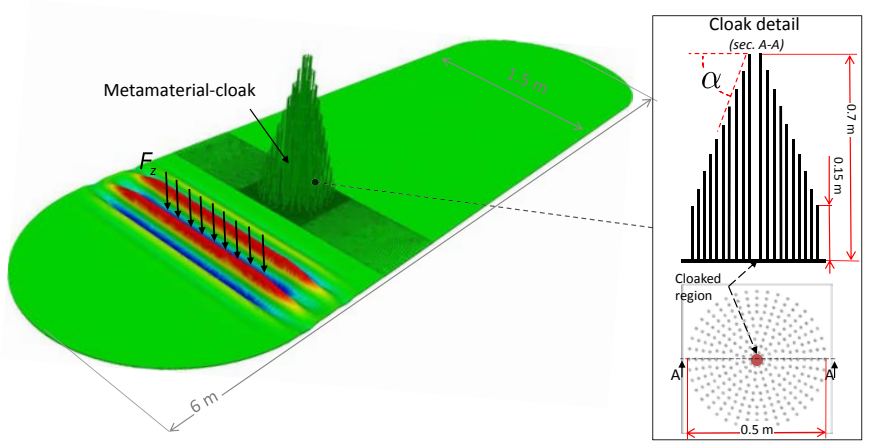


Fig. 6.15. (a) The 6-mm-thick plate and the cloak made with the metamaterial used in the simulations. The adaptive meshing technique is shown around the cloak region. A plane wave is generated by a Ricker source time function acting along the vertical direction, exciting the $A0$ modes. (b) The detailed structure of the cloak made of 10 rings of beams that are almost equally spaced, to maintain a constant density. A cross section view (Sec. A–A) shows the different shapes of the cloak profile defined by the angle α. (Modified from Colombi *et al.*, 2014).

type considered, the resonator lengths vary with a smooth gradient, from the inner to the outer layer (Fig. 6.15(b)). The plate is elongated in the longitudinal direction to allow long propagation and to ensure that the first wave-train propagates for a sufficient distance unperturbed by boundary reflections. The rods are arranged as a closely spaced cluster, and they are made of the same material as the plate; this cluster occupies a circular surface of $\sim$0.5 m diameter (see Fig. 6.15(b)). There is a 5-cm-diameter cavity (i.e., a resonator-free region) at the center of the cluster, and this is referred to as the 'cloaked' region.

Spectral element simulations using the highly parallelized software SPECFEM3D and the meshing tool CUBIT [https://cubit. sandia.gov] were applied in a similar fashion to that described in [Colombi *et al.*, 2014(a)]. The plate was excited with a plane wave centered at 4 kHz using a broadband vertical Ricker point source (Fig. 6.15(a)) that

extended from 1 kHz to 8 kHz. As the present analysis focuses on the plate surface and not along the resonators, only the vertical component of the displacement u_z on the plate is considered. The wavefield is recorded at any corner point of the finite elements for the whole simulation length, which allows quick access for post-processing and data analysis. As in Fig. 6.11, each simulation calculates a 20-ms-long signal that is sampled at 100 kHz, and is band-pass filtered between 1 kHz and 8 kHz, to eliminate numerical noise.

Despite the recent advances in the field of elastic-wave cloaking, the coexistence of both protection and cloaking regimes (although conceptually possible [Norris, 2008; Chen *et al.*, 2010; Brun *et al.*, 2009; Guenneau *et al.*, 2011; Norris and Shuvalov, 2011; Parnell and Norris, 2012; Milton *et al.*, 2006, Diatta and Guenneau, 2014]) did not turn into an actual application and only a few laboratory demonstrations have been realized [Stenger *et al.*, 2012; Scherrer *et al.*, 2013; Xu *et al.*, 2015]. We introduce a directional cloak that combines the two concepts for the backward-scattered field. Backward means in the counter direction with respect to the incident field (as shown with arrows in Fig. 6.16). The more complex case of a full cloak (which controls both backward and forward scattering) will be addressed in a future study. Thus, a directional cloak with a circular layout made of 10 concentric rings of resonators where each ring is characterized by a series of resonators of identical length (i.e. the resonators's length is a function of the radius only) is introduced. As the effective properties of the metamaterial are given by local resonances and not by Bragg scattering, the resonator layout does not need to be ordered. However, the spatial arrangement of resonators locally modifies the refraction index which usually leads to bending or lensing effect and may even generate negative index of refraction [Kaina *et al.*, 2015]. In particular, the choice of a circular shape has the advantage of being isotropic, and hence being independent of the azimuth of the incident field. While the (x, y) positions of the resonators in the circular layout remain fixed, a gradient of resonator length is applied from the outer ring (with L = 40 cm) to the inner ring (with L = 70 cm, Fig. 6.15(b)).

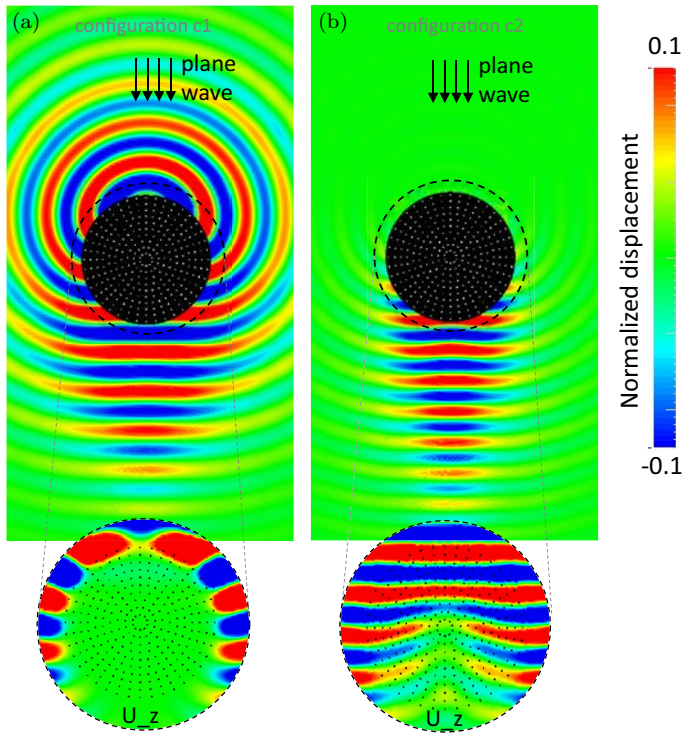


Fig. 6.16. (a) Differential field for a frequency band in the first bandgap (configuration c1) at t = 2 ms. Arrows indicate the approaching direction of the wave front. The cloak region is darkened to facilitate the interpretation. The vertical component of the field is represented in the detailed view of the cloak. (b) As for (a), but for the optimal directional cloak in the second bandgap (configuration c2).

In Fig. 6.16, the differential field $\tilde{u}_z - u_z$, where $\tilde{u}_z$ is the reference field without the metamaterial and $u_z$is the field in the presence of the metamaterial, is represented for two frequency bands inside the first and second bandgaps around frequency $f_1 = 2.5$ kHz and $f_2 = 5.6$ kHz, respectively. As well as the obvious absence of back-scattering with the f_2 excitation, the difference between the total field for waves centered at f_1 and f_2, as depicted in the close-up sections of Fig. 6.16, should be noted. In configuration f_2, the waves penetrate deep within the cloak structure, undergoing an acceleration. At the same time, waves do not propagate inside the central region. In configuration

f_1, the incident wave undergoes total reflection. In other words, the radially-varying dispersion relationship prevents the backscattered field to be retransmitted in the plate. This energy contribution bends around the cloak and is finally scattered in the forward direction.

Thus, a novel type of metamaterial for two-dimensional plates with resonators of varying lengths arranged in a concentric fashion around a circular region of a plate is numerically shown to cancel the back-scattering and to protect the cloaked region from an incident flexural wave over a large 300-Hz bandwidth within a range of stop-band frequencies. One way to interpret these numerical results is through an effective refractive index that vanishes at the inner boundary of the cloak, in agreement with earlier studies [Farhat *et al.*, 2008; Fahrat *et al.*, 2009; Stenger *et al.*, 2012]. One way to improve the cloaking (i.e., to also cancel the forward scattering) would be to use a radially-dependent plate thickness to achieve an anomalous effective refraction index [Climente *et al.*, 2014].

6.4. Part III: From "Acoustic" to "Elastic Plate" Metamaterial

In this section, we study the effect of the plate thickness on the dispersion curves obtained inside a locally-resonant metamaterial. Looking at the dispersion curve measured at the core of a multi-rods metamaterial in the left panel of Fig. 6.17, a wide bandgap between 2 kHz and 4 kHz associated with the compressional resonance of one rod is observed, a result comparable to the analog experiment described in Fig. 6.7. Indeed, there are strong similarities between the present experimental set-up and that studied previously on the effect of a cluster of 60-cm long metallic rods glued at a sub-wavelength scale to a 6-mm thick metallic plate (Fig. 6.1). The only difference lies in the thickness of the plate, which is reduced in the present case to only 2 mm, while the dimensions of the rods and their random spatial distribution remain identical. As a consequence, new phenomena are observed here that were not visible in the previous studies with the thicker plate. As the flexural rigidity of a plate is proportional to the cube of its thickness, this difference has the consequence of rendering

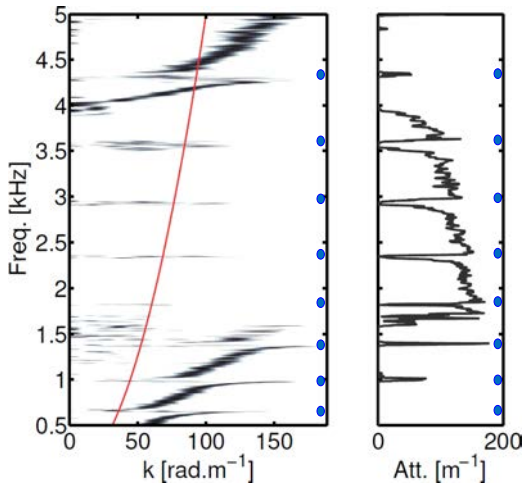


Fig. 6.17. (a) Frequency-wavenumber representation of the dispersion curve inside the metamaterial determined from experimental measurements of the vertical component of the field performed with a 3-mm thick flexible plate, to be compared to the equivalent dispersion curve obtained in the 6-mm rigid plate (Fig. 6.7). As with the rigid plate, a band structure is revealed which is different from the dispersion curve of flexural waves in the free plate (red curve labelled A_0). However, propagative bands appear inside the first bandgap that correspond to the flexural resonances of one single rod and confirm the coupling between the vertical and horizontal components of the field. (b) Imaginary part of the k-vector measured inside the metamaterial. The strong attenuation inside the first band gap [1700 Hz − 4000 Hz] associated to the longitudinal resonance of one rod gives rise to transmission bands at each flexural resonance of one rod. The frequencies associated to flexural resonances are represented as blue dots in each panel (Modified from Rupin *et al.*, 2015b).

the present 2-mm thick plate more flexible than the previous 6-mm plate. This is of critical importance for the influence of the flexural resonances of the rods. While their effects were marginal with the 6-mm plate, where the hybridization of the wave-field was mainly due to the compressional resonances of the rods, here they become influent enough to be studied in detail.

First, three narrow transmitted bands in the frequency interval of the compressional bandgap are clearly associated with the flexural

resonances of one rod whose frequencies are computed and represented as blue dots in Figs. 6.17(a)–(b). Second, before the bandgap at 2 kHz, a few narrow forbidden bands are seen that are also connected with flexural resonances. Indeed, the continuous bending branch that was observed below 2 kHz with the 6-mm plate in Fig. 6.7 is transformed for the 2-mm plate into a succession of propagative bands that have an asymptotically flat band at every flexural resonance frequency.

This is the polariton behavior that is expected when the free-space dispersion relation crosses a flexural resonance frequency, which was not observed with the more rigid plate. Note that flexural resonances for a vertical rod attached to a plate have a high quality factor compared to compressional ones when the plate is excited by vertically-polarized $A0$ mode [Rupin *et al.*, 2014a]. Indeed, the coupling between the out-of-plane vertical displacement in the plate naturally excites each compressional resonance, inducing a strong leakage from the plate to the bar, and is orthogonal to the motion associated to flexural resonances in the rod. Flexural resonances with high quality factor induce narrow bandgaps in the locally resonant metamaterial.

Thus, the signature of flexural resonances of the rods from the out-of-plane vertical displacements measured on the flexible plate inside the metamaterial confirm the strong coupling between vertical and horizontal components of the field associated to the local interactions of the $A0$-type vertically polarized incident waves and the cluster of resonators [Rupin *et al.*, 2015b]. This result is confirmed by numerical simulations on a large frequency band with two plates of different thicknesses (Fig. 6.18) on which were attached the same number of identical vertical rods. We also note that the width and frequency interval of the bandgaps are different in these two configurations, which seem at first sight in contradiction with the idea that bandgaps are solely determined by the resonance and anti-resonance frequencies of one rod (see Eq. (6.1)). In practice, however, these resonance and anti-resonance frequencies are not only determined by the length and elastic properties of one rod but also by the coupling

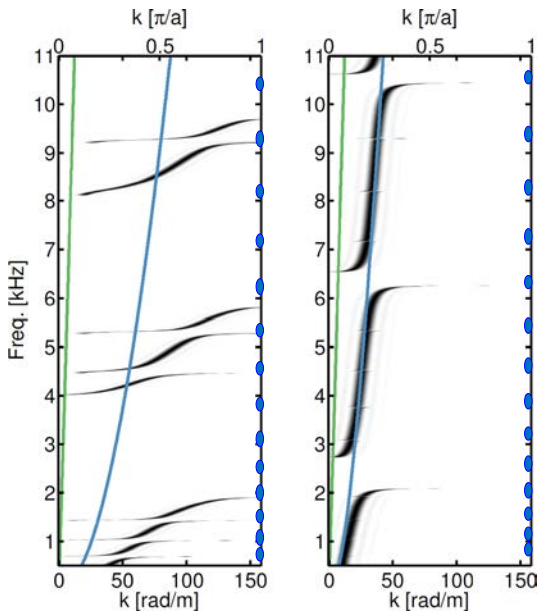


Fig. 6.18. Frequency-wavenumber representation of the dispersion curve inside the same metamaterial determined from numerical simulations of the vertical component of the field performed with two plates of (a) 6-mm and (b) 30-mm thickness. Both S_0 (green) and A_0 (blue) modes for the free plate are represented. In (a), this numerical simulation magnifies the role of flexural resonance as forbidden frequencies inside the three propagative bands (below 2 kHz, from 4 kHz to 6 kHz, from 8 kHz to 10 kHz). Because of noise and damping, this effect was barely observed in experimental measurements for a plate of the same thickness (see Fig. 6.7). In (b), note that the band gaps are much narrower which means that they are not solely determined by the resonator properties. The effect of flexural resonances on the propagative A_0 mode in the free plate are nearly invisible for a thick and rigid plate. The frequencies associated to flexural resonances are represented as blue dots in each panel.

between the rod and the plate (assumed to be perfectly clamped in Eq. (6.1)), which is strongly dependent on the local rigidity of the plate and, thus, directly associated to the plate thickness.

Finally, Fig. 6.19 also confirms with a 1-D numerical simulation the presence of transmission bands inside the band gap at the flexural frequencies of the resonators for a thin flexible plate. These peaks

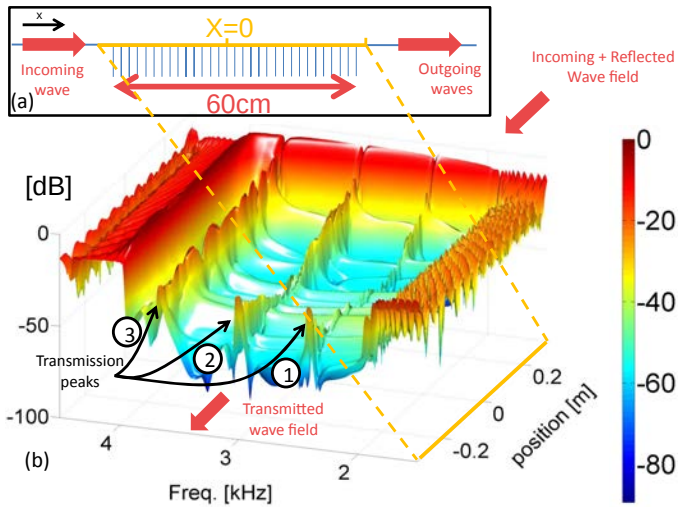


Fig. 6.19. (a) Schematic of the numerical 1-D configuration tested with vertical resonators attached to a horizontal beam. The beam and resonators elastic parameters are identical to the 2-D plate characteristics in Fig. 6.1. (b) Spatial-vs-frequency representation of the vertical field measured inside the metamaterial. Inside the first bandgap [2000 kHz – 4000 kHz, see Fig. 6.18(a)], three transmission peaks are observed at frequencies that correspond to the flexural resonances of the vertical resonators.

correspond to the transmission peaks observed in the experimental data between 2 kHz and 4 kHz (Fig. 6.17).

6.4.1. *From the plate physics to the semi-infinite half-space geophysics*

In the second part of this section, we continue to expand the width of the plate up to infinity, transforming the rod and plate mechanical system into a rod and semi-infinite half space geophysics medium. Although on the microscale, previous work on micropillar resonators on an elastic half space had already proven the existence of sub-wavelength bandgaps induced by local resonance [Khelif *et al.*, 2012; Achaoui *et al.*, 2013], we move here to larger geophysics scale and the Hz's regime exploring the thesis that resonances in trees result in forests acting as locally resonant metamaterials for Rayleigh surface

waves [Aki and Richards, 1980] in the geophysics context. A geophysical experiment demonstrates that Rayleigh wave, propagating in soft sedimentary soil at frequencies lower than 150 Hz, experiences strong attenuation, when interacting with a forest, over two separate large frequency bands. This numerical experiment is interpreted using finite element simulations that demonstrate the observed attenuation is due to bandgaps when the trees are arranged at the subwavelength scale with respect to the incident Rayleigh wave. The repetitive bandgaps are generated by the coupling of the successive longitudinal resonances of trees with the vertical component of the Rayleigh wave. For wavelengths down to 5 meters, the resulting bandgaps are remarkably large and strongly attenuating when the acoustic impedance of the trees matches the impedance of the soil. Since longitudinal resonances of a vertical resonator are inversely proportional to its length, a man-made engineered array of resonators that attenuates Rayleigh waves at frequency $\leq$ 10 Hz could be designed starting from vertical pillars coupled to the ground with height larger than 30 m. The quest for a seismic metamaterial able to manipulate seismic waves is a new and challenging topic that involves only a few prior studies [Woods, 1968; Brule *et al.*, 2013; Sheng *et al.*, 2014; Kroedel *et al.*, 2015; Colombi *et al.*, 2016; Ungureanu *et al.*, 2016]. We point out that research papers on Rayleigh wave attenuation in marble quarry with air holes displaying kHz stop bands [Meseguer *et al.*, 1999] and similar filtering effects in microstructured piezoelectric for MHz surface waves [Benchabane *et al.*, 2006] triggered the interest in control of Rayleigh surface waves in the phononic crystal community. However, the present model of locally resonant metamaterial offers wave control in a subwavelength fashion and lower target frequency. Both characteristics are well suited for potential future seismic applications. Interestingly, one can achieve some protection against Rayleigh waves with columns of soft material buried in the soil, by extending the concept of conformal optics [Leonhardt, 2006] to geophysics [Colombi *et al.*, 2016]: Indeed, four Luneberg lenses designed with a spatially varying soil's density according to the transformation seismology approach, and arranged in a checkerboard

fashion, detour Rayleigh waves around a cloaked region without any reflection. This new path to seismic cloaking, which does not require stop bands, is based upon an effective medium approach akin to the Maxwell-Garnett theory.

Because of the deeply subwavelength microstructure of locally resonant metamaterials, it is essential to explore the wavefield within the resonator array with spatio-temporal details that would require thousands of seismometers in the present geophysical configuration [Rost, 2002]. The physics of the subwavelength structure is accurately analyzed in this paper through time domain spectral element simulations. This method has been successfully applied to study the case of the metamaterial plate and rods [Colombi *et al.*, 2014(a), 2015]. By restricting the analysis to Rayleigh waves we reduce the complexity to a 2D halfspace (hence 2D simulations) and consider a linear, isotropic and homogeneous elastic medium with a linear array of 30 trees (Fig. 6.20). The simulations are performed with SPECFEM2D [see Komatitsch and Vilotte, 1998, for further details]. Perfectly matched layer conditions [so called PML, Komatitsch and Martin, 2007] are applied on the bottom and vertical boundaries of the halfspace which is otherwise traction-free. The 30-m thick halfspace is characterized by a homogeneous material with shear velocity $v_s^g = 500$ m/s and density $\rho_g = 1300$ kg/m^3. For a Poisson ratio typical of soil, the Rayleigh wave speed $v_r^g \approx v_s^g$. While these parameters are representative for an average soil, the results discussed next are not strictly limited to this wave speed but they can be generalized to very soft soils featuring $v_s^g < 300$ m/s.

Each tree is represented as a homogeneous elastic vertical resonator with constant thickness (hence constant cross-section) characterized by both longitudinal and flexural modes [Chopra, 1995, Ewins, 2000]. We use trees of both constant and random size as well as random spacing between trees to evaluate the effects of the variability that characterizes natural forests; the heights are drawn from a uniform distribution with mean 14 m varying between ± 2.5 m. The regular and random configurations are shown in Figs. 6.21(a)–(b). Typical mechanical properties of wood can be found in [Green *et al.*,

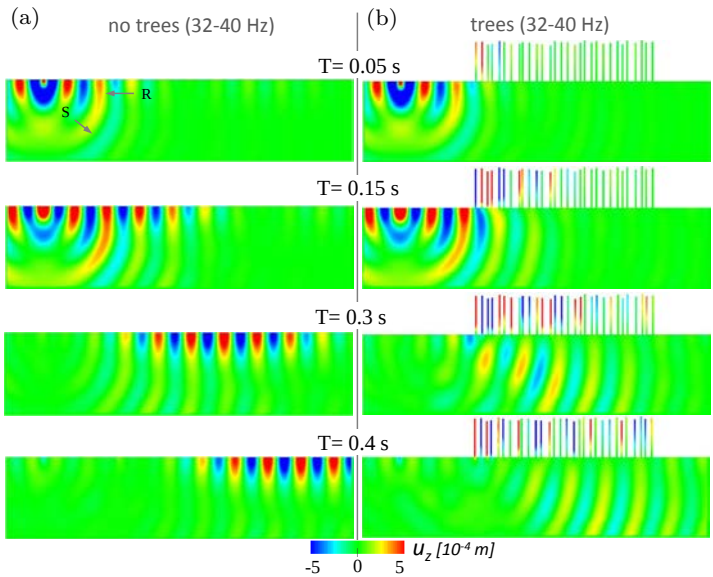


Fig. 6.20. Snapshots at different times of the vertical component of the field u_z propagating from left to right into the metamaterial made of 14-m long randomly-spaced trees on a semi-infinite ground. The panels (a) correspond to the reference medium in the absence of metamaterial. R and S indicate surface Rayleigh and body S-waves. The field was filtered in the first bandgap (see Fig. 6.21) for which the Rayleigh wave is damped very fast inside the forest of trees. (Modified from Colombi *et al.*, 2016).

1999] and vary widely depending on species, fluid content, and age. From this data, a reasonable approximation, is density, ρ_t of 450 kg/m^3. Using elastic velocities $v_s^t = 1200$ m/s and $v_p^t = 2200$ m/s, the Lamé parameters are well within the ranges given for wood.

The interactions between surface waves and trees underlying the experimental results are captured by analyzing different snapshots of the incident field propagating, from left to right, through the forest [Fig. 6.20(a)–(b)]. The vertical force is driven by a Ricker source time function centered on frequency 60 Hz. Most body wave energy generated by the vertical force disappears through the bottom boundary, as it should, leaving only horizontally propagating Rayleigh waves. A reference simulation of the halfspace without trees

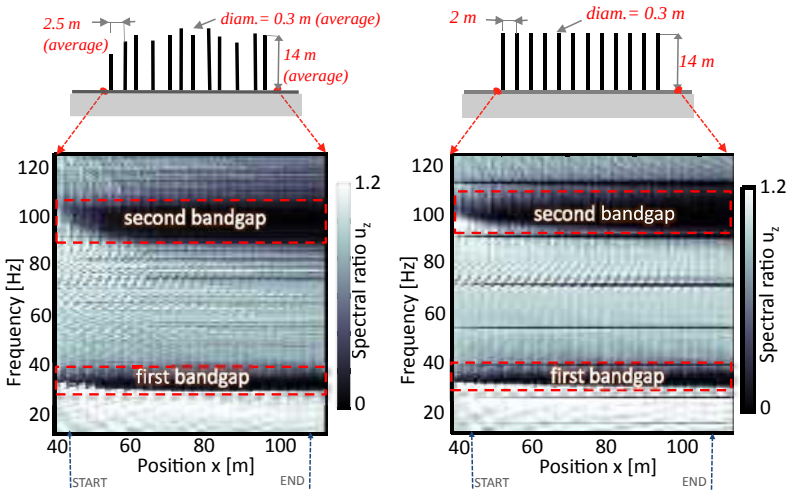


Fig. 6.21. Spectral ratio computed from numerical measurements inside versus outside of the metamaterial and shown here for two different forest configurations. The start and end points of the measurements are marked on the x-axis. (a) Random distribution of trees with different height and spacing; (b) equally-distributed trees with the same height (14 m). The first and second band gaps are indicated by the dashed red lines. For the regular arrangement of trees, note the presence of narrow forbidden frequencies between the two bandgaps associated to the flexural resonance of one tree. (Modified from Colombi *et al.*, 2016).

is used to compute a spectral ratio, ensuring the frequency spectra in Figs. 6.21(a)–(b) do not depend on the source spectrum.

In Fig. 6.21 is depicted the behavior of the spectral ratio as we penetrate deeper into the array of trees. Notice the two bandgaps appearing progressively after a so-called skin layer. As we proceed towards the tail of the array (at $x \sim 115$ m) the bandgap width no longer changes. Because of the absence of a depth dependent velocity gradient in the ground, waves diving toward depth do not return to the surface. Hence, the bandgaps persist behind the trees. It turns out that, as in the plate case, longitudinal resonance and anti-resonance frequencies determine start and end points of the bandgap. This is clearly shown by considering an array of equally sized and spaced trees [Fig. 6.21(b)]. The absence of height, and thickness variability

result in two main bandgaps with the same width as the random configuration [Fig. 6.21(a)] and narrower bandgaps due to flexural resonances.

A sensitivity analysis on the acoustic impedance of the ground $\rho_g\, v_s^g$ and the trees $\rho_t\, v_l^t$ where v_l^t is the tree's longitudinal wave velocity is performed; parameters chosen for the simulation in Fig. 6.20 are for a unitary ratio between the two impedances and a fluctuation of the ratio between 0.5 and 2 (well within admissible material parameter ranges) results in very similar bandgaps. For stronger impedance mismatch however, they tend to disappear progressively, confirming the mechanical coupling between ground and trees is the driving parameter.

As expected, the so-called hybridization phenomenon [Kaina *et al.*, 2013a, Fano, 1961, Cowan *et al.*, 2011] drives local resonances and bandgaps in this type of locally-resonant metamaterial. Indeed, the compressional resonances of tree-like resonators, excited by the vertical component (u_z) of the Rayleigh waves, introduces a phase shift of π on the incident waves causing a reflection of the wavefield around the resonant frequencies. At anti-resonance, the point of attachment between ground and tree ($z = 0$, the forcing point) is at rest [Williams *et al.*, 2015, Ewins, 2000] and thus $u_z = 0$. Because the trees are arranged on a subwavelength scale, the cumulative effect of several trees over a wavelength interferes constructively thus creating a band gap between resonance and anti-resonance [Fig. 6.21(a)]. In this physical model, the role played by the narrower flexural resonances of the tree-like resonator is marginal, despite the expected coupling with the Rayleigh wave horizontal component, yet visible with forest made of trees of the same length [Fig. 6.21(b)].

6.5. Conclusion

To conclude, we have experimentally studied ordered and disordered Lamb wave metamaterials built from long metallic rods perpendicularly attached to a thin metallic plate as a typical example of locally-resonant metamaterial. Spatiotemporal maps of the wave field inside and outside the metamaterials were measured on a frequency

spectrum that spanned about a decade. Through this experiment and the equivalent numerical simulations, these composite media were shown to support sub- and supra-wavelength modes, as well as wide band gaps. In the most general case, the metamaterial physics is explained through hybridizations between the $A0$ and $S0$ plate modes with the flexural and compressional resonances of the rods. Fitting the experimental dispersion curve with an analytical formulation and numerical simulations, we have confirmed, in the particular case of a rigid plate, the scalar interaction between the incident $A0$ Lamb mode and the compressional resonances of the rods. The extent of the bandgap is then governed by the longitudinal resonance and anti-resonance of the resonator that is further confirmed using Floquet-Bloch theory. The key parameter of the successive hybridizations is the spacing between the rods, which gives the fundamental upper limit of the reachable wave numbers (π/a). A decrease in the intensity in the metamaterial goes along with the hybridization effect, which we show through spatial distribution maps of the wave-field at different frequencies. As classically observed with locally-resonant metamaterials, the wave field properties do not depend on the (ordered or disordered) spatial arrangement of the resonators inside the metamaterial. Using an advanced single-channel focusing technique, we have demonstrated the direct connection between the highest wavenumbers and the optimal focal spot. Introducing a defect inside the metamaterial, we also showed that energy carried by elastic waves can be confined on a sub-wavelength scale from the far field. Using two-dimensional plates with resonators of varying lengths arranged in a concentric fashion, a directional cloak was numerically shown to cancel the back-scattering and to protect the cloaked region from an incident flexural wave over a large bandwidth within a range of stop-band frequencies. Finally, expanding the size of locally-resonant metamaterial to the geophysics scale, we demonstrated conclusively that a forest of trees arranged on a sub-wavelength scale induce large frequency bandgaps for Rayleigh waves at tens of Hz. Interestingly, the variability in size and positions of trees produces larger bandgaps than for uniform configuration. All these results suggest

that numerous applications in electromagnetic metamaterials can be translated to Lamb waves in plates or surface waves in semi-infinite elastic media. In particular, this study should bridge exciting wave phenomena unveiled in photonic and platonic metamaterials with surface seismic waves propagation in soils structured on a large scale by natural, or man-made, collections of resonators.

Acknowledgements

ISTerre is part of Labex OSUG@2020. A. Colombi was supported by the Marie Curie Fellowship "Metacloak". E. Williams was supported by the U. S. Office of Naval Research. F. Lemoult, M. Rupin and G. Lerosey are partly supported by LABEX WIFI under references ANR-10-LABX-24 and ANR-10-IDEX-0001-02 PSL* and by ANR grant under reference ANR-13-JS09-0001-01.

Bibliography

[1] Achaoui, Y., Laude, V., Benchabane, S. and Khelif, A. (2013). Local resonances in phononic crystals and in random arrangements of pillars on a surface, *J. Appl. Phys.* **114**.

[2] Aki, K. and Richards, P. G. (1980). *Quantitative Seismology, Theory and Methods, vol. I and II* (W. H. Freeman).

[3] Antonakakis, T., Craster, R. V. (2012), High-frequency asymptotics for microstructured thin elastic plates and platonics, *Proc. R. Soc. A* **468**(2141), pp. 1408–1427.

[4] Assouar, M., Senesi, M., Oudich, M., Ruzzene, M. and Hou, Z. (2012). Broadband plate-type acoustic metamaterial for low-frequency sound attenuation, *Phys. Rev. Lett.* **101**(17), p. 173505.

[5] Belov, P. A., Hao, Y. and Sudhakaran, S. (2006). Subwavelength microwave imaging using an array of parallel conducting wires as a lens, *Phys. Rev. B* **73**, p. 033108.

[6] Benchabane, S., Khelif, A., Rauch, J.-Y., Robert, L. and Laude, V. (2006). Evidence for complete surface wave band gap in a piezoelectric phononic crystal, *Phys. Rev. E* **73**, p. 065601.

[7] Brule, S., Javelaud, E. H., Enoch, S. and Guenneau, S. (2014). Experiments on seismic metamaterials: Molding surface waves, *Phys. Rev. Lett.* **112**, p. 133901.

[8] Brun, M., Guenneau, S. and Movchan, A. B. (2009). Achieving control of in-plane elastic waves, *Appl. Phys. Lett.* **94**(6), p. 061903.

[9] Bunimovich, L. A. (1979). On the ergodic properties of nowhere dispersing billiards, *Commun. Math. Phys.* **3**, p. 295.

[10] Capon, J. (1973). Signal processing and frequency-wavenumber spectrum analysis for a large aperture seismic array, *Methods in Computational Physics* **13**, pp. 1–59.
[11] Chen, H., Chan, C. and Sheng, P. (2010). Transformation optics and metamaterials, *Nat. Mater.* **9**, pp. 387–396.
[12] Choi, M. S. H., Lee, Y., Kim, S. B., Kang, J., Shin, M. H., Kwak, K. Y., Kang, Y. H., Lee, Park, N. and Min, B. (2011). A terahertz metamaterial with unnaturally high refractive index, *Nature* **470**, pp. 369–373.
[13] Chopra, A. K. (1995). *Dynamics of Structures: Theory and Applications to Earthquake Engineering*, 4th edition. Prentice Hall, Englewood Cliffs, NJ.
[14] Christensen, J. and de Abajo, F. J. G. (2012). Anisotropic metamaterials for full control of acoustic waves, *Phys. Rev. Lett.* **108**(12), p. 124301.
[15] Climente, A., Torrent, D. and Sanchez-Dehesa, J. (2014). Gradient index lenses for flexural waves based on thickness variations, *Appl. Phys. Lett.* **105**, p. 064101.
[16] Colombi, A., Roux, P. and Rupin, M. (2014a). Sub-wavelength energy trapping of elastic waves in a metamaterial, *J. Acoust. Soc. Am.* **136**, p. EL192.
[17] Colombi, A., Boschi, L., Roux, P. and Campillo, M. (2014b). Green's function retrieval through crosscorrelations in a two-dimensional complex reverberating medium, *J. Acoust. Soc. Am.* **135**(3), pp. 1034–1043.
[18] Colombi, A., Roux, P., Guenneau, S. and Rupin, M. (2015). Directional cloaking of flexural waves in a plate with a locally resonant metamaterial, *J. Acoust. Soc. Am.* **137**, pp. 1783–1789.
[19] Colombi, A., Guenneau, S., Roux, P. and Craster, R. V. (2016). Transformation seismology: Composite soil lenses for steering surface elastic Rayleigh waves, *Scientific Reports* **6**, p. 25320.
[20] Cowan, M. L., Page, J. H. and Sheng, P. (2011). Ultrasonic wave transport in a system of disordered resonant scatterers: Propagating resonant modes and hybridization gaps, *Phys. Rev. B* **84**.
[21] Craster, R. V. and Guenneau, S. (2012). *Acoustic Metamaterials: Negative Refraction, Imaging, Lensing and Cloaking* (Springer, London).
[22] Deymier, P. A. (2013). *Acoustic Metamaterials and Phononic Crystals*, *Vol. 173* (Springer, New York).
[23] Diatta, A. and Guenneau, S. (2014). Controlling solid elastic waves with spherical cloaks, *Appl. Phys. Lett.* **105**(2), p. 021901.
[24] Draeger, C. and Fink, M. (1997). One-channel time reversal of elastic waves in a chaotic 2D-silicon cavity, *Phys. Rev. Lett.* **79**(3), pp. 407–410.
[25] Dubois, M., Farhat, M., Bossy, E., Enoch, S., Guenneau, S. and Sebbah, P. (2013). Flat lens for pulse focusing of elastic waves in thin plates, *Appl. Phys. Lett.* **103**(7), p. 071915.
[26] Engheta, N. and Ziolkowski, R. W. (2006). *Metamaterials: Physics and Engineering Explorations* (John Wiley & Sons, New York).
[27] Ewins, D. J. (2000). *Modal Testing: Theory, Practice and Application 2nd Edition.* (Research Studies Press, Baldock, UK).

[28] Fan, S., Villeneuve, P. R., Joannopoulos, J. D. and Haus, H. A. (1998). Channel drop tunneling through localized states, *Phys. Rev. Lett.* **80**, pp. 960–963.
[29] Fang, N. (2005). Sub-diffraction-limited optical imaging with a silver superlens, *Science* **308**, pp. 534–537.
[30] Fang, N., Xi, D., Xu, J., Ambati, M., Srituravanich, W., Sun, C. and Zhang, X. (2006). Ultrasonic metamaterials with negative modulus, *Nature Mater* **5**, pp. 452–456.
[31] Fano, U. (1961). Effects of configuration interaction on intensities and phase shifts, *Phys. Rev.* **124**, pp. 1866–1878.
[32] Farhat, M., Enoch, S., Guenneau, S. and Movchan, A. B. (2008). Broadband cylindrical acoustic cloak for linear surface waves in a fluid, *Phys. Rev. Lett.* **101**, p. 134501.
[33] Farhat, M., Guenneau, S. and Enoch, S. (2009). Ultrabroadband elastic cloaking in thin plates, *Phys. Rev. Lett.* **103**, p. 024301.
[34] Farhat, M., Guenneau, S., Enoch, S., Movchan, A. B. and Petursson, G. G. (2010). Focusing bending waves via negative refraction in perforated thin plates, *Appl. Phys. Lett.* **96**, p. 081909.
[35] Fink, M. (1997). Time reversed acoustics, *Phys. Today* **50**(3), p. 34.
[36] Gallot, T., Catheline, S., Roux, P. and Campillo, M. (2012). A passive inverse filter for Green's function retrieval, *J. Acoust. Soc. Am.* **131**, p. EL21.
[37] Gilbert, K. E.(1983). A propagator matrix method for periodically stratified media, *J. Acoust. Soc. Am.* **73**, p. 137.
[38] Gracia-Salgado, R., Torrent, D. and Sanchez-Dehesa, J. (2012). Quasi two-dimensional acoustic metamaterials with negative bulk modulus, *New J. Phys.* **14**, p. 103052.
[39] Green, D. W., Winandy, J. E. and Kretschmann, D. E. (1999). Mechanical properties of wood, General technical report FPL GTR-113, USDA Forest Service, Forest Products Laboratory.
[40] Guenneau, S., Movchan, A., Pétursson, G. and Ramakrishna, S. A. (2007). Acoustic metamaterials for sound focusing and confinement, *New Journal of Physics* **9**(11), p. 399.
[41] Guenneau, S., McPhedran, R., Enoch, S., Movchan, A., Farhat, M. and Nicorovici, N.-A. P. (2011). The colors of cloaks, *J. Optics* **13**, p. 024014.
[42] Gutzwiller, M. C. (1990). *Chaos in Classical and Quantum Mechanics* (New York: Springer-Verlag).
[43] Hsu, J.-C. (2013). Effects of elastic anisotropy in phononic band-gap plates with two-dimensional lattices, *Journal of Physics D: Applied Physics* **46**(1), p. 015301.
[44] Joannopoulos, J. D., Johnson, S. G., Winn, J. N. and Meade, R. D. (2011). *Photonic Crystals: Molding the Flow of Light* (Princeton University Press, Princeton, NJ).

[45] John, S. (1987), Strong localization of photons in certain disordered dielectric superlattices, *Phys. Rev. Lett.* **58**(23), pp. 2486–2489.
[46] Kadic, M., Buckmann, T., Schittny, T. and Wegener, M. (2013). Metamaterials beyond electromagnetism, *Rep. Prog. Phys.* **76**, p. 126501.
[47] Kaina, N., Fink, M. and Lerosey, G. (2013a). Composite media mixing Bragg and local resonances for highly attenuating and broad bandgaps, *Sci. Rep.* **3**, p. 3240.
[48] Kaina, N., Lemoult, F., Fink, M. and Lerosey, G. (2013b). Ultrasmall mode volume defect cavities in spatially ordered and disordered metamaterials, *Appl. Phys. Lett.* **102**(14), p. 144104.
[49] Kaina, N., Lemoult, F., Fink, M. and Lerosey, G. (2015). Negative refractive index and acoustic superlens from multiple scattering in single negative metamaterials, *Nature* **525**, pp. 77–81.
[50] Khelif, A., Choujaa, A., Djafari-Rouhani, B., Wilm, M., Ballandras, S. and Laude, V. (2003). Trapping and guiding of acoustic waves by defect modes in a full-band-gap ultrasonic crystal, *Phys. Rev. B* **68**(21), p. 214301.
[51] Khelif, A., Choujaa, A., Benchabane, S., Djafari-Rouhani, B. and Laude, V. (2004). Guiding and bending of acoustic waves in highly confined phononic crystal waveguides, *Appl. Phys. Lett.* **84**(22), pp. 4400–4402.
[52] Khelif, A., Djafari-Rouhani, B., Vasseur, J. O. and Deymier, P. A. (2003). Transmission and dispersion relations of perfect and defect-containing waveguide structures in phononic band gap materials, *Phys. Rev.* **68**, p. 024302.
[53] Khelif, A., Achaoui, Y., Benchabane, S., Laude, V. and Aoubiza, B. (2010). Locally resonant surface acoustic wave band gaps in a two-dimensional phononic crystal of pillars on a surface, *Phys. Rev. B* **81**, p. 214303.
[54] Khelif, A., Achaoui, Y. and Aoubiza, B. (2012). Surface acoustic waves in pillars-based two-dimensional phononic structures with different lattice symmetries, *J. Appl. Phys.* **112**.
[55] Kittel, C. and McEuen, P. (1976). *Introduction to Solid State Physics*, *Vol. 8* (Wiley, New York).
[56] Komatitsch, D. and Vilotte, J.-P. (1998). The spectral element method: An efficient tool to simulate the seismic response of 2d and 3d geological structures, *Bull. Seism. Soc. Am.* **88**, pp. 368–392.
[57] Komatitsch, D. and Martin, R. (2007). An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation, *Geophysics* **72**, pp. SM155–SM167.
[58] Koschny, T., Kafesaki, M., Economou, E. N. and Soukoulis, C. M. (2004). Effective medium theory of left-handed materials, *Phys. Rev. Lett.* **93**, p. 107402.
[59] Kroedel, S., Thome, N. and Daraio, C. (2015). Wide band-gap *seismic* metastructures, *Extreme Mechanics Letters* **4**, pp. 111–117.
[60] Ktorza, L., Ceresoli Enoch, S., Guenneau, S. and Abdeddaim, R. (2015). Single frequency microwave cloaking and subwavelength imaging with curved wired media, *Optics Express* **23**(8), pp. 10319–10326.

[61] Lagendijk, A. (1993). Vibrational Relaxation Studied with Light. In Bron, W. (ed.) *Ultrashort Processes in Condensed Matter*, vol. 314 of NATO ASI Series, pp. 197–236 (Springer US).
[62] Lamb, H. (1904). On Group — Velocity, *Proc. London Math. Soc.* **1**, p. 473.
[63] Lagarrigue, C., Groby, J. P. and Tournat, V. (2013). Sustainable sonic crystal made of resonating bamboo rods, *J. Acoust. Soc. Am.* **133**, p. 247.
[64] Lemoult, F., Lerosey, G., de Rosny, J. and Fink, M. (2010). Resonant metalenses for breaking the diffraction barrier, *Phys. Rev. Lett.* **104**, p. 203901.
[65] Lemoult, F., Fink, M. and Lerosey, G. (2011a). Acoustic resonators for far-field control of sound on a subwavelength scale, *Phys. Rev. Lett.* **107**, p. 064301.
[66] Lemoult, F., Fink, M. and Lerosey, G. (2011b). Revisiting the wire medium: An ideal resonant metalens, *Waves Random Complex Media* **21**, p. 591.
[67] Lemoult, F., Kaina, N., Fink, M. and Lerosey, G. (2013). Wave propagation control at the deep subwavelength scale in metamaterials, *Nature Phys.* **9**, pp. 55–60.
[68] Leonhardt, U. (2006). Optical conformal mapping, *Science* **312**, p. 1777.
[69] Lerosey, G., De Rosny, J., Tourin, A. and Fink, M. (2007). Focusing beyond the diffraction limit with far-field time reversal, *Science* **315**, p. 1120.
[70] Leroy, V., Strybulevych, A., Scanlon, M. and Page, J. H. (2009). Transmission of ultrasound through a single layer of bubbles, *European Physics Journal E* **29**, pp. 123–130.
[71] Li, J. and Chan, C. T. (2004). Double-negative acoustic metamaterial, *Phys. Rev. E* **70**(5), p. 055602.
[72] Liu, Z., Zhang, X., Mao, Y., Zhu, Y. Y., Yang, Z., Chan, C. T. and Sheng, P. (2000b). Locally resonant sonic materials, *Science* **289**, p. 1734.
[73] Martinez-Sala, R., Sancho, J., Sanchez, J. V., Linres, J. and Mesegure, F. (1995). Sound attenuation by sculpture, *Nature* **378**, p. 241.
[74] Maznev, A. A., Wright, O. B. and Matsuda, O. (2011). Mapping the band structure of a surface phononic crystal, *New Journal of Physics* **13**(1), p. 013037.
[75] Meseguer, F., Holgado, M., Caballero, D., Benaches, N., Sanchez-Dehesa, J., Lopez, C. and Llinares, J. (1999). Rayleigh-wave attenuation by a semi-infinite two-dimensional elastic-band-gap crystal, *Phys. Rev. B* **59**, pp. 12169–12172.
[76] Milton, G., Briane, M. and Willis, J. (2006). On cloaking for elasticity and physical equations with a transformation invariant form, *New J. Phys.* **8**(248), p. 014301.
[77] Miroshnichenko, A. E., Flach, S. and 236 Kivshar, Y. S. (2010). Fano resonances in nanoscale structures, *Rev. Mod. Phys.* **82**, pp. 2257–2298.
[78] Norris, A. (2008). Acoustic cloaking theory, *Proc. R. Soc. A* **464**, pp. 2411–2434.
[79] Norris, A. and Shuvalov, A. (2011). Elastic cloaking theory, *Wave Motion* **48**, pp. 525–538.

[80] Notomi, M. (2000). Theory of light propagation in strongly modulated photonic crystals: Refraction like behavior in the vicinity of the photonic band gap, *Phys. Rev. B* **62**(16), pp. 10696–10705.

[81] Oudich, M., Senesi, M., Assouar, M., Ruzenne, M., Sun, J., Vincent, B., Hou, Z. and Wu, T.-T. (2011). Experimental evidence of locally resonant sonic band gap in two-dimensional phononic stubbed plates, *Phys. Rev. B* **84**, p. 165136.

[82] Page, J. H., Sukhovich, A., Yang, S., Cowan, M., Biest, F. V. D., Tourin, A., Fink, M., Liu, Z., Chan, C. and Sheng, P. (2004). Phononic crystals, *Phys. Stat. Sol.* (*b*) **241**, pp. 3454–3462.

[83] Page, J. H., Lee, E. J. S. and Croenne, C. (2013). Anomalous ultrasonic transport in phononic crystals with overlapping Bragg and hybridization gaps, *J. Acoust. Soc. Am.* **134**, p. 4026.

[84] Park, J., Park, B., Kim, D. and Park, J. (2012). Determination of effective mass density and modulus for resonant metamaterials, *J. Acoust. Soc. Am.* **132**(4), pp. 2793–2799.

[85] Parnell, W. and Norris, A. (2012). Hyperelastic cloaking theory: Transformation elasticity with pre-stressed solids, *Proc. R. Soc. A* **468**, pp. 2881–2903.

[86] Pendry, J. B., Holden, A. J., Robbins, D. J. and Stewart, W. J. (1999). Magnetism from conductors and enhanced nonlinear phenomena, *IEEE Trans. Microwave Theory Tech.* **47**, p. 2075.

[87] Pendry, J. B. (2000). Negative refraction makes a perfect lens, *Phys. Rev. Lett.* **85**, pp. 3966–3969.

[88] Pendry, J. B., Schurig, D. and Smith, D. R. (2006). Controlling electromagnetic fields, *Science* **312**, p. 1780.

[89] Peter, D., Komatitsch, D., Luo, Y., Martin, R., Le Goff, N., Casarotti, E., Le Loher, P., Magnoni, F., Liu, Q., Blitz, C., Nissen-Meyer, T., Basini, P. and Tromp, J. (2011). Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes, *Geophys. J. Int.* **186**(2), pp. 721–739.

[90] Pennec, Y., Djafari Rouhani, B., Larabi, H., Akjouj, A., Gillet, J. N., Vasseur, J. O. and Thabet, G. (2009). Phonon transport and waveguiding in a phononic crystal made up of cylindrical dots on a thin homogeneous plate, *Phys. Rev. B* **80**(14), p. 144302.

[91] Pierre, J., Boyko, O., Belliard, L., Vasseur, J. O. and Bonello, B. (2010). Negative refraction of zero order flexural Lamb waves through a two-dimensional phononic crystal, *Appl. Phys. Lett.* **97**(12), p. 121919.

[92] Poulton, C. G., Movchan, A. B., McPhedran, R. C., Nicorovici, N. A., Antipov, Y. A. (2003), Eigenvalue problems for doubly periodic elastic structures and phononic band gaps, *Proceedings of the Royal Society of London A* **456**, pp. 2543–2559.

[93] Psarobas, I. E., Modinos, A., Sainidou, R. and Stefanou, N. (2002). Acoustic properties of colloidal crystals, *Phys. Rev. B* **65**, p. 064307.

[94] Ramakrishna, S. and Grzegorcyk, T. (2008). *Physics and Applications of Negative Refractive Index Materials* (CRC Press, Boca Raton, FL).
[95] Rost, S. (2002). Array seismology: Methods and applications, *Reviews of Geophysics* **40**.
[96] Royer, D. and Dieulesaint, E. (2000). *Elastic Waves in Solids I: Free and Guided Propagation* (Berlin: Springer).
[97] Rupin, M., Lemoult, F., Lerosey, G. and Roux, P. (2014a). Experimental demonstration of ordered and disordered multiresonant metamaterials for lamb waves, *Phys. Rev. Lett.* **112**, p. 234301.
[98] Rupin, M., Roux, P. and Catheline, S. (2014b). One-channel inverse filter: Spatio-temporal control of a complex wave-field from a single point, *Appl. Phys. Lett.* **104**, p. 244101.
[99] Rupin, M., Catheline, S. and Roux, P. (2015a). Super-resolution experiments on lamb waves using a single emitter, *Appl. Phys. Lett.* **106**, p. 24103.
[100] Rupin, M., Roux, P., Lerosey, G. and Lemoult, F. (2015b). Symmetry issues in the hybridization of multi-mode waves with resonators: An example with Lamb waves metamaterial, *Scientific Reports* **5**, p. 13714.
[101] Scherrer, G., Hofman, M., Smigaj, W., Kadic, M., Chang, T.-M., Melique, X., Lippens, D., Vanbesien, O., Cluzel, B., de Fornel, F., Guenneau, S. and Gralak, B. (2013). Photonic crystal carpet: Manipulating wave fronts in the near field at 1.55 lm, *Phys. Rev. B* **88**, p. 115110.
[102] Shen, Y., Ye, D., Celanovic, I., Johnson, S. G., Joannopoulos, J. D. and Soljačić, M. (2014). Optical broadband angular selectivity, *Science* **343**(6178), pp. 1499–1501.
[103] Schoenberg, M. and Sen, P. N. (1983). Properties of a periodically stratified acoustic half-space and its relation to a Biot fluid, *J. Acoust. Soc. Am.* **73**, 61.
[104] Shelby, R. A., Smith, D. R. and Schultz, S. (2001). Experimental verification of a negative index of refraction, *Science* **292**(5514), pp. 77–79.
[105] Shen, J. T., Catrysse, P. B. and Fan, S. (2005). Mechanism for designing metallic metamaterials with a high index of refraction, *Phys. Rev. Lett.* **94**, p. 197401.
[106] Sheng, P. (2014). A step towards a seismic cloak, *Physics* **34**.
[107] Sigalas, M. and Economou, E. N. (1993). Band structure of elastic waves in two dimensional systems, *Solid State Commun.* **86**(3), pp. 141–143.
[108] Sigalas, M. M., Kushwaha, M. S., Economou, E. N., Kafesaki, M., Psarobas, I. E. and Steurer, W. (2005). Classical vibrational modes in phononic lattices: Theory and experiment, *Zeitschrift für Kristallographie-Crystalline Materials* **220**, pp. 765–809.
[109] Silveirinha, M. and Engheta, N. (2006). Tunneling of electromagnetic energy through subwavelength channels and bends using e-near-zero materials, *Phys. Rev. Lett.* **97**.
[110] Smith, D. R., Padilla, W. J., Vier, D. C., Nemat-Nasser, S. C. and Schultz, S. (2000). Composite medium with simultaneously negative permeability and permittivity, *Phys. Rev. Lett.* **84**, pp. 4184–4187.

[111] Smith, D. R., Pendry, J. B. and Wiltshire, M. C. K. (2004). Metamaterials and negative refractive index, *Science* **305**, pp. 788–792.

[112] Smith, D. R., Padilla, W. J., Vier, D. C., Nemat-Nasser, S. C. and Schultz, S. (2000). Composite medium with simultaneously negative permeability and permittivity, *Phys. Rev. Lett.* **84**, pp. 4184–4187.

[113] Stenger, N., Wilhelm, M. and Wegener, M. (2012). Experiments on elastic cloaking in thin plates, *Phys. Rev. Lett.* **108**, p. 014301.

[114] Sukhovich, A., Merheb, B., Muralidharan, K., Vasseur, J., Pennec, Y., Deymier, P. and Page, J. (2009). Experimental and theoretical evidence for subwavelength imaging in phononic crystals, *Phys. Rev. Lett.* **102**, p. 154301.

[115] Sukhovich, A., Jing, L. and Page, J. H. (2008). Negative refraction and focusing of ultrasound in two-dimensional phononic crystals, *Phys. Rev. B* **77**, p. 014301.

[116] Tanter, M., Aubry, J.-F., Gerber, J., Thomas, J.-L. and Fink, M. (2001). Optimal focusing by spatio-temporal inverse filter. I. Basic principles, *J. Acoust. Soc. Am.* **110**(1), pp. 37–47.

[117] Torrent D. and Sanchez-Dehesa, J. (2006). Effective parameters of clusters of cylinders embedded in a nonviscous fluid or gas, *Phys. Rev. B* **74**, p. 224305.

[118] Ungureanu, B., Achaoui, Y., Enoch, S., Brulé, S. and Guenneau, S. (2016). Auxetic-like metamaterials as novel earthquake protections, *EPJ Applied Metamaterials* **2**, p. 17.

[119] Vasseur, J. O., Deymier, P. A., Djafari-Rouhani, B., Pennec, Y. and Hladky-Hennion, A. C. (2008). Absolute forbidden bands and waveguiding in two-dimensional phononic crystal plate, *Phys. Rev. B: Condens. Matter* **77**(8), p. 085415.

[120] Veselago, V. G. (1968). The electrodynamics of substances with simultaneously negative values of ε and μ, *Soviet Physics Uspekhi* **10**, p. 509.

[121] Wegener, M. (2013). Metamaterials beyond optics, *Science* **342**(6161), pp. 939–940.

[122] Williams, E. G., Roux, P., Rupin, M. and Kuperman, W. A. (2015). Theory of multiresonant metamaterials for A0 Lamb waves, *Physical Review B* **91**, pp. 1–12.

[123] Woods, R. D. (1968). Screening of surface wave in soils, *Journal of the Soil Mechanics and Foundations Division* **94**, pp. 951–980.

[124] Wu, Y., Lai, Y. and Zhang, Z.-Q. (2007). Effective medium theory for elastic metamaterials in two dimensions, *Phys. Rev. B* **76**, p. 205313.

[125] Wu, T.-C., Wu, T.-T. and Hsu, J.-C. (2009). Waveguiding and frequency selection of Lamb waves in a plate with periodic stubbed surface, *Phys. Rev. B* **79**(10), p. 104306.

[126] Wu, T. T., Chen, Y. T., Sun, J. H., Lin, S. C. S. and Huang, T. J. (2011). Focusing of the lowest antisymmetric Lamb wave in a gradient-index phononic crystal plate, *App. Phys. Lett.* **98**(17), p. 171911.

[127] Xiao, Y., Wen, J. and Wen, X. (2012). Flexural wave band gaps in locally resonant thin plates with periodically attached spring–mass resonators, *J. Phys. D: Appl. Phys.* **45**(19), p. 195401.

[128] Xu, J., Jiang, X., Fang, N., Georget, E., Abdeddaim, R., Geffrin, J. M., Farhat, M., Sabouroux, P., Enoch S. and Guenneau, S. (2015). Molding acoustic, electromagnetic and water waves with a single cloak, *Scientific Reports* **5**, p. 10678.

[129] Yablonovitch, E. (1987). Inhibited spontaneous emission in solid-state physics and electronics, *Phys. Rev. Lett.* **58**, p. 2059.

[130] Yang, M., Ma, G., Wu, Y., Yang, Z. and Sheng, P. (2014). Homogenization scheme for acoustic metamaterials, *Phys. Rev. B* **89**, p. 064309.

[131] Zhu, J., Christensen, J., Jung, J., Martin-Moreno, L., Yin, X., Fok, L., Zhang, X. and Garcia-Vidal, F. J. (2011). A holey-structured metamaterial for acoustic deep-subwavelength imaging, *Nature Phys.* **7**, p. 52.

CHAPTER 7

Seismic Metamaterials: Controlling Surface Rayleigh Waves Using Analogies with Electromagnetic Metamaterials

STÉPHANE BRÛLÉ*

Dynamic Soil Laboratory, Ménard, France

STEFAN ENOCH

Institut Fresnel, France

SÉBASTIEN GUENNEAU

Institut Fresnel and Aix-Marseille Université, France

RICHARD CRASTER

Imperial College London, London, UK

*Corresponding author: stephane.brule@menard.mail.com

7.1. Introduction

Seismic risk assessment leads civil engineers to define the ground motion for the design of structures, which are self-adapting to this key parameter. We shall try to convince the reader that a promising way to revisit seismic risk assessment would consist of creating an artificially structured soil and thereby counteract the most devastating components of seismic signals.

In the past, a few authors [Woods (1968), Banerjee *et al.* (1988), Semblat *et al.* (2009)] obtained significant results with vibration screening in the soil itself for a local source such as industrial vibratory machines located on concrete slab for example. We will illustrate in this chapter how to extend this wave interaction concept to a broader range of frequencies, including seismic waves generated by earthquakes.

One needs to keep in mind that most of the vibration energy affecting nearby structures is carried by Rayleigh surface waves and the horizontal component of bulk waves traveling from the source of vibration. Effective screening of earth-waves can be achieved by proper interception, scattering, diffraction of surface waves or resonance effects using devices such as trenches in soil, additional inclusions, walls and the like.

7.2. From the Nano-scale to the Metre-scale

The transposition of concepts of photonic and phononic crystals, that originate from the fields of Optics and Acoustics, open new avenues of research for modelling the interaction of seismic signals with buried structures (empty or filled) including special foundations for buildings. Our main objective is to understand what the structured soils introduce as complementary effects on the soil dynamic response for seismic disturbances: Bragg reflexion, band gap, directivity, etc.

Since the end of the last century, researchers have taken advantage of technological improvements in structuring matter to achieve control over the flow of light. Photonic crystals are periodic man-made structures that have been proposed to inhibit emission thanks to photonic band gaps [Yablonovitch (1987)] and localize light [John

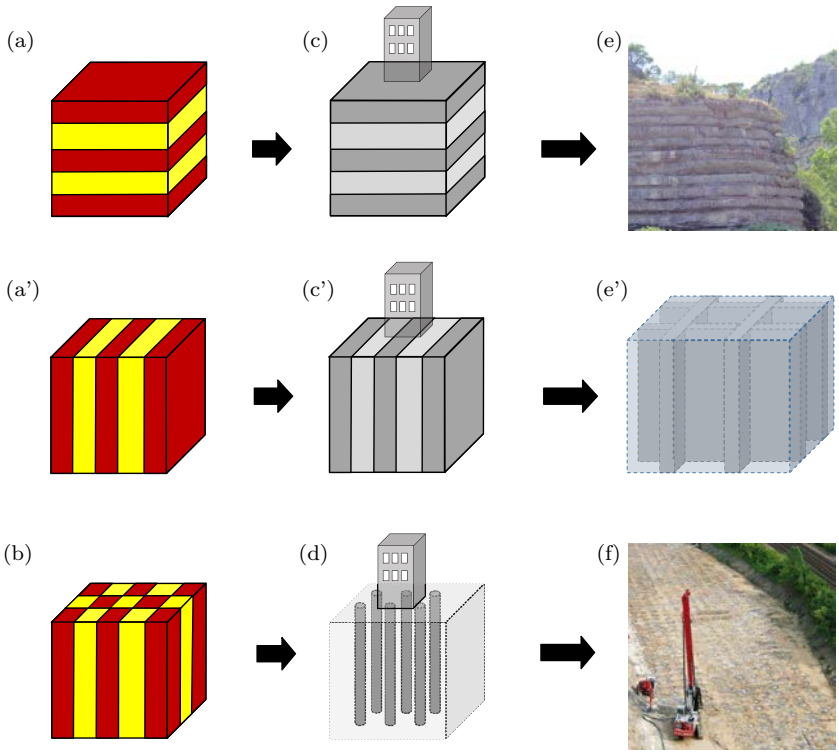


Fig. 7.1. The change of scale concept. Panels (a) and (b) suggest the 1D and 2D geometrical shape of photonic or phononic crystals [Joannopoulos *et al.* (2008)]. Middle panels are a description of the principle of a building with its foundations lying respectively on 1D structured soil with vertical strata (c) or lying on a 2D metamaterial of soil reinforced with vertical cement columns (d). At metric scale, panel (e) shows an outcrop with natural 2D alternation of rocky calcareous and soft clay sedimentary layers (photography courtesy of S. Brûlé). Panel (e) illustrates cells of vertical panels of concrete or grout mixed with the soils. Panel (f) illustrates a 2D full-scale soil reinforcement with vertical and cylindrical lime columns. The grid is around 2×2 m (photography courtesy of Ménard).

(1987)]. Photonic crystals are periodically structured electromagnetic media, generally possessing photonic band gaps, i.e. ranges of frequency in which light cannot propagate through the structure. We show some structured media in Fig. 7.1 with photonic crystals in Fig. 7.1 (a,b), the vision of translating these concepts to ground vibration using structured soils is shown in Fig. 7.1 (c,d). Fig. 7.1

(e,f) show natural occuring formations or artificially structured soils that have similar properties. As an aside Fig. 7.1(e) shows horizontal layering suitable for protection from deep shear or compressional waves whereas the vertical layering of Fig. 7.1(c) is more appropriate for protection from Rayleigh waves.

The ideas of structuring media in photonics were then taken a stage further by introducing metamaterials. In 2001, the word "metamaterial" was coined by R. M. Walser [Walser (2001)], who gave the following definition: macroscopic composite having a manmade, periodic cellular architecture designed to produce an optimized combination, not available in nature, of two or more responses to specific excitation. The concept of metamaterial arose from nano-scale world and electromagnetism. Typically, a metamaterial uses periodic arrangements of elements with size much smaller than the considered wavelength (typically hundreds of nanometers) that acquire effective properties of materials with negative optical index [Pendry (2000)], or highly anisotropic materials such as hyperbolic metamaterials [Iorch *et al.* (2013)] or invisibility cloaking devices [Pendry *et al.* (2006)]. The transition from electromagnetic to acoustic waves is possible thanks to phononic crystals, which are artificial handcrafted structures. They range in scale from a few meters down to hundreds of nanometers or less. At this scale, matter appears as continuous and the laws of classical mechanics can be applied. The search for structures with complete phononic band gaps began in 1992 with work by Sigalas and Economou [Economou and Sigalas (1993)]. They showed that an infinite 2D array of high-density parallel cylinders embedded in a low-density host material should possess a complete band gap in two dimensions (Figure 7.2). Unaware of this work, Kushwaha and co-authors [Kushwaha *et al.* (1993)] reported the existence of photonic band gaps for polarized elastic waves in 2D elastic systems. Interestingly, as unveiled in [O'Brien and Pendry (2002)] high-permittivity dielectric rods display stop bands induced by low frequency localized modes, which are associated with artificial magnetism. In a similar way, high-density rods display unique features upon resonance, such a negative effective density.

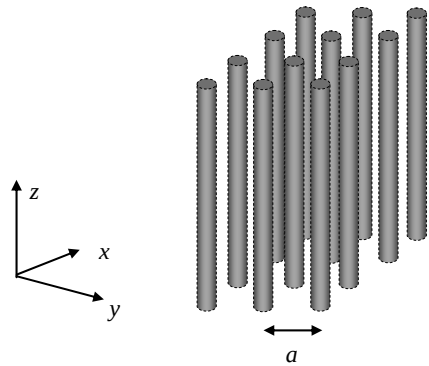


Fig. 7.2. Example of 2D square-lattice (O,x,y) of high-density parallel cylinders embedded in a lower density host material. The centre-to-centre spacing of the cylinders is a. Note that 2D arrays of high-permittivity dielectric rods in electromagnetism lead to low frequency stop bands associated with artificial magnetism [O'Brien and Pendry (2002)], which is an essential ingredient of electromagnetic wave control.

The origin of band gaps is explained by considering the interference of waves multiply scattered within a phononic crystal. A band gap appears when scattered waves interfere destructively depending on the wave frequency and on the phononic crystal geometry. The first application suggested by Kushwaha and colleagues was to acoustic insulation for human hearing, which is sensitive to frequencies between 20 and 20 000 Hz. In 2005, Hu and Chan [Hu and Chan (2005)] proposed a breakwater device for surface ocean waves as a potential application of photonic crystals at the meter-scale domain. In the same spirit, some of us envisioned in 2008 to reroute ocean waves around a region of still water surrounded by concentric arrays of pillars [Farhat *et al.* (2008)]; non-overtopping dykes for ocean waves can be also envisaged with meter scale invisibility carpets for water waves [Dupont *et al.* (2015)]. The theoretical concept of a seismic 2D grid of inclusions in the soil interacting with a part of the earthquake signal was translated to reality in [Brûlé *et al.* (2014)].

"Initially, seismic metamaterials only concern structured soils with cylindrical and vertical elements. We could call them Seismic Earth-Metamaterials and this is the theme of this chapter 7 [Achaoui *et al.* (2016)].

One can note that the definition then progressively extended to buried and tuned resonators [Finocchio *et al.* (2014), Krodel *et al.* (2015), Achaoui *et al.* (2016)] and very recently, to surface resonators [Colombi *et al.* (2015, 2016), Brûlé *et al.* (2017)]."

We can imagine as well a 1D or 2D photonic or phononic crystal as a 1D or 2D structured soil (natural or artificial) supporting the weight of a building on its free surface (Figure 7.1) (note in passing that 3D structured soils can be also envisaged to shield both surface and bulk seismic waves of all polarization [Achaoui *et al.* (2016)]). In 2012, with the aim of demonstrating the feasibility of the concept with field data, two full-scale seismic tests were held by the Ménard company in France. The first experiment used a grid of vertical empty cylindrical holes with a 50 Hertz source [Brûlé *et al.* (2014)] and the second test, presented in [Brûlé *et al.* (2017)], uses seismic conditions closer to those that are generated during an earthquake, thanks to a multi-frequency source generated by the impact of a mass on the soil. Interestingly, Rayleigh wave attenuation was achieved back in 1999 in marble quarry with air holes displaying kHz stop bands [Meseguer *et al.* (1999)].

7.3. A Tempting Analogy

No-one can dispute the validity of the physics of mechanical wave propagation in all type of materials, at all range of frequency, but one can wonder, legitimately, if components of the soils (sand, clay, rocks) could be compared so freely and directly to materials used for laboratory testing.

A first important stage to validate these intuitive analogies between concepts arising in optics and seismology [Brûlé *et al.* (2012)] is to compare the physical and mechanical characteristics of perfectly elastic man-made materials used for laboratory tests or numerical simulation and geophysical materials. Industrial materials (epoxy resin, copper, iron, etc.) have perfectly known properties, which is not at all the case for soils which are "triphasic" material (grains with pores of water and gas). In Table 7.1, we present the density ρ, the longitudinal, C_l, and transverse, C_t, velocities of Duralium (90%

Table 7.1. Density and velocities of Duralium and epoxy resin [Vasseur *et al.* (1998)].

	ρ kg.m^{-3}	C_l m.s^{-1}	C_t m.s^{-1}	ρC_l^2 10^{10} N.m^{-2}	ρC_t^2 10^{10} N.m^{-2}
Duralium	2 799	6 342	3 095	11.26	2.681
Epoxy	1 142	2 569	1 139	0.754	0.148

Table 7.2. Density and velocities of geophysical materials and weathered concrete in soil [Brûlé *et al.* (2015)].

	ρ kg.m^{-3}	v	C_l m.s^{-1}	C_t m.s^{-1}	ρC_l^2 10^{10} N.m^{-2}	ρC_t^2 10^{10} N.m^{-2}
Clay or sand	1 800	0.3	673	360	0.082	0.023
Weathered concrete in soil	2 200	0.2	3 500	2 143	2.695	1.011

of aluminium, 3 to 5% of copper, magnesium and manganese) and Epoxy used for an experiment [Vasseur *et al.* (1998)]. The test was held on cubic samples ($10 \times 10 \times 10$ cm) made of Duralium cylinders, arranged on square and centred rectangular lattices, embedded in an epoxy matrix resin. The source varied in the range of 0 to 500 kH, showing band gap frequencies. At this stage, we only consider elastic properties of soils (Table 7.2).

From Tables 7.1 and 7.2, we deduce the Young's Modulus ratio $\mathrm{E_{duralium}/E_{epoxy}} = 18.1$ for the phononic man-made material test versus the ratio $\mathrm{E_{concrete}/E_{soil}} = 40$ for soils and concrete: It is of fundamental importance that these ratios are of the same order of magnitude. This means that, with geophysical materials, we can obtain comparable contrasts with those metamaterials made in the laboratory and with steel inclusions, rather than concrete, this ratio is much higher.

With regard to the geometry of devices, the laboratory sample is constituted of arrays of 25 parallel cylinders of Duralium. The cylinder diameter, d, is 16 mm and the centre-to-centre spacing, a, of the cylinders is 20 mm. Assuming a square lattice, the filling

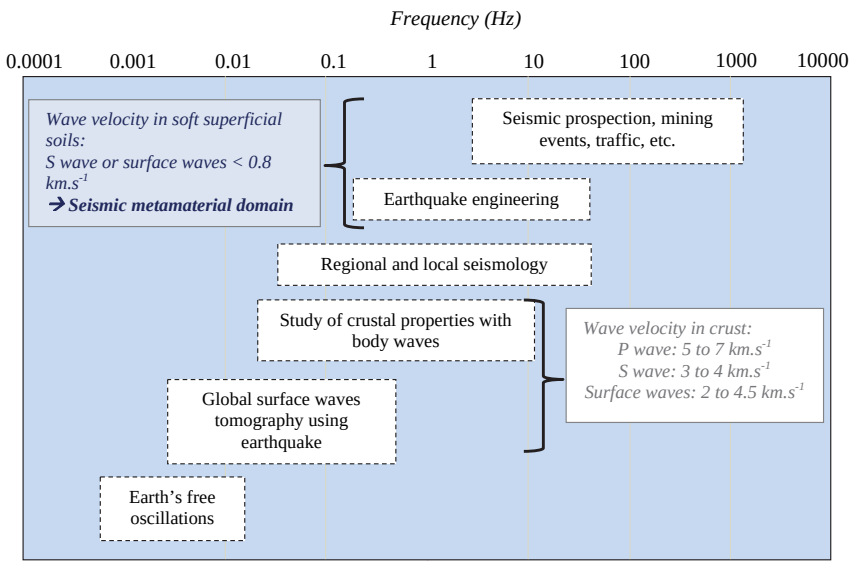
fraction $\pi d^2/4a^2$gives rise to a high value filling ratio of 50.3%. Ground reinforcement works in civil engineering, commonly provided for an improvement of the bearing capacity of the soil or for a reduction of the settlement under the load applied by the building, use concrete, wood or steel vertical inclusions, cylindrical or not. The common length of these inclusions is in the range 0 to 20 m [Chu *et al.* (2009)], but recent equipment can reach 50 m (Menard's world depth record in Louisiana clays, 2014). The diameters are in the range of 0.25 to 0.5 m, exceptionally more (1 m). On work sites, hundreds to thousands of inclusions are used for specific treatment under buildings or different type of storages (oil, natural liquefied gas, etc.). Considering the typical inclusion grid spacing (1.3 to 3 m), the filing ratio rarely exceeds 10%.

Next, we examine the order of magnitude of the signal. The source in the laboratory tests, a transducer, generates frequencies in the range of 0 to 500 kHz. That means the minimal wavelength $\lambda_{\min} = 0.0051$ m for $C_{l\,\text{epoxy}}$ and $\lambda/a > 0.0257$. Up to 128.45 kHz, the grid spacing is sub-wavelength and band gap frequencies were observed in this range of frequency [Vasseur *et al.* (1998)]. Those authors point to a band gap between 55 and 85 Hz, i.e. $1.51 < \lambda/a < 2.34$.

In Seismology, the frequency range of interest is presented in Table 7.3. We usually consider a frequency between 0.1 and 50 Hz and more precisely 0.1 to 12 Hz. For a given spacing of inclusions, range of frequencies and substitution rate (i.e. realistic), reaching $\lambda/a > 2$ needs to consider very low values of shear velocities, V_s, thus requiring soft to very soft soils ($V_s < 150$ m/s).

To conclude, we can definitively enact an analogy between millimetric phononic crystals and meter-length seismic metamaterials. However, at this stage of knowledge, we also have to keep in mind that the density of vertical elements is important for the 2D metamaterial. Different avenues may be explored, as for example the role of the anchoring of inclusion in the bedrock or a smart design of the artificial anisotropy, by varying density of element per square meter, arrangement and type of inclusions [Guenneau *et al.* (2015), Achaoui *et al.* (2016, 2017)].

Table 7.3. Frequency range of interest for seismic metamaterials: It is noted that earthquake engineering is mainly concerned with frequencies from 0.1 to 50 Hz.



7.4. Overview of Photonic, Phononic and Platonic Crystals and Links to Seismic Metamaterials

In this short section, we discuss the advent of platonic crystals and their advantages, in particular as representative soil models for parametric studies of seismic wave propagation.

In 1987, the groups of E. Yablonovitch and S. John reported the discovery of stop band structures for light [Yablonovitch (1987)] and [John (1987)]. Photonic crystals (PCs) have, since then, found numerous applications ranging from nearly perfect mirrors for incident waves whose frequencies are in stop bands of the PCs, to high-q cavities for PCs with structural defects [Srinivasan and Painter (2002)]. The occurrence of stop bands in PCs also leads to anomalous dispersion whereby dispersion curves can have a negative or vanishing group velocity. Dynamic artificial anisotropy, also known as all-angle-negative-refraction [Zengerle (1987), Notomi (2000), Gralak

et al. (2000), Luo *et al.* (2002)], allows for focusing effects through a PC, as envisioned in the 1968 paper of V. Veselago [Veselago (1968)]. With the advent of electromagnetic metamaterials [Pendry *et al.* (1999)] and [Smith *et al.* (2000)], J. Pendry pointed out that the image through the V. Veselago lens can be deeply subwavelength [Pendry (2000)], and exciting effects such as simultaneously negative phase and group velocity of light [Dolling *et al.* (2006)], invisibility cloaks [Schurig *et al.* (2006)] and tailored radiation phase pattern in epsilon near zero metamaterials were demonstrated [Alù *et al.* (2007)] and [Enoch *et al.* (2002)]. One of the attractions of platonic crystals, which are the elastic plate analogue of photonic and phononic crystals, is that much of their physics can be translated into platonics. There are mathematical subtleties in the analysis, and numerics, of the scattering of flexural waves [Antonakakis and Craster (2012)] owing to the fourth-order derivatives in the plate equations, versus the usual second-order derivatives for the wave equation of optics, involved in the governing equations; even waves within a perfect plate have differences from those of the wave equation as they are not dispersionless. Nonetheless, drawing parallels between platonics and photonics helps to achieve similar effects to those observed in electromagnetic metamaterials, such as the time dependent subwavelength resolution through a platonic flat lens [Dubois *et al.* (2013)].

In parallel, research papers in phononic crystals provided numerical and experimental evidence of filtering [Martinez-Sala *et al.* (1995)] and focusing properties [Sukhovich *et al.* (2009)] of acoustic waves. Localized resonant structures for elastic waves propagating within three-dimensional cubic arrays of thin coated spheres [Liu *et al.* (2000)] and fluid filled Helmholtz resonators [Fang *et al.* (2006)] paved the way towards acoustic analogues of electromagnetic metamaterials [Christensen and García De Abajo (2012), Craster and Guenneau (2013)], including elastic cloaks [Milton *et al.* (2006), Brun *et al.* (2009)] and [Norris and Shuvalov (2011)]. The control of elastic wave trajectories in thin plates was reported numerically [Farhat *et al.* (2012)] and experimentally in 2012 [Brûlé *et al.* (2014)] for surface seismic waves [Brûlé *et al.* (2012)] in civil engineering

applications. In fact, Rayleigh waves are generated by anthropic sources such as an explosion or a tool impact or vibration (sledge-hammer, pile driving operations, vibrating machine footing, dynamic compaction, etc.). In 1968, R.D. Woods created *in situ* tests with a 200 to 350 Hz source to show the effectiveness of isolating circular or linear empty trenches [Woods (1968)], with the same geometry, these results were compared in 1988 with numerical modeling studies provided by P.K. Banerjee [Banerjee *et al.* (1988)]. The main thrust of this Chapter is to point out the possibility to create seismic metamaterials not only for high frequency anthropic sources but for the earthquakes' frequency range i.e. 0.1 to 12 Hz. First, we have to address a fundamental question: What is the range of wavelengths taken into account for seismic building design?

We answer this fundamental question in the next section and illustrate the order of magnitude of the seismic incoming signal, sketching out what the differences are, with respect to idealised numerical simulations of wave propagation or experiments in laboratory on small samples made of perfect elastic materials.

7.5. Principles of Seismology

We now identify the components of the seismic signal with which the seismic metamaterial may interact.

Seismology is the scientific study of mechanical vibrations of the Earth. Quantitative seismology is based on seismograms, which are recordings of vibrations, that in turn can be caused artificially by man-made explosions, or caused naturally by earthquakes and volcanic eruptions [Aki and Richards (2002)]. Everyday there are dozens earthquakes that are strong enough to be felt locally, and every few days an earthquake occurs that is capable of damaging structures. We place our analysis strictly in areas concerned with far field earthquake motion to satisfy the linear elastic hypothesis for soils; near the epicenter strong nonlinearities are observed.

With the aim of distinguishing wave types propagating from the epicenter of an earthquake to the surface, we illustrate wave propagation using ray theory (rays are defined as the normal to

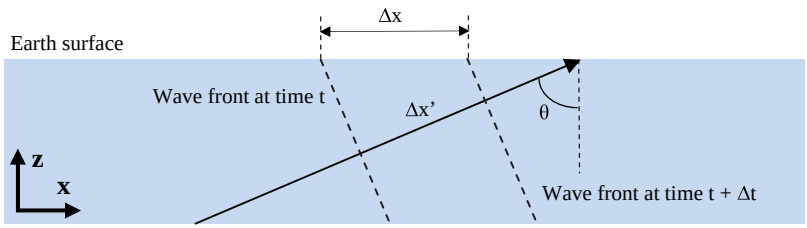


Fig. 7.3. A plane wave incident from below on a horizontal surface. The ray angle from the vertical is termed the incidence angle θ.

the wavefronts) based on optics, with Huygens' and Fermat's principles (Figure 7.3). Due to the remoteness of the epicenter, we can approximate the spherical wave-front as a plane wave-front. In smoothly varying layered media, ray theory can provide useful approximate solution. The method is applicable to media in which the characteristic dimensions of inhomogeneities are considerably larger than the prevailing wavelength of considered waves.

The estimation of seismic site effects is a key concept in seismic hazard assessment; the amplification of seismic waves at the free surface, namely "site effects" may strengthen the impact of an earthquake in specific areas (e.g. Mexico 1985). Indeed, when seismic waves propagate through sediment layers (Figure 7.4) or scatter off strong topographic irregularities, refraction/scattering phenomena can strongly increase the amplitude and the duration of the ground motion. It is then possible to observe stronger motions far away from the epicenter. On the scale of a sediment-filled basin, that often contain buildings, seismic effects involve various phenomena as wave trapping (Figure 7.5), resonance of the whole basin, propagation in heterogeneous media, surface waves bouncing from the edges of the basin and creating interference patterns [Hobiger (2011)].

Structural damage due to seismic excitation is often directly correlated to local site condition in the form of motion amplification and/or soil liquefaction inducing ground deformation. Seismologists distinguish the so-called "one" to "two or three-dimensional effects", which depend mainly on the valley geometry and soil and rock properties [Hobiger (2011)]. The most important geometry parameter is the

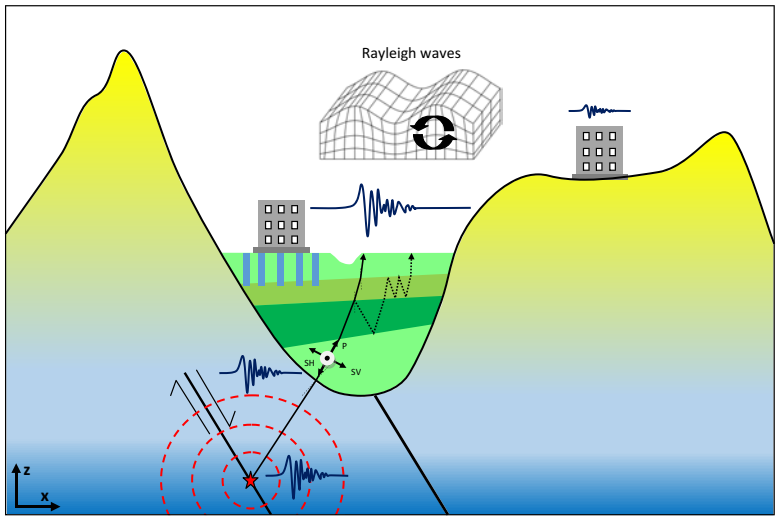


Fig. 7.4. Seismic site effect characterized by a local amplification of the signal in the sedimentary basin in comparison with signal in surrounding rocky mountains.

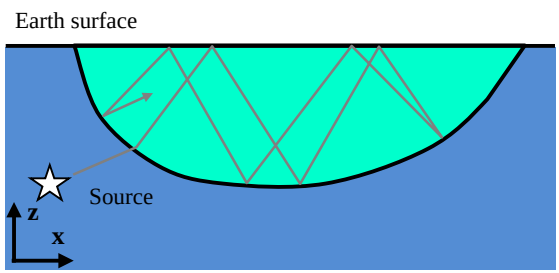


Fig. 7.5. Wave trapping in U-shaped basin (ray tracing).

ratio between the horizontal dimensions of the valley and the sediment thickness. If this ratio is large, "one-dimensional" site effects predominate and these consist in standing waves in vertical direction [Semblat *et al.* (2009)]. "Two-dimensional" or "three dimensional" site effects occur in presence of more complex structures, in other words waves arrive with various incidence due to the structure itself.

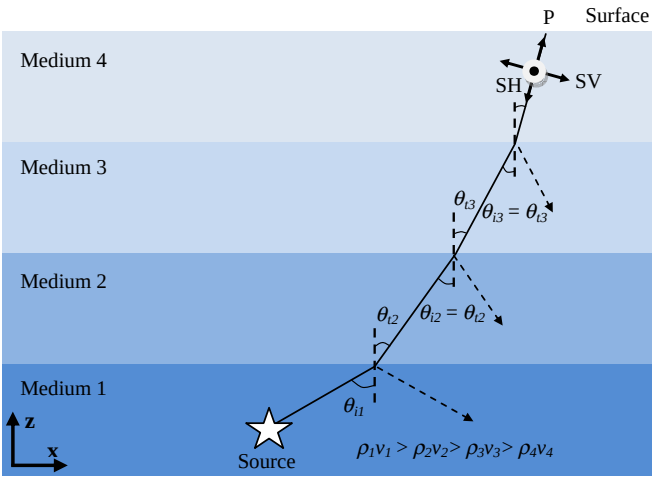


Fig. 7.6. Schematic representation of verticalization and polarization of P, SH and SV waves propagating in more and more low impedance $\rho_i v_i$ layer up to the surface. θ_i and θ_t are respectively the incident and refracted angle of the seismic ray.

By making the simplistic hypothesis that the product of the density by the velocity of each sediment layer decreases from the seismic substratum up to the surface (Figure 7.6), the seismic ray becomes more and more vertical by analogy with a decreasing refraction index in optics. This configuration leads to a specific polarization of body waves: pressure waves (P) oscillate in an almost vertical plane and shear waves (S) are polarized in a sub-horizontal plane.

It is important to note that buildings are very sensitive to the horizontal component of ground motion, less so to the vertical one. Earthquake engineering is also concerned with surface waves. Although other surface waves such as Stoneley waves might play some role in buildings' damage, we focus our analysis on Rayleigh waves (Figure 7.4) mainly generated by the conditions of the full-scale experiment dropping a heavy weight. On the Earth's surface, the coupled propagation of P and S_v waves (shear wave vertical component) is possible, generating Rayleigh waves.

It is fundamental to distinguish long period surface waves (low frequency, i.e. <1 Hz) travelling along the Earth's surface, on the crust, and short ones (<10 Hz) mainly generated by the site effects

described above or by human activities at the Earth surface. An important fact is the low value of surface wave velocity V_s, generated by natural seismic source or construction work activities, in superficial and under-consolidated recent subgrade: less than 100 m/s to 300 m/s. In these soil layers, considering the 0.1 to 10 Hz frequency range, wavelengths of induced surface waves are shorter than direct P and S waves: from few meters to hundreds of meters; this order of wavelength is similar to those of buildings. This is why we can expect a building's resonance to be triggered with some soils in the case of earthquakes, which also makes it possible to conceive of seismic metamaterials whose size could be similar to those of the building project [Brûlé *et al.* (2014)].

At this stage of the chapter, we have defined the conditions of validity for structured soils made of vertical elements (holes, concrete or steel columns). The metamaterial can interact with the horizontal component of the seismic signal in the plane (x, y) and site effect may strengthen the horizontal component of the seismic motion.

In the case of seismic site effects, the wavelengths are of the order of a few meters to hundreds of meters.

7.6. Characterization of the Seismic Signal

A way of expressing the "size" of an earthquake is the "magnitude" of an earthquake which can be determined from the logarithm of the amplitude of waves recorded by seismograph stations.

Earthquakes with magnitude of about 2.0 or less are usually called "microearthquakes"; they are not commonly felt by people and are generally recorded only on local seismographs. Events with magnitudes of about 4.5 or greater — there are several thousand such shocks annually — are strong enough to be recorded by sensitive seismographs all over the world.

The amplitudes of seismic motion can be expressed in terms of acceleration, velocity and displacement. The first data reported from an earthquake record is generally the peak ground acceleration (PGA) which expresses the tip of the maximum spike of the acceleration ground motion (Figure 7.7). Although useful to express the relative intensity of the ground motion (i.e.,

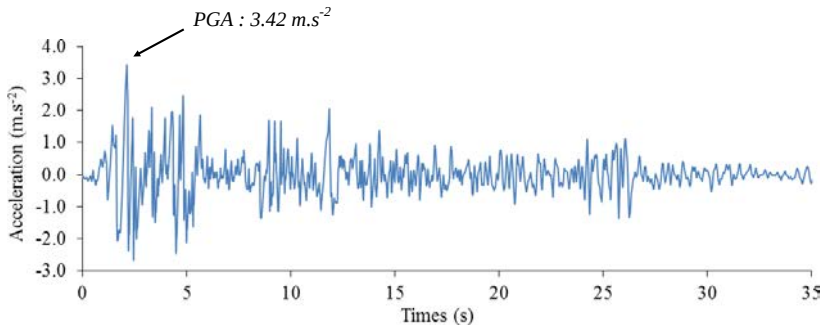


Fig. 7.7. Imperial Valley Earthquake. Peak ground acceleration (3.42 $m.s^{-2}$), N-S component, recorded at El Centro Earthquake 1940-05-19 04:36:41 UTC. Station: El Centro, California — Array Station 9; Imperial Valley Irrigation District — 302 Commercial. Station Owner: USGS. Time History Data File Reference: Pacific Earthquake Engineering Research Center (http://peer.berkeley.edu/research/motions/). Original data source: CIT-SMARTS (http://www.eerl.caltech.edu/smarts/smarts.html).

small, moderate or large), the PGA does not give any information regarding the frequency (or period) content that influences the amplification of building motion due to the cyclic ground motion. In other words, tall buildings with long fundamental periods of vibration will respond differently than short buildings with short periods of vibration. By contrast, response spectra provide these characteristics.

A seismic metamaterial will have to interact with significant earthquakes, but are we able to propose a clear definition of a strong-motion disturbance?

Estimating strong-motion characteristics for seismic risk assessment and earthquake-resistant design requires the clear definition of parameters that reflect the destructive potential of the motion [Bommer and Martinez-Pereira (2000)]. Strong earthquake ground motion is a complex natural phenomenon associated with abrupt energy release during fault rupture.

The intensity of the seismic event can be described in terms of the perceived effects of ground motion according to different intensity scales. The availability of strong ground motion records allows the

definition of quantitative indices based on the amplitude, duration and frequency content of the earthquake records.

For design, the duration of the design basis earthquake generally is not an issue since for response spectrum, equivalent static, and even time history methods of analyses, only the peaks of the response are used in developing design seismic forces. However, the duration of strong earthquake motion is a significant parameter for earthquake damage potential when considering low cycle fatigue, soil liquefaction, soil settlement, and inelastic structural response.

Many definitions of strong-motion duration have been presented in the literature, all of which attempt to isolate a certain portion of the accelerogram during which the strongest motion occurs. We select the significant duration, in seconds, of the seism T_d, which is determined from the Husid plot [Husid (1969)], usually based on the interval during which a certain portion of the total Arias intensity [Arias (1970)] (between the 5% and the 95% of a defined threshold [Trifunac and Brady (1975)]). T_d could vary from 5 to 20 s for earthquakes in the magnitude of 5 to 7.7 [Dowrick (2009)]. The duration of the recorded earthquake and the duration of the faulting could be different due to the complexity of wave propagation. The example of the Imperial Valley Earthquake is presented in Figure 7.8, the significant duration T_d is 24.42 s, it had a moment magnitude M_w of 6.9 and a maximum perceived intensity of X (*Extreme*) on the Mercalli intensity scale.

The Arias Intensity I_a [Arias (1970)] is defined in a conventional way as follows:

$$I_a = \frac{\pi}{2g} \int_0^T a^2(\tau) d\tau \tag{7.1}$$

Husid plotted normalized Arias Intensity, which he denoted as $h(t)$ versus time:

$$h(t) = \frac{\int_0^t a^2(\tau) d\tau}{\int_0^T a^2(\tau) d\tau} = \frac{\pi}{2gI_a} \int_0^t a^2(\tau) d\tau \tag{7.2}$$

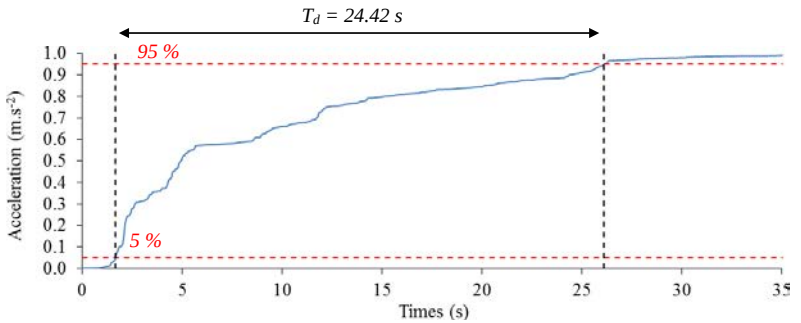


Fig. 7.8. Normalized Husid's function applied to Imperial Valley Earthquake recorded at Al Centro, station 9 (same references as Figure 7.7). T_d is the significant duration of the seism calculated between 5 and 95% of the normalized Husid's function.

Where a is the acceleration-time history in units of g, and g is the acceleration gravity. T is the total duration of the acceleration time-history.

The effect of an earthquake on the Earth's surface is called the intensity. The intensity scale consists of a series of certain key responses such as people awakening, movement of furniture, damage to chimneys, and finally — total destruction. Although numerous intensity scales have been developed over the last several hundred years to evaluate the effects of earthquakes, the one currently used is the Modified Mercalli (MM) Intensity Scale. It was developed in 1931 by the American seismologists Harry Wood and Frank Neumann. This scale, composed of increasing levels of intensity that range from imperceptible shaking to catastrophic destruction, is designated by Roman numerals (I to XII). It does not have a mathematical basis; instead it is an arbitrary ranking based on observed effects. The Modified Mercalli Intensity value assigned to a specific site after an earthquake has a more meaningful measure of severity to the non-scientist than the magnitude because intensity refers to the effects actually experienced at that place. The intensity is classified "strong" from the value V.

It can be clearly seen in Figure 7.9 that for the motion to be potentially damaging to engineered structures (Modified Mercalli

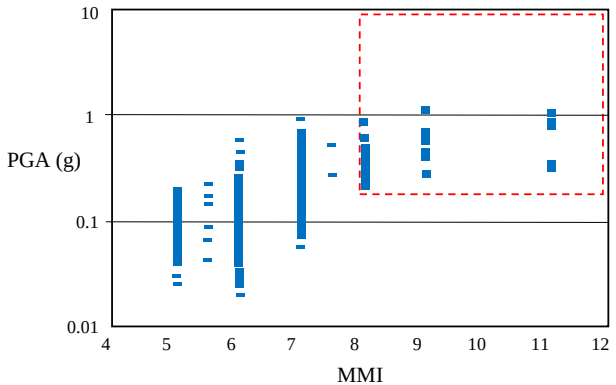


Fig. 7.9. Peak ground acceleration (PGA) versus Modified Mercalli Intensity (MMI). Original diagram from [Bommer and Martinez-Pereira (2000)]. Red dotted line charts the contour of a domain for which the MMI is greater than VIII and PGA is superior to 0.2 g.

Intensity VIII) the value of PGA must be at least 0.2 g (2 m.s^{-2}). However, we can also observe that there are clearly many accelerograms with PGA greater than 0.2 g but which are clearly not damaging, being associated with intensity of shaking as low as VI or even V on the Modified Mercalli scale. The ability of any single parameter to measure the damage potential of the ground motion will depend to a significant degree on which features of the motion (amplitude, frequency content, duration and energy) it reflects.

For more details on this methods we refer the readers towards books specialized in quantitative seismology, for instance, [Aki and Richards (2002)].

With the aim of comparing the order of length of a strong motion earthquake and an anthropic source dedicated to the simulation of ground motion during full-sale experiment, for instance, in Saint-Priest in 2012 using dynamic compaction, we present the comparison of different parameters in Table 7.4.

To conclude this paragraph, we point out that the first experiments on-site had shown the need to use a source bringing enough energy and a low-frequency signature (mean frequency around 8 Hz). Free-fall mass of few tens of tons seems to be a good source for these

Table 7.4. Comparison a strong-motion earthquake [Stover *et al.* (1981)] and dynamic compaction test by means of free-fall punch ponder.

	Imperial Valley Earthquake, CA, USA Station 9 (1940/05/19)	Dynamic compaction test Saint-Priest (2012/09/26)
Duration T_d	24.42 s	0.104 s Nota: duration in agreement with a M_w magnitude 2–3
"Epicentral distance"	10.5 km	5 m
Peak Ground Acceleration PGA	3.42 $m.s^{-2}$	8.7 $m.s^{-2}$
Energy calculated from magnitude	$3.5.10^{14}$ J $M_w = 6.9$	$1.03.10^{7}$ J $M_w \sim 2$

experiments in terms of PGA, but the characteristic duration of the signal is very short (a few tenths of a second) compared to those of earthquakes (a few seconds). There is a need to explore a new type of source, for instance, the seismic noise recording [Brûlé and Javelaud (2014b)]; it is a large-band signal (1 to more than 50 Hz), with long duration, free, it is continuously emitted, but energy-poor.

7.7. Different Ways of Representing the Seismic Wave Propagation Before Writing Equations

During the 1980s, earthquake engineering tried to describe the significant ground motions observed on Earth's surface thanks to the amplification of incident signals inside soft alluvium basins. To illustrate this concept, Figure 7.10(a), shows how the phenomenon is modelled in 1D (hypothesis of infinite horizontal layers), considering shear waves with pure vertical incidence and incident and reflected waves at free surface, propagating in phase. Most of the time, the media 1, of thickness H is constituted of a succession of layers but it is arbitrary simplified with equivalent parameters (ρ_1, μ_1). A major

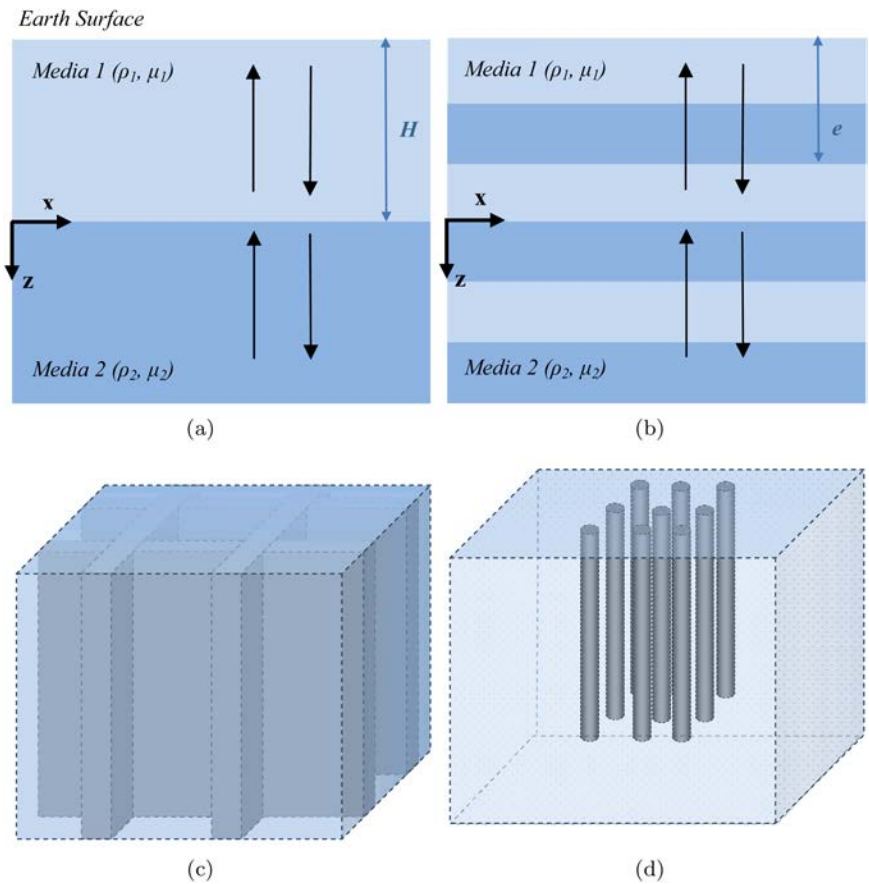


Fig. 7.10. 1D soil-model with incident and reflected S-waves in phase in the media 1 (a). Dispersive multilayered 1D media with a geometric pattern made off two layers. e is the thickness of the elementary elements (b). 2D structured soils with vertical inclusions (d) and cells of vertical panels of concrete or grout mixed with the soils (c).

advance brought by metamaterials [Brûlé *et al.* (2016)] is to imagine that the successive layers could be considered as a potentially dispersive 1D media (Figure 7.10(b)) or to consider structured soils with vertical inclusions as a 2D media (Figure 7.10(d)). Typical methodologies of civil engineering are to design foundations only in the sense that they have to resist pseudo-static loads, leading

sometimes to thick structures (Figure 7.10(c)), that ignore the potential benefit of resonating elements as concrete columns buried in the soil [Achaoui *et al.*, (2016) and Achaoui *et al.* (2017)].

In earthquake engineering terminology, this seismic materials may act on the kinematic and inertial effects, i.e. the soil-structure interaction [Brûlé and Cuira (2017)].

7.8. Wave Propagation Theory in Artificially Anisotropic Media

We have reviewed the fundamentals of seismic signal interaction with layered soils. Let us now proceed more generally and consider a linear partial differential operator L, with a spatially varying, time dependent and isotropic, coefficient m, acting on a scalar or vector field u, and a source (or forcing) term S such that:

$$\boldsymbol{L}[\boldsymbol{m}(\boldsymbol{x, y, z, t}); \boldsymbol{u}] = \boldsymbol{S}(\boldsymbol{x, y, z, t}) \tag{7.3}$$

If we make a change of variables in this governing equation (L could be for instance the Maxwell curl operator in which case m stands for the inverse of permittivity and u for the magnetic vector field or in elastodynamics L would be the Navier operator and m the rank-4 elasticity tensor): $(x, y, z, t) \rightarrow (x', y', z', t)$ with x', y' and z' functions of variables x, y, z, the above equation takes, in most cases, the following form:

$$\boldsymbol{L}[\boldsymbol{M}(\boldsymbol{x}', \boldsymbol{y}', \boldsymbol{z}', \boldsymbol{t}); \boldsymbol{u}] = \boldsymbol{S}(\boldsymbol{x}', \boldsymbol{y}', \boldsymbol{z}', \boldsymbol{t}) \tag{7.4}$$

where M has become anisotropic with coefficients that can be deduced from a product of Jacobian matrices (more precisely, in electromagnetics, M is the product the product of the Jacobian of the transform by its transpose divided by the determinent of the Jacobian [Nicolet *et al.* (1994)], whereas in elastodynamics the transformed elasticity tensor has a more complex expression [9,43,45]. The transformation of coordinates is at the core of transformation optics, which has been a useful and insightful tool in optics.

As a first illustration of the usefulness of transformation optics, let us follow Pendry's proposal for the design of an invisibility cloak, which is based upon the coordinate change $(r, \theta, z) \to (r', \theta', z)$, with $r' = R_1 +$ r $(R_2 - R_1)/R_2$ and $\theta' = \theta$. This maps a disc of radius R_2 to a corona of radii R_1 and R_2 $(R_1 < R_2)$.

For a source S which oscillates periodically in time, we first consider the time harmonic Maxwell operator $L_M = \text{curl}(m \text{ curl } u) + \omega^2/c^2 u$ where ω is the angular wave frequency and c the wavespeed of light in vacuum (approximately 3.00×10^8 meters per second), and the transformed operator looks like

$$\text{curl}(M \,\text{curl}\, u) + M^{-1} \frac{\omega^2}{c^2} u \tag{7.5}$$

where the matrix M (in fact the representation of a rank-2 tensor) in the first term (resp. second term), is a representation of the inverse of permittivity (resp. permeability) tensor, for a longitudinal magnetic field unknown $u = (0, 0, H)$. The diagonalised expression of M in a cylindrical basis is

$$M_{r''r'} = \frac{r' - R_1}{r'}, \quad M_{\theta'\theta'} = \frac{r'}{r' - R_1}, \quad M_{zz} = \left(\frac{R_2}{R_2 - R_1}\right)^2 \frac{r' - R_1}{r'} \tag{7.6}$$

One should note that on the inner boundary of the corona at $r' = R_1$, the first and third eigenvalues $M_{r'r'}$ and M_{zz} vanish, whereas the second eigenvalue $\mathrm{M}_{\theta'\theta'}$ becomes infinite (we shall see in the sequel that in the elastodynamic case the transformed elasticity tensor presents similar features). Physically, this means that one needs infinite anisotropy on the cloak's inner boundary to avoid any phase delay for the electromagnetic signal in forward scattering (behind the cloak). This makes ideal invisibility cloaks for light unfeasible as the wavespeed of light c in vacuum is a strong, unbreakable, upper limit. One way to get around this limit is to consider a cloak surrounded by a medium with refractive index $n > 1$, within which the wavespeed of light is $v = c/n$ (e.g. for visible light, the refractive index of glass is $n = 1.5$). Another challenge is the fact that the

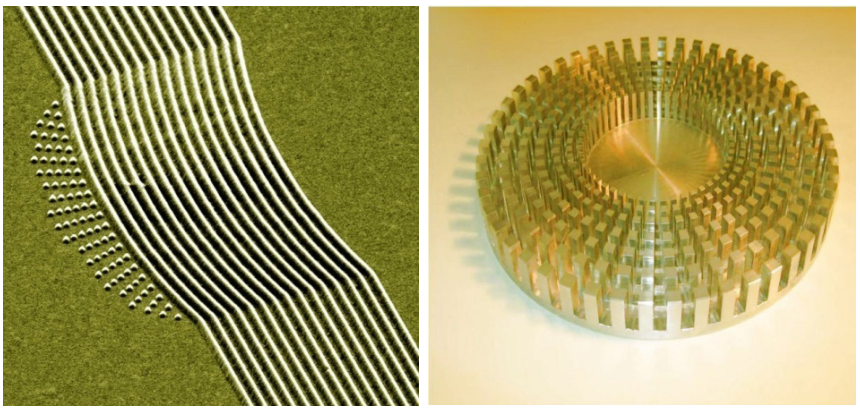


Fig. 7.11. Invisibility carpet for surface plasmon polaritons (left panel) and invisibility cloak for linear surface water waves (right panel). The carpet was designed at Institut Fresnel (Aix-Marseille Université/CNRS/Ecole Centrale Marseille), fabricated and characterized at ICFO (Romain Quidant's group) from 700 nm to 1000 nm; the cloak was designed and characterized at Institut Fresnel for linear surface water waves from 8 to 15 Hertz. The cloak was also characterized for microwaves from 3 to 7 GHz (Institut Fresnel) and for pressure waves from 4 to 8 KHz (MIT, Nicholas' Fang group). Photos courtesy of ICFO (Romain Quidant) and Photothèque CNRS (Stefan Enoch).

transformed Maxwell operator has not only some anisotropic permittivity, but it also has some anisotropic permeability hence artificial magnetism is required. We show in Figure 7.11 realizations of a carpet cloak (left panel) and a cloak (right panel) working for respectively surface plasmon in optics (from 700 to 900 nm) and microwaves (from 3 to 7 GHz). Note that the latter also works for linear water waves (from 8 to 15 Hz) and sound waves (from 4 to 8 KHz).

We show in Figure 7.12 a numerical simulation for a time-harmonic source generating concentric waves a couple of wavelengths away from the ideal cylindrical cloak with parameters as in Eq. 7.5 (left panel) and the structured cloak tested in practice (right panel). The diffracting F obstacle in the center of the cloak has Neumann boundary conditions that are a good model for an infinite conducting obstacle for microwaves (in transverse electric polarization), and a rigid obstacle for linear water waves and sound waves. One notes that

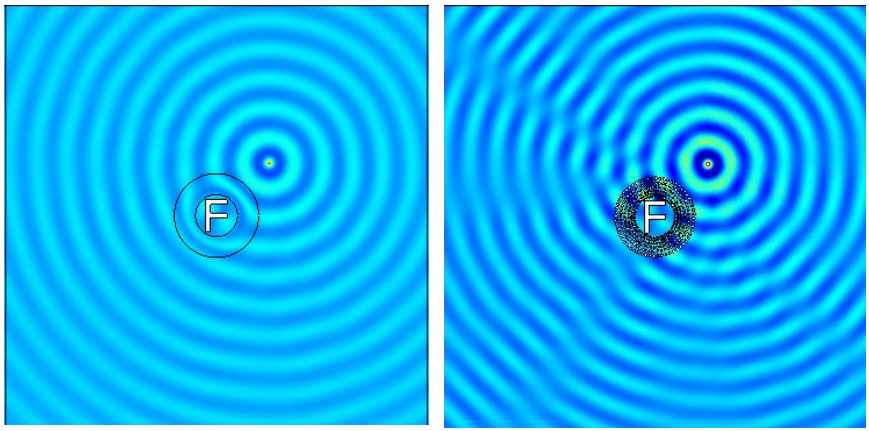


Fig. 7.12. Numerical simulations for linear surface water waves generated by an acoustic source in presence of a rigid F obstacle surrounded by an ideal invisibility cloak (left panel) and the approximated cloak (right panel). A slight phase delay is observed in wavefronts behind the cloak in right panel.

the structured medium in the right panel is described by inclusions with Neumann boundary conditions with the same geometry as for right panel in Figure 7.11.

This simulation holds for microwaves, linear water waves and pressure waves. There are almost no disturbances observed in backward scattering in Figure 7.12 (right panel) but the slight deformation of the wavefront in forward scattering can be attributed to the phase delay induced by the longer wave trajectory followed by the wave compared to a wave in free space. It turns out that similar cloaking can be observed for flexural waves in a thin-structured plate with perforations of same geometry as in Figure 7.12 (right panel), see [Farhat *et al.* (2012)] for the numerical analysis and [Stenger *et al.* (2012)] for experimental validation by the group of Martin Wegener at Karlsruhe Institute for Technology. Similar results also hold for Rayleigh waves in soils structured with the same kind of perforations (boreholes with a cross-section as in Figure 7.12 right panel).

Let us now focus our attention on mechanical waves, which are governed by the Navier equations of elastodynamics (in the framework of small deformations i.e. linear elasticity). We consider an

isotropic medium described by an elasticity tensor C (rank 4), which is fully symmetric, and a density, for a time harmonic elastodynamic wave, the displacement field u is solution of

$$\nabla.(C : \nabla u) + \rho\omega^2 u = S \tag{7.7}$$

Note that here ∇u is a rank-2 tensor, so that C: ∇u is a rank-2 tensor. Let us once again apply Pendry's transformation, then the transformed Navier operator $L_n = \nabla.(C' : \nabla u) + \rho\omega^2 u$ has a rank-4 tensor C' expressed in cylindrical coordinates as [Brun *et al.* (2009)]

$$\begin{gathered} C'_{r'r'r'r'} = \frac{r' - R_1}{r'}(\lambda + 2\mu), \quad C'_{\theta'\theta'\theta'\theta'} = \frac{r'}{r' - R_1}(\lambda + 2\mu), \\ C'_{r'r'\theta'\theta'} = C'_{\theta'\theta'r'r'} = \lambda, \quad C'_{r'\theta'r'\theta'} = \frac{r' - R_1}{r'}\mu, \\ C'_{\theta'r'\theta'r'} = \frac{r'}{r' - R_1}\mu, \quad C'_{r'\theta'\theta'r'} = C'_{\theta'r'r'\theta'} = \mu \end{gathered} \tag{7.8}$$

One notes that C' does not have the minor symmetries, as $C'_{r'\theta'r'\theta'}$ and $C'_{r'\theta'\theta'r'}$, which are equal to the shear modulus μ in the isotropic homogeneous medium in original space, have after transform a factor $(r' - R_1)/r'$ and its inverse, that leads to antagonistic behaviour when r' goes to R_1 ($C'_{r'\theta'r'\theta'}$ goes to zero while $C'_{r'\theta'\theta'r}$ goes to infinity). More precisely, at the inner boundary of the cloak there is once again a divergence of the eigenvalues of the tensor, just as in the electromagnetic case, however here it seems that the symmetry breaking is only concerned with acceleration of the shear polarization of the coupled shear and pressure in-plane waves. Nevertheless, there is no longer the problem of the constant wavespeed of light in vacuum, here it is fairly easy to accelerate the tangential speed of a mechanical wave inside the cloak with respect to the wavespeed in the bulk medium surrounding the cloak. One can thus avoid phase delay (deformed wavefront) behind the cloak. However, the challenge is to achieve metamaterials described by an asymmetric elasticity tensor, this requires some kind of elastic chirality that goes beyond the scope of the chapter.

7.9. Full-Scale Experiment

To build a theory of seismic cloaks based on transformational physics, these theoretical concepts need to be checked against experimental data. A test zone consisting of a regular mesh of vertical cylindrical voids was carried out near the alpine city of Grenoble (France) in August 2012 (Figure 7.15). The preliminary objective of this seismic field test was to point out the analogies with the control of electromagnetic and acoustic waves by a quantitative approach. In theory, it seems realistic to influence seismic waves passing through an artificial anisotropic medium. However, soils possess particular characteristics: non-elastic behavior, high rate of signal attenuation, large-scale heterogeneity, etc. These various uncertainties and the objective of realistic values for modeling require *in situ* tests to adjust soil's parameters as shear modulus, quality factor, etc. The measurement of the velocity of Rayleigh waves V_R is given by a preliminary seismic test, using the wave time arrival at various offset from the source. The Ménard civil engineering team of one of us, Stéphane Brûlé, obtained $V_R = 78$ m/s. The tested soil is a homogeneous silty clay. The thickness of the basin with similar deposits is up to 200 m. The length of columns is about 5 m and the grid spacing is smaller than 2 m: 1.73 m. The columns' mean diameter is 320 mm. A numerical simulation with finite elements shown in Figure 7.14 predicts a stop band for elastic surface waves around 50 Hertz.

The seismic metamaterial test is challenging to model in full (3D Navier equations in unbounded heterogeneous media). We therefore opted for an asymptotic model, which only captures the wave physics at the air-soil interface. We do not claim to have a complete understanding of the elastic wave propagation with such a model, rather it is guidance for the experiments. However, there is a qualitatively good agreement between our numerical simulations, in Figure 7.14, and Ménard's company experiments, in Figure 7.13. Note that the seismic source frequency at 50 Hz in Figure 7.13 is located inside the first partial stop band in Figure 7.13, which confirms the stop band origin of the wave reflection.

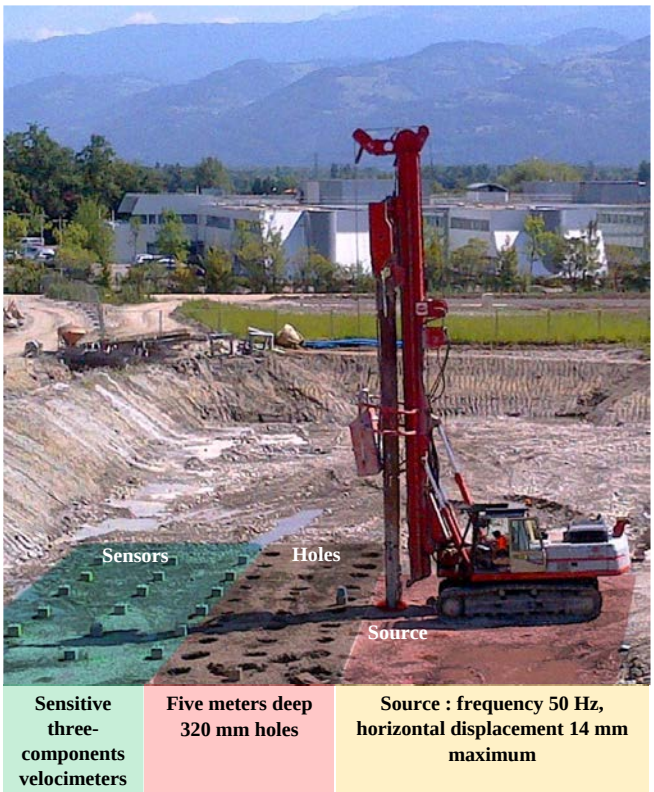


Fig. 7.13. Photo of the experimental setup of the seismic field test in Grenoble (Summer 2012), where the region highlighted in green is protected from surface seismic waves generated by a source (crane with rotating head at 50 Hz) thanks to the array of boreholes (region highlighted in blue). Boxes on the soil are velocimeters.

We consider the following approximate plate model for surface flexural waves

$$\rho^{-1}\nabla.\left[E^{\frac{1}{2}}\nabla\rho^{-1}\nabla.\left(E^{\frac{1}{2}}\nabla\Psi\right)\right]-\omega^2\frac{\rho h}{D}\psi=0 \tag{7.9}$$

where $\Psi(x,y)$ is the amplitude of displacement along the z-axis which is orthogonal to the air-soil interface in Figure 7.13 and ρ is the heterogeneous density of the plate: in the soil the density $\rho = 1500$ kg/m^3 and in boreholes $\rho = 1.2$ kg/m^3. Moreover, the spectral

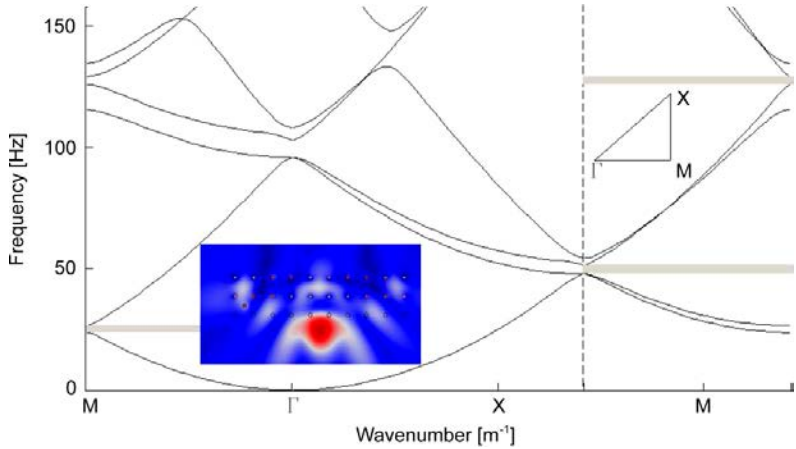


Fig. 7.14. Numerical simulation showing the band diagram of an infinite periodic structure (boreholes in soil) and the shielding effect for time-harmonic surface (bending) waves generated by a source at a frequency within the second partial Stop band.

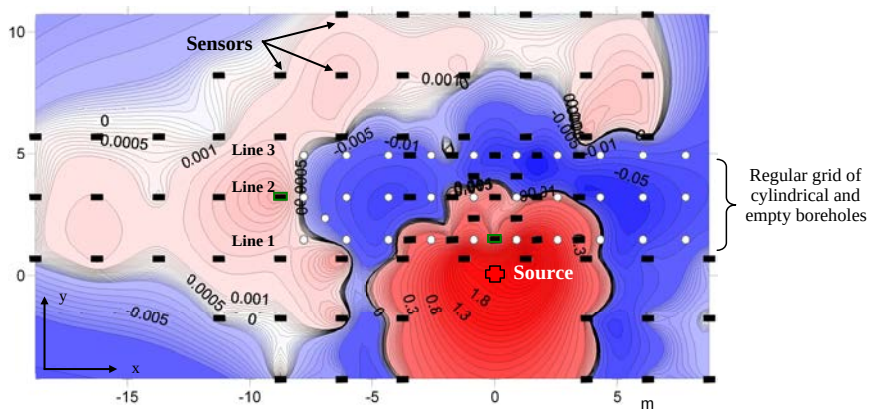


Fig. 7.15. Experimental result map after interpolation between sensors: Difference of the measured energy's field (arbitrary units) after and before carrying out the boreholes. Black rectangles symbolise sensors.

parameter $\beta^4 = \omega^2 \rho h / D$, with ω the angular flexural wave frequency, h the plate thickness (assumed to be 5 m), and D the plate rigidity, which in the present case is $D = Eh^3/(12(1-\nu^2))$. From the Ménard company characterization of the soil, we chose $E = 100$ MPA and

$\nu = 0.3$. In the finite element model, solved with COMSOL Multiphysics software, we set Floquet-Bloch boundary conditions on either sides of a periodic cell (in the xy-plane) of sidelength $d = 1.73$ m in order to compute the dispersion diagram shown in Figure 7.14. We also set perfectly matched layers (PMLs) on either boundaries of the domain in order to account for the unbounded domain in the xy-plane for the simulation shown in the inset of Figure 7.14. These PMLs are deduced from a geometric transform in Eq. 7.9 that lead to a Young's modulus similar to matrix M in Eq. 7.5.

Regarding the experimental results shown in Figure 7.15, some word of caution is in order for the experimental setup. Twenty sensors were used simultaneously with a common time-base. In order to map completely the energy's field, the sensors were set four times on site (green, blue, pink and orange grids in Figure 7.13) before and after carrying out the boreholes. The fact that the elastic energy is 2.3 times larger at the source point when it is in presence of the metamaterial in Figure 7.15 is reminiscent of the Local Density of States obtained for a source placed near a mirror in optics.

The frequency of the vibrating source in the experiments, 50 Hz, with 14 mm of lateral amplitude in xy-plane, should therefore lead to very strong reflection of surface elastic waves by the large scale metamaterial (see inset in Figure 7.14). The experimental grid is made of three discontinuous lines of ten boreholes 320 mm in diameter. Sensors are three component velocimeters (x, y, z) with a corner frequency of 4.5 Hz (-3 dB at 4.5 Hz) electronically corrected to 1 Hz. Twenty sensors were used simultaneously with a common time-base. In order to map completely the energy's field, the sensors were set four times on site before and after carrying out the boreholes.

The sensors A1 and B1 were kept at the same position during the whole experiment. They are the two sensors in Figure 7.3 that have green frames: A1 is the sensor right above the source; B1 is just on the left side of the boreholes. Due to the strong soil attenuation, the probe is close to the grid (1.5 m). Each record is about two

minutes long and can be divided into three parts: a pre-experiment part, record of ambient vibration noise, followed by the probe starting phase during which the frequency of the generated signal increases up to 50 Hz and the probe is set in place; the experimental phase itself, about one minute long and a post-experiment part, during which the mechanical probe stops, and that end by ambient vibration noise.

For each grid, the effective experimental phase's length is determined for the A1 fixed sensor. It is defined as the aforementioned experimental length reduced by ten seconds at both ends to ensure that any effect of the probe's starting and ending are well removed. The signals' average energy per second is then computed for each sensor. They are consecutively normalized by the A1's energy per second, so as to reconstruct a uniform energy field over the whole experimental area.

The boreholes' effect on the energy field is given in Figure 7.15 where the difference of the measured energy's field after and before carrying out the boreholes is displayed with interpolation between sensors. Before we finish our story-telling of the first large scale experiment of a seismic metamaterial, we would like to point out that performing such a test required a healthy dose of courage as both the civil engineering and geophysics communities were initially taken aback by the concept of a soil periodically structured on a scale of a few meters that could somewhat counteract the devastating effects of earthquakes in sedimentary basins.

7.10. Inertial Resonators for Smart Buildings

An alternative to structured soils to prevent seismic disaster is the design of smart buildings, e.g. fitted with inertial resonators like those shown in Figure 7.16. However, this approach requires tuning resonators at the eigenfrequencies of the building, and the wave protection cannot be broadband. An obvious pitfall of such smart buildings is that once the building is erected it seems fairly hard to change the resonators (note that alternative solutions include passive mass dampers in the building's foundations, and water tanks, share the same pitfall).

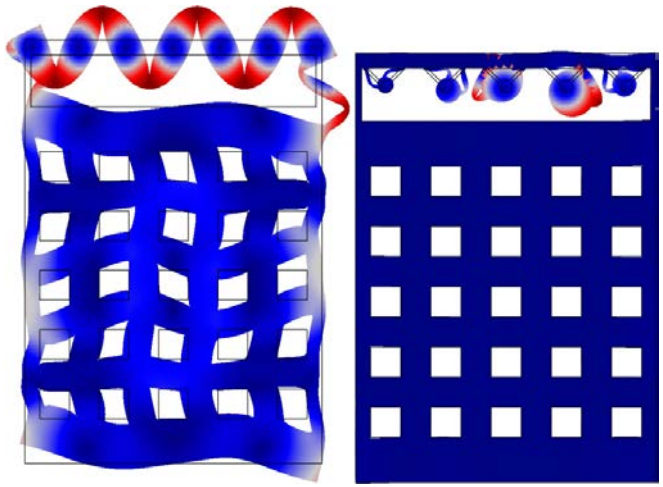


Fig. 7.16. Numerical simulation illustrating the concept of a meta-building with the deformation of a five-storey building at eigenfrequency near 5 Hertz (left panel) and its suppression with inertial resonators tuned to vibrate at 5 Hertz (right panel). Color scale ranges from dark blue (vanishing displacement) to red (maximum displacement). The same concept can be applied to suppress unwanted vibrations in bridges and other urban structures.

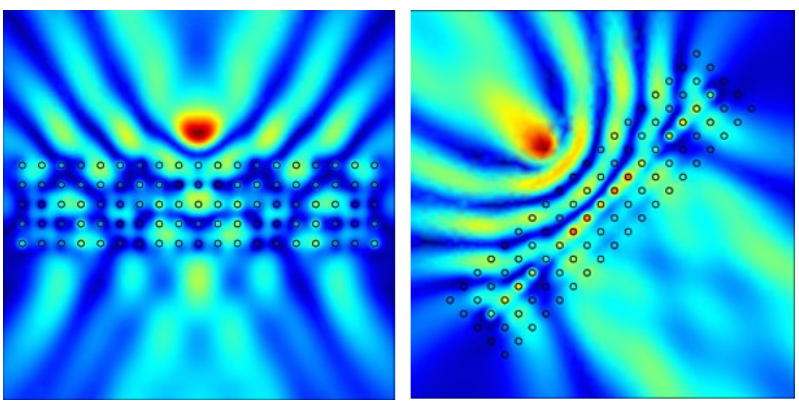


Fig. 7.17. Numerical simulation for lensing (left panel) and ultra-directive (right panel) effects for columns of concrete in soil. A large scale experiment has confirmed similar effects through an array of boreholes, about 2 meters in diameter, 5 meters in depth and with a grid spacing of 7 meters for a signal characterized by a mean frequency value about 8 Hz with a range of frequencies going from 3 to 20 Hz [Brûlé *et al.* (2017)].

7.11. Concluding Remarks

We have conclusively shown that one can enhance the interactions of surface seismic waves with structured soils by making use of analogies between optics and solid mechanics.

The possibility to conceive new seismic defences (even invisibilty cloacks) to improve the building's design and reflect/control seismic wave.

Bibliography/References

[1] Achaoui, Y., Antonakakis, T., Brulé, S., Craster, R. V., Enoch, S., and Guenneau, S. (2016) Seismic Defence Structures, (GB Patent filed on 21/10/2016).

[2] Achaoui, Y., Ungureanu, B., Brûle, S., Enoch, S. and Guenneau, S. (2016). Seismic wave damping with arrays of inertial resonators, *Extreme Mechanics Letters*, (http://dx.doi.org/10.1016/j.eml.2016.02.004).

[3] Achaoui, Y., Antonakakis, T., Brulé, S., Craster, R. V., Enoch, S., and Guenneau, S. (2017) Clamped seismic metamaterials: Ultra-low broad frequency stop-bands, *New J. Phys* **19**, p. 063022.

[4] Ahmad, S. and Al-Hussaini, T. M. (1991). Simplified design for vibration screening by open and infilled trenches, *J. Geotech. Eng.* **117**(1), pp. 67–82.

[5] Aki, K. and Richards, P. G. (2002). (ed. 2). *Quantitative Seismology: Theory and Methods* (University Science Books, California).

[6] Antonakakis, T. and Craster, R. V. (2012). High frequency asymptotics for microstructured thin elastic plates and platonics, *Proc. R. Soc. Lond. A* **468**, p. 1408.

[7] Alù A., Silveirinha, M., Salandrino, A. and Engheta, E. N. (2007). Epsilon-near-zero metamaterials and electromagnetic sources: Tailoring the radiation phase pattern, *Phys. Rev. B* **75**, p. 155410.

[8] Arias, A. (1970). A measure of earthquake intensity. In Hansen R. J., editor. *Seismic Design for Nuclear Power Plants*, pp. 438–483. Cambridge, Massachusetts: MIT Press. 23.

[9] Banerjee, P.K., Ahmad, S. and Chen, K. (1988). Advanced application of BEM to wave barriers in multi-layered three-dimensional soil media, *Earthquake Eng. & Structural Dynamics* **16**, pp. 1041–1060.

[10] Bommer, J.J. and Martinez-Pereira, A. (2000). Strong motion parameters: Definition, usefulness and predictability, *Proceedings 12WCEE*, Auckland, New Zeland, Paper No. 206.

[11] Brun, M., Guenneau, S. and Movchan, A. B. (2009). Achieving control of in-plane elastic waves, *Appl. Phys. Lett.* **94**, p. 061903.

[12] Brûlé, S., Javelaud, E. H., Enoch, S., Guenneau, S. and Komatitsch, D. (2012). Seismic metamaterials. *Proceedings of the 9th International*

Conference of the Association for Electrical, Transport and Optical Properties of Inhomogeneous Media, September 2–7, 2012, Marseille, France.

[13] Brûlé, S., Javelaud, E. H., Enoch, S. and Guenneau, S. (2014). Experiments on seismic metamaterials: Molding surface waves, *Phys. Rev. Lett.* **112**, p. 133901.

[14] Brûlé, S. and Javelaud, E. H. H/V method in geotechnical engineering. Application to a two-layers model, *Revue Française de Géotechnique, n° 142*.

[15] Brûlé, S., Guenneau, S. and Enoch, S. (2015). Structured soils under seismic disturbances, *Proceedings of Association Française du Génie Parasismique, AFPS'15*, 30 November–2 December 2015, Marne La Vallée, France.

[16] Brûlé, S., Enoch, S. and Guenneau, S. (2017). Sols structurés sous sollicitation dynamique: des métamatériaux en géotechnique, *Rev. Fr. Geotech.* **151**, p. 4.

[17] Brûlé, S., Enoch, S. and Guenneau, S. Flat lens for seismic waves, *Preprint* arXiv:1602.04492.

[18] Brûlé, S., Ungureanu, B., Achaoui, Y., Diatta, A., Aznavourian, R., Antonakakis, T., Craster, R.V., Enoch, S. and Guenneau, S. (2017). Metamaterial-like Transformed Urbanism. Innov. Infrastruct. Solut. 2:20.

[19] Brûlé, S. and Cuira, F. (2017). Pratique de l'interaction sol-structure sous séisme: Application aux Fondations et aux Soutènements. Editions Afnor, ISBN 978-2-12-46560064.

[20] Chu, J., Varaksin, S., Ulrich, K. and Mengé, P. (2009). Construction Processes, State of the Art Report, *Proceedings of the 17^{th} International Conference on Soil Mechanics and Geotechnical Engineering*, TC17 meeting ground improvement, Alexandria, Egypt.

[21] Christensen, J. and García De Abajo, F. J. (2012). Anisotropic metamaterials for full control of acoustic waves, *Phys. Rev. Lett.* **108**, p. 124301.

[22] Colombi, A., Guenneau, S., Guéguen, P. and Craster, R. V. (2016). Forest as natural seismic metamaterial: Rayleigh wave band gaps induced by local resonances, *Sci. Rep.* **6**, p. 19238.

[23] Colombi, A., Roux, P., Guenneau, S. and Rupin, M. (2015). Directional cloaking of flexural waves in a plate with a locally resonant metamaterial, *The Journal of The Acoustical Society of America* **137**, pp. 1783–1789.

[24] Craster, R. and Guenneau, S. (2013). *Acoustic Metamaterials: Negative Refraction, Imaging, Lensing and Cloaking*, vol. 166, R. Craster and S. Guenneau, Eds. Springer Verlag.

[25] Dolling, G., Enkrich, C., Wegener, M., Soukoulis, C. M. and Linden, S. (2006). Observation of simultaneous negative phase and group velocity of light, *Science* **312**, p. 892.

[26] Dowrick, D. (2009). *Earthquake Resistant Design and Risk Reduction*, 2^{nd} Edition, Wiley.

[27] Dubois, M., Bossy, E., Enoch, S., Guenneau, S., Lerosey, G. and Sebbah, P. (2013). Time drive super oscillations with negative refraction, arXiv:1303.3022v2.

[28] Dupont, G., Kimmoun, O., Molin, B., Guenneau, S. and Enoch, S. (2016). Numerical and experimental study of an invisibility carpet in a water channel, *Phys. Rev. E* **91**, p. 023010.
[29] Economou, E. N. and Sigalas, M. M. (1993). Classical wave propagation in periodic structures: Cermet versus network topology, *Phys. Rev. B* **48**, pp. 13434–13438.
[30] Enoch, S., Tayeb, G., Sabouroux, P., Guérin, N. and Vincent, P. (2002). A metamaterial for directive emission, *Phys. Rev. Lett.* **89**, p. 213902.
[31] Fang, N., Xi, D., Xu, J., Ambati, M., Srituravanich, W., Sun, C. and Zhang, X. (2006). *Nat. Mater.* **5**, p. 452.
[32] Farhat, M., Enoch, S., Guenneau, S. and Movchan, A. B. (2008). Broadband cylindrical acoustic cloak for linear surface waves in a fluid, *Phys. Rev. Lett.* **101**, p. 1345011.
[33] Farhat, M., Guenneau, S. and Enoch, S. (2012). Broadband cloaking of bending waves via homogenization of multiply perforated radially symmetric and isotropic thin elastic plates, *Phys. Rev. B* **85**, p. 020301 R.
[34] Finocchio, G., Casablanca, O., Ricciardi, G., Alibrandi, U., Garesci, F., Chiappini, M. and Azzerboni, B. (2014). Seismic metamaterials based on isochronous mechanical oscillators, *Appl. Phys. Lett.* **104**, p. 191903.
[35] Goldstein, H. (1950). *Classic Mechanics*, Addison Wesley, Reading, Massachussets.
[36] Gralak, B., Enoch, S. and Tayeb, G. (2000). Anomalous refractive properties of photonic crystals, *J. Opt. Soc. Am. A* **17**, pp. 1012–1020.
[37] Guenneau, S., Enoch, S. and Brûlé, S. (2015). Invisibility within the framework of transformational physics, *Proceedings of AFPS'15*, 30 novembre au 2 décembre 2015, Marne La Vallée, France.
[38] Hobiger, M. (2011). Polarization of surface waves: Characterization, inversion and application to seismic hazard assessment. Doctoral Thesis. Grenoble University.
[39] Hu, X. and Chan, C. T. (2005). Refraction of water waves by periodic cylinder arrays, *Phys. Rev. Lett.* **95**, p. 154501.
[40] Husid, L. R. (1969). Características de terremotos — análisis general, *Revista del IDIEM* 8, 21–42, Santiago, Chile.
[41] Iorch, I. V., Mukhin, I. S., Shadrivov, I. V., Belov, P. A. and Kivshar, Y. S. (2013). Hyperbolic metamaterials based on multilayer graphene structures, *Phys. Rev. B* **87**, pp. 075416–6.
[42] Joannopoulos, J. D., Johnson, S. G., Winn, J. and Meade, R. D. (2008). *Photonic Crystals: Molding the flow of Light*. Princeton University Press, ed. 2.
[43] John, S. (1987). Strong localization of photons in certain disordered dielectric superlattices, *Phys. Rev. Lett.* **58**, pp. 2486–2489.
[44] Kausel, E. and Roesset, J. M. (1981). Stiffness matrices for layered soils, *Bull. Seism. Soc. America* **71**(6), pp. 1743–1761.
[45] Krodel, S., Thome, N. and Daraio, C. (2015). Wide band-gap seismic metastructure, *Ex. Mech. Letters* **4**, pp. 111–117.

[46] Kushwaha, M. S., Halevi, P., Dobrzynski, L. and Djafari-Rouhani, B. (1993). Acoustic band structure of periodic elastic composites, *Phys. Rev. Lett.* **71**, pp. 2022–2025.

[47] Liu, Z., Zhang, X., Mao, Y., Zhu, Y., Yang, Z., Chan, C. T. and Sheng, P. (2000). Locally resonant sonic materials, *Science* **289**, p. 1734.

[48] Luo, C., Johnson, S. G., Joannopoulos, J. D. and Pendry, J. B. (2002). All-angle negative refraction without negative effective index, *Phys. Rev. B* **65**, p. 201104.

[49] Martinez-Sala, R., Sancho, J., Sanchez, J. V., Gomez, V., Llinares, J. and Meseguer, F. (1995). *Nature* **378**, p. 241.

[50] Meseguer, F., Holgado, M., Caballero, D., Benaches, N., Sanchez-Dehesa J., Lopez, C. and Llinares, J. (1999). Rayleigh-wave attenuation by a semi-infinite two-dimensional elastic-band-gap crystal, *Phys. Rev. B* **59**, p. 12169.

[51] Milton, G. W., Briane, M. and Willis, J. R. (2006). On cloaking for elasticity and physical equations with a transformation invariant form, *New J. Phys.* **8**, p. 248.

[52] Nicolet, A., Remacle, J. F., Meys, B., Genon, A. and Legros, W. (1994). Transformation methods in computational electromagnetics, *J. Appl. Phys.* **75**(10), pp. 6036–6038.

[53] Norris, A. and Shuvalov, A. L. (2011). Elastic cloaking theory, *Wave Motion*, **48**(6), pp. 525–538.

[54] Notomi, M. (2000). Theory of light propagation in strongly modulated photonic crystals: Refractionlike behaviour in the vicinity of the photonic band gap, *Phys. Rev. B* **62**, pp. 10696–10705.

[55] O'Brien, S. and Pendry, J. B. (2002). Photonic band-gap effects in dielectric composites, *J. Phys. Cond. Matt.* **14**, pp. 4035–4044.

[56] Pendry, J. B., Holden, A. J., Robbins, D. J. and Stewart, W. J. (1999). Magnetism from conductors and enhanced nonlinear phenomena, *IEEE Transactions on Microwave Theory and Techniques* **47**(11), p. 2075.

[57] Pendry, J. B. (2000). Negative refraction makes a perfect lens, *Phys. Rev. Lett.* **85**, pp. 3966–3969.

[58] Pendry, J. B., Schurig, D. and Smith, D. R. (2006). Controlling electromagnetic fields, *Science* **312**(5781), pp. 1789–1782.

[59] Schurig, D., Mock, J. J., Justice, B. J., Cummer, S. A., Pendry, J. B., Starr, A. F. and Smith, D. R. (2006). Metamaterial electromagnetic cloak at microwave frequencies, *Science* **314**(5801), pp. 977–980.

[60] Semblat, J. F. and Pecker, A. (2009). Waves and vibrations in soils: Earthquakes, traffic, shocks, construction works, Istituto Universitario di Studi Superiori IUSS Press, Pavia.

[61] Semblat, J. F. (2011). Modeling seismic wave propagation and amplification in 1D/2D/3D linear and nonlinear unbounded media, *Int. J. Geomechanics (ASCE)*, **11**(6), pp. 440–448.

[62] Smith, D. R., Padilla, W. J., Vier, V. C., Nemat-Nasser, S. C. and Schultz, S. (2000). *Phys. Rev. Lett.* **84**, p. 4184.

[63] Srinivasan, K. and Painter, O. (2002). Momentum space design of high-q photonic crystal optical cavities, *Opt. Express* **10**, pp. 670–684.
[64] Stenger, N., Wilhelm, M. and Wegener, M. (2012). Experiments on elastic cloaking in thin plates, *Phys. Rev. Lett.* **108**, p. 014301.
[65] Stover, C. W., Minsch, C. H., Reagor, B. G. and Smith, P. K. (1981). Earthquakes in the United States, January — March 1980, Geological Survey Circular, 853 — A.
[66] Sukhovich, A., Merheb, B., Muralidharan, K., Vasseur, J. O., Pennec, Y., Deymier, P. A. and Page, J. H. (2009). Experimental and theoretical evidence for subwavelength imaging in phononic crystals, *Phys. Rev. Lett.* **102**, p. 154301.
[67] Trifunac, M. D. and Brady, A. G. (1975). A study on the duration of strong earthquake ground motion, *Bull. Seism. Soc. Am.* **65**, pp. 581–626.
[68] Vasseur, J. O., Deymier, P. A., Frantziskonis, G., Hong, G., Djafari-Rouhani, B. and Dobrzynski, L. (1998). Experimental evidence for the existence of absolute acoustic band gaps in two-dimensional periodic composite media, *J. Phys.: Condens. Matter* **10**, pp. 6051–6064.
[69] Veselago, V. G. (1968). The electrodnamics of substances with simultaneously negative values of ε and μ, *Soviet Physics Uspekhi* **10**(4), pp. 509–514.
[70] Walser, R. M. (2001). Electromagnetic metamaterials, paper presented at the International Society for Optical Engineering (SPIE), vol. 4467, pp. 1–165.
[71] Woods, R. D. (1968). Screening of surface waves in soils, The University of Michigan Industry Program of the College of Engineering. IP-804.
[72] Yablonovitch, E. (1987). Inhibited spontaneous emission in solid-state physics and electronics, *Phys. Rev. Lett.* **58**, p. 2059.
[73] Zengerle, R. (1987). Light propagation in singly and doubly periodic waveguides, *J. Mod. Opt.* **34**, pp. 1589–1617.

CHAPTER 8

Active Cloaking and Resonances for Flexural Waves

JANE O'NEILL, ÖZGÜR SELSIL, ROSS C. McPHEDRAN, ALEXANDER B. MOVCHAN* and NATASHA V. MOVCHAN

University of Liverpool, UK

8.1. Background

The cloaking transformation, referred to as geometrical optics, is well known and well discussed in the literature (see [Dolin (1961)], [Leonhardt (2006)], [Pendry *et al.* (2006)], [Greenleaf *et al.* (2003)] for the seminal papers). The transformation of coordinates is non-degenerate but singular at the origin, and in contrast with conformal maps, the Laplace's operator changes its structure. This has been exploited in the context of modelling anisotropic media. In particular, so-called "invisibility cloaks" have been designed and constructed for problems of acoustics and out-of-plane shear elasticity [Leonhardt (2006)], [Pendry *et al.* (2006)].

The work [Milton *et al.* (2006, 2007)] and [Pendry *et al.* (2006)] has put cloaking in the framework of a new research direction, leading

*Corresponding author: abm@liv.ac.uk

to the theoretical design and implementation of "metamaterials". These are understood as multi-scale design materials that possess certain properties, such as dynamic anisotropy and negative refraction amongst many others.

Cloaking for vector elasticity and for flexural waves in elastic plates is more challenging than for the case of acoustics. The formal reason is that the "cloaking transformation" leads to new equations, which require additional physical interpretation and understanding. For example, for plane strain elasticity problems, several gauge conditions have been identified [Norris and Shuvalov (2011)] to refer to the Cosserat type materials [Brun *et al.* (2009)] or to materials where the mass is considered as a tensor quantity rather than a scalar [Milton *et al.* (2006)]. Substantial discussion and analysis of these challenges have been published by [Milton and Willis (2007)], [Norris *et al.* (2012)], [Norris *et al.* (2013)], [Farhat *et al.* (2009)]. The paper [Farhat *et al.* (2014)] poses an interesting and important question of passive scattering cancellation for bending waves in plates. It was proposed to introduce a coating around an elastic inclusion, and choose its parameters appropriately to reduce the scattering cross-section. The implementation presented by [Farhat *et al.* (2014)] is for a low-frequency regime, for a special case of "decoupling" transmission conditions, and it presents a fascinating idea of high potential impact.

The implementation of exact structured cloaks for vector elasticity or thin Kirchhoff plates is impossible, and practical approximate realisations are challenging. Active cloaking offers an efficient alternative which therefore deserves full attention. The work by Miller [Miller (2006)] has introduced the general concept of active cloaking and also discussed challenges linked to extending the cloaking action across a wide frequency range. Acoustic cancellations and their modelling in the context of active cloaks have been also discussed by [Friot and Bordier (2004)], [Friot *et al.* (2006)]. The idea of active sound filtering and noise cancellation goes back to the early eighties; the fundamental theoretical concepts and numerical simulations were discussed by [Ffowcs Williams (1984)] and by [Nelson and Elliott (1992), pp. 290–293].

In recent papers by [Guevara Vasquez *et al.* (2009(a)], [Guevara Vasquez *et al.* (2009(b)] a rigorous variational approach has been developed to achieve a broadband active exterior cloaking for time-harmonic waves governed by the Helmholtz equation. This was followed by the work of [Norris *et al.* (2012)], [Norris *et al.* (2013)] which addressed the evaluation of active source amplitudes for exterior cloaking in an inverse problem setting.

The paper [O'Neill *et al.* (2015(a)] presented a novel development of active cloaking for problems involving flexural waves in Kirchhoff plates. The challenge of resonance modes was accounted for by introducing a novel design that incorporates placing a specially chosen coating around the inclusion. This work covers broadband regimes and provides a constructive method to handle cases of resonant localisation.

The recent paper by [Futhazar *et al.* (2015)] follows the earlier work by [Guevara Vasquez *et al.* (2009(a)], [Guevara Vasquez *et al.* (2009(b)] and the paper by [Norris *et al.* (2012)], [Norris *et al.* (2013)] to approximate the active cloaking "devices" by multipole expansions for the case of Kirchhoff plates. The algorithm of [Futhazar *et al.* (2015)] has poor convergence, which implies the importance of resonance regimes for the cloaking of inclusions in Kirchhoff plates. The significance of resonance phenomena was also noted in [Guevara Vasquez *et al.* (2011)].

Practical applications and viability of passive cloaking have been discussed by [Monticone and Alù (2013)], who asked the pragmatic question "Do Cloaked Objects Really Scatter Less?", to which the answer was not straightforward. It was demonstrated that practical broadband cloaking benefits strongly from an active input generated by external sources.

The work by [Jones *et al.* (2015)] emphasises the importance of boundary conditions on the interior contours of cloaking coatings. In this paper a counterexample was constructed for the case of Kirchhoff plates, where cloaking was not possible — the notion of a "cloaking illusion" was also introduced and discussed for the cases of membrane waves as well as flexural waves in Kirchhoff plates.

8.2. Outline of Results

In the recent papers [O'Neill *et al.* (2015(a)] and [O'Neill *et al.* (2015(b)], we presented the first investigation into active cloaking for flexural waves in Kirchhoff plates. The elegant and explicit analytical approach to biharmonic active cloaking, based on multipole approximations, was motivated by the earlier investigations of the interaction of flexural waves with structured systems of scatterers ([Norris and Vemula (1995)], [Evans and Porter (2007)], [Movchan *et al.* (2007)], [McPhedran *et al.* (2009)] and [Movchan *et al.* (2009)]).

One of the simple but important features of solutions to flexural wave problems is that their general representations involve both Helmholtz and modified Helmholtz waves. Of course, all modified Helmholtz waves are evanescent, and there exists a perception that their amplitudes need not to be controlled for effective cloaking at large distances from a scatterer (typically a finite inclusion). However, these waves do contribute significantly to the boundary conditions and hence may play an important role in resonant regimes.

In [O'Neill *et al.* (2015(a)], [O'Neill *et al.* (2015(b)], the sources are located outside the scatterer and their complex wave amplitudes are chosen to eliminate propagating orders of the multipole representation of the scattered field. The cloaking method described is approximate in the sense that this elimination is of a certain order. Nonetheless, the quality of the cloaking may be improved by increasing this order.

Before presenting the analytical description of the problem and its solution, we give an outline of some of the results, and demonstrate the cloaking principles with illustrative computations. At this point, we also allude to the difference between cloaking and shielding, since there seems to be a common misconception of cloaks being identical to shields. In shielding problems, a quiet region surrounding the object is created where the wave field practically vanishes, which makes the object virtually invisible. However, this does not guarantee that the devices are undetectable and they may produce a scattering signature in the far field. In cloaking problems, the scattered field

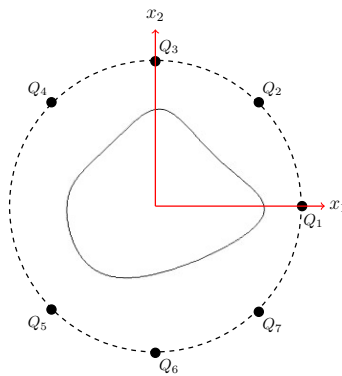


Fig. 8.1. Enlarged view of the cavity within a thin plate and the locations of the control sources surrounding the cavity.

from both the sources and the object is reduced in a region exterior to the object and sources.

Finally, we emphasise that boundary conditions of scatterers play an important role in cloaking: scatterers with Neumann boundary conditions are much "easier" to cloak compared with scatterers with Dirichlet boundary conditions. We also discuss transmission boundary conditions which are posed when an inclusion is made of a different material.

8.2.1. *Cloaking of scatterers with clamped boundaries*

Our first result is for an arbitrarily shaped scatterer in a thin elastic plate. We cloak this scatterer in the presence of an incident (plane or cylindrical) wave using seven active control sources. An enlarged view of the scatterer (cavity) and surrounding control sources is shown in Fig. 8.1.

We begin with considering the case of an incident plane flexural wave propagating in a thin plate with an arbitrarily shaped scatterer whose boundary is clamped. Fig. 8.2(a) illustrates the propagation through the plate in the absence of a scatterer; in Fig. 8.2(b) the shadow region, the most prominent signature of a defect in the medium, is clearly visible behind the scatterer; in Fig. 8.2(c) the scattered field is suppressed, creating almost perfect cloaking. We also

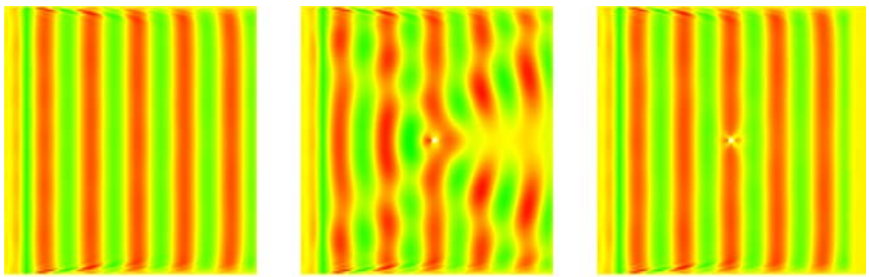


Fig. 8.2. (a): Flexural wave amplitude for a plane wave in an infinite thin plate; (b) Flexural wave amplitude in the presence of an arbitrarily shaped cavity; (c) Cloaking achieved by the use of seven control sources (sources not shown on the contour plot). To avoid reflections from the boundaries of the computational window we used perfectly matched layers.

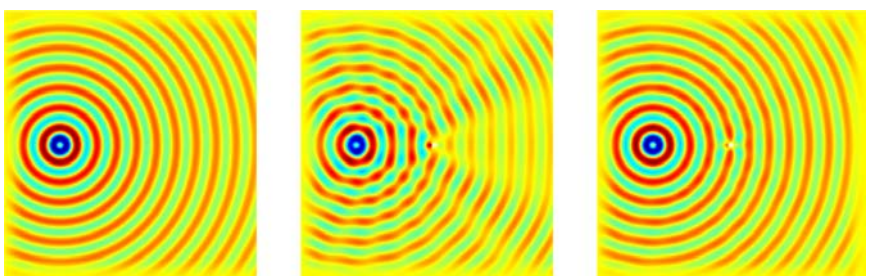
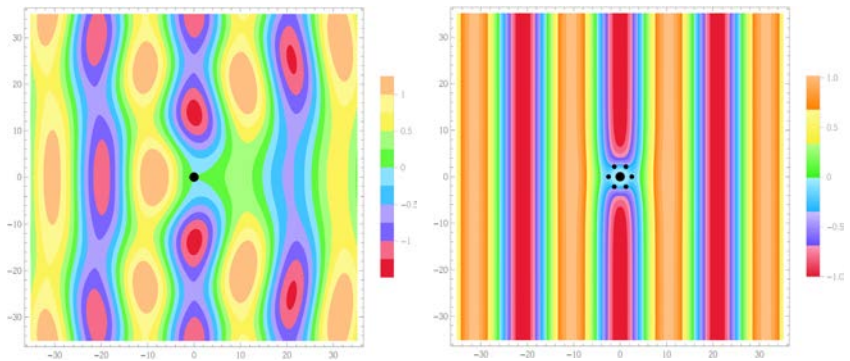


Fig. 8.3. Same as in Fig. 8.3 but this time in the presence of an incident flexural cylindrical wave. The incident wave is initiated by a remote point source.

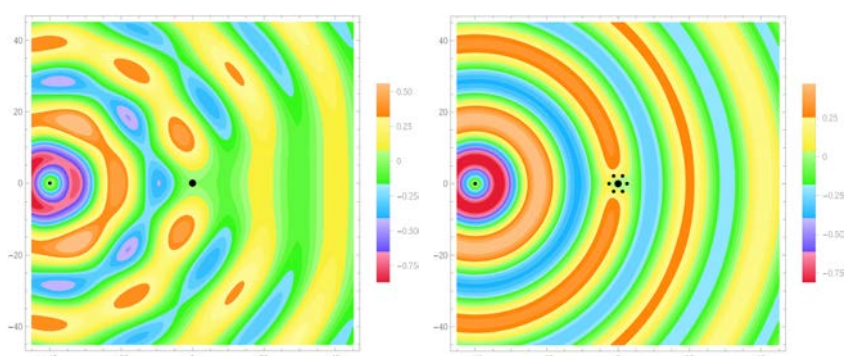
present the interaction of an incident cylindrical flexural wave with an arbitrarily shaped scatterer in Fig. 8.3; the incident wave propagates from a point source located sufficiently far away from the scatterer.

We note that when the boundary of the scatterer is clamped, even if the characteristic size of the scatterer is very small, the scattered field remains finite and considerable effort is required to suppress the shadow region behind the scatterer. On the contrary, the prescription of a free boundary condition reduces the scattered field significantly and, especially when the wavelength is larger than the characteristic



(a) (b)

Fig. 8.4. Plane incident wave: (a) Flexural wave amplitude with no control sources present. The black dot in the middle depicts the position of the circular clamped inclusion. (b) Flexural wave amplitude with six control sources. The six smaller black dots surrounding the inclusion depict the positions of the control sources.



(a) (b)

Fig. 8.5. Cylindrical incident wave: (a) and (b) as in Fig. 8.4. The small black dot on the far left depicts the point source generating the incoming wave.

size of the scatterer, the incident wave may be reconstructed behind the scatterer with very little effort.

For the case when the clamped boundary of the scatterer has a circular shape, we devise an explicit analytical algorithm based on the multipole representation of the wave amplitudes. In Figs. 8.4 and 8.5, we illustrate two different types of incident fields interacting

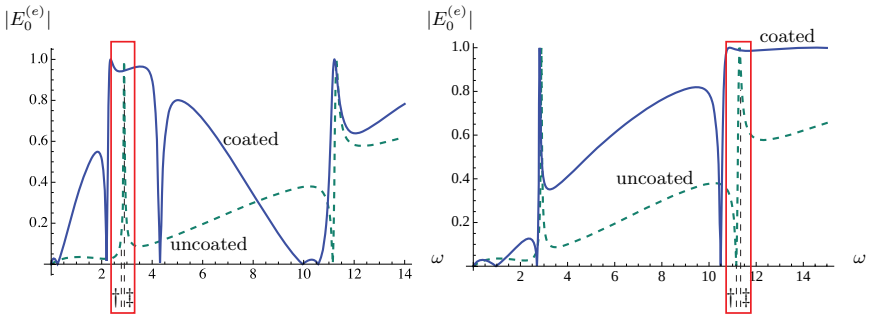


Fig. 8.6. The absolute value of the monopole term $E_0^{(e)}$ as a function of the frequency ω. Dashed/green: high contrast inclusion with no coating; solid/blue: inclusion with a suitably designed coating for (a) first resonant regime, (b) second resonant regime.

with a circular clamped scatterer: a flexural plane wave and a flexural cylindrical wave. It is demonstrated in these two figures that utilising six active sources restores the incident field in the shadow region to a high accuracy. In addition, we remark that in the immediate neighbourhood of the cloaked scatterer the displacement field is slowly varying, as commonly observed in non-resonant cases.

8.2.2. *Cloaking of elastic inclusions in resonant regimes*

Contrary to the static case, the scattered field around an elastic inclusion with a different mass density will possess a non-zero monopole term and the coefficient near such a term may change rapidly in the case of resonant vibrations. Of course, other multipole coefficients may undergo rapid variations when there is resonance but for a low-frequency resonance it is convenient to use the monopole coefficient for illustrative purposes. Consequently, in Fig. 8.6(a) and Fig. 8.6(b) we show the absolute value of the monopole coefficient $E_0^{(e)}$ as a function of the frequency ω of the incident plane wave for the case of an uncoated inclusion (dashed/green). The resonant regimes are clearly identified near the regions of rapid variation of the coefficient (see the red boxes): a low value in (a) and the first non-trivial zero of the

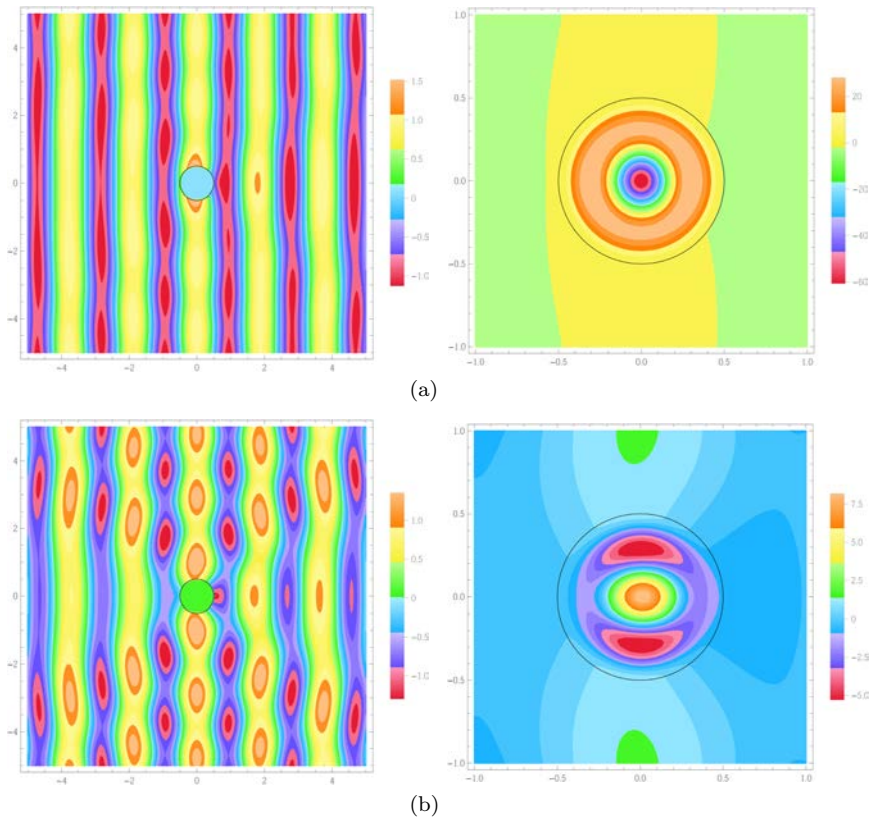


Fig. 8.7. Flexural wave amplitude in the case of an uncoated inclusion: (a) for a frequency corresponding to the first non-trivial zero of the monopole coefficient (see † in Fig. 8.6(b)), (b) for a nearby frequency corresponding to a large value of the monopole coefficient (see ‡ in Fig. 8.6(b)). In both (a) and (b), the field outside and inside the inclusion is presented on the left and right, respectively.

monopole coefficient in (b) are both denoted by †; a very small change in the frequency results in a very large variation of the monopole coefficient, denoted by ‡. These resonant regions are associated with high contrast between the inclusion and the surrounding medium. In the text below, we focus our attention on the case depicted in Fig. 8.6(b) where the variation is more drastic than in Fig. 8.6(a) for the corresponding resonant regime. In Fig 8.7, contour plots of the total wave amplitude corresponding to these two very close values of the

frequency from Fig. 8.6(b) are shown: in Fig 8.7(a) scattering is primarily due to the terms of dipole order, but the wave fronts behind the inclusion are practically straight: in Fig 8.7(b) both monopole and dipole terms are contributing to the distortion of the wavefronts behind the inclusion. The scattering patterns change dramatically even though the difference in frequency values is small.

At a frequency when the absolute value of the monopole coefficient $E_0^{(e)}$ is small but rapidly varying, active cloaking requires a rapid adjustment of the source amplitudes. To overcome this difficulty, a specially designed coating is devised with the aim of reducing this variation of the monopole coefficient for the required frequency range. Thus the coating may be considered to serve as a passive control of scattering. The two solid/blue curves in Fig. 8.6(a) and Fig. 8.6(b) illustrate the use of a suitably designed coating to tame the region of steep gradient of the monopole term in a targeted frequency interval. We note the enhancement of the scattering as a result of the increase in the monopole coefficient. This drastic enhancement of scattering corresponding to Fig. 8.6(b) is shown in Fig. 8.8: the wave amplitude in the exterior and the interior of the coated inclusion are both presented here (compare this with Fig. 8.7).

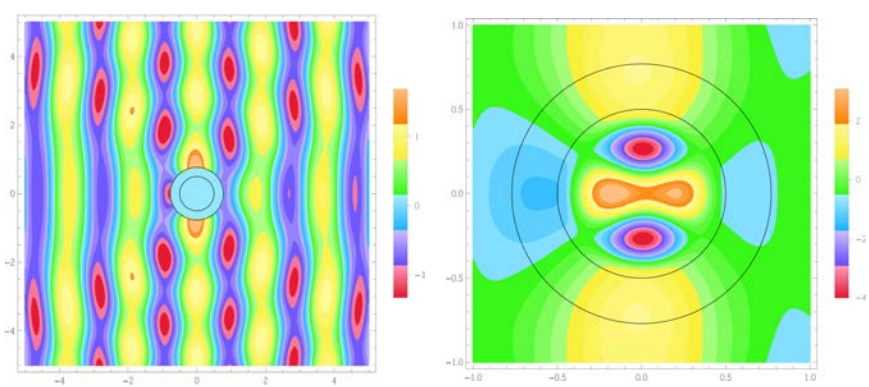


Fig. 8.8. Flexural wave amplitude: Scattering enhancement from an inclusion with a suitably designed coating and the detailed interior field (compare these with Fig. 8.7(a)).

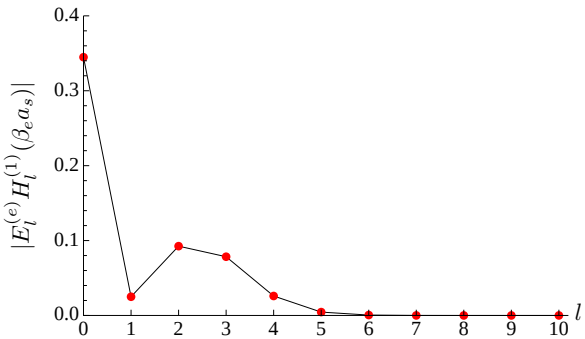


Fig. 8.9. Absolute value of the normalised outgoing wave components $E_l^{(e)} H_l^{(1)}(\beta_e a_s)$ versus the multipole order l for the resonant frequency denoted by † in Fig. 8.6(b).

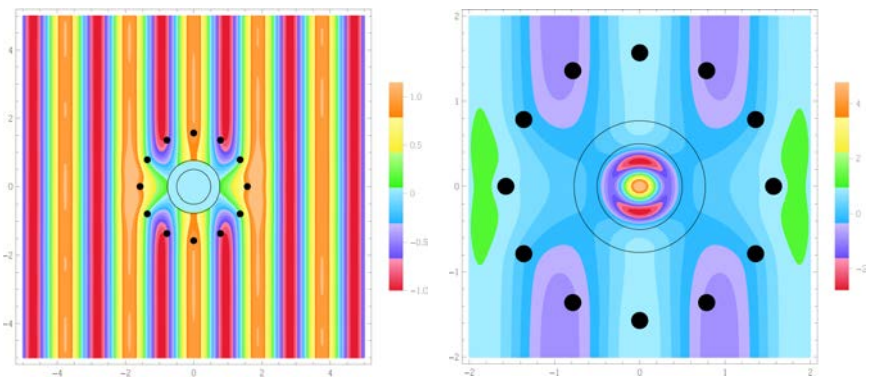
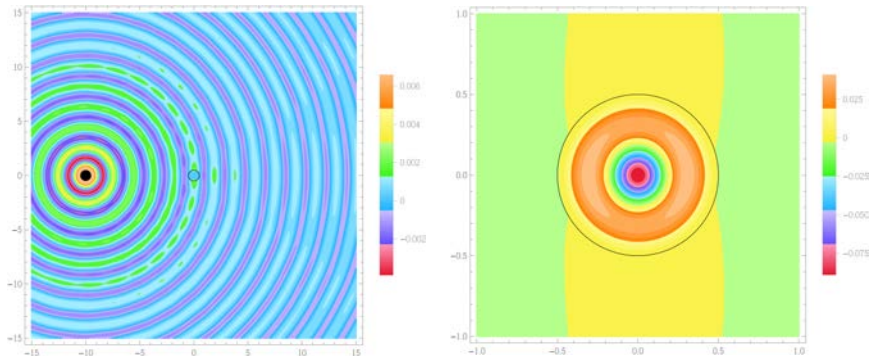


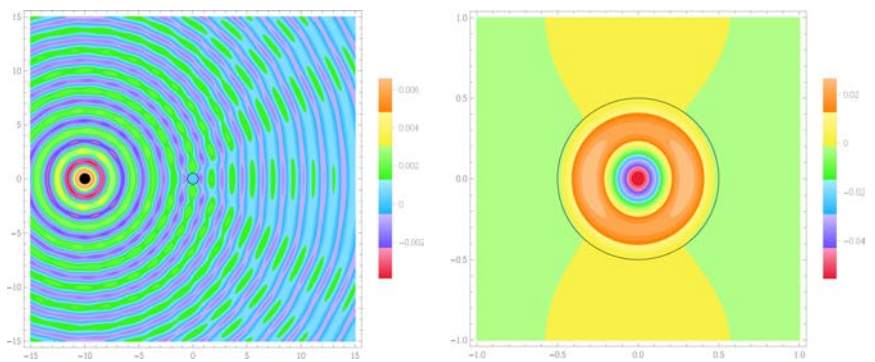
Fig. 8.10. Flexural wave amplitude using 12 active control sources: exterior and interior fields.

We note that for higher frequencies it is necessary to use more sources to achieve high-quality cloaking. In fact, the required number of control sources needed may be found by plotting the multipole field components versus the multipole order. It is evident from Fig. 8.9 that twelve sources would almost guarantee an effectively perfect cloaking. The result is presented in Fig. 8.10.

All these illustrations may be regenerated for different types of incident waves. Here, we give the results for an incident cylindrical wave in Figs. 8.11–8.13.



(a)



(b)

Fig. 8.11. (a) and (b) as in Fig. 8.7 for an incident cylindrical wave.

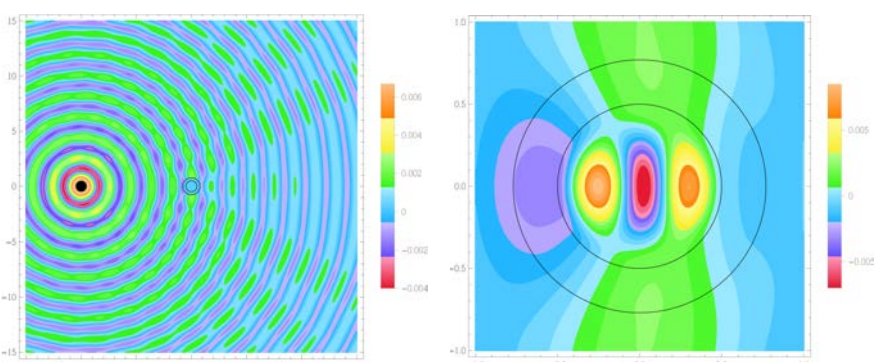


Fig. 8.12. As in Fig. 8.8 for an incident cylindrical wave.

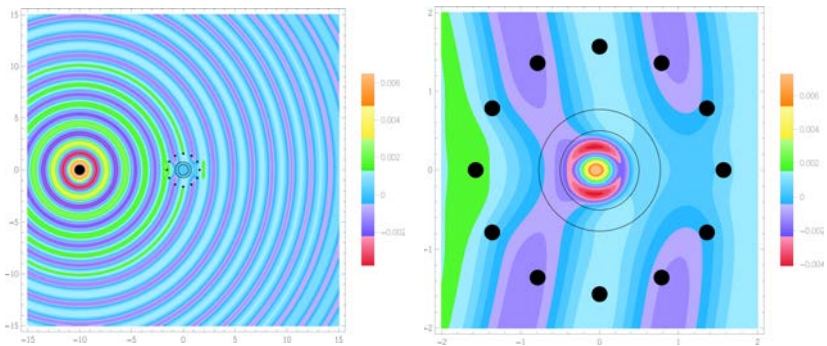


Fig. 8.13. As in Fig. 8.10 for an incident cylindrical wave.

8.3. Mathematical Model

8.3.1. *Governing equations*

For the sake of simplicity, we formulate the problem for a canonical geometry first. We note that the formulation can be extended to an arbitrary geometry with a smooth boundary (see Sec. 8.3.5). The out-of-plane elastic displacement $W(\mathbf{x};t)$ satisfies the equation of motion

$$\Delta^2 W(\mathbf{x};t) + \frac{\rho h}{D}\ddot{W}(\mathbf{x};t) + \mathcal{F}(\mathbf{x};t) = 0, \quad \mathbf{x} = (r,\theta) \in \Omega, \qquad (8.1)$$

where the region of interest Ω is a Kirchhoff plate with a clamped scatterer or an inclusion which is surrounded by an annular coating, as shown in Fig. 8.14(a) and Fig. 8.14(b) respectively; (r,θ) are the polar coordinates. We note that in what follows we use the sub/superscripts i, c and e to denote the inclusion, coating and exterior, respectively.

In Eq. (8.1), $\mathcal{F}(\mathbf{x};t)$ represents the contribution from the active point sources placed in the exterior of the clamped scatterer or the coated inclusion, Δ^2 is the biharmonic operator, a dot on the variable denotes the derivative with respect to time t, ρ is the mass density, h is the thickness of the plate, $D = Eh^3/[12(1-\nu^2)]$ is the flexural rigidity, with E the Young's modulus and ν the Poisson's ratio of the corresponding elastic material.

Assuming time-harmonic vibrations $W(\mathbf{x};t) = w(\mathbf{x})\exp(\mathrm{i}\omega t)$, outside the support of the function $\mathcal{F}(\mathbf{x};t)$, the governing equation

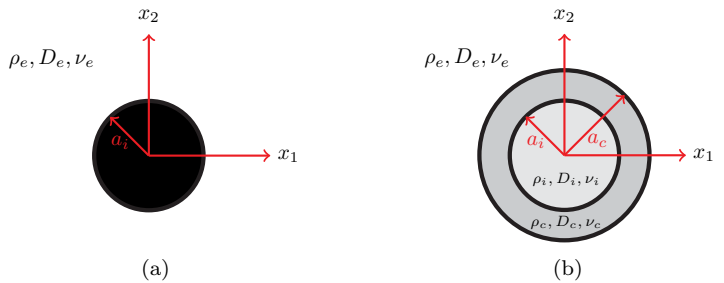


Fig. 8.14. (a) a clamped inclusion; (b) a coated inclusion. The inclusion, coating and exterior are associated with the subscripts i, c and e, respectively.

(8.1) can be reduced to

$$\left(\Delta^2 - \frac{\rho h\,\omega^2}{D}\right) w(\mathbf{x}) = (\Delta + \beta^2)(\Delta - \beta^2) w(\mathbf{x}) = 0, \quad \mathbf{x} \in \Omega. \tag{8.2}$$

Here $\beta^2 = \omega\sqrt{\rho h/D}$ is the spectral parameter. Consequently, $w(\mathbf{x})$ can be written as the sum of solutions of the Helmholtz and modified Helmholtz equations

$$w(\mathbf{x}) = w_H(\mathbf{x}) + w_M(\mathbf{x}), \tag{8.3}$$

where

$$\Delta w_H(\mathbf{x}) + \beta^2 w_H(\mathbf{x}) = 0, \quad \Delta w_M(\mathbf{x}) - \beta^2 w_M(\mathbf{x}) = 0. \tag{8.4}$$

For the clamped inclusion within the plate (see Fig. 8.14(a)), we have the boundary conditions

$$w(\mathbf{x}) = \frac{\partial w(\mathbf{x})}{\partial r} = 0 \quad \text{on} \quad r = a_i. \tag{8.5}$$

For the coated inclusion (see Fig. 8.14(b)), we pose conditions for perfect bonding along the interfaces $r = a_i$ and $r = a_c$, that is, the continuity of the displacement, its normal derivative, the moment and the transverse force. These conditions, with spatial dependence suppressed, lead to

$$w^{(c)} = w^{(l)}, \quad \frac{\partial w^{(c)}}{\partial r} = \frac{\partial w^{(l)}}{\partial r}, \tag{8.6}$$

$$D_c\left[\frac{\partial^2 w^{(c)}}{\partial r^2}+\frac{\nu_c}{r}\left(\frac{\partial w^{(c)}}{\partial r}+\frac{1}{r}\frac{\partial^2 w^{(c)}}{\partial\theta^2}\right)\right]$$
$$=D_l\left[\frac{\partial^2 w^{(l)}}{\partial r^2}+\frac{\nu_l}{r}\left(\frac{\partial w^{(l)}}{\partial r}+\frac{1}{r}\frac{\partial^2 w^{(l)}}{\partial\theta^2}\right)\right], \tag{8.7}$$

$$D_c\left[\frac{\partial}{\partial r}\Delta_{r\theta}w^{(c)}+\frac{1-\nu_c}{r^2}\frac{\partial^2}{\partial\theta^2}\left(\frac{\partial w^{(c)}}{\partial r}-\frac{w^{(c)}}{r}\right)\right]$$
$$=D_l\left[\frac{\partial}{\partial r}\Delta_{r\theta}w^{(l)}+\frac{1-\nu_l}{r^2}\frac{\partial^2}{\partial\theta^2}\left(\frac{\partial w^{(l)}}{\partial r}-\frac{w^{(l)}}{r}\right)\right]. \tag{8.8}$$

On $r=a_i$ we have $l=i$ and on $r=a_c$ we have $l=e$. In either case, we can write the solution to (8.2), outside the positions of the control sources, as

$$w(r,\theta)=\sum_{l=-\infty}^{\infty}\left[A_lJ_l(\beta r)+E_lH_l^{(1)}(\beta r)+B_lI_l(\beta r)+F_lK_l(\beta r)\right]e^{il\theta}, \tag{8.9}$$

where the terms inside the square brackets are the l-th multipole components of the flexural wave. We note that the multipole coefficients A_l, B_l are associated with the incoming wave, whereas the coefficients E_l, F_l correspond to the outgoing wave.

8.3.2. *A circular clamped inclusion*

In the case of a circular clamped inclusion, the boundary conditions (8.5) lead to the simple matrix relation

$$\begin{pmatrix}H_l^{(1)}(\beta_e a_i) & K_l(\beta_e a_i)\\ H_l^{(1)\prime}(\beta_e a_i) & K_l'(\beta_e a_i)\end{pmatrix}\begin{pmatrix}E_l^{(e)}\\ F_l^{(e)}\end{pmatrix}$$
$$=-\begin{pmatrix}J_l(\beta_e a_i) & I_l(\beta_e a_i)\\ J_l'(\beta_e a_i) & I_l'(\beta_e a_i)\end{pmatrix}\begin{pmatrix}A_l^{(e)}\\ B_l^{(e)}\end{pmatrix}, \tag{8.10}$$

whose solution may be given as

$$\begin{pmatrix} E_l^{(e)} \\ F_l^{(e)} \end{pmatrix} = \mathcal{S}_l \begin{pmatrix} A_l^{(e)} \\ B_l^{(e)} \end{pmatrix}, \tag{8.11}$$

with

$$\begin{aligned} \mathcal{S}_l = & \frac{1}{\mathcal{W}[H_l^{(1)}(\beta_e a_i), K_l(\beta_e a_i)]} \\ & \times \begin{pmatrix} \mathcal{W}[K_l(\beta_e a_i), J_l(\beta_e a_i)] & \mathcal{W}[K_l(\beta_e a_i), I_l(\beta_e a_i)] \\ \mathcal{W}[J_l(\beta_e a_i), H_l^{(1)}(\beta_e a_i)] & \mathcal{W}[I_l(\beta_e a_i), H_l^{(1)}(\beta_e a_i)] \end{pmatrix}. \end{aligned} \tag{8.12}$$

Above, we have used appropriate sub/super-scripts for the spectral parameter β and the multipole coefficients A_l, B_l, E_l and F_l, $\mathcal{W}[\cdot]$ stands for the respective Wronskian, prime, $'$, denotes the derivative with respect to the entire argument and $\mathcal{S}_l$ is the l-th order scattering matrix.

First, we consider the case when the scattered field is due to a flexural plane wave travelling along the x_1-axis. Then, the incident field can be written as

$$w_{(inc)}^{(p)}(x_1) = \exp(i\beta_e x_1) = \exp(i\beta_e r\cos\theta) = \sum_{l=-\infty}^{\infty} i^l J_l(\beta_e r) e^{il\theta}, \tag{8.13}$$

according to the Jacobi-Anger expansion. Thus, (8.11) together with (8.13) give a general expression for the scattered field as

$$\begin{aligned} w_{(sc)}^{(p)}(r,\theta) &= \sum_{l=-\infty}^{\infty} \left[E_l^{(e)} H_l^{(1)}(\beta_e r) + F_l^{(e)} K_l(\beta_e r) \right] e^{il\theta} \\ &= \mathcal{S}_0^{11}(\beta_e a_i) H_0^{(1)}(\beta_e a_i) + \mathcal{S}_0^{21}(\beta_e a_i) K_0(\beta_e a_i) \\ &\quad + \sum_{l=1}^{\infty} 2i^l [\mathcal{S}_l^{11}(\beta_e a_i) H_l^{(1)}(\beta_e a_i) \\ &\quad + \mathcal{S}_l^{21}(\beta_e a_i) K_l(\beta_e a_i)] \cos(l\theta). \end{aligned} \tag{8.14}$$

Here we used $A_l^{(e)} = i^l, B_l^{(e)} \equiv 0$ which follow directly from (8.13). Since the components of the scattering matrix are known explicitly

from (8.12), the scattered field is completely defined by the formula (8.14). To derive this well-known expression (see [Norris and Vemula (1995)]), which is very useful for confirming the efficiency of our cloaking, we applied the following symmetry relations

$$\begin{aligned} \mathcal{S}^{11}_{-l}(\beta_e a_i) &= \mathcal{S}^{11}_{l}(\beta_e a_i), \quad \mathcal{S}^{12}_{-l}(\beta_e a_i) = (-1)^l \mathcal{S}^{12}_{l}(\beta_e a_i), \\ \mathcal{S}^{21}_{-l}(\beta_e a_i) &= (-1)^l \mathcal{S}^{21}_{l}(\beta_e a_i), \quad \mathcal{S}^{22}_{-l}(\beta_e a_i) = \mathcal{S}^{22}_{l}(\beta_e a_i), \end{aligned} \tag{8.15}$$

which hold for any $l \neq 0$. We note that, for small $\beta_e r$, the first two terms of the expression (8.14) have the form

$$\begin{aligned} &- \left[H_0^{(1)}(\beta_e r) + \frac{2i}{\pi} K_0(\beta_e r) \right] \\ &\quad - \frac{2i\pi \cos\theta}{4i\gamma + \pi + 4i \log(\beta_e a_i/2)} \left[H_1^{(1)}(\beta_e r) + \frac{2i}{\pi} K_1(\beta_e r) \right], \end{aligned} \tag{8.16}$$

where the first group of terms in square brackets correspond to the normalised Green's function for the biharmonic operator and the second group is the dipole contribution.

Second, we focus our attention on the case when the scattered field is due to a point source located sufficiently far away from the scatterer. Without loss of generality, we choose this location to be $\mathbf{x}_{(inc)} = (-s, 0)$ on the x_1-axis. In this case, the incident wave is represented by

$$\begin{aligned} w^{(c)}_{(inc)}(r,\theta) &= \frac{1}{8i\beta_e^2} \left[H_0^{(1)}(\beta_e \sqrt{(r\cos\theta + s)^2 + (r\sin\theta)^2}) \right. \\ &\quad \left. + \frac{2i}{\pi} K_0(\beta_e \sqrt{(r\cos\theta + s)^2 + (r\sin\theta)^2}) \right] \\ &= \frac{1}{8i\beta_e^2} \sum_{l=-\infty}^{\infty} (-1)^l \\ &\quad \times \left[H_l^{(1)}(\beta_e s) J_l(\beta_e r) + \frac{2i}{\pi} K_l(\beta_e s) I_l(\beta_e r) \right] e^{il\theta}. \end{aligned} \tag{8.17}$$

It immediately follows that the multipole coefficients $A_l^{(e)}, B_l^{(e)}$ have the form

$$A_l^{(e)} = \frac{(-1)^l}{8i\beta_e^2} H_l^{(1)}(\beta_e s), \quad B_l^{(e)} = \frac{(-1)^l}{4\pi\beta_e^2} K_l(\beta_e s). \tag{8.18}$$

The corresponding scattered wave coefficients $E_l^{(e)}$ and $F_l^{(e)}$ can be found using (8.11) which would lead to a similar formula to (8.14) for scattered field $w_{(sc)}^{(c)}(r,\theta)$.

8.3.3. *Scattering from a circular clamped inclusion with control sources present — incident plane wave*

Figures 8.4 and 8.5 depict cloaking of a circular clamped inclusion with six active control sources. For the case of Fig. 8.4, the wave field incident on the clamped inclusion consists of a plane wave as well as the contributions from the active sources as shown in Fig. 8.15. The active sources are characterised by the Green's function for the biharmonic operator and their representations vary according to their locations $\mathbf{x}^{(s,n)} = (x_1^{(s,n)}, x_2^{(s,n)})$, $n = 1, \ldots, 6$ (with the appropriate modification of formula (8.17)). Distributing these six active sources evenly, starting from the x_1-axis, on a circle of radius a_s and labelling them as $Q_n, n = 1 \ldots 6$, we can write the l-th order coefficients for

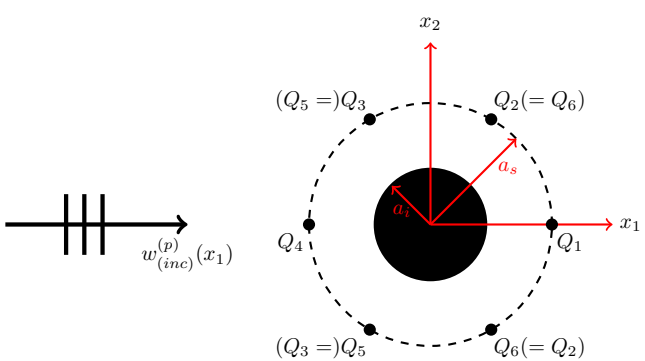


Fig. 8.15. A circular clamped inclusion surrounded by six active control sources. $w_{(inc)}^{(p)}(x_1)$ represents the incoming flexural plane wave.

the total wave incident on the clamped inclusion as

$$\begin{aligned} A_l^{(e)} &= i^l + \frac{H_l^{(1)}(\beta_e a_s)}{8i\beta_e^2} \sum_{n=1}^{6} Q_n \exp\left(\frac{(n-1)}{3} il\pi\right), \\ B_l^{(e)} &= \frac{K_l^{(1)}(\beta_e a_s)}{4\pi\beta_e^2} \sum_{n=1}^{6} Q_n \exp\left(\frac{(n-1)}{3} il\pi\right). \end{aligned} \tag{8.19}$$

We remark that, to maintain an up-down symmetry, we set $Q_2 = Q_6$ and $Q_3 = Q_5$.

Now, using (8.11), we obtain the l-th order coefficient for the $H_l^{(1)}(\beta_e r)$ term in the field for $r > a_s$ as

$$\mathcal{S}_l^{11} A_l^{(e)} + \mathcal{S}_l^{12} B_l^{(e)} + \frac{J_l(\beta_e a_s)}{8i\beta_e^2} \sum_{n=1}^{6} Q_n \exp\left(\frac{(n-1)}{3} il\pi\right). \tag{8.20}$$

Thus, we can derive four equations for the unknowns $Q_1, Q_2 = Q_6, Q_3 = Q_5$ and Q_4, by substituting the representations for $A_l^{(e)}$ and $B_l^{(e)}$ from (8.19) into the expression (8.20) and equating them to zero:

$$\begin{aligned} &\mathcal{S}_l^{11} \left[i^l + \frac{H_l^{(1)}(\beta_e a_s)}{8i\beta_e^2} \sum_{n=1}^{6} Q_n \exp\left(\frac{(n-1)}{3} il\pi\right) \right] \\ &\quad + \mathcal{S}_l^{12} \left[\frac{K_l^{(1)}(\beta_e a_s)}{4\pi\beta_e^2} \sum_{n=1}^{6} Q_n \exp\left(\frac{(n-1)}{3} il\pi\right) \right] \\ &\quad + \frac{J_l(\beta_e a_s)}{8i\beta_e^2} \sum_{n=1}^{6} Q_n \exp\left(\frac{(n-1)}{3} il\pi\right) = 0, \end{aligned} \tag{8.21}$$

$l = 0, 1, 2, 3$ (the entries of the scattering matrix are as given in (8.12)). Once the complex wave amplitudes Q_n, $n = 1 \ldots 6$ are found, the representation for the total flexural wave amplitude outside the

scatterer is given by

$$w(r,\theta) \approx w_{(inc)}^{(p)}(r,\theta) + \sum_{n=1}^{6} Q_n G(r\cos\theta + x_1^{(s,n)}, r\sin\theta + x_2^{(s,n)})$$
$$\sum_{l=-L}^{L} \left[E_l^{(e)} H_l^{(1)}(\beta_e r) + F_l^{(e)} K_l(\beta_e r) \right] e^{il\theta} \tag{8.22}$$

where the summation limit L must be chosen sufficiently large to ensure wave amplitude accuracy. The expression (8.22) of course gives a complex value and its real part represents the wave amplitude at a particular instant of time. In Fig. 8.4 we present two snapshots of $\Re[\exp(-i\Psi)w(r,\theta)]$ for a particular value of the phase $\Psi = \pi$, with no active control sources present in Fig. 8.4(a) and with six control sources in Fig. 8.4(b). The radius of the clamped inclusion is set to be $a_i = 1.0$ and the control sources are located on a circle of radius $a_s = 2.5$. We take the non-dimensional parameter $\beta_e a_i$ to be equal to 0.3. It turns out that the summation limit could be set as $L = 2$ for required accuracy. Solution of the algebraic system (8.21) gives $Q_1 = -0.848 - 1.146\,i$, $Q_2 = Q_6 = -1.114 - 0.582\,i$, $Q_3 = Q_5 = -1.114 + 0.582\,i$ and $Q_4 = -0.848 + 1.146\,i$.

8.3.4. *Scattering from a circular clamped inclusion with control sources present — incident wave generated by a point source*

Figure 8.5 was obtained by replacing the incident plane wave by a remote point source as shown in Fig. 8.16. The Green's function $\mathcal{G}(\mathbf{x};\mathbf{x}')$ in an infinite plate containing a circular clamped inclusion satisfies

$$(\Delta^2 - \beta^4)\mathcal{G}(\mathbf{x};\mathbf{x}') + \delta(\mathbf{x}-\mathbf{x}') = 0, \quad \mathbf{x} = (r,\theta) \in \Omega,$$
$$\mathcal{G}(\mathbf{x};\mathbf{x}') = \frac{\partial \mathcal{G}(\mathbf{x};\mathbf{x}')}{\partial r} = 0, \quad |\mathbf{x}| = a_i, \tag{8.23}$$

where $\delta(\mathbf{x}-\mathbf{x}')$ denotes the Dirac delta function centred at $\mathbf{x}'$. We recall that the point source generating the incoming wave is located

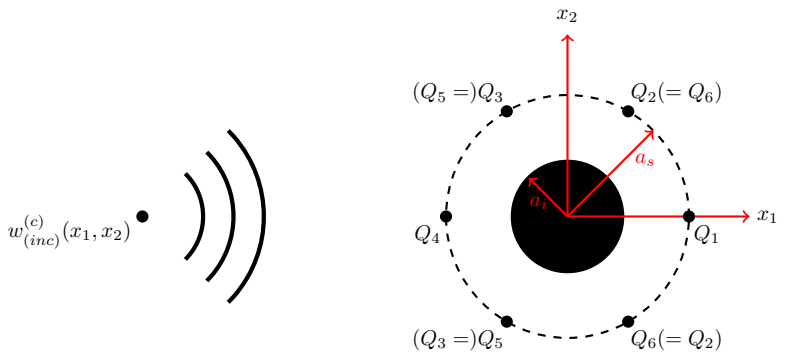


Fig. 8.16. A circular clamped inclusion surrounded by six active control sources. $w^{(c)}_{(inc)}(x_1, x_2)$ represents the incoming flexural wave generated by a remote point source.

at $\mathbf{x}_{(inc)} = (-s, 0)$. This point source, like the active control sources surrounding the clamped inclusion, is described by the Green's function for the biharmonic operator (see (8.17)).

Because $\mathcal{G}$ satisfies the boundary conditions imposed on the inclusion, for six active sources and an incident cylindrical wave, we can write

$$g(\mathbf{x}; \mathbf{x}_{(inc)}) = \mathcal{G}(\mathbf{x}; \mathbf{x}_{(inc)}) + \sum_{n=1}^{6} \mathcal{B}_n \mathcal{G}(\mathbf{x}; \mathbf{x}^{(s,n)}), \tag{8.24}$$

for the representation of the cloaking Green's function $g(\mathbf{x}; \mathbf{x}_{(inc)})$ everywhere outside the scatterer, where $\mathbf{x}^{(s,n)}$ denotes the position of the n-th active source.

Hence, we can write the l-th order coefficients for the total wave incident on the clamped inclusion as

$$\begin{aligned} A_l^{(e)} &= \frac{(-1)^l}{8i\beta_e^2} H_l^{(1)}(\beta_e s) + \frac{H_l^{(1)}(\beta_e a_s)}{8i\beta_e^2} \sum_{n=1}^{6} Q_n \exp\left(\frac{(n-1)}{3} il\pi\right), \\ B_l^{(e)} &= \frac{(-1)^l}{4\pi\beta_e^2} K_l(\beta_e s) + \frac{K_l^{(1)}(\beta_e a_s)}{4\pi\beta_e^2} \sum_{n=1}^{6} Q_n \exp\left(\frac{(n-1)}{3} il\pi\right). \end{aligned} \tag{8.25}$$

This formula can be obtained directly from (8.19) by replacing the contribution of the incident plane wave ($A_l^{(e)} = i^l, B_l^{(e)} \equiv 0$) with the contribution of the the cylindrical wave given by (8.18). Now, substituting formulae (8.25) for the incoming wave coefficients into (8.20), we can derive the equations for the unknowns $Q_1, Q_2 = Q_6, Q_3 = Q_5$ and Q_4 as

$$\mathcal{S}_l^{11}\left[\frac{(-1)^l}{8i\beta_e^2}H_l^{(1)}(\beta_e s) + \frac{H_l^{(1)}(\beta_e a_s)}{8i\beta_e^2}\sum_{n=1}^{6} Q_n \exp\left(\frac{(n-1)}{3} il\pi\right)\right]$$
$$+ \mathcal{S}_l^{12}\left[\frac{(-1)^l}{4\pi\beta_e^2}K_l(\beta_e s) + \frac{K_l^{(1)}(\beta_e a_s)}{4\pi\beta_e^2}\sum_{n=1}^{6} Q_n \exp\left(\frac{(n-1)}{3} il\pi\right)\right]$$
$$+ \frac{J_l(\beta_e a_s)}{8i\beta_e^2}\sum_{n=1}^{6} Q_n \exp\left(\frac{(n-1)}{3} il\pi\right) = 0, \tag{8.26}$$

$l = 0, 1, 2, 3$. As in the case for the incident plane wave, the representation for the total flexural wave amplitude outside the scatterer is given by

$$w(r,\theta) \approx w_{(inc)}^{(c)}(r,\theta) + \sum_{n=1}^{6} Q_n G(r\cos\theta + x_1^{(s,n)}, r\sin\theta + x_2^{(s,n)})$$
$$\sum_{l=-L}^{L}\left[E_l^{(e)}H_l^{(1)}(\beta_e r) + F_l^{(e)}K_l(\beta_e r)\right]e^{il\theta}, \tag{8.27}$$

once the complex wave amplitudes Q_n, $n = 1, \ldots, 6$ have been found.

Again, we present two snapshots of $\Re[\exp(-i\Psi)w(r,\theta)]$ for the phase $\Psi = \pi$, one with no active control sources present in Fig. 8.5(a) and one with six control sources in Fig. 8.5(b). Geometric and physical parameters associated with the illustrations have been kept the same as those for the case of an incident plane wave. Solving Eqs. (8.26) we obtain $Q_1 = -0.253 + 0.386\,i$, $Q_2 = Q_6 = 0.224 + 0.275\,i$, $Q_3 = Q_5 = 0.470 - 0.119\,i$ and $Q_4 = 0.282 - 0.292\,i$.

8.3.5. *A clamped cavity of arbitrary shape with a smooth boundary*

The algorithm of the previous sections can be generalised to cloaking an arbitrarily shaped scatterer (cavity) with a smooth boundary. Once again the method relies on annulling selected terms of the total scattered field everywhere outside the scatterer via the appropriate tuning of active source amplitudes. However, we now obtain the scattered field representations via far field approximations to two model problems: scattering of the incident field from the object and point source scattering from each of the control source locations. For the first model problem, the total wave amplitude $w(\mathbf{x})$ satisfies the following boundary value problem

$$\Delta^2 w(\mathbf{x}) - \beta^4 w(\mathbf{x}) = 0, \quad \mathbf{x} \in \mathbb{R}^2 \backslash \bar{\mathcal{D}},$$

$$w(\mathbf{x}) = \frac{\partial w(\mathbf{x})}{\partial n} = 0, \quad \mathbf{x} \in \partial\mathcal{D},$$

where $\mathcal{D}$ is the region occupied by the arbitrarily shaped scatterer. The incident wave may be a plane or a cylindrical wave (see Fig. 8.17). The field $w(\mathbf{x})$ can be written as the sum of the incident field $w_{(inc)}(\mathbf{x})$ and the scattered field $w_{(sc)}(\mathbf{x})$ as

$$w(\mathbf{x}) = w_{(inc)}(\mathbf{x}) + w_{(sc)}(\mathbf{x}).$$

If the incident field is a plane wave then we refer to (8.13) and if it is a cylindrical wave then we refer to (8.17).

The scattered field $w_{(sc)}$, either resulting from the incident plane or cylindrical wave has the asymptotic representation

$$w_{(sc)} = \sum_{l=-\infty}^{\infty} \left[E_l H_l^{(1)}(\beta_e r) + F_l K_l(\beta_e r) \right] e^{il\theta} \sim \sum_{l=-\infty}^{\infty} E_l H_l^{(1)}(\beta_e r) e^{il\theta}, \tag{8.28}$$

where E_l are constant coefficients. We note that the evanescent term in the scattered field is neglected since $K_l(\beta_e r) = O(\exp(-\beta_e r)/(\beta_e r))$ as $\beta_e r \gg 1$, (see formula 9.7.2 in [Abramowitz and Stegun (1965)]).

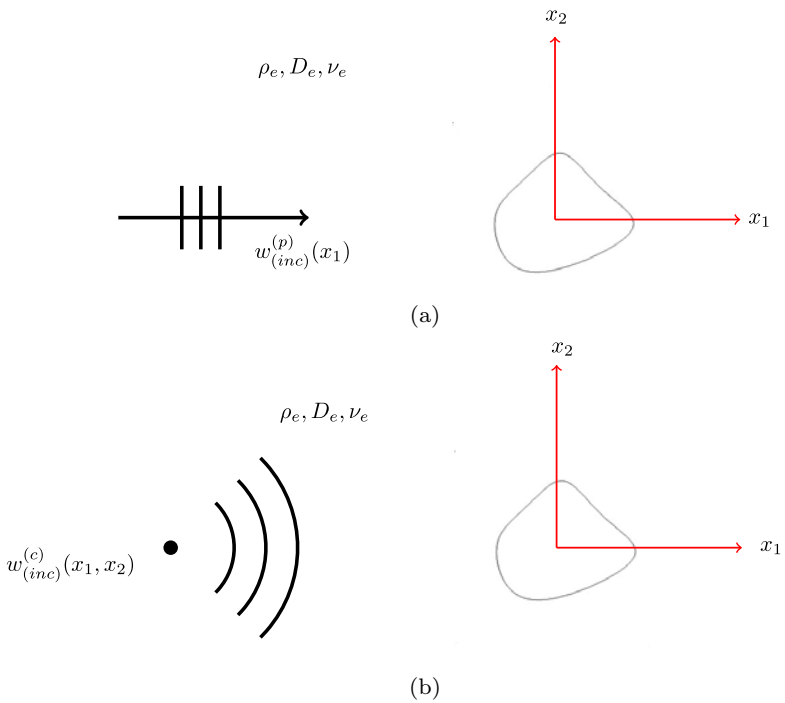


Fig. 8.17. Scattering from an arbitrarily shaped cavity: (a) incident wave is a plane wave; (b) incident wave is a cylindrical wave generated by a remote point source.

The second model problem involves scattering of a cylindrical wave from the arbitrarily shaped scatterer (cavity) (see Fig. 8.18). We do this in order to find the contribution from the active sources to the total scattered wave everywhere outside the scatterer. The total amplitude resulting from the cylindrical wave perturbation emitted from source n is denoted by $w^{(s,n)}(\mathbf{x})$ and satisfies the boundary value problem

$$\Delta^2 w^{(s,n)}(\mathbf{x}) - \beta^4 w^{(s,n)}(\mathbf{x}) + \delta(\mathbf{x} - \mathbf{x}^{(s,n)}) = 0, \quad \mathbf{x} \in \mathbb{R}^2 \backslash \bar{\mathcal{D}},$$

$$w^{(s,n)}(\mathbf{x}) = \frac{\partial w^{(s,n)}(\mathbf{x})}{\partial n} = 0, \quad \mathbf{x} \in \partial\mathcal{D},$$

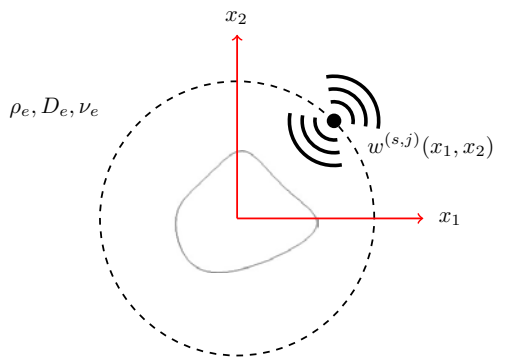


Fig. 8.18. Scattering from a control source.

where $\delta(\mathbf{x}-\mathbf{x}^{(s,n)})$ denotes the Dirac delta function, centred at $\mathbf{x}^{(s,n)}$. In fact, $w^{(s,n)}(\mathbf{x})$ admits the solution

$$w^{(s,n)}(\mathbf{x}) = w^{(s,n)}_{(inc)}(\mathbf{x}) + w^{(s,n)}_{(sc)}(\mathbf{x}) = G(\mathbf{x}-\mathbf{x}^{(s,n)}) + w^{(s,n)}_{(sc)}(\mathbf{x}) \\ \sim \sum_{l=-\infty}^{\infty} A_l^{(s,n)} H_l^{(1)}(\beta_e r) e^{il\theta}, \quad n = 1, \ldots, N, \tag{8.29}$$

where once again, exponentially small terms satisfying the modified Helmholtz equation are not shown. Here $G(\mathbf{x}-\mathbf{x}^{(s,n)})$ denotes the Green's function for the biharmonic operator, $w^{(s,n)}_{(sc)}(\mathbf{x})$ is the scattered field due to the unit source at $\mathbf{x}^{(s,n)}$, and $A_l^{(s,n)}$ are constant coefficients.

Assuming that the coefficients E_l and $A_l^{(s,n)}$ are given, active cloaking is achieved by introducing a set of N control sources with complex amplitudes Q_n placed at the points $\mathbf{x}^{(s,n)}$ in the exterior of the scatterer $\mathcal{D}$ (see Fig. 8.1 for a configuration with seven active sources). After the truncation to order L in the expansions (8.28), (8.29), we choose $N = 2L+1$, so that the total amplitude $w_{(total)}(\mathbf{x})$ is approximately equal to the incident field $w_{(inc)}(\mathbf{x})$,

that is

$$w_{(total)}(\mathbf{x}) = w(\mathbf{x}) + \sum_{n=1}^{2L+1} Q_n w^{(s,n)}(\mathbf{x}) \approx w_{(inc)}(\mathbf{x}). \tag{8.30}$$

To find the active source amplitudes Q_n, we substitute (8.28) and (8.29) into (8.30), and obtain the following system of linear algebraic equations

$$E_l + \sum_{n=1}^{2L+1} Q_n A_l^{(s,n)} = 0, \quad l = -L, \ldots, L. \tag{8.31}$$

The evaluation of Q_n relies on the solution of the model problems discussed above. For sufficiently large fixed $\beta_e r$, the asymptotic representations (8.28) and (8.29) are in fact Fourier series of which we only need to evaluate the first $N = 2L + 1$ coefficients. This can be done, for example, numerically using a FEM package such as COMSOL. Once the necessary coefficients E_l and $A_l^{(s,n)}$ for $l = -L, \ldots, L$ are known, we solve the system (8.31) for the amplitudes Q_n of the control sources. Note here, that the number of sources determines the value of L in Eq. (8.31).

The unperturbed plane wave propagating horizontally and the plane wave interacting with the arbitrarily shaped clamped scatterer are shown in Fig. 8.2(a) and Fig. 8.2(b), respectively. A shadow region is clearly visible behind the scatterer in Fig. 8.2(b). Once again, we note that perfectly matched layers are used on the exterior boundary of the computational domain to provide non-reflective boundary conditions.

To set up an active cloak, seven control sources are positioned around the scatterer, as shown in Fig. 8.1. This means that we have seven Fourier coefficients, which is sufficient to successfully approximate series (8.28) and (8.29).

The resulting computation in Fig. 8.2(c) indicates that the unperturbed plane wave has emerged behind the scatterer, so the cloaking by active sources successfully eliminates the shadow region as required.

We obtained similar results for an incoming cylindrical wave which are illustrated in Figs. 8.3(a), (b) and (c).

8.3.6. *Multipole representation for a plane wave scattered by a coated circular inclusion*

Now, we focus our attention towards a coated inclusion, as shown in Fig. 8.14(b). The general solution outside the position of the control sources is given in (8.9) with appropriate indices i, c or e, depending on the respective region. Using the appropriate representations, Eqs. (8.6)–(8.8) give us two matrix relations

$$\mathcal{A}^{(ci)} \begin{pmatrix} A_l^{(c)} \\ E_l^{(c)} \\ B_l^{(c)} \\ F_l^{(c)} \end{pmatrix} = \mathcal{B}^{(ci)} \begin{pmatrix} A_l^{(i)} \\ B_l^{(i)} \end{pmatrix} \quad \text{for} \quad r = a_i, \tag{8.32}$$

where $\mathcal{A}^{(ci)}$ and $\mathcal{B}^{(ci)}$ are 4×4 and 4×2 matrices, respectively; and

$$\mathcal{A}^{(ec)} \begin{pmatrix} A_l^{(e)} \\ E_l^{(e)} \\ B_l^{(e)} \\ F_l^{(e)} \end{pmatrix} = \mathcal{B}^{(ec)} \begin{pmatrix} A_l^{(c)} \\ E_l^{(c)} \\ B_l^{(c)} \\ F_l^{(c)} \end{pmatrix} \quad \text{for} \quad r = a_c, \tag{8.33}$$

where $\mathcal{A}^{(ec)}$ and $\mathcal{B}^{(ec)}$ are both 4×4 matrices. The representations of these four matrices are given in Appendix A.1. At this stage, it is important to emphasise that reducing the matrices $\mathcal{A}^{(ci)}$ and $\mathcal{A}^{(ec)}$ to block-diagonal structure is essential to pursue the derivation of an analytical formula for the scattering matrix $\mathcal{S}_l$. We note that in what follows $\mathcal{A}^{(ci,*)}, \mathcal{A}^{(ec,*)}$ denote the required matrices in block-diagonal form and $\mathcal{B}^{(ci,*)}, \mathcal{B}^{(ec,*)}$ are the associated matrices, with all these matrices being given in Appendix A.1.

We use the four matrices $\mathcal{A}^{(ci,*)}$, $\mathcal{B}^{(ci,*)}$, $\mathcal{A}^{(ec,*)}$ and $\mathcal{B}^{(ec,*)}$ to construct two further matrices $\mathcal{C}^{(ci,*)}$ and $\mathcal{C}^{(ec,*)}$ as follows:

$$\mathcal{C}^{(ci,*)} = {\mathcal{A}^{(ci,*)}}^{-1} \mathcal{B}^{(ci,*)}, \quad \mathcal{C}^{(ec,*)} = {\mathcal{A}^{(ec,*)}}^{-1} \mathcal{B}^{(ec,*)}, \tag{8.34}$$

where the expressions which would represent the results of the matrix inversions in the previous equation are too cumbersome to be given explicitly. The final step in the construction is to cascade via a matrix product the progression from the interface between the inclusion and the coating to the interface between the coating and the exterior, yielding the explicit transfer matrix $\mathcal{C}^{(ei,*)}$, which incorporates the interface conditions at both interfaces:

$$\mathcal{C}^{(ei,*)} = \mathcal{C}^{(ec,*)}\mathcal{C}^{(ci,*)}. \tag{8.35}$$

The transfer matrix $\mathcal{C}^{(ei,*)}$ can then be used to obtain the wave coefficients $A_l^{(e)}$, $B_l^{(e)}$, $E_l^{(e)}$ and $F_l^{(e)}$ in the exterior region in terms of the wave coefficients $A_l^{(i)}$, $B_l^{(i)}$ inside the inclusion as:

$$\begin{pmatrix} A_l^{(e)} \\ E_l^{(e)} \\ B_l^{(e)} \\ F_l^{(e)} \end{pmatrix} = \mathcal{C}^{(ei,*)} \begin{pmatrix} A_l^{(i)} \\ B_l^{(i)} \end{pmatrix}. \tag{8.36}$$

The final step in this procedure is to eliminate the wave coefficients in the inclusion to determine the 2×2 scattering matrix $\mathcal{S}_l$ which relates outgoing wave coefficients to incoming wave coefficients in the exterior region. This was done for a circular clamped inclusion in Sec. 8.3.2. Even though the matrix relation (8.11) remains the same, the new scattering matrix $\mathcal{S}_l$ is much more complicated for a coated inclusion. The details are given in Appendix A.2. We note that the discussion in Sec. 8.3.2 is valid, though we replace the old scattering matrix by this newly established one.

8.3.7. *Active cloaking of a coated inclusion*

As we briefly discussed in Sec. 8.2.2, the multipole coefficients associated with the propagating part of the scattered field in the multipole representation may vary rapidly with frequency in the case of resonant vibrations. We first consider an uncoated inclusion to identify resonant regimes where our active cloaking method requires

rapid adjustments. Figure 8.6 (see the dashed/green curve in both (a) and (b)) illustrates exactly this for a high contrast inclusion. The following geometric and material properties are used to obtain the dashed/green curve depicting the absolute value of $E_0^{(e)}$ as a function of the frequency ω: the radius of the inclusion $a_i = 0.50$; the densities of the inclusion and exterior $\rho_i = 0.05, \rho_e = 1.0$; the flexural rigidities $D_i = 2.5 \times 10^{-4}, D_e = 1.0$; the Poisson's ratios $\nu_i = \nu_e = 0.3$. As we mentioned earlier, our focus is on the second peak of the monopole term for which the rapid variation is more prominent. Hence we direct the reader's attention to Fig. 8.6(b). This resonant region is highlighted by a red box in Fig. 8.6(b): within this red box, $E_0^{(e)}$ has a minimum at the value of $\omega = 11.15$ (denoted by †) which is followed by a peak at $\omega = 11.30$ (denoted by ‡). This interval of steep gradient causes dramatic changes in both the spatial distribution around the inclusion and in its amplitude variation, as shown in Fig. 8.7(a) for $\omega = 11.15$ and in Fig. 8.7(b) for $\omega = 11.30$.

To tame the large variation in this resonant regime, we use a specially designed coating, whose effect is illustrated in Fig. 8.6(b) by the solid/blue curve. It is clear from the figure that the absolute value of the monopole coefficient $E_0^{(e)}$ has a flat profile within the targeted frequency range. The properties of the coating are chosen to be as follows: the radius $a_c = 0.77$; the density $\rho_c = 0.005$; the flexural rigidity $D_c = 2.5 \times 10^{-2}$; the Poisson's ratio $\nu_c = 0.3$. Comparing Figs. 8.7(a) and 8.8, we see that for $\omega = 11.15$, the designed coating enhances scattering. The active cloaking algorithm described in detail in Sec. 8.3.3 for a clamped inclusion is also suitable for a coated inclusion, even with enhanced scattering within the resonant interval.

Figure 8.9 shows that at least 10 control sources are required for the cloaking algorithm to produce effective cloaking at the frequency $\omega = 11.15$. This is because for $l \geq 5$, the contributions to the propagating part of the scattered field, $|E_l^{(e)} H_l^{(1)}(\beta_e a_s)|$, are sufficiently small. However, in order to see the near perfect cloaking

(as presented in Fig. 8.10) we use twelve sources. The material and geometric properties of the coated inclusion are as before with the frequency of the incident plane wave being $\omega = 11.15$.

We obtain similar results for an incoming cylindrical wave which are illustrated in Figs. 8.11–8.13. These results were obtained for the same coated inclusion and the same frequency value within the resonant regime discussed above.

8.3.8. *A note on controlling the quasi-static regimes*

It is a well-known result that scattering reduction can be achieved by the addition of a coating whose density is chosen to annul the leading-order behaviour of the monopole coefficient of the scattered field. It is also well-known that for a Kirchhoff plate containing a circular elastic inclusion the monopole and quadrupole terms are of the same order $O((\beta_e a_i)^2)$ when $\beta_e a_i \to 0$:

$$\begin{aligned} |E_0^{(e)}| &\sim \frac{\pi}{4}\left[\frac{\rho_i}{2\rho_e} + \frac{D_e}{D_e(1-\nu_e) + D_i(1+\nu_i)} - 1\right](\beta_e a_i)^2, \\ |E_2^{(e)}| &\sim \frac{\pi}{4}\left[\frac{D_e(1-\nu_e) - D_i(1-\nu_i)}{2(D_e(3+\nu_e) + D_i(1-\nu_i))}\right](\beta_e a_i)^2. \end{aligned} \tag{8.37}$$

We note that when the shear moduli of the inclusion and of the exterior matrix are the same, the quadrupole coefficient vanishes, and the monopole term becomes dominant.

The notion of neutrality for coated inclusions is well developed in mechanics of static composites (see, for example, [Milton (2002)]). In quasi-static regimes, as $\beta_e a_i \to 0$, a finite coating may be introduced around an inclusion, so that $E_0^{(e)}$ and $E_2^{(e)}$ vanish. An elementary example when $E_2^{(e)}$ is zero corresponds to the case when the shear moduli of the inclusion, coating and exterior matrix are all equal. For a general case of a coated inclusion, closed form asymptotic formulae for $E_0^{(e)}$ and $E_2^{(e)}$ are cumbersome. For an illustration, we note that the monopole coefficient $E_0^{(e)}$ is asymptotically $o((\beta_e a_i)^2)$ if the following relation for the parameters of the coating is

satisfied:

$$\begin{aligned}&[2D_e\mathfrak{a}_1(\mathfrak{b}_1+2D_c-2D_e)\beta_e^4+\mathfrak{a}_1\mathfrak{b}_1\mathfrak{b}_2\beta_c^2\beta_e^2-2D_c\mathfrak{a}_1\mathfrak{b}_1\beta_c^4]\,a_c^4\\&\quad+2\mathfrak{a}_2\mathfrak{b}_2[D_c\beta_c^4-D_i\beta_i^4]\,a_i^4\\&\quad+[4D_e\mathfrak{a}_2(\mathfrak{b}_2+D_e)\beta_e^4+\mathfrak{a}_2\mathfrak{b}_2^2\beta_c^2\beta_e^2-2D_i\mathfrak{a}_1\mathfrak{b}_1\beta_i^4\\&\quad+2D_c(\mathfrak{a}_1\mathfrak{b}_1-\mathfrak{a}_2\mathfrak{b}_2)\beta_c^4]\,a_i^2a_c^2=0,\end{aligned}\tag{8.38}$$

where

$$\begin{aligned}&\mathfrak{a}_1=D_i(1+\nu_i)+D_c(1-\nu_c),\quad \mathfrak{a}_2=D_i(1+\nu_i)-D_c(1+\nu_c),\\&\mathfrak{b}_1=D_c(1+\nu_c)+D_e(1-\nu_e),\quad \mathfrak{b}_2=D_c(1-\nu_c)-D_e(1-\nu_e).\end{aligned}$$

Here, D,ν denote the flexural rigidity and the Poisson's ratio for the inclusion, the coating and the exterior with appropriate subscripts. By assuming identical material parameter values (for example of the coating and exterior matrix), the above formula reduces to the well-known result derived by Konenkov [Konenkov (1964)] and Torrent [Torrent *et al.* (2014)] (see (8.37)) for the case of an uncoated inclusion.

It is also important to note that in the very special case of $D_i = D_c = D_e$, $\nu_i = \nu_c = \nu_e$ but $\rho_i \neq \rho_c \neq \rho_e$, Eq. (8.38) reduces to the classical mass-compensation equation

$$\rho_c = \frac{\rho_e - \rho_i(a_i/a_c)^2}{1-(a_i/a_c)^2}.\tag{8.39}$$

However, the above simple ideas of scattering reduction in the quasi-static regimes fail entirely for higher frequencies, where asymptotic formulae (8.37) are not valid. In such cases of advanced scattering, and possible resonances, the active cloaking, described previously is the way forward.

8.4. Concluding Remarks

We have presented a novel efficient asymptotic method to actively cloak clamped scatterers in the presence of incident plane and cylindrical flexural waves in Kirchhoff plates. If the geometry of the

scatterer is circular, then the results are given in closed analytical form using multipole expansions. The method is generic and works equally well for cloaking arbitrarily shaped scatterers with piecewise smooth boundaries. We note that cloaking clamped scatterers is a more challenging problem than cloaking scatterers with a free boundary. This is due to a finite monopole term for the former case, even in the low-frequency regime.

A natural extension is to cloak elastic inclusions from flexural waves. The most interesting case corresponds to a high contrast between the inclusion itself and the surrounding matrix leading to resonances for specific frequencies. To overcome difficulties associated with these resonant regimes, we have combined passive and active cloaking techniques by designing a coating for a circular elastic inclusion which controls rapid variation of multipole coefficients.

The results, as shown in a series of illustrative examples within the text, show a significant suppression of the shadow region behind the scatterer.

Acknowledgments

J. O'Neill would like to greatly acknowledge the support from the EPSRC through the grant EP/L50518/1. Ö. Selsil, R.C. McPhedran and N.V. Movchan acknowledge the financial support of the European Commission's Seventh Framework Programme under the contract number PIAPP-GA-284544-PARM-2. A.B. Movchan acknowledges the support from EPSRC Programme Grant EP/L024926/1.

Appendix

A.1. Kirchhoff Plates: The Representations for the Matrices $\boldsymbol{\mathcal{A}}^{(ci)}, \boldsymbol{\mathcal{B}}^{(ci)}$, $\boldsymbol{\mathcal{A}}^{(ec)}, \boldsymbol{\mathcal{B}}^{(ec)}$ and their Reduced Forms $\boldsymbol{\mathcal{A}}^{(ci,*)}, \boldsymbol{\mathcal{B}}^{(ci,*)}, \boldsymbol{\mathcal{A}}^{(ec,*)}, \boldsymbol{\mathcal{B}}^{(ec,*)}$

Here we give a detailed description of how to obtain the entries of the matrices $\mathcal{A}_{pq}^{(ci)}$, $p, q = 1, 2, 3, 4$; $\mathcal{B}_{pq}^{(ci)}$, $p = 1, 2, 3, 4; q = 1, 2$ and their

reduced forms $\mathcal{A}_{pq}^{(ci,*)}$, $p,q = 1,2,3,4$; $\mathcal{B}_{pq}^{(ci,*)}$, $p = 1,2,3,4; q = 1,2$ used in Section 8.3.6.

A.1.1. *Transmission conditions at the interior interface of the coating*

We first consider the interface between the inclusion and coating positioned at $r = a_i$. Substitution of the relevant field representations (8.9) into the two conditions (8.6) of perfect bonding (corresponding to the continuity of the displacement and its normal derivative across the interface), for each multipole order l, gives

$$A_l^{(c)} J_l(\beta_c a_i) + E_l^{(c)} H_l^{(1)}(\beta_c a_i) + B_l^{(c)} I_l(\beta_c a_i) + F_l^{(c)} K_l(\beta_c a_i)$$
$$= A_l^{(i)} J_l(\beta_i a_i) + B_l^{(i)} I_l(\beta_i a_i), \tag{A.1}$$

$$\beta_c a_i [A_l^{(c)} J_l{}'(\beta_c a_i) + E_l^{(c)} H_l^{(1)\prime}(\beta_c a_i)$$
$$+ B_l^{(c)} I_l{}'(\beta_c a_i) + F_l^{(c)} K_l{}'(\beta_c a_i)]$$
$$= \beta_i a_i [A_l^{(i)} J_l{}'(\beta_i a_i) + B_l^{(i)} I_l{}'(\beta_i a_i)], \tag{A.2}$$

where above, and in what follows, the prime denotes the derivative with respect to the entire argument. With the help of equations (A.1), (A.2), we can write condition (8.7), which corresponds to the continuity of the moment across the interface, as

$$\beta_c^2 a_i^2 [A_l^{(c)} J_l{}''(\beta_c a_i) + E_l^{(c)} H_l^{(1)\prime\prime}(\beta_c a_i) + B_l^{(c)} I_l{}''(\beta_c a_i) + F_l^{(c)} K_l{}''(\beta_c a_i)]$$
$$= \frac{D_i}{D_c} \beta_i^2 a_i^2 [A_l^{(i)} J_l{}''(\beta_i a_i) + B_l^{(i)} I_l{}''(\beta_i a_i)]$$
$$+ \left(\frac{D_i}{D_c} \nu_i - \nu_c \right) \beta_i a_i [A_l^{(i)} J_l{}'(\beta_i a_i) + B_l^{(i)} I_l{}'(\beta_i a_i)]$$
$$- \left(\frac{D_i}{D_c} \nu_i - \nu_c \right) l^2 [A_l^{(i)} J_l(\beta_i a_i) + B_l^{(i)} I_l(\beta_i a_i)]. \tag{A.3}$$

Next, we deal with condition (8.8), which enforces the continuity of the transverse force across the interface. We first note that

$$\frac{\partial}{\partial r}\Delta_{r\theta}w + \frac{1-\nu_k}{r^2}\frac{\partial^2}{\partial\theta^2}\left(\frac{\partial w}{\partial r} - \frac{w}{r}\right) = \frac{\partial^3 w}{\partial r^3} - \frac{1}{r^2}\frac{\partial w}{\partial r} + \frac{1}{r}\frac{\partial^2 w}{\partial r^2} + \frac{2-\nu_c}{r^2}\frac{\partial^3 w}{\partial r\partial\theta^2} - \frac{3-\nu_c}{r^3}\frac{\partial^2 w}{\partial\theta^2}. \tag{A.4}$$

Hence this condition can be written as

$$\begin{aligned}
&\beta_c^3 a_i^3[A_l^{(c)} J_l'''(\beta_c a_i) + E_l^{(c)} {H_l^{(1)}}'''(\beta_c a_i) + B_l^{(c)} I_l'''(\beta_c a_i) \\
&\quad + F_l^{(c)} K_l'''(\beta_c a_i)] \\
&\quad - \beta_c a_i[A_l^{(c)} J_l'(\beta_c a_i) + E_l^{(c)} {H_l^{(1)}}'(\beta_c a_i) + B_l^{(c)} I_l'(\beta_c a_i) \\
&\quad + F_l^{(c)} K_l'(\beta_c a_i)] \\
&\quad + \beta_c^2 a_i^2[A_l^{(c)} J_l''(\beta_c a_i) + E_l^{(c)} {H_l^{(1)}}''(\beta_c a_i) + B_l^{(c)} I_l''(\beta_c a_i) \\
&\quad + F_l^{(c)} K_l''(\beta_c a_i)] \\
&\quad - (2-\nu_c)\beta_c l^2 a_i[A_l^{(c)} J_l'(\beta_c a_i) + E_l^{(c)} {H_l^{(1)}}'(\beta_c a_i) + B_l^{(c)} I_l'(\beta_c a_i) \\
&\quad + F_l^{(c)} K_l'(\beta_c a_i)] \\
&\quad - (3-\nu_c) l^2[A_l^{(c)} J_l(\beta_c a_i) + E_l^{(c)} H_l^{(1)}(\beta_c a_i) + B_l^{(c)} I_l(\beta_c a_i) \\
&\quad + F_l^{(c)} K_l(\beta_c a_i)] \\
&= \frac{D_i}{D_c}\Big\{\beta_i^3 a_i^3[A_l^{(i)} J_l'''(\beta_i a_i) + B_l^{(i)} I_l'''(\beta_i a_i)] \\
&\quad - \beta_i a_i[A_l^{(i)} J_l'(\beta_i a_i) + B_l^{(i)} I_l'(\beta_i a_i)] \\
&\quad + \beta_i^2 a_i^2[A_l^{(i)} J_l''(\beta_i a_i) + B_l^{(i)} I_l''(\beta_i a_i)] \\
&\quad - (2-\nu_i)\beta_i l^2 a_i[A_l^{(i)} J_l'(\beta_i a_i) + B_l^{(i)} I_l'(\beta_i a_i)] \\
&\quad + (3-\nu_i) l^2[A_l^{(i)} J_l(\beta_i a_i) + B_l^{(i)} I_l(\beta_i a_i)]\Big\}.
\end{aligned} \tag{A.5}$$

Now, using (A.1), (A.2) and (A.3), equality (A.5) can be reduced to

$$\begin{aligned}
&\beta_c^3 a_i^3 [A_l^{(c)} J_l'''(\beta_c a_i) + E_l^{(c)} H_l^{(1)'''}(\beta_c a_i) + B_l^{(c)} I_l'''(\beta_c a_i) \\
&\qquad + F_l^{(c)} K_l'''(\beta_c a_i)] \\
&= \frac{D_i}{D_c} \beta_i^3 a_i^3 [A_l^{(i)} J_l'''(\beta_i a_i) + B_l^{(i)} I_l'''(\beta_i a_i)] \\
&\qquad - \left\{ \frac{D_i}{D_c} \nu_i - \nu_c + \frac{D_i}{D_c} \left[1 + (2 - \nu_i) l^2\right] - \left[1 + (2 + \nu_c) l^2\right] \right\} \beta_i a_i \\
&\qquad \times \left[A_l^{(i)} J_l'(\beta_i a_i) + B_l^{(i)} I_l'(\beta_i a_i) \right] \\
&\qquad + 3l^2 \left(\frac{D_i}{D_c} - 1 \right) \left[A_l^{(i)} J_l(\beta_i a_i) + B_l^{(i)} I_l(\beta_i a_i) \right]. \qquad \text{(A.6)}
\end{aligned}$$

A.1.2. *Reduction of the system of transmission conditions on $r = a_i$ to block diagonal form*

To deal with the second and third-order derivatives in (A.3) and (A.6) respectively, we consider the Bessel equations

$$z^2 \mathcal{W}_\pm''(z) + z \mathcal{W}_\pm'(z) \pm (z^2 \mp l^2) \mathcal{W}_\pm(z) = 0,$$

and their derivatives with respect to z, rearrange and obtain

$$\beta_c^2 a_i^2 \mathcal{W}_\pm''(\beta_c a_i) = -\beta_c a_i \mathcal{W}_\pm'(\beta_c a_i) \pm (\beta_c^2 a_i^2 \mp l^2) \mathcal{W}_\pm(\beta_c a_i), \qquad \text{(A.7)}$$

$$\begin{aligned}
\beta_c^3 a_i^3 \mathcal{W}_\pm'''(\beta_c a_i) = &\left(2 \mp \beta_c^2 a_i^2 + l^2\right) \beta_c a_i \mathcal{W}_\pm'(\beta_c a_i) \\
&\pm \left(\beta_c^2 a_i^2 \mp 3l^2\right) \mathcal{W}_\pm(\beta_c a_i), \qquad \text{(A.8)}
\end{aligned}$$

where we replaced z by $\beta_c a_i$. Here $\mathcal{W}_+$ corresponds to $J_l, H_l^{(1)}$ and $\mathcal{W}_-$ to I_l, K_l, respectively. Using (A.7) and (A.8), we re-write equality (A.3) as

$$\begin{aligned}
&A_l^{(c)} \left[-\beta_c a_i J_l'(\beta_c a_i) - \left(\beta_c^2 a_i^2 - l^2\right) J_l(\beta_c a_i) \right] \\
&\qquad + E_l^{(c)} \left[-\beta_c a_i H_l^{(1)'}(\beta_c a_i) - \left(\beta_c^2 a_i^2 - l^2\right) H_l^{(1)}(\beta_c a_i) \right]
\end{aligned}$$

$$
\begin{aligned}
&+B_l^{(c)}\left[-\beta_c a_i I_l{}'(\beta_c a_i)+\left(\beta_c^2 a_i^2+l^2\right) I_l(\beta_c a_i)\right]\\
&+F_l^{(c)}\left[-\beta_c a_i K_l{}'(\beta_c a_i)+\left(\beta_c^2 a_i^2+l^2\right) K_l(\beta_c a_i)\right]\\
&=A_l^{(i)}\Bigg\{-\left[\frac{D_i}{D_c}(1-\nu_i)+\nu_c\right]\beta_i a_i J_l{}'(\beta_i a_i)\\
&+\left[-\frac{D_i}{D_c}\beta_i^2 a_i^2+l^2\left[\frac{D_i}{D_c}(1-\nu_i)+\nu_c\right]\right] J_l(\beta_i a_i)\\
&+B_l^{(i)}\Bigg\{-\left[\frac{D_i}{D_c}(1-\nu_i)+\nu_c\right]\beta_i a_i I_l{}'(\beta_i a_i)\\
&+\left[\frac{D_i}{D_c}\beta_i^2 a_i^2+l^2\left[\frac{D_i}{D_c}(1-\nu_i)+\nu_c\right]\right] I_l(\beta_i a_i), \qquad \text{(A.9)}
\end{aligned}
$$

and equality (A.6) as

$$
\begin{aligned}
&A_l^{(c)}\left[\left(2-\beta_c^2 a_i^2+l^2\right)\beta_c a_i J_l{}'(\beta_c a_i)+\left(\beta_c^2 a_i^2-3l^2\right) J_l(\beta_c a_i)\right]\\
&\qquad +E_l^{(c)}[(2-\beta_c^2 a_i^2+l^2)\beta_c a_i H_l^{(1)'}(\beta_c a_i)\\
&\qquad +(\beta_c^2 a_i^2-3l^2)H_l^{(1)}(\beta_c a_i)]\\
&+B_l^{(c)}[(2+\beta_c^2 a_i^2+l^2)\beta_c a_i I_l{}'(\beta_c a_i)-(\beta_c^2 a_i^2+3l^2)I_l(\beta_c a_i)]\\
&+F_l^{(c)}[(2+\beta_c^2 a_i^2+l^2)\beta_c a_i K_l{}'(\beta_c a_i)-(\beta_c^2 a_i^2+3l^2)K_l(\beta_c a_i)]\\
&=A_l^{(i)}\Bigg\{\left[\tfrac{D_i}{D_c}\left[(1-\nu_i)(1-l^2)-\beta_i^2 a_i^2\right]+1+\nu_c+(2-\nu_c)l^2\right]\\
&\qquad \times\beta_i a_i J_l{}'(\beta_i a_i)+\left(\tfrac{D_i}{D_c}\beta_i^2 a_i^2-3l^2\right) J_l(\beta_i a_i)\Bigg\}\\
&+B_l^{(i)}\Bigg\{\left[\tfrac{D_i}{D_c}\left[(1-\nu_i)(1-l^2)+\beta_i^2 a_i^2\right]+1+\nu_c+(2-\nu_c)l^2\right]\\
&\qquad \times\beta_i a_i I_l{}'(\beta_i a_i)+\left(-\tfrac{D_i}{D_c}\beta_i^2 a_i^2-3l^2\right) I_l(\beta_i a_i)\Bigg\}. \qquad \text{(A.10)}
\end{aligned}
$$

Equations (A.1), (A.2), (A.9) and (A.10) finally lead us to the matrix equation (8.32) with their left-hand and right-hand sides forming the rows of $\mathcal{A}_{pq}^{(ci)}$ and $\mathcal{B}_{pq}^{(ci)}$ respectively. The components of the matrices

$\mathcal{A}_{pq}^{(ci)}$, $p,q=1,2,3,4$; $\mathcal{B}_{pq}^{(ci)}$, $p=1,2; q=1,2,3,4$ are given below:

$$\mathcal{A}_{11}^{(ci)} = J_l(\beta_c a_i), \quad \mathcal{A}_{12}^{(ci)} = H_l^{(1)}(\beta_c a_i),$$

$$\mathcal{A}_{13}^{(ci)} = I_l(\beta_c a_i), \quad \mathcal{A}_{14}^{(ci)} = K_l(\beta_c a_i),$$

$$\mathcal{A}_{21}^{(ci)} = \beta_c a_i J_l{}'(\beta_c a_i), \quad \mathcal{A}_{22}^{(ci)} = \beta_c a_i H_l^{(1)'}(\beta_c a_i),$$

$$\mathcal{A}_{23}^{(ci)} = \beta_c a_i I_l{}'(\beta_c a_i), \quad \mathcal{A}_{24}^{(ci)} = \beta_c a_i K_l{}'(\beta_c a_i),$$

$$\mathcal{A}_{31}^{(ci)} = -(\beta_c^2 a_i^2 - l^2) J_l(\beta_c a_i) - \beta_c a_i J_l{}'(\beta_c a_i),$$

$$\mathcal{A}_{32}^{(ci)} = -(\beta_c^2 a_i^2 - l^2) H_l^{(1)}(\beta_c a_i) - \beta_c a_i H_l^{(1)'}(\beta_c a_i),$$

$$\mathcal{A}_{33}^{(ci)} = (\beta_c^2 a_i^2 + l^2) I_l(\beta_c a_i) - \beta_c a_i I_l{}'(\beta_c a_i),$$

$$\mathcal{A}_{34}^{(ci)} = (\beta_c^2 a_i^2 + l^2) K_l(\beta_c a_i) - \beta_c a_i K_l{}'(\beta_c a_i),$$

$$\mathcal{A}_{41}^{(ci)} = (\beta_c^2 a_i^2 - 3l^2) J_l(\beta_c a_i) + (-\beta_c^2 a_i^2 + l^2 + 2)\beta_c a_i J_l{}'(\beta_c a_i),$$

$$\mathcal{A}_{42}^{(ci)} = (\beta_c^2 a_i^2 - 3l^2) H_l^{(1)}(\beta_c a_i) + (-\beta_c^2 a_i^2 + l^2 + 2)\beta_c a_i H_l^{(1)'}(\beta_c a_i),$$

$$\mathcal{A}_{43}^{(ci)} = -(\beta_c^2 a_i^2 + 3l^2) I_l(\beta_c a_i) + (\beta_c^2 a_i^2 + l^2 + 2)\beta_c a_i I_l{}'(\beta_c a_i),$$

$$\mathcal{A}_{44}^{(ci)} = -(\beta_c^2 a_i^2 + 3l^2) K_l(\beta_c a_i) + (\beta_c^2 a_i^2 + l^2 + 2)\beta_c a_i K_l{}'(\beta_c a_i);$$

$$\mathcal{B}_{11}^{(ci)} = J_l(\beta_i a_i), \quad \mathcal{B}_{12}^{(ci)} = I_l(\beta_i a_i),$$

$$\mathcal{B}_{21}^{(ci)} = \beta_i a_i J_l{}'(\beta_i a_i), \quad \mathcal{B}_{22}^{(ci)} = \beta_i a_i I_l{}'(\beta_i a_i),$$

$$\mathcal{B}_{31}^{(ci)} = \left\{ -\frac{D_i}{D_c}\beta_i^2 a_i^2 + l^2 \left[(1-\nu_i)\frac{D_i}{D_c} + \nu_c \right] \right\} J_l(\beta_i a_i)$$
$$- \left[(1-\nu_i)\frac{D_i}{D_c} + \nu_c \right] \beta_i a_i J_l{}'(\beta_i a_i),$$

$$\mathcal{B}_{32}^{(ci)} = \left\{ \frac{D_i}{D_c}\beta_i^2 a_i^2 + l^2 \left[(1-\nu_i)\frac{D_i}{D_c} + \nu_c \right] \right\} I_l(\beta_i a_i)$$
$$- \left[(1-\nu_i)\frac{D_i}{D_c} + \nu_c \right] \beta_i a_i I_l{}'(\beta_i a_i),$$

$$\mathcal{B}_{41}^{(ci)} = \left(\frac{D_i}{D_c}\beta_i^2 a_i^2 - 3l^2\right) J_l(\beta_i a_i) + \left\{\frac{D_i}{D_c}\left[(1-\nu_i)(1-l^2) - \beta_i^2 a_i^2\right] + 1 + \nu_c + (2-\nu_c)l^2\right\} \times \beta_i a_i J_l{}'(\beta_i a_i),$$

$$\mathcal{B}_{42}^{(ci)} = -\left(\frac{D_i}{D_c}\beta_i^2 a_i^2 + 3l^2\right) I_l(\beta_i a_i) + \left\{\frac{D_i}{D_c}\left[(1-\nu_i)(1-l^2) + \beta_i^2 a_i^2\right] + 1 + \nu_c + (2-\nu_c)l^2\right\} \times \beta_i a_i I_l{}'(\beta_i a_i).$$

The matrices $\mathcal{A}^{(ci,*)}, \mathcal{B}^{(ci,*)}$ are obtained by diagonalising the matrices (using row operations) $\mathcal{A}^{(ci)}, \mathcal{B}^{(ci)}$, respectively:

$$-(\beta_c^2 a_i^2 + l^2)R_1 + R_2 + R_3 \longrightarrow R_1^*,$$
$$2l^2 R_1 - (\beta_c^2 a_i^2 + l^2 + 1)R_2 + R_3 + R_4 \longrightarrow R_2^*,$$
$$(\beta_c^2 a_i^2 - l^2)R_1 + R_2 + R_3 \longrightarrow R_3^*,$$
$$2l^2 R_1 + (\beta_c^2 a_i^2 - l^2 - 1)R_2 + R_3 + R_4 \longrightarrow R_4^*,$$

$$\mathcal{A}_{11}^{(ci,*)} = -2\beta_c^2 a_i^2 J_l(\beta_c a_i), \quad \mathcal{A}_{12}^{(ci,*)} = -2\beta_c^2 a_i^2 H_l^{(1)}(\beta_c a_i),$$
$$\mathcal{A}_{13}^{(ci,*)} = \mathcal{A}_{14}^{(ci,*)} = 0, \quad \mathcal{A}_{21}^{(ci,*)} = -2\beta_c^3 a_i^3 J_l{}'(\beta_c a_i),$$
$$\mathcal{A}_{22}^{(ci,*)} = -2\beta_c^3 a_i^3 H_l^{(1)'}(\beta_c a_i), \quad \mathcal{A}_{23}^{(ci,*)} = \mathcal{A}_{24}^{(ci,*)} = 0,$$
$$\mathcal{A}_{31}^{(ci,*)} = \mathcal{A}_{32}^{(ci,*)} = 0, \quad \mathcal{A}_{33}^{(ci,*)} = 2\beta_c^2 a_i^2 I_l(\beta_c a_i),$$
$$\mathcal{A}_{34}^{(ci,*)} = 2\beta_c^2 a_i^2 K_l(\beta_c a_i), \quad \mathcal{A}_{41}^{(ci,*)} = \mathcal{A}_{42}^{(ci,*)} = 0,$$
$$\mathcal{A}_{43}^{(ci,*)} = 2\beta_c^3 a_i^3 I_l{}'(\beta_c a_i), \quad \mathcal{A}_{44}^{(ci,*)} = 2\beta_c^3 a_i^3 K_l{}'(\beta_c a_i).$$

$$\mathcal{B}_{1k}^{(ci,*)} = -(\beta_c^2 a_i^2 + l^2)\mathcal{B}_{1k}^{(ci)} + \mathcal{B}_{2k}^{(ci)} + \mathcal{B}_{3k}^{(ci)},$$
$$\mathcal{B}_{2k}^{(ci,*)} = 2l^2 \mathcal{B}_{1k}^{(ci)} - (\beta_c^2 a_i^2 + l^2 + 1)\mathcal{B}_{2k}^{(ci)} + \mathcal{B}_{3k}^{(ci)} + \mathcal{B}_{4k}^{(ci)},$$

$$\mathcal{B}_{3k}^{(ci,*)} = (\beta_c^2 a_i^2 - l^2)\mathcal{B}_{1k}^{(ci)} + \mathcal{B}_{2k}^{(ci)} + \mathcal{B}_{3k}^{(ci)},$$
$$\mathcal{B}_{4k}^{(ci,*)} = 2l^2\mathcal{B}_{1k}^{(ci)} + (\beta_c^2 a_i^2 - l^2 - 1)\mathcal{B}_{2k}^{(ci)} + \mathcal{B}_{3k}^{(ci)} + \mathcal{B}_{4k}^{(ci)}, k = 1, 2.$$

A.1.3. *Transmission conditions at the exterior interface of the coating*

To obtain the entries of the matrices $\mathcal{A}_{pq}^{(ec)}$; $\mathcal{B}_{pq}^{(ec)}$, $p, q = 1, 2, 3, 4$, and their reduced forms $\mathcal{A}_{pq}^{(ec,*)}$; $\mathcal{B}_{pq}^{(ec,*)}$, $p, q = 1, 2, 3, 4$, we use a similar method to that described above, however, at the interface between the coating and exterior medium, the conditions for perfect bonding are equated on the radius $r = a_c$. By substituting the corresponding field representations into the two conditions (8.6), we obtain

$$A_l^{(e)} J_l(\beta_e a_c) + E_l^{(e)} H_l^{(1)}(\beta_e a_c) + B_l^{(e)} I_l(\beta_e a_c) + F_l^{(e)} K_l(\beta_e a_c)$$
$$= A_l^{(c)} J_l(\beta_c a_c) + E_l^{(c)} H_l^{(1)}(\beta_c a_c) + B_l^{(c)} I_l(\beta_c a_c) + F_l^{(c)} K_l(\beta_c a_c), \tag{A.11}$$

$$\beta_e a_c[A_l^{(e)} J_l{}'(\beta_e a_c) + E_l^{(e)} H_l^{(1)'}(\beta_e a_c) + B_l^{(e)} I_l{}'(\beta_e a_c) + F_l^{(e)} K_l{}'(\beta_e a_c)]$$
$$= \beta_c a_c[A_l^{(c)} J_l{}'(\beta_c a_c) + E_l^{(c)} H_l^{(1)'}(\beta_c a_c) + B_l^{(c)} I_l{}'(\beta_c a_c)$$
$$+ F_l^{(c)} K_l{}'(\beta_c a_c)]. \tag{A.12}$$

We can now use (A.11) and (A.12) to write the condition (8.7) on $r = a_c$ as

$$\beta_e^2 a_c^2[A_l^{(e)} J_l{}''(\beta_e a_c) + E_l^{(e)} H_l^{(1)''}(\beta_e a_c) + B_l^{(e)} I_l{}''(\beta_e a_c)$$
$$+ F_l^{(e)} K_l{}''(\beta_e a_c)]$$
$$= \frac{D_c}{D_e}\beta_c^2 a_c^2[A_l^{(c)} J_l{}''(\beta_c a_c) + E_l^{(c)} H_l^{(1)''}(\beta_c a_c) + B_l^{(c)} I_l{}''(\beta_c a_c)$$
$$+ F_l^{(c)} K_l{}''(\beta_c a_c)]$$
$$+ \left(\frac{D_c}{D_e}\nu_c - \nu_e\right)\beta_c a_c[A_l^{(c)} J_l{}'(\beta_c a_c) + E_l^{(c)} H_l^{(1)'}(\beta_c a_c)$$
$$+ B_l^{(c)} I_l{}'(\beta_c a_c) + F_l^{(c)} K_l{}'(\beta_c a_c)]$$

$$- \left(\frac{D_c}{D_e}\nu_c - \nu_e\right) l^2 [A_l^{(c)} J_l(\beta_c a_c) + E_l^{(c)} H_l^{(1)}(\beta_c a_c)$$
$$+ B_l^{(c)} I_l(\beta_c a_c) + F_l^{(c)} K_l(\beta_c a_c)]. \tag{A.13}$$

In order to write condition (8.8) on $r = a_c$, we use equation (A.4) as well as equations (A.11), (A.12) and (A.13), grouping together zeroth and first-order derivatives on the left-hand side. This gives

$$\begin{aligned}
&\beta_e^3 a_c^3 [A_l^{(e)} J_l'''(\beta_e a_c) + E_l^{(e)} {H_l^{(1)}}'''(\beta_e a_c) + B_l^{(e)} I_l'''(\beta_e a_c) \\
&\quad + F_l^{(e)} K_l'''(\beta_e a_c)] \\
&= \frac{D_c}{D_e} \beta_c^3 a_c^3 [A_l^{(c)} J_l'''(\beta_c a_c) + E_l^{(c)} {H_l^{(1)}}'''(\beta_c a_c) \\
&\quad + B_l^{(c)} I_l'''(\beta_c a_c) + F_l^{(c)} K_l'''(\beta_c a_c)] \\
&\quad - \left\{ \frac{D_c}{D_e}\nu_c - \nu_e + \frac{D_c}{D_e}\left[1 + (2-\nu_c) l^2\right] - \left[1 + (2-\nu_e) l^2\right] \right\} \beta_c a_c \\
&\quad \times [A_l^{(c)} J_l'(\beta_c a_c) + E_l^{(c)} {H_l^{(1)}}'(\beta_c a_c) \\
&\quad + B_l^{(c)} I_l'(\beta_c a_c) + F_l^{(c)} K_l'(\beta_c a_c)] \\
&\quad + 3l^2 \left(\frac{D_i}{D_c} - 1\right) [A_l^{(c)} J_l(\beta_c a_c) + E_l^{(c)} H_l^{(1)}(\beta_c a_c) \\
&\quad + B_l^{(c)} I_l(\beta_c a_c) + F_l^{(c)} K_l(\beta_c a_c)].
\end{aligned} \tag{A.14}$$

A.1.4. *Reduction of the system of transmission conditions on $r = a_c$ to block diagonal form*

We use the relevant form of equations (A.7) and (A.8) to replace the second and third order derivatives in equations (A.13) and (A.14). Equation (A.13) simplifies to

$$\begin{aligned}
&A_l^{(e)} \left[-\beta_e a_c J_l'(\beta_e a_c) - \left(\beta_e^2 a_c^2 - l^2\right) J_l(\beta_e a_c)\right] \\
&\quad + E_l^{(e)} \left[-\beta_e a_c {H_l^{(1)}}'(\beta_e a_c) - \left(\beta_e^2 a_c^2 - l^2\right) H_l^{(1)}(\beta_e a_c)\right]
\end{aligned}$$

$$
\begin{aligned}
&+ B_l^{(e)} \left[-\beta_e a_c I_l{}'(\beta_e a_c) + \left(\beta_e^2 a_c^2 + l^2\right) I_l(\beta_e a_c) \right] \\
&+ F_l^{(e)} \left[-\beta_e a_c K_l{}'(\beta_e a_c) + \left(\beta_e^2 a_c^2 + l^2\right) K_l(\beta_e a_c) \right] \\
&= A_l^{(c)} \Bigg\{ - \left[\frac{D_c}{D_e}(1-\nu_c) + \nu_e \right] \beta_c a_c J_l{}'(\beta_c a_c) \\
&\quad + \left[-\frac{D_c}{D_e}\beta_c^2 a_c^2 + l^2 \left[\frac{D_c}{D_e}(1-\nu_c) + \nu_e \right] \right] J_l(\beta_c a_c) \\
&\quad + E_l^{(c)} \Bigg\{ - \left[\frac{D_c}{D_e}(1-\nu_c) + \nu_e \right] \beta_c a_c H_l^{(1)\prime}(\beta_c a_c) \\
&\quad + \left[-\frac{D_c}{D_e}\beta_c^2 a_c^2 + l^2 \left[\frac{D_c}{D_e}(1-\nu_c) + \nu_e \right] \right] H_l^{(1)}(\beta_c a_c) \\
&\quad + B_l^{(c)} \Bigg\{ - \left[\frac{D_c}{D_e}(1-\nu_c) + \nu_e \right] \beta_c a_c I_l{}'(\beta_c a_c) \\
&\quad + \left[\frac{D_c}{D_e}\beta_c^2 a_c^2 + l^2 \left[\frac{D_c}{D_e}(1-\nu_c) + \nu_e \right] \right] I_l(\beta_c a_c) \\
&\quad + F_l^{(c)} \Bigg\{ - \left[\frac{D_c}{D_e}(1-\nu_c) + \nu_e \right] \beta_c a_c K_l{}'(\beta_c a_c) \\
&\quad + \left[\frac{D_c}{D_e}\beta_c^2 a_c^2 + l^2 \left[\frac{D_c}{D_e}(1-\nu_c) + \nu_e \right] \right] K_l(\beta_c a_c), \qquad \text{(A.15)}
\end{aligned}
$$

and (A.14) becomes

$$
\begin{aligned}
&A_l^{(e)} \left[\left(2 - \beta_e^2 a_c^2 + l^2\right) \beta_e a_c J_l{}'(\beta_e a_c) + \left(\beta_e^2 a_c^2 - 3l^2\right) J_l(\beta_e a_c) \right] \\
&\quad + E_l^{(e)} [\left(2 - \beta_e^2 a_c^2 + l^2\right) \beta_e a_c H_l^{(1)\prime}(\beta_e a_c) \\
&\quad + \left(\beta_e^2 a_c^2 - 3l^2\right) H_l^{(1)}(\beta_e a_c)] \\
&\quad + B_l^{(c)} \left[\left(2 + \beta_e^2 a_c^2 + l^2\right) \beta_e a_c I_l{}'(\beta_e a_c) - \left(\beta_e^2 a_c^2 + 3l^2\right) I_l(\beta_e a_c) \right] \\
&\quad + F_l^{(e)} [\left(2 + \beta_e^2 a_c^2 + l^2\right) \beta_e a_c K_l{}'(\beta_e a_c) - \left(\beta_e^2 a_c^2 + 3l^2\right) K_l(\beta_e a_c)] \\
&= A_l^{(c)} \Bigg\{ \left[\frac{D_c}{D_e} \left[(1-\nu_c)(1-l^2) - \beta_c^2 a_c^2 \right] + 1 + \nu_e + (2-\nu_e) l^2 \right]
\end{aligned}
$$

$$\times \beta_c a_c J_l{}'(\beta_c a_c) + \left(\frac{D_c}{D_e}\beta_c^2 a_c^2 - 3l^2\right) J_l(\beta_c a_c)\Bigg\}$$
$$+ E_l^{(c)}\Bigg\{\left[\frac{D_c}{D_e}\left[(1-\nu_c)(1-l^2) - \beta_c^2 a_c^2\right] + 1 + \nu_e + (2-\nu_e)l^2\right]$$
$$\times \beta_c a_c H_l^{(1)'}(\beta_c a_c) + \left(\frac{D_c}{D_e}\beta_c^2 a_c^2 - 3l^2\right) H_l^{(1)}(\beta_c a_c)\Bigg\}$$
$$+ B_l^{(c)}\Bigg\{\left[\frac{D_c}{D_e}\left[(1-\nu_c)(1-l^2) + \beta_c^2 a_c^2\right] + 1 + \nu_e + (2-\nu_e)l^2\right]$$
$$\times \beta_c a_c I_l{}'(\beta_c a_c) - \left(\frac{D_c}{D_e}\beta_c^2 a_c^2 + 3l^2\right) I_l(\beta_c a_c)\Bigg\}$$
$$+ F_l^{(c)}\Bigg\{\left[\frac{D_c}{D_e}\left[(1-\nu_c)(1-l^2) + \beta_c^2 a_c^2\right] + 1 + \nu_e + (2-\nu_e)l^2\right]$$
$$\times \beta_c a_c K_l{}'(\beta_c a_c) - \left(\frac{D_c}{D_e}\beta_c^2 a_c^2 + 3l^2\right) K_l(\beta_c a_c)\Bigg\}. \tag{A.16}$$

The matrices $\mathcal{A}^{(ec)}$ and $\mathcal{B}^{(ec)}$ in equation (8.33) can be written explicitly, and their elements are given below:

$$\mathcal{A}_{11}^{(ec)} = J_l(\beta_e a_c), \quad \mathcal{A}_{12}^{(ec)} = H_l^{(1)}(\beta_e a_c), \quad \mathcal{A}_{13}^{(ec)} = I_l(\beta_e a_c),$$
$$\mathcal{A}_{14}^{(ec)} = K_l(\beta_e a_c), \quad \mathcal{A}_{21}^{(ec)} = \beta_e a_c J_l{}'(\beta_e a_c), \quad \mathcal{A}_{22}^{(ec)} = \beta_e a_c H_l^{(1)'}(\beta_e a_c),$$
$$\mathcal{A}_{23}^{(ec)} = \beta_e a_c I_l{}'(\beta_e a_c), \quad \mathcal{A}_{24}^{(ec)} = \beta_e a_c K_l{}'(\beta_e a_c),$$
$$\mathcal{A}_{31}^{(ec)} = -(\beta_e^2 a_c^2 - l^2) J_l(\beta_e a_c) - \beta_e a_c J_l{}'(\beta_e a_c),$$
$$\mathcal{A}_{32}^{(ec)} = -(\beta_e^2 a_c^2 - l^2) H_l^{(1)}(\beta_e a_c) - \beta_e a_c H_l^{(1)'}(\beta_e a_c),$$
$$\mathcal{A}_{33}^{(ec)} = (\beta_e^2 a_c^2 + l^2) I_l(\beta_e a_c) - \beta_e a_c I_l{}'(\beta_e a_c),$$
$$\mathcal{A}_{34}^{(ec)} = (\beta_e^2 a_c^2 + l^2) K_l(\beta_e a_c) - \beta_e a_c K_l{}'(\beta_e a_c),$$
$$\mathcal{A}_{41}^{(ec)} = (\beta_e^2 a_c^2 - 3l^2) J_l(\beta_e a_c) + (-\beta_e^2 a_c^2 + l^2 + 2)\beta_e a_c J_l{}'(\beta_e a_c),$$

$$\mathcal{A}_{42}^{(ec)} = (\beta_e^2 a_c^2 - 3l^2) H_l^{(1)}(\beta_e a_c) + (-\beta_e^2 a_c^2 + l^2 + 2)\beta_e a_c H_l^{(1)\prime}(\beta_e a_c),$$
$$\mathcal{A}_{43}^{(ci)} = -(\beta_e^2 a_c^2 + 3l^2) I_l(\beta_e a_c) + (\beta_e^2 a_c^2 + l^2 + 2)\beta_e a_c I_l{}'(\beta_e a_c),$$
$$\mathcal{A}_{44}^{(ci)} = -(\beta_e^2 a_c^2 + 3l^2) K_l(\beta_e a_c) + (\beta_e^2 a_c^2 + l^2 + 2)\beta_e a_c K_l{}'(\beta_e a_c);$$

$$\mathcal{B}_{11}^{(ec)} = J_l(\beta_c a_c), \quad \mathcal{B}_{12}^{(ec,*)} = H_l^{(1)}(\beta_c a_c), \quad \mathcal{B}_{13}^{(ec)} = I_l(\beta_c a_c),$$
$$\mathcal{B}_{14}^{(ec,*)} = K_l(\beta_c a_c), \quad \mathcal{B}_{21}^{(ec)} = \beta_c a_c J_l{}'(\beta_c a_c), \quad \mathcal{B}_{22}^{(ec)} = \beta_c a_c H_l^{(1)\prime}(\beta_c a_c),$$
$$\mathcal{B}_{23}^{(ec)} = \beta_c a_c I_l{}'(\beta_c a_c), \quad \mathcal{B}_{24}^{(ec)} = \beta_c a_c K_l{}'(\beta_c a_c),$$
$$\mathcal{B}_{31}^{(ec)} = \left\{ l^2 \left[\frac{D_c}{D_e}(1-\nu_c) + \nu_e \right] - \frac{D_c}{D_e}\beta_c^2 a_c^2 \right\} J_l(\beta_c a_c)$$
$$- \beta_c a_c \left[\frac{D_c}{D_e}(1-\nu_c) + \nu_e \right] J_l{}'(\beta_c a_c),$$
$$\mathcal{B}_{32}^{(ec)} = \left\{ l^2 \left[\frac{D_c}{D_e}(1-\nu_c) + \nu_e \right] - \frac{D_c}{D_e}\beta_c^2 a_c^2 \right\} H_l^{(1)}(\beta_c a_c)$$
$$- \beta_c a_c \left[\frac{D_c}{D_e}(1-\nu_c) + \nu_e \right] H_l^{(1)\prime}(\beta_c a_c),$$
$$\mathcal{B}_{33}^{(ec)} = \left\{ l^2 \left[\frac{D_c}{D_e}(1-\nu_c) + \nu_e \right] + \frac{D_c}{D_e}\beta_c^2 a_c^2 \right\} I_l(\beta_c a_c)$$
$$- \beta_c a_c \left[\frac{D_c}{D_e}(1-\nu_c) + \nu_e \right] I_l{}'(\beta_c a_c),$$
$$\mathcal{B}_{34}^{(ec)} = \left\{ l^2 \left[\frac{D_c}{D_e}(1-\nu_c) + \nu_e \right] + \frac{D_c}{D_e}\beta_c^2 a_c^2 \right\} K_l(\beta_c a_c)$$
$$- \beta_c a_c \left[\frac{D_c}{D_e}(1-\nu_c) + \nu_e \right] K_l{}'(\beta_c a_c),$$
$$\mathcal{B}_{41}^{(ec)} = \left(\frac{D_c}{D_e}\beta_c^2 a_c^2 - 3l^2 \right) J_l(\beta_c a_c)$$
$$+ \beta_c a_c \left\{ \frac{D_c}{D_e} \left[1 - \beta_c^2 a_c^2 + l^2(\nu_c - 1) - \nu_c \right] \right.$$
$$\left. + \nu_e + 1 + l^2(2 - \nu_e) \right\} J_l'(\beta_c a_c),$$

$$\mathcal{B}_{42}^{(ec)} = \left(\frac{D_c}{D_e}\beta_c^2 a_c^2 - 3l^2\right) H_l^{(1)}(\beta_c a_c) + \beta_c a_c \left\{\frac{D_c}{D_e}\left[1 - \beta_c^2 a_c^2 + l^2(\nu_c - 1) - \nu_c\right] + \nu_e + 1 + l^2(2 - \nu_e)\right\} H_l^{(1)\prime}(\beta_c a_c),$$

$$\mathcal{B}_{43}^{(ec)} = \left(-\frac{D_c}{D_e}\beta_c^2 a_c^2 - 3l^2\right) I_l(\beta_c a_c) + \beta_c a_c \left\{\frac{D_c}{D_e}\left[1 + \beta_c^2 a_c^2 + l^2(\nu_c - 1) - \nu_c\right] + \nu_e + 1 + l^2(2 - \nu_e)\right\} I_l'(\beta_c a_c),$$

$$\mathcal{B}_{44}^{(ec)} = \left(-\frac{D_c}{D_e}\beta_c^2 a_c^2 - 3l^2\right) K_l(\beta_c a_c) + \beta_c a_c \left\{\frac{D_c}{D_e}\left[1 + \beta_c^2 a_c^2 + l^2(\nu_c - 1) - \nu_c\right] + \nu_e + 1 + l^2(2 - \nu_e)\right\} K_l'(\beta_c a_c).$$

A.1.5. *Further algebraic simplification*

In order to obtain the matrices $\mathcal{A}_{pq}^{(ec,*)}$; $\mathcal{B}_{pq}^{(ec,*)}$, $p, q = 1, 2, 3, 4$, (where $\mathcal{A}_{pq}^{(ec,*)}$ is block diagonal), we initially perform the following row operations on the matrix $\mathcal{A}^{(ec)}$ which gives

$$\begin{aligned}
(\beta_e^2 a_c^2 + l^2)R_1 - R_2 - R_3 &\longrightarrow R_1^*,\\
-2l^2 R_1 + (\beta_e^2 a_c^2 + l^2 + 1)R_2 - R_3 - R_4 &\longrightarrow R_2^*,\\
(\beta_e^2 a_c^2 - l^2)R_1 + R_2 + R_3 &\longrightarrow R_3^*,\\
2l^2 R_1 + (\beta_e^2 a_c^2 - l^2 - 1)R_2 + R_3 + R_4 &\longrightarrow R_4^*.
\end{aligned}$$

The reduced form matrix $\mathcal{A}_{pq}^{(ec,*)}$, $p, q = 1, 2, 3, 4$, can be obtained as

$$\mathcal{A}_{11}^{(ec,*)} = 2\beta_e^2 a_c^2 J_l(\beta_e a_c), \quad \mathcal{A}_{12}^{(ec,*)} = 2\beta_e^2 a_c^2 H_l^{(1)}(\beta_e a_c),$$

$$\mathcal{A}_{13}^{(ec,*)} = \mathcal{A}_{14}^{(ec,*)} = 0, \quad \mathcal{A}_{21}^{(ec,*)} = 2\beta_e^3 a_c^3 J_l{}'(\beta_e a_c),$$
$$\mathcal{A}_{22}^{(ec,*)} = 2\beta_e^3 a_c^3 H_l^{(1)'}(\beta_e a_c), \quad \mathcal{A}_{23}^{(ec,*)} = \mathcal{A}_{24}^{(ec,*)} = 0,$$
$$\mathcal{A}_{31}^{(ec,*)} = \mathcal{A}_{32}^{(ec,*)} = 0, \quad \mathcal{A}_{33}^{(ec,*)} = 2\beta_e^2 a_c^2 I_l(\beta_e a_c),$$
$$\mathcal{A}_{34}^{(ec,*)} = 2\beta_e^2 a_c^2 K_l(\beta_e a_c), \quad \mathcal{A}_{41}^{(ec,*)} = \mathcal{A}_{42}^{(ec,*)} = 0,$$
$$\mathcal{A}_{43}^{(ec,*)} = 2\beta_e^3 a_c^3 I_l{}'(\beta_e a_c), \quad \mathcal{A}_{44}^{(ec,*)} = 2\beta_e^3 a_c^3 K_l{}'(\beta_e a_c).$$

We perform the same row operations on matrix $\mathcal{B}_{pq}^{(ec)}$, $p, q = 1, 2, 3, 4$. It is clear that its reduced form reads as follows

$$\mathcal{B}_{1k}^{(ec,*)} = (\beta_e^2 a_c^2 + l^2)\mathcal{B}_{1k}^{(ec)} - \mathcal{B}_{2k}^{(ec)} - \mathcal{B}_{3k}^{(ec)},$$
$$\mathcal{B}_{2k}^{(ec,*)} = -2l^2\mathcal{B}_{1k}^{(ec)} + (\beta_e^2 a_c^2 + l^2 + 1)\mathcal{B}_{2k}^{(ec)} - \mathcal{B}_{3k}^{(ec)} - \mathcal{B}_{4k}^{(ec)},$$
$$\mathcal{B}_{3k}^{(ec,*)} = (\beta_e^2 a_c^2 - l^2)\mathcal{B}_{1k}^{(ec)} + \mathcal{B}_{2k}^{(ec)} + \mathcal{B}_{3k}^{(ec)},$$
$$\mathcal{B}_{4k}^{(ec,*)} = 2l^2\mathcal{B}_{1k}^{(ec)} + (\beta_e^2 a_c^2 - l^2 - 1)\mathcal{B}_{2k}^{(ec)} + \mathcal{B}_{3k}^{(ec)} + \mathcal{B}_{4k}^{(ec)},$$

$k = 1, 2, 3, 4$.

A.2. Kirchhoff Plate: Explicit Representations for the Entries of the Scattering Matrix $\mathcal{S}_l$ and the Coefficients $A_l^{(i)}, B_l^{(i)}, A_l^{(c)}, E_l^{(c)}, B_l^{(c)}, F_l^{(c)}, E_l^{(e)}$ and $F_l^{(e)}$

Splitting the matrix $\mathcal{C}^{(ei,*)}$ into two parts we can write

$$\begin{pmatrix} A_l^{(e)} \\ B_l^{(e)} \end{pmatrix} = \begin{pmatrix} \mathcal{C}_{11}^{(ei,*)} & \mathcal{C}_{12}^{(ei,*)} \\ \mathcal{C}_{31}^{(ei,*)} & \mathcal{C}_{32}^{(ei,*)} \end{pmatrix} \begin{pmatrix} A_l^{(i)} \\ B_l^{(i)} \end{pmatrix},$$
$$\begin{pmatrix} E_l^{(e)} \\ F_l^{(e)} \end{pmatrix} = \begin{pmatrix} \mathcal{C}_{21}^{(ei,*)} & \mathcal{C}_{22}^{(ei,*)} \\ \mathcal{C}_{41}^{(ei,*)} & \mathcal{C}_{42}^{(ei,*)} \end{pmatrix} \begin{pmatrix} A_l^{(i)} \\ B_l^{(i)} \end{pmatrix}.$$

Now eliminating $A_l^{(i)}, B_l^{(i)}$ from these two equations, we obtain

$$\begin{pmatrix} E_l^{(e)} \\ F_l^{(e)} \end{pmatrix} = \mathcal{S}_l \begin{pmatrix} A_l^{(e)} \\ B_l^{(e)} \end{pmatrix},$$

where $\mathcal{S}_l$ is the scattering matrix with the elements

$$\mathcal{S}_l^{11} = \frac{\mathcal{C}_{21}^{(ei,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{22}^{(ei,*)}\mathcal{C}_{31}^{(ei,*)}}{\mathcal{C}_{11}^{(ei,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{12}^{(ei,*)}\mathcal{C}_{31}^{(ei,*)}}, \quad \mathcal{S}_l^{12} = \frac{\mathcal{C}_{11}^{(ei,*)}\mathcal{C}_{22}^{(ei,*)} - \mathcal{C}_{12}^{(ei,*)}\mathcal{C}_{21}^{(ei,*)}}{\mathcal{C}_{11}^{(ei,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{12}^{(ei,*)}\mathcal{C}_{31}^{(ei,*)}},$$

$$\mathcal{S}_l^{21} = \frac{\mathcal{C}_{32}^{(ei,*)}\mathcal{C}_{41}^{(ei,*)} - \mathcal{C}_{31}^{(ei,*)}\mathcal{C}_{42}^{(ei,*)}}{\mathcal{C}_{11}^{(ei,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{12}^{(ei,*)}\mathcal{C}_{31}^{(ei,*)}}, \quad \mathcal{S}_l^{22} = \frac{\mathcal{C}_{11}^{(ei,*)}\mathcal{C}_{42}^{(ei,*)} - \mathcal{C}_{12}^{(ei,*)}\mathcal{C}_{41}^{(ei,*)}}{\mathcal{C}_{11}^{(ei,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{12}^{(ei,*)}\mathcal{C}_{31}^{(ei,*)}}.$$

We note that using the relation (8.36), all coefficients appearing in the representation (8.9) can be obtained as follows:

$$A_l^{(i)} = \frac{\mathcal{C}_{32}^{(ei,*)}}{\mathcal{C}_{11}^{(ei,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{12}^{(ei,*)}\mathcal{C}_{31}^{(ei,*)}}\, A_l^{(e)},$$

$$B_l^{(i)} = -\frac{\mathcal{C}_{31}^{(ei,*)}}{\mathcal{C}_{11}^{(ei,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{12}^{(ei,*)}\mathcal{C}_{31}^{(ei,*)}}\, A_l^{(e)}; \tag{A.17}$$

$$A_l^{(c)} = \frac{\mathcal{C}_{11}^{(ci,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{12}^{(ci,*)}\mathcal{C}_{31}^{(ei,*)}}{\mathcal{C}_{11}^{(ei,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{12}^{(ei,*)}\mathcal{C}_{31}^{(ei,*)}}\, A_l^{(e)},$$

$$B_l^{(c)} = \frac{\mathcal{C}_{31}^{(ci,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{32}^{(ci,*)}\mathcal{C}_{31}^{(ei,*)}}{\mathcal{C}_{11}^{(ei,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{12}^{(ei,*)}\mathcal{C}_{31}^{(ei,*)}}\, A_l^{(e)}, \tag{A.18}$$

$$E_l^{(c)} = \frac{\mathcal{C}_{21}^{(ci,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{22}^{(ci,*)}\mathcal{C}_{31}^{(ei,*)}}{\mathcal{C}_{11}^{(ei,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{12}^{(ei,*)}\mathcal{C}_{31}^{(ei,*)}}\, A_l^{(e)},$$

$$F_l^{(c)} = \frac{\mathcal{C}_{41}^{(ci,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{42}^{(ci,*)}\mathcal{C}_{31}^{(ei,*)}}{\mathcal{C}_{11}^{(ei,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{12}^{(ei,*)}\mathcal{C}_{31}^{(ei,*)}}\, A_l^{(e)}; \tag{A.19}$$

$$E_l^{(e)} = \frac{\mathcal{C}_{21}^{(ei,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{22}^{(ei,*)}\mathcal{C}_{31}^{(ei,*)}}{\mathcal{C}_{11}^{(ei,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{12}^{(ei,*)}\mathcal{C}_{31}^{(ei,*)}}\, A_l^{(e)},$$

$$F_l^{(e)} = \frac{\mathcal{C}_{41}^{(ei,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{42}^{(ei,*)}\mathcal{C}_{31}^{(ei,*)}}{\mathcal{C}_{11}^{(ei,*)}\mathcal{C}_{32}^{(ei,*)} - \mathcal{C}_{12}^{(ei,*)}\mathcal{C}_{31}^{(ei,*)}}\, A_l^{(e)}. \tag{A.20}$$

Bibliography

[1] Dolin, L. S. (1961). To the possibility of comparison of three-dimensional electromagetic system with non-uniform anisotropic filling, *Izv. Vyssh. Uch. Zav. Radiofizika* **4**(5), pp. 964–967.

[2] Leonhardt, U. (2006). Optical conformal mapping, *Science* **312**, p. 1777.

[3] Pendry, J. B., Schurig, D. and Smith, D. R. (2006). Controlling electromagnetic fields, *Science* **312**, p. 1780.

[4] Greenleaf, A., Lassas, M. and Uhlmann, G. (2003). Anisotropic conductivities that cannot be detected by EIT, *Physiol. Meas.* **24**, p. 413.

[5] Norris, A. N. and Shuvalov, A. L. (2011). Elastic cloaking theory, *Wave Motion* **48**(6), p. 525.

[6] Brun, M., Guenneau, S. and Movchan, A. B. (2009). Achieving control of in-plane elastic waves, *Appl. Phys. Lett.* **94**, p. 061903.

[7] Milton, G. W., Briane, M. and Willis, J. R. (2006). On cloaking for elasticity and physical equations with a transformation invariant form, *New Journal of Physics* **8**, 248.

[8] Milton, G. W. and Willis, J. R. (2007). On modifications of Newton's second law and linear continuum elastodynamics, *Proc. Roy. Soc. A* **463**, p. 855.

[9] O'Neill, J., Selsil, Ö. McPhedran, R. C., Movchan, A. B. and Movchan, N. V. (2015). Active cloaking of inclusions for flexural waves in thin elastic plates, *Quarterly Journal of Mechanics and Applied Mathematics* **68**(3), p. 263.

[10] O'Neill, J., Selsil, Ö. McPhedran, R. C., Movchan, A. B. and Movchan, N. V. (2015). Active cloaking of resonant coated inclusions for waves in membranes and Kirchhoff plates, http://arxiv.org/abs/1505.06138.

[11] Futhazar, G., Parnell, W. J. and Norris, A. N. (2015). Active cloaking of flexural waves in thin plates, *Journal of Sound and Vibration* **356**, p. 1.

[12] Norris, A. N., Amirkulova, F. A. and Parnell, W. J. (2012). Source amplitudes for active exterior cloaking, *Inverse Problems* **28**(10), p. 105002.

[13] Norris, A. N., Amirkulova, F. A. and Parnell, W. J. (2013). Active elastodynamic cloaking, *Mathematics and Mechanics of Solids* **19**, p. 603.

[14] Farhat, M., Guenneau, S., Enoch, S. and Movchan, A. B. (2009). Cloaking bending waves propagating in thin elastic plates, *Physical Review B* **79**, p. 033102.

[15] Miller, D. A. (2006). On perfect cloaking, *Optics Express* **14**, p. 12457.

[16] Friot, E. and Bordier, C. (2004). Real-time active suppression of scattered acoustic radiation, *Journal of Sound and Vibration* **278**, p. 563.

[17] Friot, E., Guillermin, R. and Winninger, M. (2006). Active control of scattered acoustic radiation: A real-time implementation for a three-dimensional object, *Acta Acustica united with Acustica* **92**, p. 278.

[18] Ffowcs Williams, J. E. (1984). Review lecture: Anti-sound, *Proceedings of the Royal Society London A* **395**, p. 63.

[19] Nelson, P. A. and Elliott, S. J. (1992). *Active Control of Sound*, (Academic Press).

[20] Guevara Vasquez, F., Milton, G. W. and Onofrei, D. (2009). Active exterior cloaking for the 2D laplace and helmholtz equations, *Physcial Review Letters* **103**, p. 073901.
[21] Guevara Vasquez, F., Milton, G. W. and Onofrei, D. (2009). Broadband exterior cloaking, *Optical Express* **17**, p. 14800.
[22] Guevara Vasquez, F., Milton, G. W. and Onofrei, D. (2011). Exterior cloaking with active sources in two dimensional acoustics, *Wave Motion* **48**, p. 515.
[23] Monticone, F. and Alù, A. (2013). Do cloaked objects really scatter less?, *Phys. Rev. X* **3**, p. 041005.
[24] Jones, I. S., Brun, M., Movchan, N. V. and Movchan, A. B. (2015). Singular perturbations and cloaking illusions for elastic waves in membranes and Kirchhoff plates, *International Journal of Solids and Structures* **69–70**, p. 498.
[25] Farhat, M., Chen, P.-Y., Bağcı, H., Enoch, S., Guenneau, S. and Alu; A. (2014). Platonic scattering cancellation for bending waves in a thin plate, *Scientific Reports* **4**, p. 4644.
[26] Norris, A. N. and Vemula, C. (1995). Scattering of flexural waves on thin plates, *J. Sound Vib.* **181**(1), p. 115.
[27] Evans, D. V. and Porter, R. (2007). Penetration of flexural waves through a periodically constrained thin elastic plate floating on water, *J. Eng. Math.* **58**, p. 317.
[28] Movchan, A. B., Movchan, N. V. and McPhedran, R. C. (2007). Bloch-Floquet bending waves in perforated thin plates, *Proc. R. Soc. A* **463**, p. 2505.
[29] McPhedran, R. C., Movchan, A. B. and Movchan, N. V. (2009). Platonic crystals: Bloch bands, neutrality and defects, *Mech. Mater.* **41**, p. 356.
[30] Movchan, N. V., McPhedran, R. C., Movchan, A.B. and Poulton, C. G. (2009). Wave scattering by platonic grating stacks, *Proc. R. Soc. A* **465**, p. 3383.
[31] Abramowitz, M. and Stegun, I. A. (1965). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Dover Publications).
[32] Milton, G. W. (2002). *The Theory of Composites* (Cambridge University Press).
[33] Konenkov, Yu. K. (1964). Diffraction of a flexural wave by a circular obstacle in a plate, *Soviet Physics-Acoustics* **10**(2), p. 153.
[34] Torrent, D., Pennec, Y. and Djafari-Rouhani, B. (2014). Effective medium theory for elastic metamaterials in thin elastic plates, *Phys. Rev. B* **90**, p. 104110.
[35] Norris, A. N. and Vemula, C. (1995). Scattering of flexural waves on thin plates, *J. Sound Vib.* **181**(1), p. 115.

CHAPTER 9

Cloaking in Water Waves

RICHARD PORTER

University of Bristol, UK
richard.porter@bristol.ac.uk

9.1. Introduction

This review article is written with two goals in mind. The first is to provide researchers working in the area of water waves an insight into ideas emerging from the physics community on the use of metamaterials in controlling waves in unusual ways. The second is to give physicists already familiar with these ideas an overview of the theory of water waves, the common approximations that are used in developing solutions, and how metamaterial concepts may be implemented within this framework.

In particular, whilst there has been an expansive crossover of ideas developed primarily from optics and electromagnetics into the acoustics and elasticity research communities, much less has been realised in terms of establishing such a connection in the water wave community.

There is a good reason why this might be. Water waves are quite different to waves in electromagnetics, acoustics or elasticity in that there is a special direction (the vertical) which makes simple connections between theories less easy to establish: the limitations imposed

by a free surface and a fluid depth are a recurring theme throughout this article. The applications of metamaterials in water waves are also perhaps a little harder to imagine. On the other hand wave phenomena are easily visualised on the surface of water and this makes it an attractive medium in which to develop ideas from other areas of physics.

My own interest in this subject was stimulated after becoming aware of the publication of the work of Ref. 11. Immediately evident was the widespread influence of the work of Refs. 19 and 30 on invisibility cloaking in optics and other extensions to areas of science including acoustics and elasticity. Although I've not fully engaged with metamaterial science I've slowly picked up many of the intruiging developments that have been made over the last 15–20 years, especially those with applications in water waves. Broadly speaking, a metamaterial is a medium in which properties of a field can be propagated in a manner not normally found in natural materials. They are most often comprised of microstructures much smaller than the natural lengthscales intrinsic to the underlying field variables in such a way that their macroscopic effect on the field allows complex phenomena to be exhibited. These include negative refraction in which oblique waves bend backwards as they enter the metamaterial and perfect lensing, e.g. Refs. 28 and 36. In the context of water waves, see Ref. 13.

One of the most fascinating areas to emerge from the science of metamaterials is invisibility cloaking in which obstacles are rendered undetectable to the observer and it is principally on this topic which the current article will focus. It turns out that the capacity to cloak in water waves leads to a reduction (to zero for a perfect cloak) of the so-called mean drift force. This is a second-order effect in water waves and hence generally smaller in magnitude than the primary oscillatory forces due to water waves but, unlike those wave forces, it is steady and has important consequences in the design of marine structures such as the foundations of offshore wind turbines.

The layout of the article is as follows. In the remainder of this section, we shall introduce the underlying equations of motion which

govern small-amplitude water waves and apply it to our canonical problem of the scattering by a vertical cylinder. In Section 9.2, a series of approximations to the full governing equation in which the depth dependence is removed will be introduced and discussed in the context of cloaking cylinders. Section 9.3 describes the transformation media approach and how it applies to depth-averaged models. This includes a discussion of how to design the metamaterials needed to practically implement the wave control needed to cloak. In Section 9.4 we consider examples of cloaking in the unapproximated full linear theory and the work is summarised in Section 9.5.

9.1.1. *Linearised theory of water waves*

There are many textbooks which carefully outline the derivation of the linearised theory of water waves which we summarise below. It is hard to beat the account in the book of Ref. 24 and we refer to this extensively throughout.

Throughout this article we take z to be the vertical axis with $z = 0$ set to the level of the undisturbed free surface. A fluid of density ρ under the influence of gravity g lies below $z = 0$ in equilibrium and is bounded below by an impermeable bed at $z = -h(x, y)$ where (x, y) are the horizontal coordinates. It is assumed in deriving conditions at the free surface that a fluid (such as the air) having density much less than ρ and in which the pressure is constant lies above $z = 0$ in equilibrium. The effect of this fluid on the motion of the water is then negligible.

There are many other assumptions used in the derivation of classical water wave theory. The first is that the fluid is inviscid. This is a good assumption in the bulk but fails close to boundaries, including on the free surface — see [Ref. 24, Section 9]. However, it is a good approximation overall provided any experimental realisation of the theory is performed on a large enough scale. Another assumption is that the flow is irrotational (implying there is no vorticity in the fluid). In order that theory can be realised in experiments this often implies avoiding geometrical structures with sharp corners or edges which induce vortex shedding. Irrotationality implies velocity of the

fluid $\mathbf{u}(x, y, z, t)$ satisfies the condition $\nabla \times \mathbf{u} = 0$ which allows us to write $\mathbf{u} = \nabla\Phi(x, y, z, t)$ where Φ is a scalar potential (called the velocity potential) and $\nabla = (\partial_x, \partial_y, \partial_z)$. Further, it is routine to assume that the water is incompressible and this is expressed mathematically as $\nabla \cdot \mathbf{u} = 0$. Consequently, Φ satisfies

$$\nabla^2\Phi = 0, \quad \text{in } \mathcal{V} \tag{9.1}$$

where $\mathcal{V}$ represents the domain occupied by the fluid and $\nabla^2 = \nabla \cdot \nabla$ is the three-dimensional Laplacian.

Euler's equation expressing conservation of momentum for an inviscid fluid allows the pressure to be determined in the fluid. Setting this constant on the free surface and linearising on the assumption of small steepness of the surface elevation defined by $z = \zeta(x, y, t)$ about the mean level furnishes the dynamic boundary condition

$$\Phi_t + g\zeta = 0, \quad \text{on } z = 0. \tag{9.2}$$

Similarly, a linearised version of the kinematic condition on the surface, which states that the surface moves with the fluid, gives

$$\zeta_t = \Phi_z, \quad \text{on } z = 0. \tag{9.3}$$

The two equations (9.2) and (9.3) can be combined to eliminate ζ so that

$$\Phi_{tt} + g\Phi = 0, \quad \text{on } z = 0 \tag{9.4}$$

expresses a combined kinematic and dynamic boundary condition. The condition that there is no flow through all fixed wetted boundaries $\mathcal{S}$ with given unit normal $\mathbf{n}$ out of the fluid implies that

$$\Phi_n \equiv \mathbf{n} \cdot \nabla\Phi = 0, \quad \text{on } (x, y, z) \in \mathcal{S}. \tag{9.5}$$

One section of $\mathcal{S}$ is the sea bed where $\mathbf{n} = (h_x, h_y, -1)$ and (9.5) reduces to

$$\Phi_z + \nabla_h h \cdot \nabla_h \Phi = 0, \quad \text{on } z = -h(x, y) \tag{9.6}$$

where $\nabla_h \equiv (\partial_x, \partial_y)$ is the gradient projected onto the two-dimensional horizontal plane. The bed is often assumed to be flat

outside some bounded region of the (x, y)-plane, and given by $h = h_0$, a constant, and there (9.6) reduces to

$$\Phi_z = 0, \quad \text{on } z = -h_0. \tag{9.7}$$

In order to fully specify a problem which can be solved (numerically or analytically) one must also impose initial conditions. These most often imply stating the initial surface elevation and velocity at a reference time $t = 0$. However advantage can be (and most often is) taken of the fact that the governing equation for Φ and the boundary condition it satisfies are linear and so any time domain solution can be inferred via inverse Fourier transforms of frequency domain solutions. Thus, we write $\Phi(x, y, z, t) = \Re\{\phi(x, y, z)\mathrm{e}^{-\mathrm{i}\omega t}\}$ and $\zeta(x, y, t) = \Re\{\eta(x, y)\mathrm{e}^{-\mathrm{i}\omega t}\}$ where ω is the assumed radian frequency and now ϕ and η are frequency-dependent complex-valued functions incorporating information about the amplitude and the phase of the fluid motion.

It follows from (9.1), (9.4) and (9.5) that ϕ now satisfies

$$\nabla^2\phi = 0, \quad \text{in } \mathcal{V} \tag{9.8}$$

with

$$\phi_z - K\phi = 0, \quad \text{on } z = 0 \tag{9.9}$$

where $K = \omega^2/g$ and

$$\phi_n = 0, \quad \text{on } \mathcal{S} \tag{9.10}$$

which includes (9.6) in the form

$$\phi_z + \nabla_h h \cdot \nabla_h \phi = 0, \quad \text{on } z = -h(x, y) \tag{9.11}$$

simplifying (9.7) to $\phi_z = 0$ on $z = -h_0$ where the bed is flat. Also (9.2) translates to

$$\eta(x, y) = (\mathrm{i}\omega/g)\phi(x, y, 0) \tag{9.12}$$

representing the time-independent free-surface elevation.

9.1.2. *The character of waves*

Waves propagate on the surface of the water (they are akin to guided, interface or surface waves in other physical disciplines) with exponential decay in the direction of increasing depth away from the surface. Waves of amplitude A over a flat bed of depth h_0 from infinity in a direction θ_0 w.r.t. the positive x-axis are given by the potential

$$\phi_{inc} = \frac{-\mathrm{i}gA}{\omega}\frac{\cosh k_0(z+h_0)}{\cosh k_0 h_0}\mathrm{e}^{\mathrm{i}k_0(x\cos\theta_0+y\sin\theta_0)}. \tag{9.13}$$

This satisfies (9.8) and (9.11) on $z=-h_0$ and (9.9) provided that the unique positive real root k of

$$K \equiv \frac{\omega^2}{g} = k\tanh kh \tag{9.14}$$

corresponding to $h=h_0$ is assigned to k_0. The equation (9.14) is called the water wave dispersion relation and encodes information on how waves of different frequencies travel at different speeds and how both relate to the fluid depth. As usual, the phase speed is given by $c=\omega/k$ and the group velocity has magnitude $c_g=\mathrm{d}\omega/\mathrm{d}k$.

We shall refer to the shallow water regime as being when $kh \ll 1$ or $\lambda \gg h$ where $\lambda = 2\pi/k$ is the wavelength. In this case (9.14) shows that $k \approx \omega/\sqrt{gh}$ and $c \approx \sqrt{gh}$ whilst $c_g \approx c$. Water waves are therefore non-dispersive in the shallow water approximation. On the other hand in the deep water regime, $kh \gg 1$ or $\lambda \ll h$, (9.14) shows that $k \approx \omega^2/g$ and $c \approx g/\omega$ with $c_g \approx \frac{1}{2}c$.

In many applications of linear wave theory one is interested in how an incident plane wave as described by (9.13) interacts with fixed or moving marine structures such as the legs of an oil rig, an offshore breakwater, varying bathymetry, the coast, ships or wave energy absorbing devices.

9.1.3. *Scattering by a vertical circular cylinder*

Even with the simplifying assumptions that comprise the linearised theory of water waves, the three-dimensional boundary-value problem is complicated and substantial analytic progress is often

restricted to a small class of problems with a simple geometric configuration. Many of these are posed in a two-dimensional reduction of the water wave problem in which plane parallel-crested waves are incident on a geometry which has constant cross-section in one horizontal direction. In such problems wave propagation is essentially one-dimensional and confined to the other perpendicular horizontal direction. Thus, wave scattering problems are reduced to determining a reflection and a transmission coefficient.

Undoubtably the simplest non-trivial problem to consider analytically in three dimensions involves the scattering by a circular cylinder of constant cross section protruding vertically from a bed of constant depth h_0 through the surface of the fluid. This problem forms the basis of much of what is to come in the rest of this article and it is instructive to use it as an example of the theory introduced so far.

It is natural to use cylindrical coordinates $x = r\cos\theta$, $y = r\sin\theta$. Then it can be shown that

$$\begin{aligned}\phi_{inc} &= \mathrm{e}^{\mathrm{i}k_0 r\cos(\theta-\theta_0)}\cosh k_0(z+h_0)\\ &= \cosh k_0(z+h_0)\sum_{n=-\infty}^{\infty}\mathrm{i}^n J_n(k_0 r)\mathrm{e}^{\mathrm{i}n(\theta-\theta_0)}.\end{aligned}\tag{9.15}$$

where J_n are Bessel functions. Note that here and henceforth we dismiss the constant prefactor in (9.13) for clarity as the problem is linear and solutions can be scaled *a posteriori* as necessary.

The fact that cylinder surface (radius a, say) is aligned with the vertical and extends throughout the depth means that the scattered wave response to the incident wave has the same depth dependence as the incident wave (we elaborate on this in a moment). The total diffracted field is $\phi = \phi_{inc} + \phi_s$ where

$$\phi_s = \cosh k_0(z+h_0)\sum_{n=-\infty}^{\infty} a_n H_n^{(1)}(k_0 r)\mathrm{e}^{\mathrm{i}n(\theta-\theta_0)}\tag{9.16}$$

and a_n are as yet unknown Fourier-Bessel coefficients. The potential in (9.16), derived by separation of variables, satisfies (9.8), (9.9)

and (9.11) where $H_n^{(1)}(z) = J_n(z) + \mathrm{i}Y_n(z)$ is the Hankel function of the first kind. This choice, rather than any other combination of Bessel functions, ensures that the scattered potential represents outgoing waves at infinity. This so-called radiation condition is a previously unstated requirement in the specification of the water wave boundary-value problem posed in the frequency domain.

Upon satisfaction of the cylinder boundary condition (9.10), which here is $\phi_r = 0$ on $r = a$, for all $-h_0 < z < 0$, $-\pi < \theta \leq \pi$ determines

$$a_n = -\mathrm{i}^n \frac{J_n'(k_0 a)}{H_n^{(1)'}(k_0 a)} \tag{9.17}$$

and the solution is complete. The solution $\phi = \phi_{inc} + \phi_s$ formed from (9.15), (9.16) and (9.17) is attributed to Ref. 22 amongst researchers in the water waves although it is familiar in other areas of physics.

As already indicated the geometry in this problem, which is comprised of a flat bed and vertically-walled scatterers, means that $\cosh k_0(z + h_0)$ can be factorised from the potential as

$$\phi(x, y, z) = \psi(x, y) \cosh k_0(z + h_0) \tag{9.18}$$

and (9.8) reduces to the two-dimensional wave equation

$$(\nabla_h^2 + k_0^2)\psi = 0 \tag{9.19}$$

which shares solutions common to scalar TE- and TM-polarised waves in electromagnetics and two-dimensional inviscid low-Mach number acoustics. Unlike those physical disciplines, in water waves where Neumann boundary conditions represent fixed impermable boundaries, there is no physical realisation of a boundary which has a Dirichlet condition imposed upon it. Boundary conditions of the form $\phi_n + \alpha\phi = 0$, $\alpha \in \mathbb{C}$ are used in fluids to represent dissipative surfaces such as rough walls or porous membranes.

As a brief aside, if the cylinder were truncated and did not extend throughout the entire fluid depth, the reduction made in (9.19) would not be possible. Instead, one would have to expand the scattering potential outside the cylinder $r > a$ in a complete set of depth

modes (see Ref. 25, for example) which are defined by the infinite sequence of imaginary roots of (9.14). The additional depth modes in the series correspond to spatially evanescent or localised waves which decay exponentially away from the scatterer. As shown in Ref. 25 this solution would have to be matched to a series representation of the potential in the fluid region in $r < a$ no longer occupied by the vertical cylinder.

Although we shall make some small diversions along the way, the main thrust of this article will concern how to cloak a circular cylinder extending through the depth. By cloak, we shall mean that an observer suitably far away from the cylinder cannot detect the presence of the cylinder by monitoring the wave field. In other words, we require that there is no energy scattered in any direction away from the cylinder.

Returning to (9.16), (9.17) and letting $k_0 r \to \infty$ we have

$$\phi_s \sim \sqrt{\frac{2}{\pi k_0 r}} \mathrm{e}^{\mathrm{i}k_0 r - \mathrm{i}\pi/4} \mathcal{A}(\theta;\theta_0) \cosh k_0(z+h_0) \tag{9.20}$$

representing an outgoing circular wave field with amplitude in the θ direction due to an incident wave propagating in the direction θ_0 given by

$$\mathcal{A}(\theta;\theta_0) = -\sum_{n=-\infty}^{\infty} \frac{J_n'(k_0 a)}{H_n^{(1)'}(k_0 a)} \mathrm{e}^{\mathrm{i}n(\theta-\theta_0)} \tag{9.21}$$

after using the large argument asymptotics of the Hankel function. The total scattered energy — or scattering cross-section — is defined here as

$$\sigma = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\mathcal{A}(\theta;\theta_0)|^2 \, \mathrm{d}\theta = -\Re\{\mathcal{A}(\theta_0;\theta_0)\}. \tag{9.22}$$

The final equality, originally derived by Ref. 21 in the water wave context, is familiar in physics and known as the optical theorem. Either of the two definitions of σ in (9.22) can by used with (9.20)

to show that

$$\sigma_{cyl} = \sum_{n=-\infty}^{\infty} \left| \frac{J_n'(k_0 a)}{H_n^{(1)'}(k_0 a)} \right|^2 \tag{9.23}$$

and this is never zero for any $k_0 a$, i.e. a cylinder will always scatter energy (e.g. see Fig. 9.3(a)).

We now consider a range of possibilities for cloaking a cylinder. Following the pioneering papers of Refs. 19 and 30 the principle idea in cloaking is to alter the material parameters of a property within a domain exterior to the object being cloaked in such a way to bend waves around that object. There are a number of possibilities that will be considered in this regard. A natural starting point is to exploit the fact that waves refract under a change in depth; one can observe oblique waves straighten up as the approach the shoreline of a shallow beach.

Much of the complication we encounter lies in the fact that the water wave problem is inherently three-dimensional with one special direction (the vertical) and has boundaries defined by the free surface and the sea bed.

Thus, initially we shall consider a number of reduced models — approximations to the governing equations — which remove this complication.

9.2. Two-Dimensional Approximations to the Full Theory

9.2.1. *Ray theory*

We can consider an approximation to the three-dimensional water wave scattering problem based on ray theory (or geometric optics). In this approach, the effects of refraction due to changes in the depth are captured but diffraction is not. The basis of the approximation is that the wavelength is assumed to be much smaller than the length scale of horizontal bottom variations. In other words, $|\nabla_h h|/kh \ll 1$. Then (following Ref. 24, Sections 3.1, 3.2 for example), a multiple scales approach can be used to derive the Eikonal equation familiar

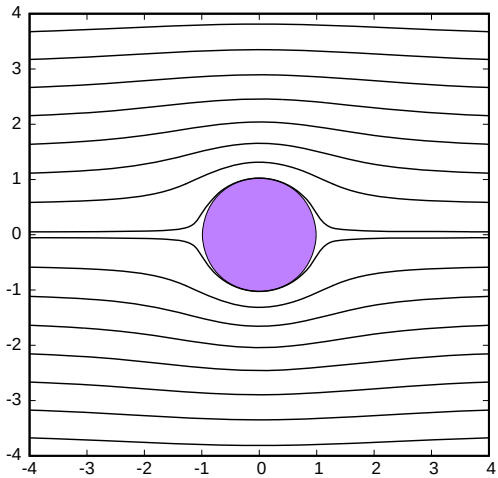


Fig. 9.1. Ray paths according to (9.26) around the cylinder radius $a = 1$.

in optics and acoustics:

$$S_x^2 + S_y^2 = k^2 \tag{9.24}$$

where k is the wavenumber satisfying (9.14). Here, $S(x, y) = constant$ represent lines of constant phase and wave rays follow paths that are perpendicular to these lines. Normally one might prescribe a depth variation $h(x, y)$ which, through (9.14) would define a wavenumber variation $k = k(h)$ and then solve (9.24) for S and hence determine the ray paths — see [Ref. 24, Sections 3.3, 3.4] for examples of how this done.

However we can also consider an inverse approach whereby the function $S(x, y)$ is prescribed and we infer a depth variation from (9.14). To cloak a cylinder of radius a, rays will need to be bent around the cylinder by changes in depth and so we will require $S \to k_0 x$ at large distances and $S_r = 0$ on $r = a$. That is, the rays are to be parallel and aligned with the x direction at large distances where the depth is assumed to tend to h_0 whilst lines of constant phase are perpendicular to the cylinder on its surface. The simplest (but not the only) choice of function satisfying these requirements is, in polar

coordinates,

$$S(r,\theta) = k_0 r\cos\theta + k_0 a^2\cos\theta/r \tag{9.25}$$

which is equivalent to the potential for the streaming flow of an inviscid irrotational fluid past a circle. Since S is a harmonic function the conjugate function provides us with the ray paths. Thus these are given by

$$k_0 r\sin\theta - k_0 a^2\sin\theta/r = constant. \tag{9.26}$$

It follows from using (9.25) in (9.24) that

$$k^2 = k_0^2\left(1 + \frac{a^4}{r^4} - \frac{2a^2}{r^2}\cos 2\theta\right) \equiv k_0^2 F^2. \tag{9.27}$$

Although there is no restriction on the range of values of kh in this theory, if we make the assumption $kh \ll 1$ everywhere in the domain then we can use the shallow water dispersion relation to give

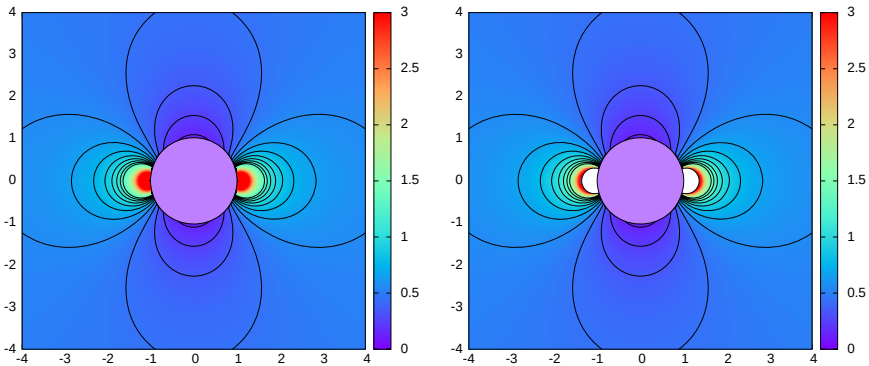
$$\frac{h}{h_0} = \frac{1}{F^2}. \tag{9.28}$$

This formula predicts infinite depth at the 'singular points' $(\pm a, 0)$ fore and aft of the cylinder. On physical reasoning this is to be expected and these are also points singularities in other solutions to cloaking problems including those proposed by Ref. 30. The main difficulty here is that $|\nabla_h h| \not\ll kh$ nor $kh \not\ll 1$ in large areas fore and aft of the cylinder and the basis of the approximation is violated. See Fig. 9.2(a) where the red coloured lobes represent depths greater than six times the depth in the far-field.

Using the full dispersion relation (9.14) instead of the shallow water version allows us to extend the range of values kh can take and then

$$h = \frac{1}{k_0 F}\tanh^{-1}\left(\frac{\tanh k_0 h_0}{F}\right). \tag{9.29}$$

Now h is undefined for values of (r,θ) such that $F < \tanh k_0 h_0$. These emerge fore and aft of the cylinder — in Fig. 9.2(b) they are represented by the two white lobes.



(a) (b)

Fig. 9.2. The depth profiles around the cylinder radius $a = 1$, $h_0 = \frac{1}{2}$ under ray theory assuming: (a) shallow water defined by (9.28); (b) unrestricted depth defined by (9.29) with $k_0 = 1$.

Other more exotic versions of the phase function can be used (S is certainly not required to be harmonic) provided they satisfy the two requirements stated earlier. Numerical experiments suggest that issues of large gradients and regions of undefined depth that invalidate these solutions cannot be overcome.

9.2.2. *Shallow water equations*

There are two common ways of deriving the linearised shallow water equations (or linearised long wave equations). One is to return to principles of fluid dynamics and make a shallow depth approximation from the outset. We shall adopt this approach later. The other is to apply the approximation directly to the full three-dimensional governing equations presented in Section 9.1.1. For a formal derivation based on this latter approach one can follow Ref. 24, Section 4.1 who rescales coordinates on physical lengthscales and then performs an asymptotic expansion of the potential in a small parameter $|\nabla_h h|/kh$. This is all performed under the assumption $kh \ll 1$ so that (9.14) is replaced locally by $k^2 \approx \omega^2/(gh)$ and gives rise to the linear Shallow Water Equation

$$\nabla_h \cdot (h \nabla_h \eta) + \frac{\omega^2}{g} \eta = 0 \tag{9.30}$$

for the surface wave elevation, $\eta(x, y)$ defined by (9.12). If the depth is constant and $h = h_0$ then (9.30) reduces to the two-dimensional wave equation

$$\nabla_h^2 \eta + k_0^2 \eta = 0 \tag{9.31}$$

which we have seen in (9.18) can be applied at any depth, not just when $k_0 h_0 \ll 1$ and waves propagating in the positive x direction are given by $\eta = \mathrm{e}^{\mathrm{i}k_0 x}$.

By writing $\eta = h^{-1/2}\overline{\eta}$ we can transform (9.30) into its canonical form

$$\nabla_h^2 \overline{\eta} + k_0^2 n^2 \overline{\eta} = 0 \tag{9.32}$$

where

$$n^2 = \frac{h_0}{h}\left(1 + \frac{|\nabla_h h|^2}{4k^2h^2} - \frac{\nabla_h^2 h}{2k^2 h}\right) \tag{9.33}$$

acts as a refractive index dependent upon the depth.

9.2.3. *Conformal mapping*

Following Ref. 19 we can introduce a conformal mapping from the physical (x, y) plane into a new (u, v) plane via $\beta = f(\xi)$ where $\beta = u + \mathrm{i}v$ and $\xi = x + \mathrm{i}y$. Under the transformation,

$$\nabla_h^2 \equiv \partial_\xi \partial_{\xi^*} = |f'(\xi)|^2 \partial_\beta \partial_{\beta^*} \equiv |f'(\xi)|^2 \widetilde{\nabla}_h^2 \tag{9.34}$$

where $\widetilde{\nabla}_h^2 = \partial_{uu} + \partial_{vv}$. Thus (9.34) is mapped to

$$\widetilde{\nabla}_h^2 \widetilde{\eta} + k_0^2 \widetilde{n}^2 \widetilde{\eta} = 0 \tag{9.35}$$

where $\widetilde{\eta}(u, v) \equiv \overline{\eta}(x, y)$ and

$$\widetilde{n}^2(u, v) = n^2(x, y)/|f'(\xi)|^2 \tag{9.36}$$

and so the transformation preserves the structure of the Shallow Water Equation whilst the refractive index is transformed by the mapping.

In terms of cloaking a cylinder, we can introduce a conformal transformation

$$f(\xi) = \xi + a^2/\xi \tag{9.37}$$

which maps the cylinder $|\xi| = a$ onto the line $-2a < u < 2a$, $v = 0$ in the β-plane. Thus, if $h = h_0$ in the mapped system so that $\widetilde{n} = 1$ with the solution $\widetilde{\eta}(u, v) = \mathrm{e}^{\mathrm{i}k_0 u}$ representing waves propagating without distortion past the line $-2a < u < 2a$, $v = 0$, then the mapping (9.37) gives us the Shallow Water Equation (9.35) for waves deflected past a cylinder of radius a with a depth profile h defined by

$$|1 - a^2/\xi^2|^2 = \frac{h_0}{h}\left(1 + \frac{|\nabla_h h|^2}{4k^2h^2} - \frac{\nabla_h^2 h}{2k^2 h}\right). \tag{9.38}$$

This nonlinear PDE would have to be solved for h outside $r = a$. It is not clear (at least to the author) how this would be done, and it may not be possible to define a function h which satisfies (9.38) everywhere.

We note in passing that if one chooses to return to (9.33) and make it more tractable by neglecting derivatives of h (justified, perhaps, because it shallow water approximation demands small gradients) then the right-hand side term in (9.38) becomes h_0/h and so with $\xi = r\mathrm{e}^{\mathrm{i}\theta}$, (9.38) is the same as (9.28) which we arrived at by taking a shallow water approximation to ray theory. This is an unsurprising outcome.

As was observed in Section 2.1 the difficulty here is that the map (9.37) which transforms a line in Cartesian coordinates to the exterior of the circle creates variations in the depth which violate the assumptions of the model.

We remark however that this method can be useful in other applications where one wishes to control waves. Recently, for example, Ref. 40 have used this method to bend, focus and directionally radiate waves. Also recently, Ref. 5 has used conformal mappings of parallel waveguides into meandering waveguides. Here the distortion to the bed created by the mapping is small and the argument that higher order terms in (9.38) can be approximately neglected is valid. Thus

$h = h_0/|f'(\xi)|^2$ is exactly the mapping of Ref. 5 to show that waves can be perfectly transmitted with no reflection along a meandering waveguide.

We return to the Shallow Water Equations later.

9.2.4. *The Mild-Slope Equations*

This is a good point at which to briefly mention an alternative reduced model which is more sophisticated than the shallow water approximation but retains its structure. Crucially, it is valid for all kh although still restricted by $|\nabla_h h|/kh \ll 1$. Thus the limitation $kh \ll 1$ of shallow water theory is removed. As in shallow water theory, the basis of the approximation is that bed gradients are small enough to suppose the bed is locally flat. Thus, whilst h varies globally, locally we assume the separable representation, inspired by the fact that (9.18) is exact for a flat bed,

$$\phi(x, y, z) \approx \psi(x, y) \cosh k(z + h) \tag{9.39}$$

where $k = k(h(x, y))$ is now determined locally by the dispersion relation (9.14). There are many derivations, variants and extensions of the Mild-Slope Equations (MSEs), originally attributed to Refs. 3 and 37. This partly reflects the rather ad hoc nature of the approximation. See Refs. 24, Section 3.5 for a derivation which shows the depth averaging process in action. For a more formal approach which is underpinned by a variational principle, see Ref. 7. That work resulted in the so-called Modified Mild-Slope Equations (MMSE) which retained terms proportional to the gradient and curvature of the bed neglected in previous derivations. Thus, the MMSE is given by

$$\nabla_h \cdot (u_0 \nabla_h \psi) + (k^2 + u_1 |\nabla_h h|^2 + u_2 \nabla_h^2 h)\psi = 0 \tag{9.40}$$

where $u_0 = cc_g$ (the product of phase speed and group velocity), u_1 and u_2 are defined in Ref. 7. and restated in Ref. 24, Section 3.5 in terms of the local depth h; the surface elevation η is a scaled function of ψ. As in Section 9.2.2 a transformation into canonical form can be

achieved by writing $\psi = u_0^{-1/2}\overline{\psi}$ so that (9.40) becomes

$$\nabla_h^2\overline{\psi} + k_0^2 n^2 \overline{\psi} = 0 \tag{9.41}$$

in which the refractive index n now given by

$$n^2 = \frac{k^2}{k_0^2}\left(1 + A|\nabla_h h|^2 + B\nabla_h^2 h\right) \tag{9.42}$$

and A and B are known but complicated functions of h (see Ref. 14 for example). Comparison of (9.32), (9.33) with (9.41), (9.42) shows that the structure of the Shallow Water Equation and the MMSE are identical and the equations differ only in the complexity of the coefficients multiplying the higher order terms. In principle, this means that any problem that is considered under the Shallow Water Equations should also be considered under a mild-slope approximation where the range of values of kh is unrestricted.

We shall also introduce another form the the MMSE later on.

9.3. Transformation Media Approach

The approach taken by Ref. 30 to cloaking has since seen widespread use. As in the case of Ref. 19 it is also based on a mapping between a space where waves propagate uninterrupted to a distorted space surrounding an object. It had previously been shown in Ref. 41 how Maxwell's equations of electromagnetics were invariant under a coordinate transformation, provided material parameters (permittivity and permeability in this case) could be interpreted as tensors encoding a spatially-varying anisotropic medium. We shall shortly see how the same approach can be applied in the water wave problem and the difficulties it introduces.

Soon after the paper of Refs. 30, 41 showed how such material parameters could be achieved using sub-wavelength split ring resonators; this was experimentally demonstrated a year later in Ref. 38. Thus the modern science of metamaterials was born and it has developed rapidly since. We shall discuss such structures in the water wave

context shortly — the point is that it is no good devising a cloak if it cannot be realised.

We shall work with the Shallow Water Equations initially but will comment on the MMSE towards the end, having already made the point in Section 9.2.4 that what you can do for $kh \ll 1$ should extend to all kh under an equivalent MSE model.

9.3.1. *The Shallow Water Equations revisited*

It helps to return to the derivation of the Shallow Water Equations from first principles (see Ref. 24, Section 3.5). Thus, if the wavelength is long compared to the depth and the bed gradients are small compared to the wavelength it is resonable to assume that the fluid velocity vector is approximately two-dimensional, or $\mathbf{u}(x, y, z, t) \approx (\mathbf{v}(x, y, t), 0)$ where $\mathbf{v} = (u(x, y, t), v(x, y, t))$. That is, there is negligible dependence on the depth and negligible vertical velocity. If the depth is $h(x, y)$ and the surface elevation is $\zeta(x, y, t)$ and $|\zeta| \ll h$ then conservation of mass is expressed as

$$\rho\zeta_t + \nabla_h \cdot (\rho h\mathbf{v}) = 0 \tag{9.43}$$

where ρ is the fluid density, $\rho h\mathbf{v}$ is the flux and the momentum equation is

$$\rho\mathbf{v}_t = -\nabla_h(\rho g\zeta). \tag{9.44}$$

and the pressure is approximately hydrostatic. Combining (9.43) and (9.44) gives

$$\rho\zeta_{tt} = \nabla_h \cdot (h\nabla_h(\rho g\zeta)) \tag{9.45}$$

and after assuming time harmonic motion with angular frequency ω with $\rho g\zeta = \Re\{\eta \mathrm{e}^{-\mathrm{i}\omega t}\}$ we return to

$$\nabla_h \cdot (h\nabla_h\eta) + \frac{\omega^2}{g}\eta = 0 \tag{9.46}$$

as in (9.30). Now imagine waves travelling in different directions are able to experience different material properties. For example, imagine

the fluid depth h to be multi-valued so that waves travelling in the x-direction experience a depth $h_1(x, y)$ whilst waves travelling in the y-direction experience a depth $h_2(x, y)$. The same arguments could apply to either gravity or the density ρ. To illustrate the ideas let us assume that ρ and g are constant and h is allowed to be anisotropic in the manner suggested. Now mass conservation (9.43) needs to be modified to reflect the fluxes in x and y directions are different. Thus we can write

$$\zeta_t + \nabla_h \cdot (\mathsf{h}\mathbf{v}) = 0 \tag{9.47}$$

where h is a rank-2 tensor given by

$$\mathsf{h} = \begin{pmatrix} h_1 & 0 \\ 0 & h_2 \end{pmatrix} \tag{9.48}$$

and following through the derivation as before results in the Shallow Water Equation modified for anisotropic depth:

$$\nabla_h \cdot (\mathsf{h}\nabla_h \eta) + \frac{\omega^2}{g}\eta = 0. \tag{9.49}$$

We remark that full anisotropy can also be considered in which h has off-diagonal entries. This would require development of a medium in which fluid flow in one direction will induce mass flux in a perpendicular direction. Realisations of fluid metamaterials are considered in Section 9.3.4.

9.3.2. *Mapping*

We shall work in plane polar coordinates from now on. This means that anisotropy will apply in radial and angular directions rather than in Cartesian directions as considered in 9.3.1.

We start with waves propagating in the direction $\theta = 0$ over a flat bed $h = h_0$ without scattering obstacles. The governing equation (9.31), with the shallow water assumption $k_0 = \omega^2/(gh_0)$, written in

polars is

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \eta}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 \eta}{\partial \theta^2}+k_0^2\eta=0 \tag{9.50}$$

and the solution is $\eta(r,\theta)=\mathrm{e}^{\mathrm{i}k_0x}=\mathrm{e}^{\mathrm{i}k_0r\cos\theta}$.

Now consider a mapping $(r,\theta)\to(\varrho,\theta)$ where $r=f(\varrho)$ for $0<r<b$ and $r=\varrho$ otherwise, such that $r=b$ is mapped to $\varrho=b$ and $r=0$ is mapped to $\varrho=a$. Then (9.50) remains unchanged for $r>b$ and for $0<r<b$ it is mapped to

$$\frac{1}{f'f}\frac{\partial}{\partial \varrho}\left(h_0\frac{f}{f'}\frac{\partial \widetilde{\eta}}{\partial \varrho}\right)+\frac{1}{f^2}\frac{\partial}{\partial \theta}\left(h_0\frac{\partial \widetilde{\eta}}{\partial \theta}\right)+K\widetilde{\eta}=0 \tag{9.51}$$

for $a<\varrho<b$ where $\widetilde{\eta}(\varrho,\theta)=\eta(r,\theta)$.

9.3.3. *A linear map*

The map suggested by Ref. 30 is defined by

$$f(\varrho)=\frac{b(\varrho-a)}{b-a}. \tag{9.52}$$

Using this in (9.51) gives

$$\frac{1}{\varrho}\frac{\partial}{\partial \varrho}\left(\varrho h_1\frac{\partial \widetilde{\eta}}{\partial \varrho}\right)+\frac{1}{\varrho}\frac{\partial}{\partial \theta}\left(h_2\frac{1}{\varrho}\frac{\partial \widetilde{\eta}}{\partial \theta}\right)+\widetilde{K}\widetilde{\eta}=0 \tag{9.53}$$

where

$$h_1=h_0(1-a/\varrho),\quad h_2=h_0/(1-a/\varrho) \tag{9.54}$$

and

$$\widetilde{K}=K(1-a/\varrho)/(1-a/b)^2. \tag{9.55}$$

We see that (9.53) is a Shallow Water Equation in the form given by (9.49) in polar coordinates where h_1 and h_2 are spatially-varying anisotropic depths experienced by waves moving in radial and angular directions respectively.

On the edge of the cloak, the conditions in the unmapped system are that η and $h_0\eta_r$ are continuous across $r = b$. Under the general mapping $r = f(\varrho)$ this implies $\widetilde{\eta}$ is continuous and

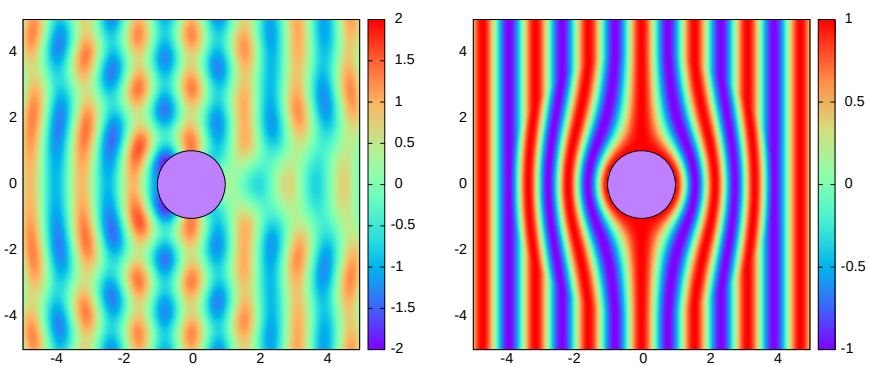
$$\frac{h_0}{f'}\widetilde{\eta}_\varrho|_{\varrho=b^-} = h_0\widetilde{\eta}_\varrho|_{\varrho=b^+}. \tag{9.56}$$

For the map (9.53) this reduces to

$$h_1\widetilde{\eta}_\varrho|_{\varrho=b^-} = h_0\widetilde{\eta}_\varrho|_{\varrho=b^+} \tag{9.57}$$

which is the exactly the physical shallow water flux condition required of $\widetilde{\eta}$ at the edge of the cloak. On the cylinder, $\varrho = a$, $h_1\widetilde{\eta}_\varrho = 0$ which confirms no flux into the cylinder in the mapped problem. A snapshot in time of the surface elevation for the linear transformation is shown in Fig. 9.3.

In the context of water waves Ref. 11 performed a transformation of the Laplacian ∇_h^2 using homogenisation theory in which the cloaking region $a < \varrho < b$ is filled with a fluid metamaterial consisting of an annular array of narrow vertical posts embedded in the fluid. They were able to show how the distribution of posts could be varied to mimic a diagonal tensor h with a given radial variation. It is not clear if Ref. 11 were using their metamaterial annular array of posts



(a) (b)

Fig. 9.3. (a) Snapshot of surface wave elevation for scattering by an uncloaked vertical cylinder ($a = 1$, $k_0 = 4$); and (b) perfect cloaking under a linear transformation with depths (9.54) and gravity implied by (9.55) with cloak size $b = 4$.

to mimic an effective anisotropic depth, or appealing to a different underlying physical mechanism to vary the phase speed within the cloak. In any case, Ref. 11 proposed a different prescription of the diagonal tensor entries to (9.54) and held their value of $\widetilde{K}$ fixed. Nevertheless, they produced experiments to show that a degree of cloaking could be achieved.

The mapped variables (9.54) and (9.55) are identical to those given in Ref. 9 who considered cloaking in two-dimensional acoustics. In this case the governing equations are analagous to the Shallow Water Equations with density replacing depth and the bulk modulus replacing gravity. They are also the same as TE-polarized waves in electromagnetics specified by Ref. 30 with permeability and permittivity taking the place of depth and the reciprocal of gravity.

9.3.4. *Metamaterial depth and gravity*

Under the linear mapping of the previous section (9.55) shows that a cloak requires an effective gravity given by

$$\widetilde{g} = g(1 - a/b)^2/(1 - a/\varrho). \tag{9.58}$$

Such an effect can be realised in water waves since gravity enters the equations of motion through one of the two conditions on the free surface. Thus, one can load the free surface to mimic the effect of changing gravity. This has been done in early papers modelling the effect of floating broken ice on the ocean surface. For example, Ref. 17 show that if the surface is loaded with a floating mass m, the term in ω^2/g occupied by gravity is replaced by

$$g(1 - m\omega^2/\rho g). \tag{9.59}$$

By allowing $m(\varrho) \geq 0$ to vary, (9.59) implies a reduction in the effective gravity. Unfortunately the specification in (9.58) requires gravity to vary above and below the reference value of g as $a < \varrho < b$.

It is also worth mentioning the work of Ref. 15 who demonstrated that doubly periodic sub-wavelength arrays of split ring resonators acting as Helmholtz resonators could be used to change the effective gravity felt by waves.

With regard to realising an anisotropic metamaterial depth, we have already referred to the method used by Ref. 11 in which homogenisation theory applied a sub-wavelength array of vertical posts can mimic the effects of anisotropic depth variations. The same methodology was applied by Ref. 10 in creating an invisiblity carpet for water waves in a channel: a device for hiding an obstacle placed in front of a plane wall from the far-field observer.

An alternative sub-wavelength realisation of a metamaterial water depth, again developed using homogenisation theory, was presented by Ref. 4. They considered how to redirect a propagating water wave through an angled junction in a parallel-walled waveguide without reflection. Crucially their transformation was volume-preserving with the implication that gravity is unaffected by the mapping. The physical realisation of their metamaterial depth was comprised of a microstructured corrugated rectangular bed profile. A similar technique was used by Ref. 8 who demonstrated how a metamaterial depth could be used to rotate an incident wave field.

A further possible realisation of the anisotropic depth, such as the one demanded by (9.54), can be performed using two interlocking arrays of thin closely-spaced vertical fins. One set of fins are arranged radially so that their height follows the prescription given by the function h_2 and the second set are arranged in a circular pattern with heights following the function h_1. Then waves travelling radially are not influenced by the radial fins but experience the depth profile of the circular fins and vice versa.

9.3.5. *A nonlinear map*

More recently, Ref. 42 suggested an alternative mapping function for the cloak in order to cope with the difficulty of having to alter gravity this arising from the linear map. Instead of (9.52) they used

$$f(\varrho) = b\sqrt{\frac{\varrho^2 - a^2}{b^2 - a^2}} \tag{9.60}$$

in (9.51) and this results in (9.53) with

$$h_1 = h_0(1 - a^2/b^2)(1 - a^2/\varrho^2), \quad h_2 = h_0(1 - a^2/b^2)/(1 - a^2/\varrho^2) \tag{9.61}$$

and now $\widetilde{K} = K$. That is, there appears to be no requirement that gravity is altered under this nonlinear mapping.

However, on the boundary of the cloak the flux matching condition (9.56) gives

$$\frac{h_0}{1 - a^2/b^2}\widetilde{\eta}_\varrho|_{\varrho=b^-} = h_0\widetilde{\eta}_\varrho|_{\varrho=b^+} \tag{9.62}$$

and the prefactor on the left-hand side is *not* $h_1(b)$ as is required for a physical flux condition, i.e. there is a mismatch in the flux at the boundary of the cloak of $(1 - a^2/b^2)$.

One way to overcome this difficulty is to rescale h_1 and h_2 in (9.61) by a factor of $(1 - a^2/b^2)$. In doing so, one must also rescale K and so we return to (9.53) with a transformed value of

$$\widetilde{K} = K/(1 - a^2/b^2). \tag{9.63}$$

This implies gravity should be rescaled in the cloak by a constant factor of $(1-a^2/b^2)$. This reduction in effective gravity can be implemented using the mass loading solution proposed in Section 9.3.4.

In Ref. 42 they argue instead that the flux mismatch factor of $(1-a^2/b^2)$ is small when $b \gg a$ and computations show that cloaking improves significantly as the b/a, the size of the cloaking increases. This will be confirmed by the calculations below.

As Ref. 42 show, the effect of the flux mismatch on scattering of waves can be analysed directly by solving the problem analytically. A plane incident wave from infinity is scattered by a cloak defined by the functions in (9.61) when the physical flux condition (9.57) is enforced. Solutions outside the cloak where the depth is h_0 are written

$$\widetilde{\eta}(\varrho, \theta) = \sum_{n=-\infty}^{\infty} (\mathrm{i}^n J_n(k_0\varrho) + a_n H_n^{(1)}(k_0\varrho))\mathrm{e}^{\mathrm{i}n\theta}, \quad \varrho > b \tag{9.64}$$

as in Section 9.1.3 for the problem of scattering by a cylinder and a_n are scattering coefficients to be determined. If the cloak were perfect, all a_n would be zero.

In $a < \varrho < b$ the most general solution to (9.53), which is bounded on the cylinder $\varrho = a$ can be written

$$\widetilde{\eta}(\varrho, \theta) = \sum_{n=-\infty}^{\infty} b_n J_n \left(k_0 b \sqrt{\frac{\varrho^2 - a^2}{b^2 - a^2}} \right) \mathrm{e}^{\mathrm{i}n\theta}. \tag{9.65}$$

The flux into the cylinder is $\lim_{\varrho \to a}(h_1 \widetilde{\eta}_\varrho) = 0$ as required. It is also interesting to note that $\widetilde{\eta}(a, \theta)$, the complex wave amplitude around the cylinder, is constant.

Matching $\widetilde{\eta}(b, \theta)$ for $-\pi < \theta < \pi$ from (9.64) and (9.65) gives

$$\mathrm{i}^n J_n(k_0 b) + a_n H_n^{(1)}(k_0 b) = b_n J_n(k_0 b). \tag{9.66}$$

Then we apply (9.57) to (9.64) and (9.65) to get

$$\mathrm{i}^n J_n'(k_0 b) + a_n H_n^{(1)\prime}(k_0 b) = b_n (1 - a^2/b^2) J_n'(k_0 b). \tag{9.67}$$

Eliminating b_n gives

$$a_n = \frac{\mathrm{i}^n (a^2/b^2) J_n'(k_0 b) J_n(k_0 b)}{-2\mathrm{i}/(\pi k_0 b) - (a^2/b^2) J_n'(k_0 b) H_n^{(1)}(k_0 b)} \tag{9.68}$$

after using a Wronskian relation for Bessel functions. If the flux mismatch were not present the factor $(1 - a^2/b^2)$, on the right-hand side of (9.67) would be replaced by 1 and this would result in $a_n = 0$ for all n whilst $b_n = \mathrm{i}^n$ — equivalent to a perfect cloak.

We remark that the cloaking solution above is not restricted to plane waves. For example, the problem of a wave source placed at $r = c$, $\theta = 0$ outside the cloak can be considered by replacing the Jacobi-Anger representation of a plane wave train in (9.15) by

$$\sum_{n=-\infty}^{\infty} H_n^{(1)}(k_0 r_>) J_n(k_0 r_<) \mathrm{e}^{\mathrm{i}n\theta} \tag{9.69}$$

using Graf's addition theorem where $r_< = \min\{r, c\}$, $r_> = \max\{r, c\}$. The result is that the factor i^n carried throughout the previous calculation is replaced by $H_n^{(1)}(k_0 c)$.

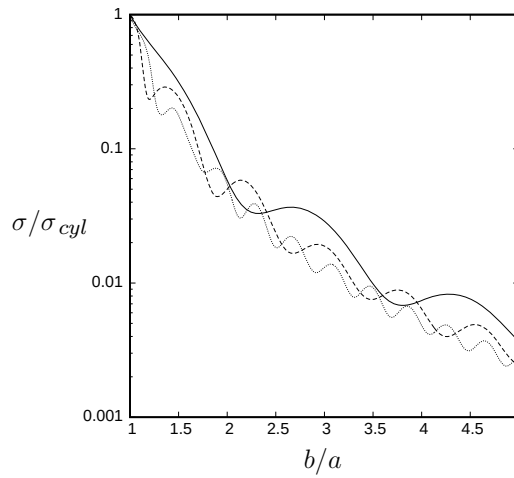


Fig. 9.4. Normalised cloaking factor σ/σ_{cyl} for the nonlinear mapping against cloak size b/a with $k_0 a = 1$ (solid), 2 (dashed) and 4 (dotted).

From (9.64), (9.22) the scattering cross-section is

$$\sigma = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} |a_n|^2 \tag{9.70}$$

and hence $\sigma = O((b/a)^{-4})$. Calculations confirm this decay rate and Fig. 9.4 shows the variation of σ normalised against the scattering cross section of an uncloaked cylinder, σ_{cyl}, given by (9.23) for different incident wavelengths. Snapshots of the surface elevation for two cloaks of sizes $b = 2a$ and $b = 4a$ with are shown in Fig. 9.5.

9.3.6. *A note on the Mild-Slope Equations*

In Section 3 we introduced the Modified Mild-Slope Equation (MMSE), given in its original form by (9.40) and in its canonical form by (9.41), (9.42). However, a different mapping of the variable $\psi(x, y)$, proportional to the free surface, was introduced by Ref. 32 using $\psi = s(h)\chi$ where $s(h)$ is defined in Ref. 32, Section 2 and results in the transformed MMSE

$$\nabla_h \cdot (k^{-2} \nabla_h \chi) + (1 - \nu(h) |\nabla_h h|^2) \chi = 0 \tag{9.71}$$

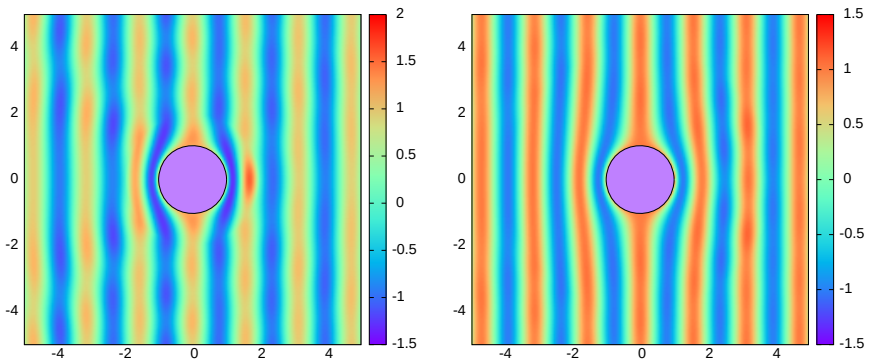


Fig. 9.5. Near-perfect cloaking under a nonlinear transformation with depths given by (9.61) and gravity unchanged, $a = 1$, $k_0 = 4$ and cloak sizes of (a) $b = 2$ and (b) $b = 4$.

where $v(h)$ is a relatively complicated, but explicit, function of the local depth. There are some advantages of this form of the MMSE. First, terms in the second derivative of h are eliminated and this means that jump conditions — required in the earlier versions of the MMSE — across discontinuities in the bed slope are redundant. Secondly, Ref. 32 calculates that $v(h) < 0.030$ for all $h > 0$ and so the final term in (9.71) is small. Moreover, as $v(h) \sim O((kh)^2)$ as $kh \to 0$ and it is easy to see that (9.71) tends to the Shallow Water Equations in this limit. When kh is large $v(h) \sim \mathrm{e}^{-kh}$ and, since $k \sim K$, (9.71) tends to a wave equation with no depth effects, as expected. As a result solutions of

$$\nabla_h \cdot (k^{-2} \nabla_h \chi) + \chi = 0 \tag{9.72}$$

are expected to be good approximations to solutions of (9.71). The advantage of using (9.72) is that its structure is aligned with the Shallow Water Equation (9.49), with k determined locally as a function of h via (9.14).

In particular the transformation media methods outlined in Section 9.3.2 can be applied to the MMSE in the form (9.72) and solutions will be valid not just in the limit of $kh \ll 1$ but for all kh.

For example, following Ref. 42 we could apply the transformation media method with the mapping (9.60) to the define a cloak with radial and angular wavenumbers

$$k_1 = \frac{k_0}{\sqrt{(1-a^2/b^2)(1-a^2/\varrho^2)}}, \quad k_2 = k_0\sqrt{\frac{1-a^2/\varrho^2}{1-a^2/b^2}}. \qquad (9.73)$$

The depths h_1 and h_2 would then be defined by inverting the dispersion relation: $h_i = \tanh^{-1}(K/k_i)/k_i$, $i = 1, 2$. It is anticipated that such a solution would apply beyond the limitations of shallow water theory although it would still suffer the flux mismatch problem discussed in Section 9.3.5 and would still be subject to the small-gradient constraint underpinning the mild-slope approximation.

9.3.7. *Other cloaking devices*

As demonstrated in the derivation of (9.46) the fluid density, whether or not it is anisotropic, is not a material parameter in the Shallow Water Equations (nor the full linear theory of Section 1). Thus spatial variations of density in the horizontal coordinates will not influence wave motion. Reference 16 suggest otherwise, although their governing equations do not appear to relate directly to the Shallow Water Equations. Vertical density stratification fundamentally alters the governing equations where it is normal to adopt the Boussinesq approximation — see Ref. 31, for example.

As previously suggested, instead of (or in addition to) varying the bathymetry, the free surface condition can be altered to control waves and a simple mass-loading model was proposed in Section 9.3.4. Another possibility is to place a thin flexible plate on the surface of the water to act as a cloak. The floating thin flexible plate model is widely used by researchers interested in the interaction between ice sheets and ocean waves; for a comprehensive review of work in this area, see Ref. 39. Problems involving waves in thin flexible plates surrounding vertical cylinders have previously been considered in the work of Ref. 6 and Ref. 20. Also see Ref. 2 for scattering problems involving flexible circular plates of varying thickness.

It is common to use Kirchhoff thin plate theory to model a flexible sheet. Even in the simpler problem of a flexible plate *in vacuo*, the fourth-order governing equations are not invariant under a coordinate transformation. The addition of a shallow water layer below the floating elastic plate complicates the underlying equations which increase in derivative order from four to six. For an elastic plate with variable thickness $D(x, y)$ over a variable bed $z = -h(x, y)$, Ref. 33 derive mild-slope equations for the reduced two-dimensional potential $\psi(x, y)$ in the fluid. Under the additional simplifying assumption of shallow water the governing equation for ψ is (Ref. [33, Eq. (4.9)])

$$(1 - \alpha + \mathcal{L})(\nabla_h \cdot ((h - d)\nabla_h \psi) + K\psi = 0 \tag{9.74}$$

where $\alpha = K(\rho_s/\rho)D(x, y)$ and ρ_s is the plate density and thickness and $z = -d(x, y)$ represents the underside of the floating plate. In the above

$$\mathcal{L}\psi = \nabla_h^2(\beta\nabla_h^2\psi) - (1 - \nu)(\beta_{xx}\psi_{xx} + \beta_{yy}\psi_{yy} - 2\beta_{xy}\psi_{xy}) \tag{9.75}$$

where $\beta = ED^3(x, y)/(12\rho g(1 - \nu^2))$ represents the bending stiffness in terms of Young's modulus, E, and Poisson's ratio, ν. Of course, one could keep D fixed and vary Young's modulus or Poisson's ratio or vary all three. We note that if the floating plate is removed so that $\alpha = 0$, $d = 0$ and $\mathcal{L} = 0$ in (9.74) then we simply return to the usual free surface Shallow Water Equation (9.46).

Reference 43 have suggested an alternative model for the floating plate equation which can be media transformed, following the earlier work of Ref. 12 on *in vacuo* thin flexible plates.

We remark that (9.74) and (9.75) are derived under the usual assumptions of isotropic elasticity. A derivation that incorporates anisotropic material properties should be performed to assess the feasibility of using a floating flexible plate as a cloak. This is currently the subject of an investigation by Zareei & Alam (personal communication).

9.4. Full Linear Theory

Thus far we have only considered approximations to full linear theory in which the complication of the depth dependence and lateral boundary conditions have been removed. Cloaking is the science of rendering objects invisible and there is no guarantee that a cloak developed under a reduced model will be effective in an unapproximated environment. One must be especially cautious when the limits on the assumptions forming the basis of an underlying model have been exceeded.

9.4.1. *Conformal mapping*

In this section the implementation of a mapping from (x, y, z) to (u, v, w) space in the full linearised water wave boundary-value problem is described. As in Section 9.2.3 a conformal mapping in the horizontal plane is introduced by writing $\beta = f(\xi)$ where $\beta = u + \mathrm{i}v$ and $\xi = x + \mathrm{i}y$. Laplace's equation in three-dimensions is preserved by a simultaneous rescaling of the vertical coordinate using $w = |f'|z$. That is $\nabla^2\phi$ is mapped into $\widetilde{\nabla}^2\widetilde{\phi} = 0$ where $\phi(x, y, z) \equiv \widetilde{\phi}(u, v, w)$ and $\widetilde{\nabla}^2 \equiv \partial_{uu} + \partial_{vv} + \partial_{ww}$. The mapped free surface boundary condition (9.9) is

$$\widetilde{\phi}_w - \widetilde{K}\widetilde{\phi} = 0, \quad \text{on } w = 0 \tag{9.76}$$

where $\widetilde{K} = K/|f'|$. The bed $z = -h(x, y)$ is mapped to the boundary $w = -\widetilde{h}(u, v) \equiv |f'|h$ and the transformation of the general bed condition (9.11) results in

$$\widetilde{\phi}_w + \{\widetilde{\nabla}_h\widetilde{h} + |f'|\widetilde{h}\widetilde{\nabla}_h(|f'|^{-1})\} \cdot \widetilde{\nabla}_h\widetilde{\phi} = 0, \quad \text{on } w = -\widetilde{h}. \tag{9.77}$$

In (x, y, z) space defined by a flat bed of depth h_0 in the presence of a thin impermeable vertical barrier $-2a < x < 2a$, $y = 0$, $-h_0 < z < 0$, plane waves propagating in the direction $\theta = 0$ are described by the potential ϕ_{inc} in (9.15). The mapping of the barrier to a cylinder in (u, v, w) space is given by

$$\xi = \beta + \frac{a^2}{\beta} \tag{9.78}$$

or $\beta = f(\xi) = \frac{1}{2}(\xi + \sqrt{\xi^2 - 4a^2})$. Then $f'(\xi) = \beta^2/(\beta^2 - a^2)$ and

$$|f'| = \frac{1}{\sqrt{1 + a^4/\varrho^4 - 2a^2 \cos 2\varphi/\varrho^2}} \tag{9.79}$$

when $\beta = \varrho e^{i\varphi}$. Apart from the singular points at $\beta = \pm a$, $\widetilde{h} = h_0|f'|$ defines a physically-realisable variable depth. Here $\widetilde{K} = K/|f'|$ represents a variation of effective gravity on the free surface which both falls and rises above the far-field value, g. Thus the mass-loading model described in Section 9.3.4 cannot be used to realise the surface condition (9.76).

When $h = h_0$ is a constant, the transformed bed condition (9.77) simplifies to

$$\widetilde{\phi}_w = 0, \quad \text{on } w = -\widetilde{h}(u, v) = -h_0|f'|. \tag{9.80}$$

This is *not* the usual condition (c.f. (9.11)) $\widetilde{\phi}_w + \widetilde{\nabla}_h \widetilde{h} \cdot \widetilde{\nabla}_h \widetilde{\phi} = 0$ required on a natural bed. Thus, in addition to the variable surface condition, one would also need to design a metamaterial bed to realise (9.80) in order to cloak a cylinder under full linear theory using the mapping proposed here. For example, such a bed could be formed by a vertical cascade of narrowly-spaced thin horizontal plates immersed within the fluid whose edges are designed to follow the profile $\widetilde{h} = h_0|f'|$ where f' is given by (9.79). Surface waves would feel the macroscopic effect of the variable bathymetry but locally the fluid would satisfy a vertically-directed no-flow condition.

The solution described above was initially formulated in Ref. 34 although the details and the description of the metamaterial bed in the paragraph above are new.

9.4.2. *A direct approach*

In this section the approach reported in Ref. 35 used to cloak a circular cylinder without relying on metamaterials is presented. Instead of designing a cloak which bends waves around the cylinder it relies on combining wave diffraction effects from the cylinder and an annular

region of variable bathymetry to destroy outgoing waves in all directions. This approach shares similarities with the one used by Ref. 1 in an electromagnetic context in which they coated a dielectric cylinder with an annular dielectric with constant properties and showed how the scattering cross-section could be reduced significantly.

The work published in Ref. 14, developed to investigate the focusing of water waves over a bathymetric lens using the MMSE, was adapted by Ref. 34 to include a vertical cylinder. The shape of the bed in the annular cloak $a < r < b$ was expanded in a weighted set of prescribed modal functions:

$$h(r,\theta) = \sum_{p=1}^{P} \sum_{q=1}^{Q} \alpha_{pq} f_p(r) \cos 2(q-1)\theta \tag{9.81}$$

in which

$$f_p(r) = T_{2p}\left(\frac{b-r}{b-a}\right) - (-1)^p \tag{9.82}$$

where T_p are Chebychev polynomials. This choice implies that $h_r = 0$ at $r = b$ so that there is no discontinuity in bed slope at the edge of the cloak, a feature which was chosen for convenience as it avoided the need to implement jump conditions in the MMSE solution. Chebychev polynomials were chosen because it was anticipated that greater resolution of the bed may be required close to the cylinder boundary $r = a$. The incident wave was assumed to be aligned with $\theta = 0$ and only even angular modes were used in the expansion (9.81) by appealing to ideas of time-reversal symmetry.

In (9.81) the $P \times Q$ coefficients α_{pq} were treated as free parameters in a numerical multi-parameter optimisation whose objective function to be minimised was the normalised "cloaking factor" C defined to be the scattering cross-section σ of the cloaked bed divided by σ_{cyl}, say, for uncloaked bed, given by (9.23).

In the studies of Ref. 35 results were presented for cloaking targeted at a particular frequency and geometry ($k_0 h_0 = 1$, $a/h_0 = \frac{1}{2}$) although other parameters were considered in numerical experiments leading to that publication.

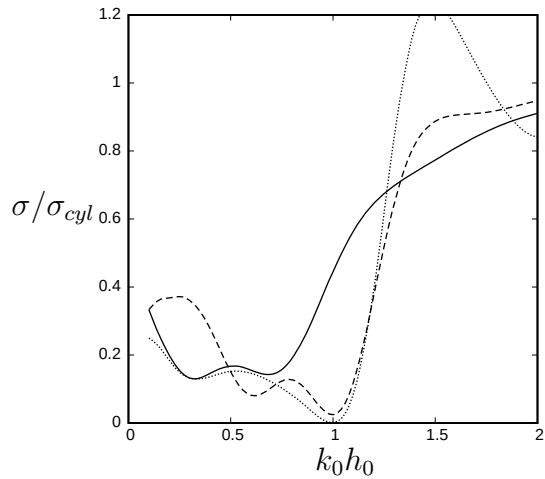


Fig. 9.6. Cloaking factor against wavenumber under full linear theory calculations for axisymmetric beds ($Q = 1$) optimised for cloaking at $k_0h_0 = 1$, $a/h_0 = \frac{1}{2}$, $b/h_0 = 5$ with $P = 2$ (solid), $P = 4$ (dashed) and $P = 8$ (dotted).

The initial work using the MMSE was presented in Ref. 34 suggested that cloaking factors could be progressively reduced towards zero as the number of degrees of freedom defining the bed were increased. A year later, Newman[26] published work adopting the same principles and methodology of Ref. 34 but using a modified version of the fully three-dimensional boundary element code, WAMIT. The work was combined in the paper of Ref. 35 where numerical results supported the hypothesis that the cloaking factor could be reduced to zero as the number of degrees of freedom in the bed were increased under full linear theory — see Fig. 9.6 for example. That work also alluded to the fact that cloaking was sensitive to small changes in the bathymetry and the comparison between MMSE and full linear theory was shown to be generally quite poor due to the large gradients predicted in cloaking beds. This comment serves to act as a cautionary note regarding the predictions made in Section 9.3 under depth-reduced models.

Figure 9.7(a) shows the effect of the size of the cloak on the cloaking factor for cloaking-optimised axisymmetric beds. Reference 35

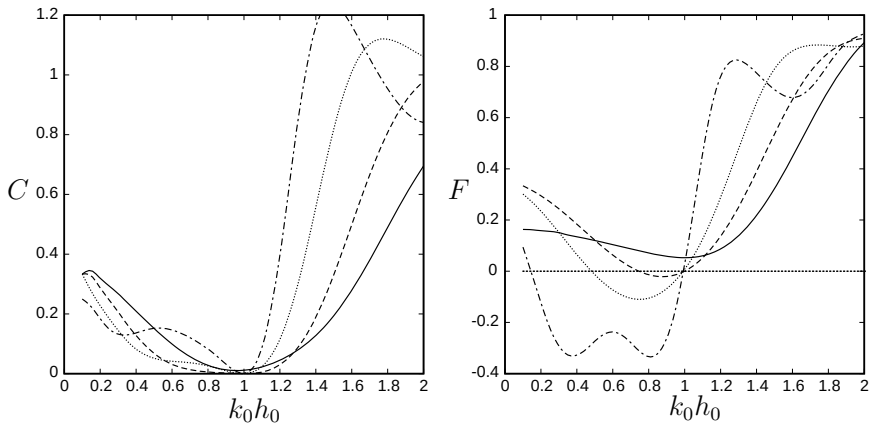


Fig. 9.7. In (a) the cloaking factor $C = \sigma/\sigma_{cyl}$ and in (b) the normalised mean drift force on the cylinder, F, both plotted against against wavenumber for an axisymmetric bed, $Q = 1$, $P = 8$ optimised to cloak at $k_0h_0 = 1$ with $a/h_0 = \frac{1}{2}$ and $b/h_0 = 2$ (solid), 3 (dashed), 4 (dotted), 5 (chained).

demonstrated that cloaking is best for $b \approx 5a$ whilst the cloaking effect can also seen to be broadbanded. Figure 9.7(b) illustrates the significant reduction in the mean second-order drift force on the cylinder (normalised against an uncloaked cylinder), especially around the cloaking wavenumber $k_0h_0 = 1$.

Reference 35 also showed that convergence towards perfect cloaking was improved by breaking axisymmetry and letting $Q > 1$. See Figs. 9.8.

9.4.3. *Extensions to the direct approach*

Newman later developed a double-precision version of the original single-precision code WAMIT to numerically investigate alternative ways of cloaking a cylinder under full linear theory.[27] This included, amongst many other examples, surrounding the vertical cylinder to be cloaked with a concentric circular ring of N fixed truncated surface-piercing circular cylinders. Again, a numerical optimisation procedure was used to tune a number of free parameters used to

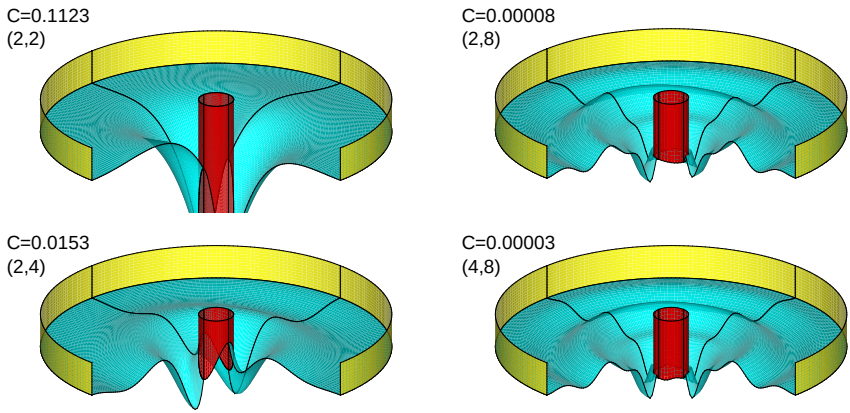


Fig. 9.8. Beds optimised to cloak at $k_0 h_0 = 1$ for different values of (Q, P) with $a/h_0 = \frac{1}{2}$, $b/h_0 = 5$. Values of cloaking factor $C = \sigma/\sigma_{cyl}$ achieved in each case are shown. The yellow boundary depicts the edge of cloak.

encode the geometric configuration of the cloak. Newman[27] was able to show that whilist an axisymmetric ring quickly reduced the cloaking factor as N increased, significant improvements to cloaking factors could be achieved by breaking the axisymmetry of the cloak, with values of $C = 1.1 \times 10^{-9}$ being reported.

The numerical evidence seemed to suggest that perfect cloaking is possible. However, M. McIver[23] provides a formal proof that, under certain conditions, perfect cloaking is not possible. The proof applies to a specific class of cloaks comprised of elements individually satisfying the so-called "John condition" (e.g. see Ref. 18), Geometrically, the John condition is satisfied if lines projected vertically downwards from all points on the free surface meet the flat bed without intersecting a body in the fluid. This is a powerful result. Unfortunately, Newman's[27] configuration of N surface-piercing truncated cylinders discussed in the previous paragraph does satisfy the John condition and so it transpires that perfect cloaking is not possible in this example despite the extremely low cloaking factors calculated. This also puts in doubt the suggestion that perfect cloaking is achievable in the earlier example of Ref. 35 involving variable bathymetry.

9.5. Summary

In this article, I have tried to present a range of different approaches that can be used to investigate the cloaking in water waves. This includes the simplified models of ray theory, shallow water theory and the mild-slope approximation in addition to the consideration of full linear wave theory. It has been shown how mapping methods may be applied within the framework of each of these models. It has also been shown how metamaterials can be designed in water wave problem to realise unnatural boundary conditions that emerge from these mappings such as anisotropic water depth and variable effective gravity.

The most successful attempts at creating a cloak for a circular cylinder seem to be: (i) using the nonlinear transformation of Ref. 42 shown in Section 9.3.5 within the framework of the shallow water approximation and (ii) numerical optimisation of bathymetry or other structures embedded in the fluid that form a cloak under full linear theory. The latter has the advantage of being exact and realisable without metamaterial parameters. The former has the advantage that they bend waves around the cylinder rather than relying on the cylinder for multiple interference effects that are integral to the cloak.

The transformation media approach and, more generally, the adoption of metamaterials in water waves can provide an interesting range of new possibilities to water wave problems. For example, one might consider mapping the well-known explicit Stokes edge wave solution (or its shallow water approximation) along a plane sloping beach under a coordinate transformation into circular coordinates to provide a solution in which waves are trapped to a circular island by a "metamaterial beach".

References

[1] Alù, A. and Engheta, N. (2005). Achieving transparency with plasmonic and metamaterial coatings. *Phys. Rev. E.* **72**, p. 016623.

[2] Bennetts, L. G., Biggs, N. R. T. and Porter, D. (2009). Wave scattering by an axisymmetric ice floe of varying thickness. *IMA. J. Appl. Math.* **74**, pp. 273–295.

[3] Berkhoff, J. C. W. (1973). Computation of combined refraction-diffraction. *Proc. 13th Conf. on Coastal Engng., July 1972, Vancouver, Canada*, vol. 2, pp. 471–490. ASCE.
[4] Berraquero, C. P., Maurel, A., Petitjeans, P. and Pagneux, V. (2013). Experimental realization of a water-wave metamaterial shifter. *Phys. Rev. E.* **88**, p. 051002.
[5] Bobinksi, T. M. (2016). *Metamateriaux pour les ondes a la surface de l'eau (Metamaterials for water waves)* (Ph.D. Thesis, Universite Pierre et Marie Curie.)
[6] Brockenhurst, P., Korobkin, A. A. and Parau, E. I. (2011). Hydroelastic wave diffraction by a vertical cylinder. *Phil. Trans. Roy. Soc. Lond. A.* **369**, pp. 2832–2851.
[7] Chamberlain, P. G. and Porter, D. (1995). The modified mild-slope equation. *J. Fluid Mech.* **291**, pp. 393–407.
[8] Chen, H., Yang, J., Zi, J. and Chan, C. T. (2009). Transformation media for linear liquid surface waves. *EuroPhys. Lett.* **85**, p. 24004.
[9] Cummer, S. A. and Schurig, D. (2007). One path to acoustic cloaking. *New J. Phys.* **9**(45), pp. 1–8.
[10] Dupont, G., Kimmoun, O., Molin, B., Guenneau, S. and Enoch, S. (2015). Numerical and experimental study of an invisibility carpet in a water channel. *Phys. Rev. E.* **91**, p. 023010.
[11] Farhat, M., Enoch, S., Guenneau, S. and Movchan, A. (2008). Broadband cylindrical acoustic cloak for linear surface waves in a fluid. *Phys. Rev. Lett.* **101**, p. 134501.
[12] Farhat, M., Guenneau, S. and Enoch, S. (2009). Ultrabroadband elastic cloak in thin plates. *Phys. Rev. Lett.* **103**, p. 024301.
[13] Farhat, M., Guenneau, S., Enoch, S. and Movchan, A. (2010). All-angle-negative-refraction and ultra-refraction for liquid surface waves in 2D phononic crystals. *J. Comp. Appl. Math.* **234**(6), pp. 2011–2019.
[14] Griffiths, L. S. and Porter, R. (2012). Focusing of surface waves by variable bathymetry. *Appl. Ocean Res.* **34**, pp. 150–163.
[15] Hu, X., Chan, C. T., Ho, K.-M. and Zi, J. (2011). Negative effective gravity in water waves by periodic resonator arrays. *Phys. Rev. Lett.* **106**, p. 174501.
[16] Iida, T. and Kashiwagi, M. (2016). Shallow water cloaking with anisotropic fluid. *Proc. 31st Int. Workshop on Water Waves and Floating Bodies, Michigan, USA*. pp. 65–68.
[17] Keller, J. B. and Goldstein, D. (1953). Water wave reflection due to surface tension and floating ice. *Trans. American Geophys. Union.* **34**(1), pp. 43–48.
[18] Kuznetsov, N., Maz'ya, V. and Vainberg, B. (2002). *Linear Water Waves.* Cambridge University Press.
[19] Leonhardt, U. (2006). Optical conformal mapping. *Science* **312**, p. 1777.
[20] Malenica, S. and Korobkin, A. A. (2003). Water wave diffraction by vertical circular cylinder in partially frozen sea. *Proc. 18th Int. Workshop on Water Waves and Floating Bodies*, Le Croisic, France.

[21] Maruo, H. (1960). The drift of a body floating on waves. *J. Ship Res.* **4**, pp. 1–10.
[22] McCamy, R. C. and Fuchs, R. A. (1954). Wave forces on a pile: A diffraction theory. *Tech. Memo.* **69**, U.S. Army Board, U.S. Army Corp. of Eng.
[23] McIver, M. (2014). The scattering properties of a system of structures in water waves. *Quart. J. Mech. Appl. Math.* **67**(4), pp. 631–639.
[24] Mei, C. C., Stiassnie, M. and Yue, D. K.-P. (2005). *Theory and Applications of Ocean Surface Waves. Part 1: Linear Aspects.* (World Scientific Publishing Co. Pte. Ltd.)
[25] Miles, J. W. and Gilbert, F. (1968). Scattering of gravity waves by a circular dock. *J. Fluid Mech.* **34**, pp. 783–793.
[26] Newman, J. N. (2012). Scattering by a cylinder with variable bathymetry. *Proc. 27th Int. Workshop on Water Waves and Floating Bodies*, Copenhagen, Denmark.
[27] Newman, J. N. (2014). Cloaking a circular cylinder in water waves. *Euro. J. Mech B/Fluids* **47**, pp. 145–150.
[28] Pendry, J. B. (2000). Negative refraction makes a perfect lens. *Phys. Rev. Lett.* **85**(18), pp. 3966–3969.
[29] Pendry, J. B., Holden, A. J., Robbins, D. J. and Stewart, W. J. (1999). Magnetism from conductors and enhanced nonlinear phenomena. *IEEE. Trans. Microwave Theory and Techniques* **47**(11), pp. 2075–2084.
[30] Pendry, J. B., Schurig, D. and Smith, D. R. (2006). Controlling electromagnetic fields. *Science* **312**, p. 1780.
[31] Phillips, O. M. (1977). *The Dynamics of the Upper Ocean (2nd Edition).* Cambridge University Press.
[32] Porter, D. (2003). The mild-slope equations. *J. Fluid Mech.* **494**, pp. 51–63.
[33] Porter, D. and Porter, R. (2005). Approximations to wave scattering by an ice sheet of variable thickness over undulating bed topography. *J. Fluid Mech.* **509**, pp. 145–179.
[34] Porter, R. (2011). Cloaking of a cylinder in waves. *Proc. 26th Int. Workshop on Water Waves and Floating Bodies, Athens, Greece.* pp. 169–173.
[35] Porter, R. and Newman, J. N. (2014). Cloaking of a vertical cylinder in waves using variable bathymetry. *J. Fluid Mech.* **750**, pp. 124–143.
[36] Shelby, R. A., Smith, D. R. and Schultz, S. (2001). Experimental verification of a negative index of refraction. *Phys. Rev. Lett.* **292**(6), pp. 77–79.
[37] Smith, R. and Sprinks, T. (1975). Scattering of surface waves by a conical island. *J. Fluid Mech.* **72**, pp. 373–384.
[38] Smith, D. R., Padilla, W. J., Vier, D. C., Nemat-Nasser, S. C. and Schultz, S. (2000). Composite medium with simultaneously negative permeability and permittivity. *Phys. Rev. Lett.* **84**(18), p. 4184.
[39] Squire, V. A. (2007). Of ocean waves and sea-ice revisited. *Cold Regions Sci. Tech.* **49**(2), pp. 110–133.
[40] Wang, Z., Zhang, P., Nie, X. and Zhang, Y. (2015). Manipulating water wave propagation via gradient index media. *Sci. Rep.* **5**, p. 16846.

[41] Ward, A. J. and Pendry, J. B. (1996). Refraction and geometry in Maxwell's equations *J. Mod. Opt.* **43**(4), pp. 773–793.
[42] Zareei, A. and Alam, M.-R. (2015). Cloaking in shallow-water waves via nonlinear medium transformation. *J. Fluid Mech.* **778**, pp. 273–287.
[43] Zareei, A. and Alam, R. (2016). Cloaking by a floating thin plate. *Proc. 31st Int. Workshop on Water Waves and Floating Bodies, Michigan, USA*. pp. 197–200.

CHAPTER 10

Molding Sound Propagation and Scattering with Acoustic Metamaterials and Metasurfaces

ROMAIN FLEURY and ANDREA ALÙ*

The University of Texas at Austin, USA

Acoustic metamaterials are artificially structured materials that are capable of interacting with acoustic waves in anomalous ways, leading to exotic physical properties, not found in natural materials. In this chapter, we review and highlight how the extreme sound interaction properties associated with acoustic metamaterials can be exploited to reach a new degree of acoustic wave control, enabling fascinating wave phenomena with large application potentials, including cloaking and sound manipulation via anomalous matching and tunneling effects. We discuss the analogue of plasmonic cloaking for acoustic waves, as well as exceptional matching phenomena that occur at the interface between a natural acoustic medium and a passive

*Corresponding author: alu@utexas.edu

acoustic metamaterial with near zero effective density. We also discuss active acoustic metasurfaces obeying parity-time symmetry conditions, which can induce anomalous tunneling effects with large reflection asymmetry, and demonstrate that such active planar surfaces have the potential to relax some of the limitations associated with bulk passive 3D metamaterials, including sensitivity to losses, bandwidth of operation and imperfect isotropy. Finally, the asymmetry in the transmission properties of acoustic meta-atoms is discussed within the context of linear nonreciprocal acoustics.

10.1. Introduction

Acoustic metamaterials are artificial materials that inherit their acoustic properties from their engineered structure rather than from their constituent materials.[1] They do not behave like anything that can be found in nature, and are described by constitutive parameters, like their density ρ or bulk modulus κ, which typically take unusual and extreme values, for instance negative,[2–9] zero,[10–12] or very large.[6] Due to their anomalous physical properties, metamaterials are associated with exotic wave phenomena, such as negative refraction,[13] subwavelength imaging,[14] and cloaking.[15,16] This chapter discusses a few fascinating phenomena associated with acoustic metamaterials, including unusual scattering features, such as cloaking and scattering cancellation, and unusual propagation features, such as anomalous tunneling and other exotic acoustic transmission phenomena, not found in conventional media. In the following, we first discuss several examples of cloaking effects based on a dominant scattering cancellation approach that leads to simple cloak designs employing either homogenous layers or ultrathin acoustic metasurfaces. Then, we also discuss how acoustic metamaterials can induce fascinating tunneling phenomena for sound, for instance acoustic supercoupling through density-near-zero acoustic channels and broadband acoustic energy funneling at the acoustic Brewster angle with a metamaterial buffer. Finally, we discuss the case of active metamaterials, and demonstrate how gain, fluid motion or the use of time-modulated systems can be used as additional degrees

of freedom to increase our ability to cloak sound and control its transmission via unique matching effects. These additional degrees of freedom will enable giant asymmetries in reflection, such as in the case of PT symmetric structures, and asymmetry in transmission, such as in the case of non-reciprocal acoustic systems.

10.2. Plasmonic and Mantle Acoustic Cloaking

Among the fascinating acoustic phenomena associated with metamaterials, cloaking is arguably one of the most intriguing. Cloaking refers to the possibility of cancelling the scattering from an object by placing it inside a metamaterial cloak. Initially, cloaking was proposed for electromagnetic waves as a way to turn objects invisible to the human eye or to radar systems,[17–20] and later it was extended to acoustic waves,[23] seeking invisibility to sonars or other kinds of acoustic imaging systems, or reduced disturbance in complex acoustic environments. There exist two dominant techniques within the science of cloaking. The first one was proposed independently by Pendry and Leonhardt in 2006,[18,19] and relies on transformation optics, a method that re-interprets wave propagation in curved geometrical spaces as propagation in inhomogeneous and anisotropic metamaterials. The second method was proposed in 2005,[17] and it was originally called plasmonic cloaking. Unlike transformation optics, plasmonic cloaking induces invisibility using a few homogenous and isotropic metamaterial layers, and it is arguably simpler to implement in practice. Later, a variant of plasmonic cloaking employing a single isotropic homogenous metasurface, called a mantle cloak, was developed.[22,23] More details on the history of cloaking techniques can be found in recent review papers on the topic, see Ref. 15 for instance for a complete overview of cloaking for both electromagnetic and acoustic waves, or Ref. 16 for a detailed comparison of all the available cloaking methods and a discussion of their respective limitations. Cloaking of acoustic waves is also thoroughly discussed in a recent review article by Norris.[23] We focus here specifically on the acoustic analogues of plasmonic and mantle cloaking,[24–28] i.e., a scattering cancellation method that is able to drastically reduce the scattering

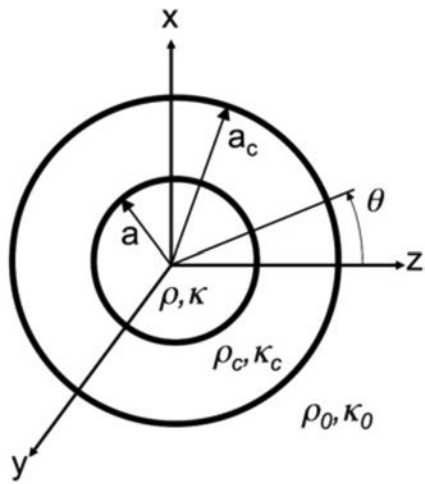


Fig. 10.1. We consider the scattering from an acoustic sphere of radius a surrounded by a spherical cloaking shell of radius a_c. Both the cloak and the object are assumed to be homogenous and isotropic. © 2006 Elsevier. Adapted with permission.[25]

of acoustic waves from a given object by using homogenous isotropic metamaterials[24–26] or metasurfaces.[27,28]

Consider the acoustic scatterer schematically shown in Fig. 10.1. It is composed of a spherical core made of a given acoustic material with density ρ and bulk modulus κ, and a spherical cover made of an acoustic medium with density ρ_c and bulk modulus κ_c. The object that we want to cloak is the spherical core, of radius a, and for this purpose we consider a metamaterial cloak in the form of an isotropic homogenous shell of external radius a_c. Assuming an incident time-harmonic plane wave of frequency ω propagating along $\vec{z}$, the pressure field scattered by this spherically-symmetry target may be written as a superposition of radiating spherical harmonics[29]

$$p_{sc} = \sum_{i=0}^{+\infty} i^n (2n+1) p_n A_n h_n^{(1)}(k_0 r) P_n(\cos\theta). \tag{10.1}$$

The coefficients p_n represent the generic (known) coefficients of the expansion of the incident wave in spherical harmonics, whereas

the coefficients A_n represent the scattering coefficients of the object, which are independent on the form of the incident field and only depend on the geometry of the object, its acoustic properties, and frequency. In Eq. (10.1), i is the imaginary number and $k_0 = \omega\sqrt{\rho_0/\kappa_0}$ is the wavenumber in the background. The scattering coefficients can be calculated by enforcing pressure and normal velocity continuity at the two spherical interfaces, yielding

$$A_n = -U_n/(U_n + iV_n), \tag{10.2}$$

with

$$U_n = \begin{vmatrix} -j_n(ka) & j_n(k_c a) & y_n(k_c a) & 0 \\ -\frac{1}{\rho} ka j_n'(ka) & \frac{1}{\rho_c} k_c a j_n'(k_c a) & \frac{1}{\rho_c} k_c a y_n'(k_c a) & 0 \\ 0 & j_n(k_c a_c) & y_n(k_c a_c) & j_n(k_0 a_c) \\ 0 & \frac{1}{\rho_c} k_c a_c j_n'(k_c a) & \frac{1}{\rho_c} k_c a_c y_n'(k_c a) & \frac{1}{\rho_0} k_0 a_c j_n'(k_0 a_c) \end{vmatrix} \tag{10.3}$$

$$V_n = \begin{vmatrix} -j_n(ka) & j_n(k_c a) & y_n(k_c a) & 0 \\ -\frac{1}{\rho} ka j_n'(ka) & \frac{1}{\rho_c} k_c a j_n'(k_c a) & \frac{1}{\rho_c} k_c a y_n'(k_c a) & 0 \\ 0 & j_n(k_c a_c) & y_n(k_c a_c) & y_n(k_0 a_c) \\ 0 & \frac{1}{\rho_c} k_c a_c j_n'(k_c a) & \frac{1}{\rho_c} k_c a_c y_n'(k_c a) & \frac{1}{\rho_0} k_0 a_c y_n'(k_0 a_c) \end{vmatrix} \tag{10.4}$$

Each scattering coefficient A_n quantifies how much scattering is expected from the object in the n-th spherical harmonic channel. The relevant metric to quantify the overall visibility is the total scattering cross-section, which equals the ratio between the total scattered power and the incident power flux

$$\sigma = \frac{4\pi}{k^2} \sum_{n=0}^{+\infty} (2n+1)|A_n|^2. \tag{10.5}$$

Mathematically speaking, a perfectly invisible object must have $\sigma = 0$, which implies that for any n, $A_n = 0$. However, for an object of

given acoustic size $k_0 a_c$, only contributions up to the order $N \approx k_0 a_c$ are non-negligible, and we can in practice truncate the sum in (10.5). This means that invisibility can be induced by canceling only a finite number of scattering coefficients, by focusing on the dominant scattering from the object. This is at the basis of plasmonic acoustic cloaking, which cancels each dominant scattering term in (10.5) by using N homogenous spherical shells. Although one can cloak arbitrarily large spherical objects with this technique, for demonstration purposes the example of Fig. 10.1 only considers a single homogenous isotropic metamaterial layer, and therefore is relevant for objects such as $k_0 a_c < 1$. Cancellation of a given harmonic occurs when the coefficient U_n in (10.2) vanishes, provided that under the same condition $V_n \neq 0$, since $V_n = 0$ corresponds to a resonance associated with enhanced scattering. In the long wavelength limit, the cancellation condition $U_n = 0$ can be found in close form to be

$$\left(\frac{a}{a_c}\right)^3 = \bar{\kappa}\frac{1-\bar{\kappa}_c}{\bar{\kappa}-\bar{\kappa}_c} \quad \text{if } n = 0, \tag{10.6}$$

$$\left(\frac{a}{a_c}\right)^3 = \frac{\bar{\rho}_c - 1}{\bar{\rho}_c - \bar{\rho}} \, \frac{n\bar{\rho}_c + (n+1)\bar{\rho}}{n\bar{\rho}_c + n + 1} \quad \text{if } n > 0. \tag{10.7}$$

Note that Eqs. (10.6) and (10.7) can also be directly found using Born approximation, as done in Ref. 30 for matter-waves, which, like acoustic waves, follow a scalar Helmholtz equation, leading to analogous cloaking conditions[31,32] realizable with semiconductors or quantum metamaterials.[33] In this long-wavelength approximation, a closed-form analytical expression of the cloaking condition can be found, and for a given cloak thickness a_c it is found that a single layer cloak can only cancel the first two scattering orders $n = 0$ and $n = 1$, which is sufficient if the object is small enough. These expressions are given in Ref. 25, and imply a full control over the cloak mass and bulk modulus, i.e., the use of metamaterial concepts.

Figure 10.2 shows a contour plot of the total scattering cross-section σ of a core-shell structure with $a_c = 1.1a$ and at a frequency such that $k_0 a = 0.5$, in a wide range of cloak parameters (top). The sphere to be cloaked has $\rho = 2\rho_0$ and $\kappa = 0.8\kappa_0$. We notice that

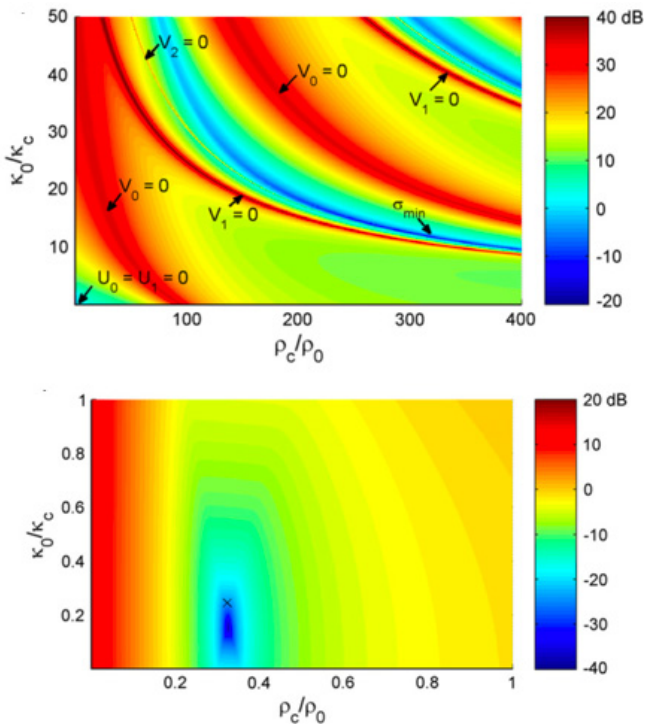


Fig. 10.2. Variation of the total scattering cross-section for a wide range of cloak constitutive parameters (top), and in the vicinity of the quasistatic cloaking condition (bottom), for $a_c = 1.1a$, $k_0 a = 0.5$, $\rho = 2\rho_0$ and $\kappa = 0.8\kappa_0$. The black cross in the bottom diagram marks the position found using the analytical quasistatic expression © 2006 Elsevier. Adapted with permission.[25]

the scattering cross-section can take any value between −20 dB and 40 dB, with the minimum and maximum values obtained when the conditions $U_n = 0$ and $V_n = 0$, respectively, are fulfilled for a given order n. Interestingly, $U_0 = 0$ and $U_1 = 0$ can be achieved with a single layer with a cloak with density and bulk modulus lower than the background (blue area in bottom plot). The quasistatic cloaking conditions (10.6) and (10.7), specialized for $n = 0, 1$, are marked by a black cross in Fig. 10.2, bottom. Here, the long wavelength approximation does not apply, yet the quasistatic theory is useful as

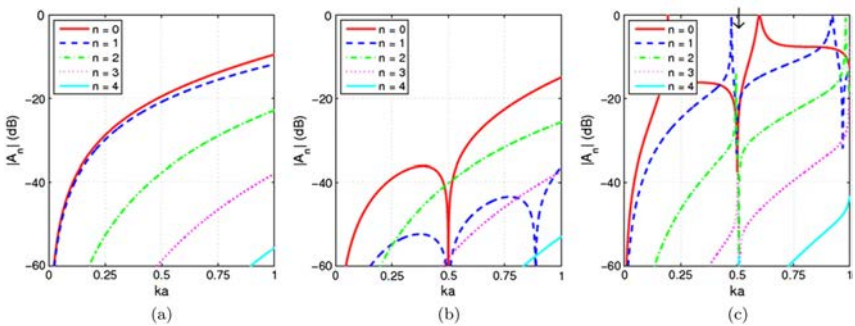


Fig. 10.3. Magnitude of the scattering coefficient versus frequency for (a) the bare object, (b) a covered object under the cloaking condition and (c) a covered object under the resonance condition © 2006 Elsevier. Adapted with permission.[25]

an estimate, and still predicts the cloaking condition with reasonable accuracy.

Figure 10.3 shows the behavior of the first five A_n coefficients versus frequency, for three different cases: (a) the uncovered core, (b) the core covered with a cloak designed to cancel the A_0 and A_1 coefficients at the frequency $k_0a = 0.5$ and (c) the core covered with a cloak designed to have a quadripolar resonance $V_2 = 0$ at the frequency $k_0a = 0.5$. It is clear that the cloaking phenomenon is indeed due to a drastic suppression of the dominant scattering of the object at the design frequency. In both cases of a cloaked (b) or resonant (c) object, the dominant scattering is of quadripolar nature, either because the monopolar and dipolar terms are negligible or because of a quadripolar resonance. The scattered field distribution around the object is shown in Fig. 10.4. The scattered field of the cloaked object (right) is compared to the one in the absence of the cloak (left). We see that, in magnitude, the cloaked object scatters much less than the bare object, and the dominant scattering is quadripolar, unlike in the case of the bare object for which the scattering is dominated by a mixture of a monopolar and a dipolar contribution. More insights into the cloaking phenomenon may be obtained by looking at a snapshot in time of the total acoustic pressure field around the object, as in Fig. 10.5. The left panel of Fig. 10.5 shows the case of

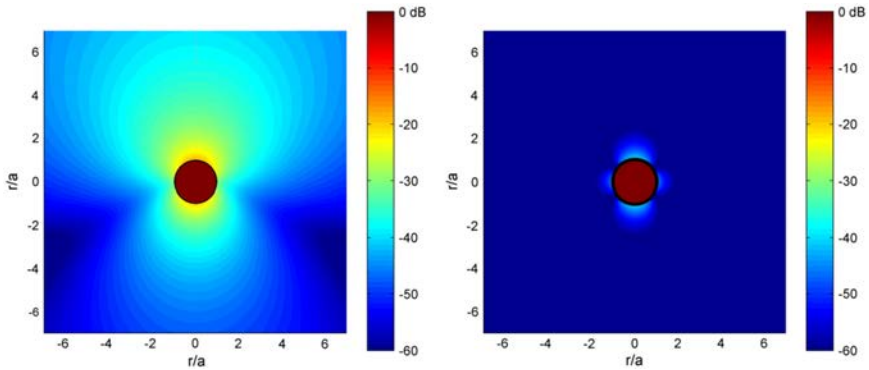


Fig. 10.4. Magnitude of the scattered acoustic pressure around the object in the cases (left) of the bare object and (right) the cloaked object. © 2006 Elsevier. Adapted with permission.[25]

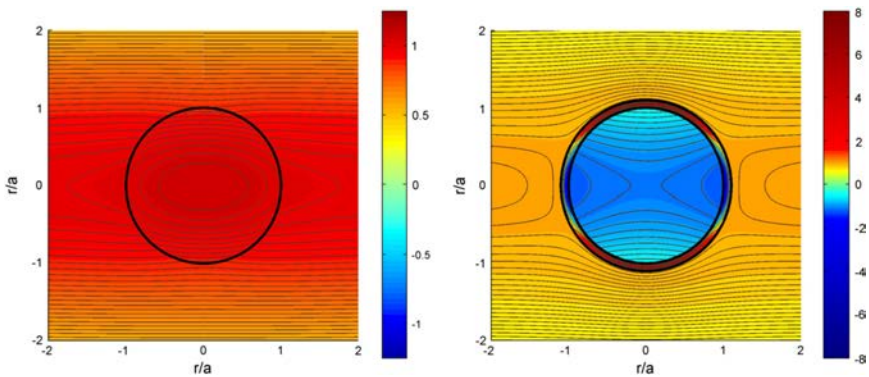


Fig. 10.5. Acoustic pressure field (snapshot in time) in the case of the bare sphere (left) and the cloaked sphere (right). © 2006 Elsevier. Adapted with permission.[25]

the bare object, to be compared with the case of the cloaked object. We see that the cloak restores the incident field as if the object were not there, with flat pressure contours, and this phenomenon is due to resonant excitation of the cloak. Figure 10.6 shows the plot of the acoustic particle velocity, on top of the acoustic pressure field distribution. The velocity field is shown with white, black and grey arrows in the background, core, and shell, respectively. It is interesting that the velocity field inside the sphere remains unchanged when

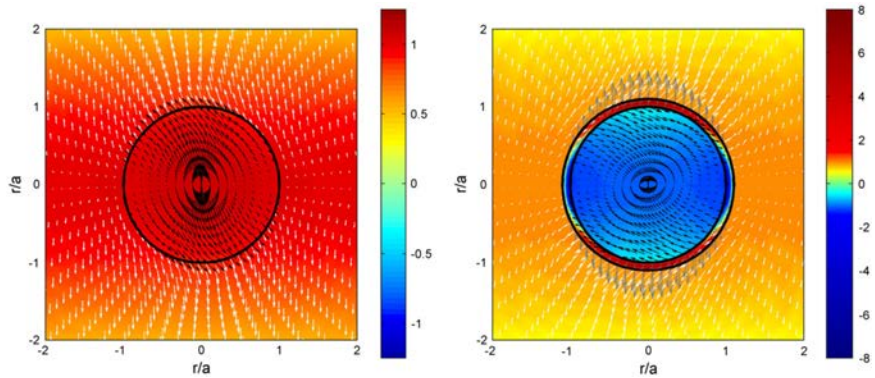


Fig. 10.6. Acoustic particle velocity field (snapshot in time) in the case of the bare sphere (left) and the cloaked sphere (right). © 2006 Elsevier. Adapted with permission.[25]

the object is cloaked. It is the extra scattering source composed of the cloak that compensates for the natural scattering from the core and induce a much smoother and uniform particle velocity field in the surrounding medium.

Similar effects can be induced by replacing the metamaterial shell with an acoustic metasurface, i.e., an ultrathin surface whose acoustic impedance is tailored to cancel the dominant scattering from the object.[27,28] Figure 10.7 shows examples of acoustic metasurfaces that may be used to induce acoustic scattering cancellation of a small acoustic sphere. Figure 10.8 shows the dependency of the total acoustic scattering cross section (SCS) σ, for a soft object with $2a = \lambda/5$ and $\kappa_r = 0.1$, on the acoustic surface reactance (panel a) and frequency (panel b). It is evident that depending on the value of the surface reactance, we may induce either a scattering resonance or a scattering cancellation. Under the scattering cancellation condition, we are able to reduce the total SCS by more than 30 dB at the frequency f_0. Panel (c) and (d) show the case of a hard object with $\kappa_r = 10$, which can also be cloaked using a surface reactance of opposite sign. Interestingly, plasmonic and mantle acoustic cloaks naturally let the incident power penetrate the cloak, a unique feature

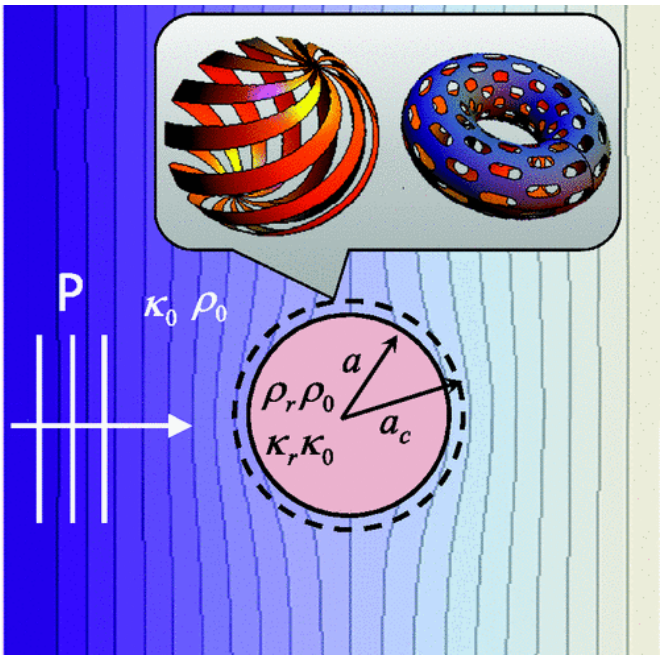


Fig. 10.7. Scattering cancellation may also be induced by using an ultrathin acoustic metasurface. © 2012 APS. Reprinted with permission.[28]

that may be used to realize cloaked acoustic sensors that can receive and sense a signal while minimally perturbing probed field.[34–40]

10.3. Density-Near-Zero Acoustic Supercoupling

Beyond cloaking and scattering manipulation, acoustic metamaterials are capable of inducing novel guided propagation features for acoustic waves, including anomalous matching and tunneling effects. This section discusses the phenomenon of supercoupling of acoustic waves through ultra narrow channels with near-zero density. In electromagnetics, a similar phenomenon is known as epsilon-near-zero (ENZ) supercoupling. In these zero-index materials, wave propagation is accompanied by very large phase velocity, producing static-like field distributions. These properties are observed, in particular, in an

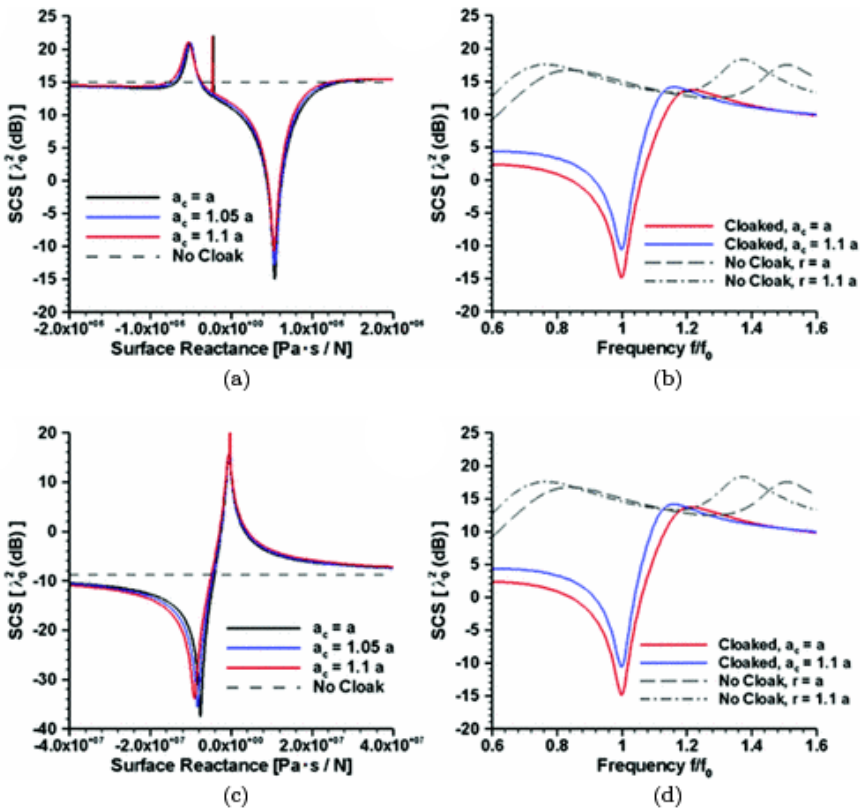


Fig. 10.8. Scattering cross section (SCS) of a soft acoustic sphere ($\kappa_r = 0.1$) with a size of a fifth of a wavelength versus the value of the surface reactance (a) and the frequency (b) for different ratios a_c/a. Panel (c) and (d) are similar to (a) and (b), but for a hard sphere ($\kappa_r = 10$) © 2012 APS. Reprinted with permission.[28]

abnormal tunneling and matching effect through very narrow waveguides filled with ENZ materials, also known as 'supercoupling'.[41,42] This matching effect is based on compensating the large impedance mismatch of waveguides with different transverse cross-section using the anomalously large wave impedance of ENZ materials.[41] Upon tunneling, the electric field is enhanced in a uniform way through the channel due to the extremely long wavelength in the zero index material, implying properties usually observed at very low frequency

in electrical wires, such as independence of the matching effect on the length of the channel, as well as immunity to twists and bends. This electromagnetic effect is interesting to build zero-delay coupling lines,[43] robust geometry-independent filters,[41] concentrating and harvesting light,[44] sensing,[45] boosting classical or quantum light emission[46,47] and non-linear effects.[48,49] Here, we discuss the acoustic analog of ENZ supercoupling, which is achieved using a metamaterial with density near zero (DNZ).[10] As in the case of ENZ supercoupling, the effect is based on the extreme value of one of the constitutive parameters in a waveguide, which totally cancels and compensate the geometrical mismatch at the connection with a much larger waveguide. The acoustic energy can therefore be squeezed in a small and narrow volume with this method.

The typical geometry is sketched in the inset of Fig. 10.9: two large acoustic waveguides with cross-section S_{wg} are connected via a narrow channel with $S_{ch} \ll S_{wg}$. The reflection coefficient at the input of the narrow tube reads[10]

$$R = \frac{(Z_{ch}^2 - Z_{wg}^2)\tan(\beta_{ch} l)}{(Z_{ch}^2 + Z_{wg}^2)\tan(\beta_{ch} l) + 2iZ_{ch}Z_{wg}} \tag{10.8}$$

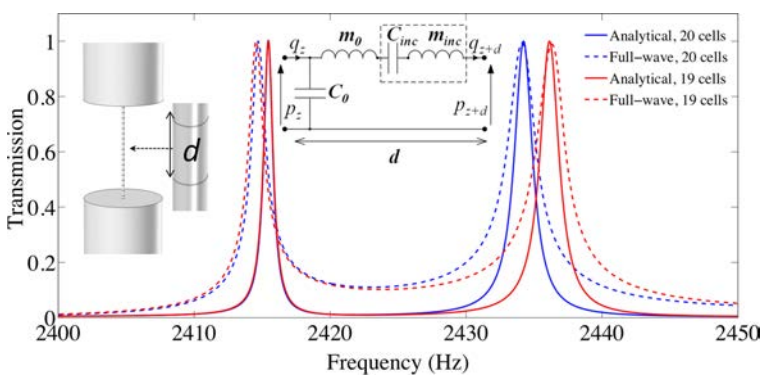


Fig. 10.9. Power transmittance between two air-filled large cylindrical waveguides connected by a much smaller waveguide with near-zero effective density. This metamaterial channel is built by periodically loading a narrow air-filled circular waveguide with polyimide membranes, effectively realizing a material with zero density at a design frequency. © 2013 APS. Reprinted with permission.[10]

where Z_{wg} and Z_{ch} are the line impedances of the outside and channel waveguide respectively, l is the channel length and β_{ch} is the channel wave number. Total transmission can occur any time the numerator in (10.8) vanishes, i.e. either under a Fabry-Perot (FP) resonance ($\beta_{ch} l = n\pi$), or under the condition $Z_{ch} = Z_{wg}$, i.e. if the line impedances happen to be matched. This matching condition, independent of the channel length, implies that:

$$\frac{\sqrt{\rho_{ch}\kappa_{ch}}}{S_{ch}} = \frac{\sqrt{\rho_{wg}\kappa_{wg}}}{S_{wg}}, \tag{10.9}$$

where ρ_{ch}, ρ_{wg} are the densities in the waveguides and κ_{ch}, κ_{wg} are the bulk moduli. Now because $S_{ch} \ll S_{wg}$, the matching condition (10.9) can be met only if the characteristic impedance of the channel $\sqrt{\rho_{ch}\kappa_{ch}}$ takes extremely low value.

For instance, when $\rho_{ch} \to 0$ (DNZ), matching occurs with total transmission associated with infinite phase velocity. This is the acoustic equivalent of ENZ supercoupling.

Such a density near zero material can be achieved in air-filled channels by periodically loading them with transverse membranes and operating near their resonance, at a specific frequency at which the inertia of the membrane compensates the one of the fluid contained within a metamaterial unit cell.[10] The result of FEM simulations of the corresponding acoustic structure and analytical transmittance calculations are shown in Fig. 10.9, showing the first two distinct transmission peaks. The first peak corresponds to DNZ supercoupling, as it is independent of the length of the channel, unlike the second peak, which is associated with the first FP resonance. Figure 10.10 offers more insights into this peculiar transmission phenomenon. The field distributions at the top of the figure correspond to the magnitude of the acoustic particle velocity in the channel in three different cases. The top plot corresponds to the case of a negative density channel, in which propagation is evanescent and transmission is zero. The middle case shows the case of the field at the supercoupling frequency, demonstrating the uniform field profile and velocity field enhancement, equal on average to the ratio of the

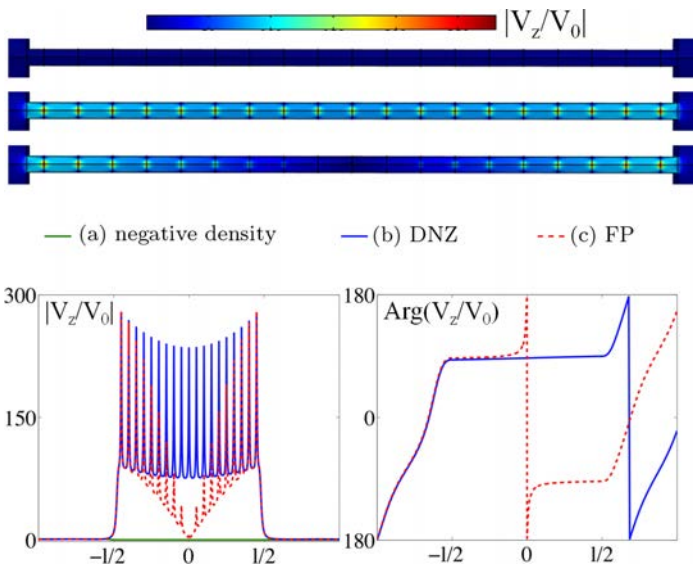


Fig. 10.10. Top: Acoustic particle velocity below the supercoupling frequency (top), at the supercoupling frequency (middle) and at the first FP resonance (bottom). Bottom: Velocity magnitude enhancement along the channel's center in the three above-mentioned scenarios (left) and corresponding phase (right). © 2013 APS. Reprinted with permission.[10]

waveguide cross-sections, consistent with conservation of the volumetric flow. This is very different from the field distribution at the first FP frequency, shown in the bottom plot, which has a node at the channel's center and a non-uniform phase throughout the channel, as confirmed by plotting the magnitude and phase of the velocity along the channel's center, done in the bottom graphs in Figure 10.10. Different from resonant transmission through a FP mode, the DNZ matching phenomenon does not rely on multiple reflections at the entrance and output boundary of the channel, and is therefore independent on the length of the channel. Even more fascinating, since the phase velocity inside the channel is extremely large, everything happens as if the channel had no acoustic length, and sound is transmitted regardless of the specific channel's topology, independently of twists and bends, as demonstrated in Fig. 10.11. Similar to electrical

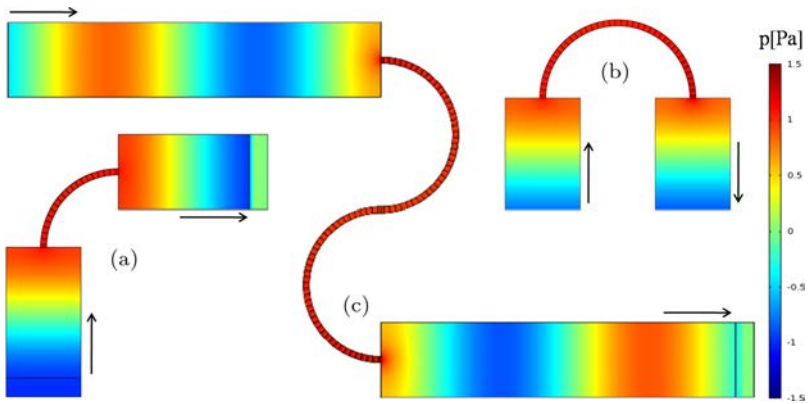


Fig. 10.11. Acoustic DNZ supercoupling is independent of twists and bends of the channel. © 2013 APS. Reprinted with permission.[10]

wires operated at DC, propagation of acoustic waves is quasi-static and it becomes insensitive to folding or bending.

10.4. Acoustic Brewster Angle and Broadband Energy Funneling

Density-near-zero supercoupling is an impedance matching effect that is inherently narrowband, due to the necessity of achieving zero mass density via material dispersion engineering. In this section, we demonstrate a different tunneling phenomenon that, on the contrary, is extremely broadband. It is the acoustic analog of Brewster transmission, i.e., an impedance matching phenomenon occurring for plane waves at a metamaterial interface, in a wide frequency band but for a specific angle of incidence.

Consider a plane boundary between two acoustic materials, 1 and 2, on which a plane wave is incident. Impedance matching occurs when the line impedances Z_1 and Z_2 of the two media are equal, which is the condition for having an angle of intromission, or acoustic Brewster angle,[50]

$$\tan^2\theta = \frac{m^2 - n^2}{n^2 - 1} \tag{10.10}$$

where $m = \rho_2/\rho_1$ is the ratio of densities and $n = c_1/c_2$ is the ratio of sound velocities. The matching condition (10.10) can only be fulfilled if $m > n > 1$ or the reciprocal case $1 > n > m$, which is not allowed by most natural materials because the speed of sound typically increases with density. However, if one of the media is a metamaterial, this may be possible. We now demonstrate that by engineering a metamaterial buffer and placing it between the two materials, it is possible to force broadband transmission and funnel acoustic power at an angle regardless of the fact the two materials are naturally extremely mismatched, and even when the intromission condition (10.10) cannot happen in the absence of the buffer. To do this, each interface of the metamaterial buffer must be matched to the corresponding surrounding half-space. Assuming a metamaterial made of hard wall material of a given thickness with periodically carved slits in it, we can equate the line impedance Z_1 of a plane wave incident from a homogeneous half space with density ρ_1 and sound velocity c_1 and the one Z_{ch} of the hard wall screen with straight sub wavelength channels filled with some acoustic material with parameters ρ_{ch} and c_{ch}. Similarly to the case of DNZ tunneling, we get the matching condition

$$Z_1 = Z_{ch} \Rightarrow \frac{\rho_1 c_1}{A_1 \cos\theta} = \frac{\rho_{ch} c_{ch}}{A_{ch}} \tag{10.11}$$

where A_1 is the length of a unit cell of the periodic buffer and A_{ch} the one of the channel, θ being the angle of incidence. We see that designing a buffer to match a given material is always possible, however matching two distinct media can only be done with buffers with different slit widths. We therefore propose the geometry of Fig. 10.12: materials 1 and 2 are connected through a buffer with tapered slits, which are filled with Material 2. We assume the case $1 < n$, $n > m$, i.e. there is no natural Brewster angle and no critical angle from medium 1 to medium 2.[50] We note l the thickness of the buffer and a its period. The lengths a_1 and a_2 are the slit widths at the interfaces with the media 1 and 2, respectively. A necessary condition for funneling is that the two interfaces of the buffer are

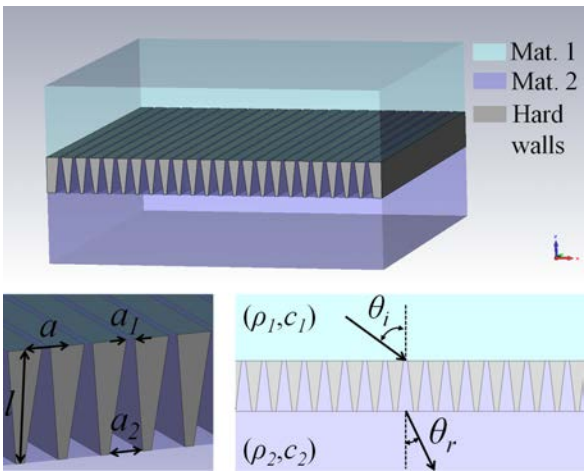


Fig. 10.12. A metamaterial made with a hard wall with tapered slits is used to funnel acoustic energy between two naturally mismatched media over an ultra large bandwidth. © 2014 ASA. Reprinted with permission.[51]

matched:

$$\frac{a_1}{a} = \frac{m}{n}\cos\theta_i, \tag{10.12}$$

$$\frac{a_2}{a} = \cos\theta_r. \tag{10.13}$$

The angles θ_i and θ_r cannot be chosen independently, and must conserve transverse momentum:

$$\sin\theta_i = n\sin\theta_r. \tag{10.14}$$

We use this strategy to match two very mismatched materials, Aluminum and Silicon Rubber. The speed of sound and density in Aluminum are $c_1 = 6420$ m/s and $\rho_1 = 2700$ kg/m^3. and $c_2 = 1006$ m/s and $\rho_2 = 990$ kg/m^3 in Silicon Rubber RTV-602.[52] We therefore have $m = 0.37$ and $n = 6.38$. Transmission between these two media is naturally inefficient, with a maximum of 1% power transmission under normal incidence.

Since $1 < n$, $n > m$, the condition (10.11) can never be approached. Choosing $\theta_i = 30$ and $a = 5$ cm guarantees single Floquet mode operation from 0 Hz to $c_2/a\sin 30° = 40$ kHz. The required values for the width at the interfaces are $a_1 = 2.4$ mm and $a_2 = 49.8$

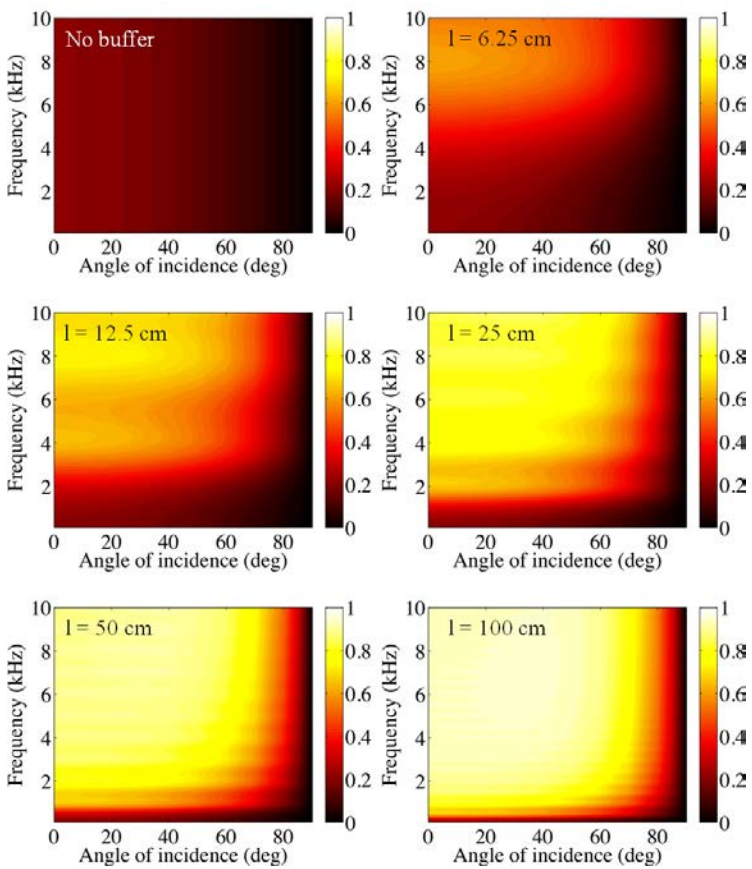


Fig. 10.13. Power transmission coefficient versus angle of incidence and frequency for a buffer of various thicknesses l inserted between Aluminum and Silicon Rubber. © 2014 ASA. Reprinted with permission.[51]

mm. The thickness of the screen needs to be sufficiently big to enable broadband adiabatic matching through the tapered apertures, and is determined via full-wave finite elements simulations, obtaining the power transmission coefficient as a function of both angle of incidence and frequency. In our simulations, the thickness is increased from 0 cm (no buffer) to 100 cm.

Figure 10.13 shows that, in the absence of buffer (top left plot), transmission between Aluminum and Silicon Rubber is very poor regardless of the frequency or the angle of incidence. When a thin

buffer is used (cases with $l = 6.25$ cm and 12.5 cm), transmission starts to increase until the buffer is thick enough to induce adiabatic matching between the materials, obtaining very high transmission over more than 30 kHz, for a buffer as thin as a few wavelengths. Remarkably, the method is not very sensitive to the angle of incidence for 100 cm thick buffers, with a high transmittance in a wide window around the design angle, starting from 0° all the way to 60°.

Figure 10.14 shows the distribution of the acoustic pressure, the particle velocity and the RMS power flow over a unit cell of the

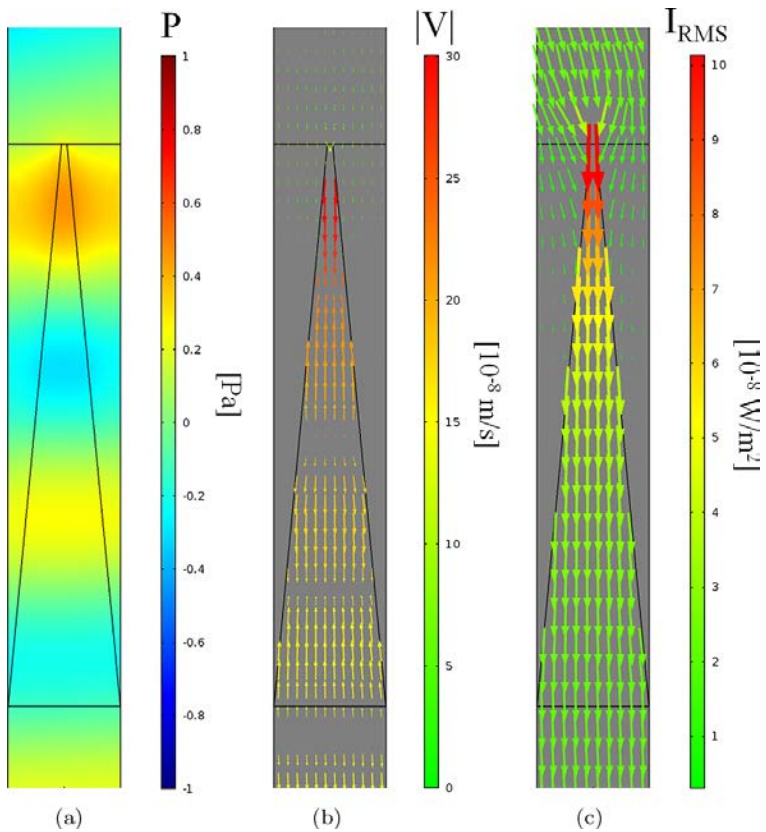


Fig. 10.14. Snapshot in time of (a) the acoustic pressure, (b) the particle velocity and (c) the power flow at the funneling condition. © 2014 ASA. Reprinted with permission.[51]

buffer at the funneling condition for the case of the 25 cm thick metamaterial made of stainless steel. For such a matching effect the thickness of the buffer can be arbitrary, and does not have to be a multiple of the half-wavelength, as it would be under a longitudinal resonance condition. The average power flow pictures the path of the acoustic energy through the buffer, confirming its ability to funnel the acoustic power with no reflection. Its value is identical on both sides of the screen, but its direction is different, consistent with momentum conservation along the buffer.

10.5. Unidirectional Invisibility Based on Parity-time Symmetry

So far, we have only considered tunneling phenomena obtained using lossless structures. However, the scattering and propagation through systems built using elements that support loss or gain can also provide surprising and unusual sound-matter interactions. Typically, gain and loss can be considered as an additional degree of freedom to engineer the flow of acoustic power and lead to unique solutions in our ability to control sound. We demonstrate here how the use of loss and gain in a Parity-Time (PT) symmetric metamaterial or metasurface can lead to fascinating effects.

PT symmetry is a special kind of space-time symmetry that describes the invariance of a system upon the combined operation of two operators. The first one, the Parity operator P, takes all the inversions of space $(x, y, z) \rightarrow (-x, -y, -z)$, and the second one, the time reversal operator T, takes the inversion of time $t \rightarrow -t$. For example, a one-dimensional PT symmetric system must be the same as its time-reversed mirror image. In practice, taking the time-reversed image of a given acoustic medium means changing loss into gain and flipping the sign of all parameters that are odd-symmetric upon T, e.g. an external magnetic field bias. Taking the mirror image means flipping the left and right side of the system using spatial inversion about its central axis. The simple electronic circuit of Fig. 10.15 is an example of a PT symmetric system. It is composed of two lumped elements separated by a portion of lossless transmission line

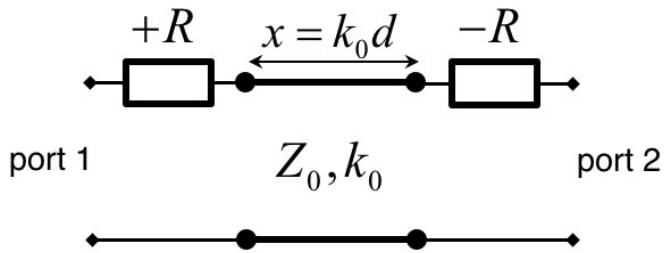


Fig. 10.15. Example of a PT symmetric circuit made of a pair of resistors with opposite values separated by a portion of lossless transmission line of length d. [53]

of length d, characteristic impedance Z_0 and propagation constant k_0, and placed in a medium of characteristic impedance Z_0. The lumped elements have purely real impedance: the one on the left is a simple resistor of value $R > 0$, while the one on the right is a negative resistor of opposite value $-R$, i.e. an active source element. Because the time-reversed image of a resistor of value R is a negative resistor of value $-R$, the right hand side and the left hand side of the two-port network shown in Fig. 10.15 are the time-reversed image of each other, i.e. the system is PT symmetric.

PT symmetric systems have attracted a lot of attention over the past few years. In quantum physics, there has been a lot of interest in studying the spectral properties of non-Hermitian Hamiltonians that commute with the parity-time (PT) operator.[54] These PT-symmetric systems, surprisingly, can exhibit entirely real spectrum eigenvalues, and may be used to define a consistent unitary extension of quantum mechanics. A fascinating property of PT symmetric systems is the notion of spontaneous symmetry breaking, i.e., the possibility for the *state* of the system to lose its PT symmetry, yielding a particular solution of the physical equations that is less symmetric than the theory itself. This generally happens when a continuous parameter quantifying the non-Hermiticity of the Hamiltonian exceeds a threshold value, and it is accompanied by a phase transition from an entirely real spectrum to a complex, or partially complex one. Interestingly, these properties can be observed

in paraxial optical systems possessing an even distribution of the real part of the refractive index along the optical axis, and an odd distribution for its imaginary part, in a balanced loss/gain configuration.[55] The perpendicular direction represents the time axis, on which the time-evolution of the system is effectively observed, yielding unique diffraction effects. Another distinctive aspect of PT symmetric system is their scattering signature.[56] For instance, the scattering matrix S of a PT-symmetric system reflects its phase transition: the S matrix is unimodular in the PT symmetric phase, with each eigenvalues being unimodular, while the broken symmetry state is characterized by pairs of eigenvalues with reciprocal moduli. These unusual scattering properties may be used to our advantage to manipulate waves in novel ways, and build novel loss-compensated metamaterials based on PT symmetry.

It is quite straightforward to understand why PT symmetric systems may induce strong asymmetry in reflection by considering the following on-dimensional example of a generic two-port system described by its scattering matrix $S = \{\{r_1, t_{12}\}, \{t_{21}, r_2\}\}$. If the two-port network is lossless, its absorption is zero whether one excites from port 1 or from port 2, and therefore $1 = |t_{21}|^2 + |r_1|^2 = |t_{12}|^2 + |r_2|^2$. If it is also reciprocal, then $t_{12} = t_{21} = t$ and energy conservation implies that $|r_1| = |r_2|$, i.e. any reflectance asymmetry between the two sides is impossible. However, if the network has a non zero absorption, the quantities $|t|^2+|r_1|^2$ and $|t|^2+|r_2|^2$ are both less than one, and the reflectivity on both side can be different if the distribution of the absorption losses is not symmetric with respect to ports 1 and 2 (the input impedances are indeed different), i.e., if the system does not have spatial inversion symmetry, i.e. P symmetry. Therefore, asymmetry in the reflectance from a reciprocal system can only be created by a using a loss or gain distribution that breaks parity symmetry.

Now we show that the system of Fig. 10.15 not only supports asymmetry in reflection, but that due to PT symmetry this asymmetry is extreme, with a side being perfectly impedance matched (transparent), while the other side is extremely reflective. In addition,

the system shows fascinating transmission properties when operated at an exceptional point. Its scattering matrix can be calculated and depends on two parameters x and r:

$$\mathbf{S} = \begin{pmatrix} \dfrac{r(r-2)\sin(x)}{\mathrm{r}^2\sin(x)+2\mathrm{j}\mathrm{e}^{jx}} & \dfrac{2j}{\mathrm{r}^2\sin(x)+2\mathrm{j}\mathrm{e}^{jx}} \\ \dfrac{2j}{\mathrm{r}^2\sin(x)+2\mathrm{j}\mathrm{e}^{jx}} & \dfrac{r(r+2)\sin(x)}{\mathrm{r}^2\sin(x)+2\mathrm{j}\mathrm{e}^{jx}} \end{pmatrix} \tag{10.15}$$

Here, $r = R/Z_0 > 0$ is the value of the resistor normalized to the line impedance and $x = k_0 d$ is the electrical distance separating the two lumped elements. Quite interestingly, we see that when the system is designed so that $r = 2$, the S matrix simplifies into

$$\mathbf{S}|_{r=2} = \begin{pmatrix} 0 & e^{jx} \\ e^{jx} & 2-2e^{2jx} \end{pmatrix}, \tag{10.16}$$

i.e. the system is completely matched from port 1 regardless of the distance x, and is highly reflective from port 2, with a reflection $|r_2|$ that can exceed unity depending on the distance. This unidirectional reflectionless behavior is accompanied by an interesting property of the system in transmission. The transmission from both port is *always unitary* and is accompanied by a phase *advance* that is exactly opposite of the phase the signal would acquire over the length x in absence of the system. The same transmission behavior is observed in a matched double negative slab, in which the backward phase velocity enables phase compensation for a propagating wave. This implies that here, one also has backward phase evolution between the lumped elements, but because the transmission line that separates the elements is right handed, this implies that the power flow is also backward between the resistors. We have verified this behavior experimentally for acoustic waves in a waveguide, implementing a stable positive/negative resistor pair at a design frequency using loudspeakers loaded with active electronic circuits.[58] This experiment demonstrated the feasibility of mimicking the properties of double

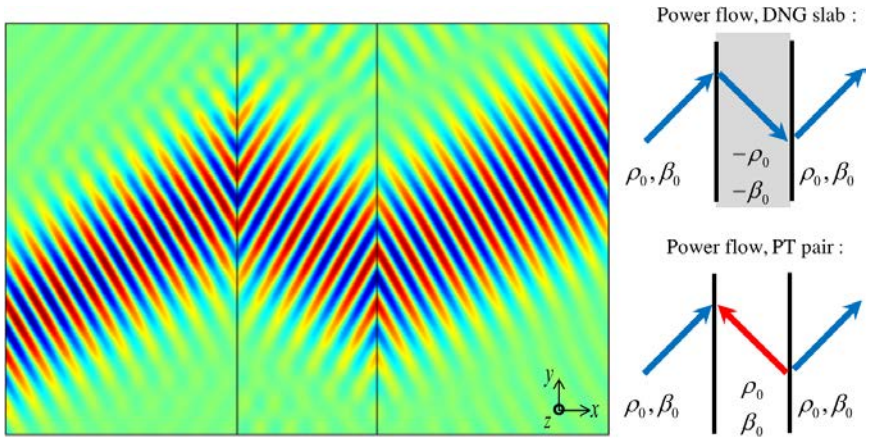


Fig. 10.16. PT symmetric metasurface pairs can induce negative refraction of acosutic beams in air and without the need of a bulk double negative metamaterial and associated drawbacks. © 2014 APS. Adapted with permission.[53]

negative bulk metamaterials with a pair of PT symmetric lumped elements.

The phase compensation capability of double negative metamaterials is at the origin of the phenomenon of negative refraction enabling planar focusing in a Veselago lens. The above results suggest that PT symmetric metasurface pairs can induce similar effects without the need for a bulk metamaterials and avoiding all the associated problems, like its imperfect isotropy, complexity, sensitivity to losses, and bandwidth limitation due to its passive nature. This is verified in Fig. 10.16, where we show that properly designed PT symmetric metasurface pairs can induce negative refraction for an acoustic Gaussian beam in a loss-immune, metamaterial-free, and potentially broadband fashion. More details about the designs and time-harmonic field animations can be found in Ref. 53. Once we have negative refraction, we can have planar focusing in the form of a Veselago lens, as demonstrated in Fig. 10.17. A similar effect was exploited in Ref. 59 so induce unidirectional cloaking of large objects by employing metasurfaces with PT symmetric gain and loss distributions.

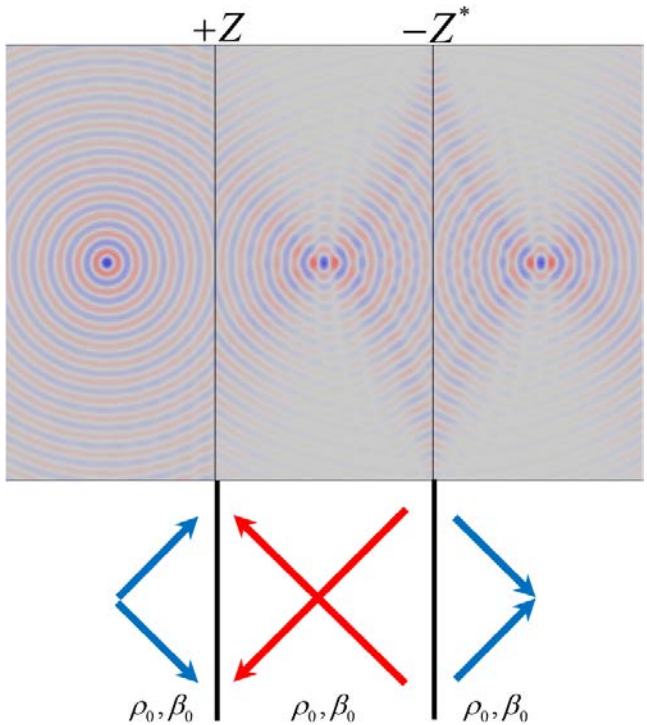


Fig. 10.17. PT symmetric metasurface pairs can induce planar focusing in air and without the need of a bulk Veselago lens and associated drawbacks. © 2014 APS. Adapted with permission.[53]

10.6. Non-reciprocal Manipulation of Acoustic Waves

We have seen in the previous section that by using a distribution of loss and gain that breaks P symmetry, but not PT symmetry, it is possible to induce giant asymmetry in the reflection properties of an acoustic system, with zero reflection on one side and a high reflection on the other. However, in all the examples previously considered in this chapter, transmission is symmetric both in phase and amplitude, leading to the question: can we also induce giant acoustic transmission asymmetry, for instance creating an acoustic power flow diode, that lets acoustic power be transmitted in one direction and blocks transmission it in the other?

In a linear system, symmetry of the transmission properties, also known as Rayleigh reciprocity, is intimately related to a general property of wave propagation: time-reversal symmetry. Breaking Rayleigh reciprocity in these systems implies the use of an external bias, such as a magnetic field, that is odd-symmetric upon time reversal.[60] In order to break reciprocity for airborne sound, which does not interact with magnetic fields, another form of odd-symmetric bias can be used: fluid motion.

Consider the situation shown in Fig. 10.18(a). A ring cavity is carved in a hard wall material such as Aluminum, and filled with air. The cavity is turned into a 3-port device by drilling three holes at 120° intervals to couple with three external waveguides. In order to break time-reversal symmetry for this acoustic system, we assume that the air inside the cavity is moving along the cavity in a circular motion and at a constant velocity v. With time-reversal symmetry broken, reciprocity does not hold anymore, and it becomes possible to transmit sound from waveguide 1 to waveguide 3, but not from waveguide 3 to waveguide 1. This behavior has been experimentally demonstrated for airborne acoustic sound around 800 Hz, by implementing the air rotation using small fans placed inside the cavity (panel (b)).[61] Panel c shows the frequency dependency of the magnitude of the transmission to port 2 and port 3 assuming excitation from port 1, assuming that the internal fluid is at rest. We see that the transmission is resonant, reaching a maximum at the ring resonance, and equally split into ports 2 and 3. By symmetry, transmission from port 3 to port 1 would be identical, and the system with a fluid at rest is not capable of breaking reciprocity. The situation drastically change when the fluid is moving at a constant velocity of 0.5 m/s (panel d).

We see that at the cavity resonance, transmission to waveguide 3 is now unitary, with identically zero power going into waveguide 2. By symmetry, from waveguide 3 the acoustic wave would only be transmitted to waveguide 2, thereby breaking Rayleigh reciprocity between waveguides 1 and 3. Sound is always transmitted in the waveguide at the left of the input, in a uni-rotational manner, creating a circulator for acoustic waves. Remarkably, this effect can

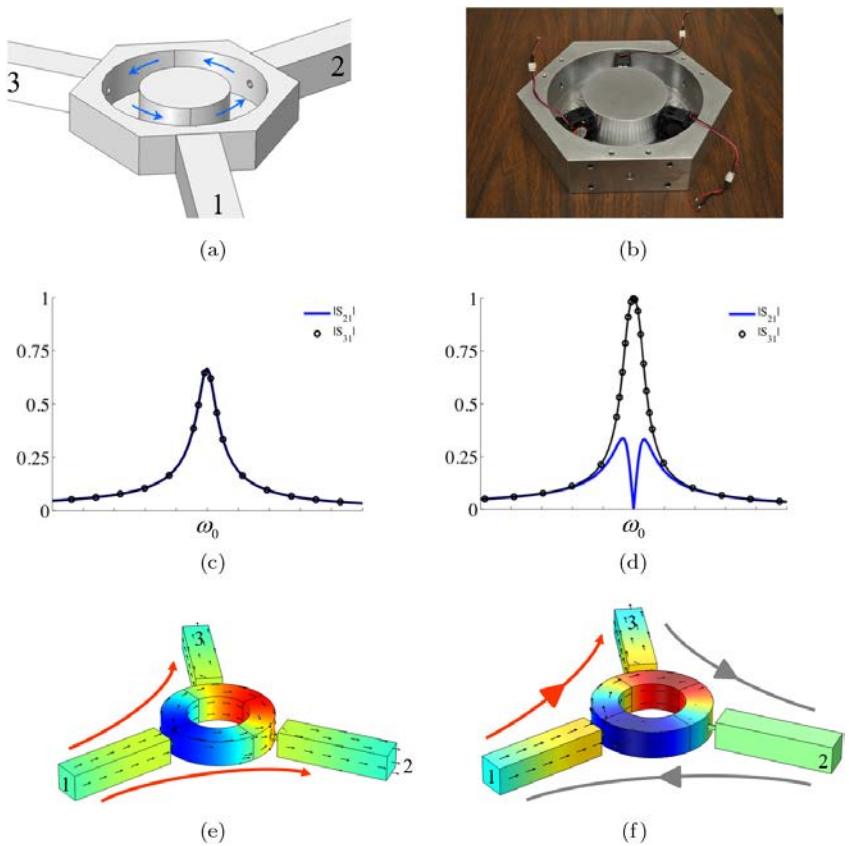


Fig. 10.18. A non-reciprocal acoustic circulator can be obtained from a ring cavity in which the air is circulated (a). (b) Implementation of the fluid motion using fans in an Aluminum cavity. (c) Magnitude of transmission to waveguides 2 and 3 upon excitation from waveguide 1 (respectively $|S_{21}|$ and $|S_{31}|$) for a cavity in absence of internal air motion. (d) Same as (c) but for in the case of internal air motion with a velocity of 0.5m/s. (e) Acoustic pressure field profile (color) and average acosutic power flow (arrows) at the ring resonance frequency in absence of internal air motion. (f) same as (e) but in the case of internal air motion with a velocity of 0.5m/s © 2014 AAAS. Adapted with permission.[61]

be obtained for extremely low fluid motion, 0.5 m/s, when compared to the speed of sound in air, 340 m/s. This is due to the resonant nature of the cavity, which enhances the interaction of the acoustic wave with the fluid in motion.

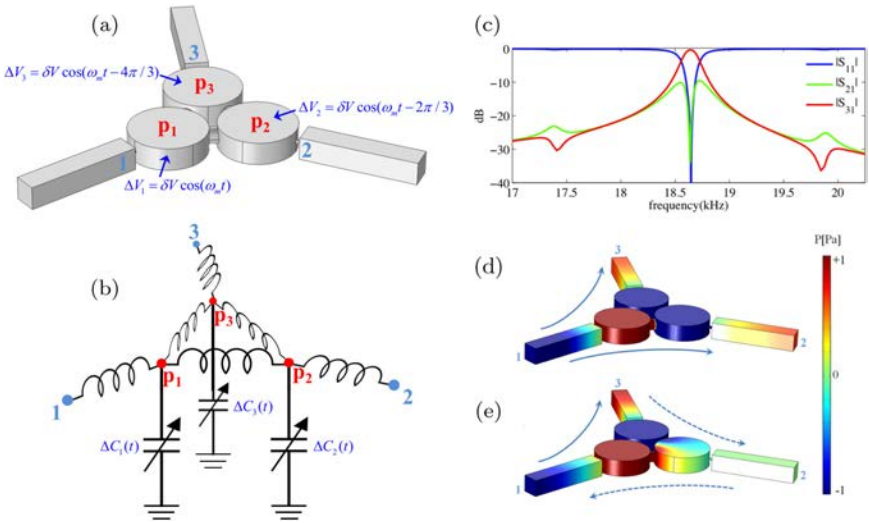


Fig. 10.19. (a) An alternative to fluid motion to break reciprocity is to apply a rotating time-harmonic modulation to an acoustic resonators. Here, we consider a resonator composed of three connected cavities and modulate the volumes of the cavities in a rotating fashion. (b) Lumped circuit equivalent of the modulated resonator. (c) Scattering parameters of the circulator. (d) acoustic pressure in the unmodulated case and (e) in the modulated case. © 2015 APS. Adapted with permission.[62]

One drawback of using fluid motion is that it creates noise on top of the acoustic signal, and the method is not easily transposable to acoustic waves in solids or at higher frequency. To go around these issues, it is possible to impart a sense of rotation to an acoustic resonator by applying a time-harmonic modulation of the resonator properties in a rotating fashion. Figure 10.19(a) shows an example of an acoustic resonator composed of three cavities that are connected to each other via small channels. The volumes of each cavity is modulated periodically in time with an amplitude δV and at a frequency ω_m, and this modulation is applied with phase differences of $2\pi/3$ between adjacent cavities, essentially imparting a sense of rotation to the structure. Figure 10.19(b) shows the lumped circuit equivalent of the structure, which can be used as a good model to determine the optimal modulation depth δV and frequency ω_m to be applied to

achieve acoustic circulation.[62] For a signal at 18.6 kHz, we found that a modulation depth of 5% at a frequency of 1200 Hz was enough to induce more than 40 dB isolation with very low insertion losses (0.3 dB), with a device whose total size does not exceed $\lambda/6$. The scattering parameters extracted from full-wave numerical simulations of the modulated structure, under these optimal conditions, are shown in panel (c), confirming the validity of the analytical modeling based on a lumped circuit in a full-wave scenario. The associated pressure field distributions in the unmodulated (panel d) and modulated (panel e) cases confirm the giant breaking of reciprocity induced by this method at the resonance frequency of the resonator.

10.7. Conclusions

In this chapter, we have discussed and reviewed the interesting acoustic properties of metamaterials, metasurfaces, and meta-atoms, which may be used to dramatically enhance our ability to control acoustic waves. Homogenous isotropic passive metamaterials and metasurfaces can be used to strongly reduce the scattering of objects, by using a dominant scattering cancellation method analogous to plasmonic cloaking of electromagnetic waves. Passive metamaterials with extreme constitutive parameters can lead to exotic acoustic tunneling phenomena such as density-near-zero supercoupling or broadband Brewster acoustic funneling, of interest to transmit acoustic waves at mismatched boundaries or in acoustic signal manipulation. Gain and loss in metamaterials and metasurfaces can provide a new degree of freedom to induce novel tunneling effects, which in the case of PT symmetric metasurfaces, lead to largely asymmetric reflection phenomena and transmission with anomalously compensated phase, a phenomenon that may be used to replace double negative bulk metamaterials to induce negative refraction and planar focusing. Finally, the dual case of giant asymmetric transmission can be induced in linear acoustic systems by breaking time-reversal symmetry via the use of a moving medium or a spatio-temporally modulated structure. Altogether, many degrees of freedom in acoustic metamaterials and metasurfaces can be exploited to control the

scattering of sound and reach a new degree of control over acoustic waves and signals.

Acknowledgments

This work has been partially supported by the Defense Threat Reduction Agency, the Air Force Office of Scientific Research, and the National Science Foundation.

References

[1] Craster, R. V. and Guenneau, S., *Acoustic Metamaterials: Negative Refraction, Imaging, Lensing and Cloaking* Springer (2012).
[2] Lee, S. H., Park, C. M., Seo, Y. M., Wang, Z. G. and Kim, C. K. (2010). *Phys. Rev. Lett.* **104**, p. 054301.
[3] Li, J. and Chan, C. T. (2004). *Phys. Rev. E* **70**, p. 055602.
[4] Fang, N., Xi, D., Xu, J., Ambati, M., Srituravanich, W., Sun, C. and Zhang, X. (2006). *Nat. Mater.* **5**, p. 452.
[5] Yang, M., Ma, G., Yang, Z. and Sheng, P. (2013). *Phys. Rev. Lett.* **110**, p. 134301.
[6] Liang, Z. and Li, J. (2012). *Phys. Rev. Lett.* **108**, p. 114301.
[7] Xie, Y., Popa, B.-I., Zigoneanu, L. and Cummer, S. A. (2013). *Phys. Rev. Lett.* **110**, p. 175501.
[8] Bongard, F., Lissek, H. and Mosig, J. R. (2010). *Phys. Rev. B* **82**, p. 094306.
[9] Seo, Y. M., Park, J. J., Lee, S. H., Park, C. M., Kim, C. K. and Lee, S. H. (2012). *J. Appl. Phys.* **111**, p. 023504.
[10] Fleury, R. and Alù, A. (2013). *Phys. Rev. Lett.* **111**, p. 055501.
[11] Liu, F., Lai, Y., Huang, X. and Chan, C. T. (2011). *Phys. Rev. B* **84**, p. 224113.
[12] Engheta, N. (2013). *Science* **340**, p. 286.
[13] Veselago, V. G. (1968). *Sov. Phys. Uspekhi* **10**, p. 509.
[14] Pendry, J. B. (2000). *Phys. Rev. Lett.* **85**, p. 3966.
[15] Fleury, R. and Alù, A. (2014). *FERMAT* **1**(7), pp. 1–24.
[16] Fleury, R., Monticone, F. and Alù, A. (2015). *Phys. Rev. Appl.* **4**, p. 037001.
[17] Alù, A. and Engheta, Nader (2005). *Phys. Rev. E* **72**, p. 016623.
[18] Pendry, J. B., Schurig, D. and Smith, D. R. (2006). *Science* **312**, p. 1780.
[19] Leonhardt, U. (2006). *Science* **312**, p. 1777.
[20] Schurig, D., Mock, J. J., Justice, B. J., Cummer, S. A., Pendry, J. B., Starr, A. F. and Smith, D. R. (2006). *Science* **314**, p. 977.
[21] Alù, A. (2009). *Phys. Rev. B* **80**, p. 245115.
[22] Chen, P.-Y. and Alù, A. (2011). *Phys. Rev. B* **84**, p. 205110.
[23] Norris, A. N. (2015). *Acoust. Today* **11**, p. 38.

[24] Guild, M. D., Alù, A. and Haberman, M. R. (2011). *J. Acoust. Soc. Am.* **129**, p. 1355.
[25] Guild, M. D., Haberman, M. R. and Alù, A. (2011). *Wave Motion* **48**, p. 468.
[26] Guild, M. D., Haberman, M. R. and Alù, A. (2012). *Phys. Rev. B* **86**, p. 104302.
[27] Chen, P.-Y., Farhat, M., Guenneau, S., Enoch, S. and Alù, A. (2011). *Appl. Phys. Lett.* **99**, p. 191913.
[28] Farhat, M., Chen, P.-Y., Guenneau, S., Enoch, S. and Alù, A. (2012). *Phys. Rev. B* **86**, p. 174303.
[29] Morse, P. M. and Ingard, K. U., *Theoretical Acoustics* (Princeton University Press, 1968).
[30] Fleury, R. and Alù, A. (2012). *Appl. Phys. A* **109**, p. 781.
[31] Fleury, R. and Alù, A. (2013). *Phys. Rev. B* **87**, p. 045423.
[32] Fleury, R. and Alù, A. (2013). *Phys. Rev. B* **87**, p. 201106.
[33] Fleury, R. and Alù, A. (2014). *Phys. Rev. B* **90**, p. 035138.
[34] Alù, A. and Engheta, N. (2009). *Phys. Rev. Lett.* **102**, p. 233901.
[35] Fan, P., Chettiar, U. K., Cao, L., Afshinmanesh, F., Engheta, N. and Brongersma, M. L. (2011). *Nat. Photonics* **6**, p. 380.
[36] Fleury, R., Soric, J. and Alù, A. (2014). *Phys. Rev. B* **89**, p. 045122.
[37] Soric, J. C., Fleury, R., Monti, A., Toscano, A., Bilotti, F. and Alu, A. (2014). *IEEE Trans. Antennas Propag.* **62**, p. 4220.
[38] Xu, T., Zhu, X.-F., Liang, B., Li, Y., Zou, X.-Y. and Cheng, J.-C. (2012). *Appl. Phys. Lett.* **101**, p. 033509.
[39] Greenleaf, A., Kurylev, Y., Lassas, M. and Uhlmann, G. (2011). *Phys. Rev. E* **83**, p. 016603.
[40] Greenleaf, A., Kurylev, Y., Lassas, M., Leonhardt, U. and Uhlmann, G. (2012). *Proc. Natl. Acad. Sci.* **109**, p. 10169.
[41] Alù, A., Silveirinha, M. G. and Engheta, N. (2008). *Phys. Rev. E* **78**, p. 016604.
[42] Edwards, B., Alù, A., Young, M. E., Silveirinha, M. and Engheta, N. (2008). *Phys. Rev. Lett.* **100**, p. 033903.
[43] Alù, A. and Engheta, N. (2010). *IEEE Transactions on Antennas and Propagation* **58**(2), p. 328.
[44] Alù, A. and Engheta, N. (2008). *Phys. Rev. B* **78**, p. 035440.
[45] Alù, A. and Engheta, N. (2008). *Phys. Rev. B* **78**, p. 045102.
[46] Alù, A. and Engheta, N. (2009). *Phys. Rev. Lett.* **103**, p. 043902.
[47] Fleury, R. and Alù, A. (2013). *Phys. Rev. B* **87**, p. 201101.
[48] Powell, D. A., Alù, A., Edwards, B., Vakil, A., Kivshar, Y. S. and Engheta, N. (2009). *Phys. Rev. B* **79**, p. 245135.
[49] Argyropoulos, C., Chen, P.-Y., D'Aguanno, G., Engheta, N. and Alù, A. (2012). *Phys. Rev. B* **85**, p. 045129.
[50] Brekhovskikh, L. M. and Godin, O. A., *Acoustics of Layered Media I: Plane and Quasi-Plane Waves* (Surendra Kumar, 1998).
[51] Fleury, R. and Alù, A. (2014). *J. Acoust. Soc. Am.* **136**, p. 2935.

[52] Folds, D. L. (1974). *J. Acoust. Soc. Am.* **56**, p. 1295.
[53] Fleury, R., Sounas, D. L. and Alù, A. (2014). *Phys. Rev. Lett.* **113**, p. 023903.
[54] Bender, C. M. and Boettcher, S. (1998). *Phys. Rev. Lett.* **80**, pp. 5243–5246.
[55] Rüter, C. E., Makris, K. G., El-Ganainy, R., Christodoulides, D. N., Segev, M. and Kip, D. (2010). *Nat. Phys.* **6**, pp. 192–195.
[56] Chong, Y. D., Ge, L. and Stone, A. D. (2011). *Phys. Rev. Lett.* **106**, p. 093902.
[57] Lin, Z., Ramezani, H., Eichelkraut, T., Kottos, T., Cao, H. and Christodoulides, D. N. (2011). *Phys. Rev. Lett.* **106**, p. 213901.
[58] Fleury, R., Sounas, D. and Alù, A. (2015). *Nat. Commun.* **6**, p. 5905.
[59] Sounas, D. L., Fleury, R. and Alù, A. (2015). *Phys. Rev. Appl.* **4**, p. 014005.
[60] Fleury, R., Sounas, D. L., Haberman, M. R. and Alù, A. (2015). *Acoust. Today* **11**, p. 14.
[61] Fleury, R., Sounas, D. L., Sieck, C. F., Haberman, M. R. and Alù, A. (2014). *Science* **343**, p. 516.
[62] Fleury, R., Sounas, D. L. and Alù, A. (2015). *Phys. Rev. B* **91**, p. 174306.

CHAPTER 11

The Carpet Cloak

FU LIU and JENSEN LI*

University of Birmingham, UK

JINGJING ZHENG and YU LUO†

Nanyang Technological University, Singapore

Carpet cloak constitutes a specific type of example in transformation optics that is without singularity. It hides an object by squashing it into a plane. The resultant devices become easier to fabricate and have a larger frequency bandwidth. In this Chapter, we will discuss two different approaches: the quasiconformal and the linear maps to realize carpet cloaks. We will also examine different devices employing such maps.

11.1. Introduction

Certain types of animals such as octopus and chameleon change their skin colours and patterns to resemble the background environment as a camouflage to hide from their predators.[1] While these are visual illusions, recent advances in metamaterials allow us to

Corresponding authors: *j.li@bham.ac.uk
†luoyu@ntu.edu.sg

obtain invisibility by directly bending light around an object, as if the light follows its original path without the object.[2–27] The invisibility achieved is thus independent of the mechanism of detection. The design of the required metamaterials to fabricate an invisibility cloak is based on a framework called transformation optics (TO),[2,3] which is driven by a coordinate transformation. Due to the form-invariance property of Maxwell's equations, a coordinate transformation induces a corresponding transformation of material parameters, which can then be realized in the notion of metamaterials. The fields are transformed, and thus are manipulated, in the way that they stick to the coordinates being transformed. In addition to invisibility cloaks, such an approach has also enabled a wide range of interesting applications, including concentrators, field rotators and optical illusion devices.[28–35] For a general coordinate transformation,[2–6] however, the induced electric permittivity and magnetic permeability are generally both anisotropic and inhomogeneous. Their values can also span a wide range if the coordinate transformation is very extreme or is possessing singularities, e.g. in the case of a cylindrical cloak. It limits the working frequency bandwidth of the resultant device as a wide range of material parameters would require resonating metamaterial atoms. In fact, such limitations can be partially circumvented if the functionality of the device can be adopted and achieved by employing maps without singularity and hence with simpler material parameters. The carpet cloaks using quasiconformal maps[7–15] and linear maps[16–27] introduced in this chapter are representative examples along this line. The quasiconformal map generates materials with minimized anisotropy, which becomes a constant near to a value one in the device. By choosing appropriate boundaries to avoid sharp corners for certain applications like a carpet cloak,[7–15] the range of the refractive indices can be made smaller as well. The quasiconformal map approach will then be useful, e.g. to design a range of transformation optical devices, including waveguide benders, couplers and surface-conformal antennas, simply based on an index profile with very small or negligible anisotropy.[33–42] On the other hand, the linear map generates materials with constant index as well,

although the anisotropy is not minimized.[16–27] It has a huge benefit that it allows the carpet cloak and other devices to be constructed by only a single kind of anisotropic materials. Based on the linear maps, analytic examples of transformation optics can be designed easily with intuitive physical insights. It then allows a straightforward realization due to its simplicity.

11.2. Transformation Optics for Wave Propagation in 2D

Here, we will concentrate on the case of wave propagation in two-dimensions. We will look into the case of in-plane wave propagation, in which both the fields and materials in the third dimension (along z) are invariant. The coordinate transformation is assumed to be two dimensional, between two sets of coordinates (x', y') and (x, y), where $z' = z$ is assumed. The coordinate transformation can be captured by the associated Jacobian matrix Λ defined by

$$\Lambda = \begin{pmatrix} \partial_{x'} x & \partial_{y'} x \\ \partial_{x'} y & \partial_{y'} y \end{pmatrix}. \tag{11.1}$$

For in-plane wave propagation where E_z and H_z polarizations are decoupled, the transformation of the medium can be written for the two polarizations separately. For E_z-polarization, the permittivity along the z-direction and the in-plane 2×2 permeability tensor are transformed according to TO[2] by

$$\begin{aligned} \epsilon &= \frac{\epsilon'}{\det \Lambda} = \epsilon' \hat{z} \cdot \nabla x' \times \nabla y', \\ \mu &= \frac{\Lambda \mu' \Lambda^T}{\det \Lambda} \quad \text{with } \sqrt{\frac{\mu_T}{\mu_L}} + \sqrt{\frac{\mu_L}{\mu_T}} = \frac{Tr(\mu)}{\sqrt{\det \mu}}, \end{aligned} \tag{11.2}$$

where μ_T and μ_L are the two principal values (eigenvalues) of the in-plane tensor μ. The quantity $Tr(\mu)/\sqrt{\det \mu}$ approaches a value of 2 for isotropic μ as a minimum. It measures the degree of anisotropy

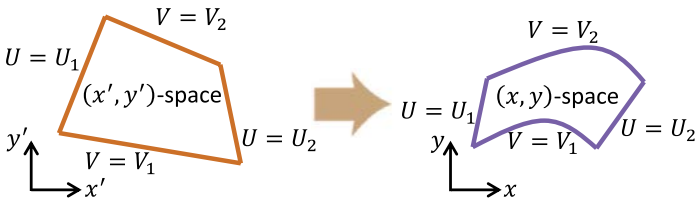


Fig. 11.1. A coordinate transformation maps corners to corners and edges to edges, through an intermediate coordinate (U, V) with common boundary values in both spaces. Quasiconformal map is defined as the one that there exists constants α and α' so that $(\alpha U, V)$ and $(\alpha' U, V)$ are conformally mapped to (x, y) and (x', y') respectively.

of μ if we are dealing with positive-definite permeability. For H_z-polarization, we only need to exchange the roles of ϵ and μ in Eq. (11.2).

11.3. The Quasi-Conformal Map

The investigation of a quasi-conformal map in TO is useful when we are limited to isotropic dielectric materials or metamaterials with small anisotropy. For convenience of discussion, a coordinate transformation in two dimensions can be described through an intermediate coordinate (U, V), as shown in Fig. 11.1. The boundaries of both spaces are generally curved, with the 4 corners in both spaces under transformation specified. Suppose we start from isotropic materials in (x', y') space (e.g. vacuum to be perceived for the case of a cloak). The materials are then sequentially transformed to (U, V), and then to (x, y) for physical realization. If we would like the maximum anisotropy of the physical medium to be minimized, it turns out that we should look for a mapping, in which it is valid to define two constants α and α' so that $(\alpha U, V)$ and $(\alpha' U, V)$ can be conformally mapped to (x, y) and (x', y') respectively. In other words, the overall map is a composition of two conformal maps and a linear map in the middle. It is technically called a Teichmüller map, a kind of quasiconformal maps which have a complex Beltrami coefficient of constant modulus and spatially varying argument. Here, we simply call these optimal maps as the quasiconformal maps for convenience because

they are orthogonal maps, just look like the familiar conformal maps (see also Refs. 7 and 43).

The quasiconformal property of the mapping can be written as

$$\begin{aligned} \alpha \nabla U &= -\hat{z} \times \nabla V, \\ \alpha' \nabla' U &= -\hat{z} \times \nabla' V. \end{aligned} \tag{11.3}$$

Taking divergence on Eq. (11.3) gives the Laplace equations on an analog potential V in both spaces:

$$\nabla^2 V = \nabla'^2 V = 0. \tag{11.4}$$

The quasiconformal map can be obtained by minimization of anisotropy directly but here we adopt the analog potential approach.[7,34] By integrating along an arbitrary path AB in the unprimed space and its corresponding transformed path in the primed space, we obtain

$$\gamma \triangleq \frac{\alpha'}{\alpha} = \frac{\int_A^B d\boldsymbol{l}' \cdot \hat{z} \times \nabla' V}{\int_A^B d\boldsymbol{l} \cdot \hat{z} \times \nabla V}, \tag{11.5}$$

which is actually the ratio of conformal moduli of the two spaces. The quasiconformal transformation medium can then be simplified from Eq. (11.2), with Eq. (11.3) and assumed isotropic μ', in the form of

$$\epsilon = \epsilon' \gamma \frac{|\nabla V|^2}{|\nabla' V|^2}, \quad \mu = \frac{\mu'}{\gamma} \hat{U}\hat{U} + \mu' \gamma \hat{V}\hat{V} \tag{11.6}$$

where the orthogonal basis vectors along the U-lines and V-lines in the unprimed space are given by $\hat{U} = \hat{V} \times \hat{z}$ and $\hat{V} = \nabla V / |\nabla V|$. The anisotropy measure defined in the last section is related to γ by $Tr(\mu)/\sqrt{\det \mu} = \gamma + 1/\gamma$, a constant in space. Similar formulas can be obtained for U to express the quasiconformal transformation medium.

The first example is a unidirectional cloak, which can hide an object in a particular direction. For illustration of the principle, we

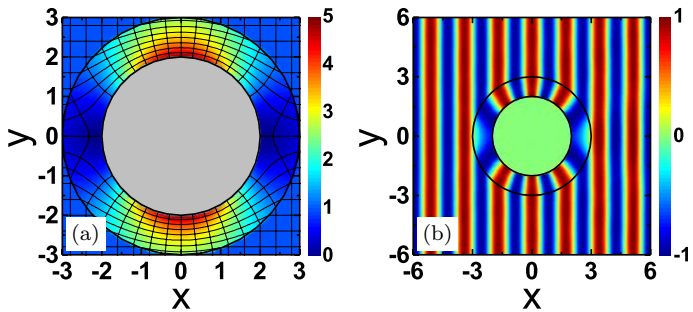


Fig. 11.2. A unidirectional cloak (from radii $a = 2$ to $b = 3$) with the quasiconformal map. (a) Permittivity profile with a square grid of the virtual space plotted in the physical space. (b) Full-wave simulation for a plane wave of wavelength 1.7 incident from the left.

take an analytic quasiconformal mapping

$$x' = \frac{b^2(r^2 + a^2)}{(b^2 + a^2)r^2}x, \quad y' = \frac{b^2(r^2 - a^2)}{(b^2 - a^2)r^2}y, \tag{11.7}$$

where $r^2 = x^2 + y^2$. It transforms a disk of radius b in the primed space to a circular ring of $a \leq r \leq b$, the cloak, in the unprimed space. An identity map is used outside the cloak for $r \geq b$. The quality of the map is shown in Fig. 11.2(a) for a square grid in the (x', y')-space. By taking $(U, V) = (x', y')$, from Eqs. (11.6) and (11.5), we can write the induced permittivity compactly as

$$\epsilon = \gamma |\nabla y'|^2, \tag{11.8}$$

with $\gamma = (b^2 - a^2)/(b^2 + a^2)$, meaning constant anisotropy throughout the whole device. $1/\gamma$ and γ are also the two principal values of magnetic permeability along the orthogonal x' and y'-lines in Fig. 11.2(a). These constitute the material parameters of the cloak. We have assumed vacuum as the background but a simple scaling on the resultant material profile can be easily applied when we change the background from vacuum to another homogeneous medium. Figure 11.2(b) shows the full-wave simulation (COMSOL Multiphysics for all numerical simulations here) when a plane wave

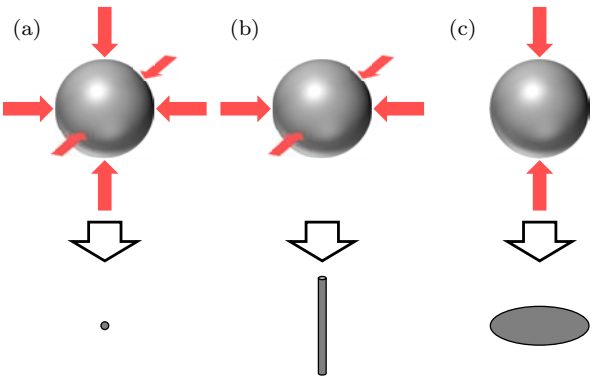


Fig. 11.3. Three ways in making an object invisible: To squash an object into (a) an infinitesimally small point, (b) an infinitesimally thin wire, and (c) an infinitesimally thin plate, corresponding to (a) spherical complete cloak, (b) cylindrical cloak and (c) carpet cloak.

impinges from the left. It passes through the cloak without being scattered.

Next, we use the quasiconformal map to design a carpet cloak. There are actually three different ways to hide an object by squashing it into a tiny point, a thin wire or a thin sheet, as shown in Fig. 11.3. The former two are related to a coordinate transformation with singularities, inducing an extremely wide range of material parameters and anisotropy required for the cloak while the third one does not require a coordinate transformation with singularity and therefore a quasiconformal map is a suitable tool to generate the required cloak. Being called a carpet cloak for this particular scheme, it can hide the scattering of a bump when the cloak is placed on top of it.

In the current example for illustration, we transform a rectangular region with dimension $w \times h (w = 3, h = 2)$ to the shape shown in Fig. 11.4(a) with its bottom boundary curved upwards. In this case, there is no straight-forward analytic solution. However, the quasiconformal map in this case (with a rectangle in the virtual space) can be very well approximated by solving Laplace equation on the coordinates.[34] By using intermediate coordinate simply as $(U, V) = (x', y')$ again, we set $y' = 0(2)$ at the bottom (top) boundary and Neumann boundary condition for the left and right

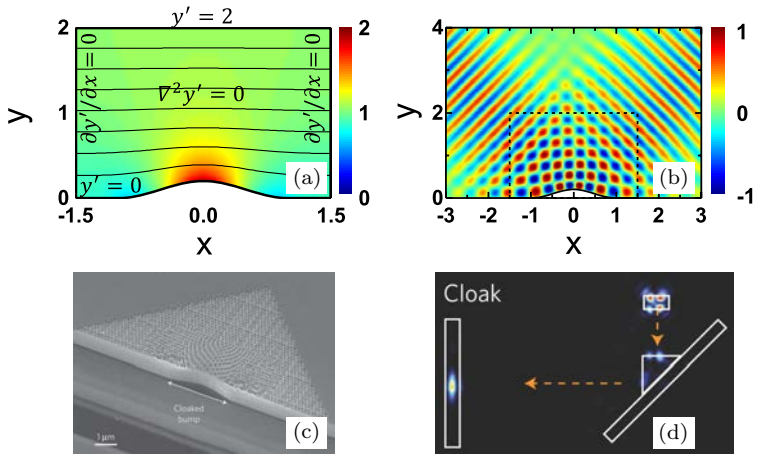


Fig. 11.4. A quasiconformal carpet cloak. (a) Induced permittivity profile of the cloak. (b) Full-wave simulation of E_z-profile with an incident beam at 45 degrees to the curved PEC surface with the cloak on top of it at a free-space wavelength of 0.4. (c) Scanning electron microscope image of the first infrared carpet cloak made from a hole lattice of varying density on a silicon-on-insulator wafer.[9] (d) Experimental result. An infrared Gaussian beam incident on the carpet cloak has reflection beam profile just like that reflected from a flat surface.

boundary to solve for "potential" y': $\nabla^2 y' = 0$ as an electrostatic analog. Figure 11.4(a) shows the numerically solved coordinate lines with constant y'. Since we have already assumed to establish a quasiconformal map, it turns out that solving y' without x' (or vice versa) will be enough to obtain the material profile. The constant γ, from Eq. (11.5), can then be easily probed out by

$$\frac{1}{\gamma} = \frac{1}{w}\int \frac{\partial y'}{\partial y} dx \tag{11.9}$$

with integration along the top-boundary of the cloak. Again, $1/\gamma$ and γ are the two orthogonal principal values of magnetic permeability, with one of them along these lines. The scalar permittivity can be obtained by $\epsilon = \gamma |\nabla y'|^2$ (Eq. 11.8). The permittivity profile, ranging from 0.65 to 1.91, is plotted in Fig. 11.4(a). For the full-wave simulation shown in Fig. 11.4(b), we have used a PEC boundary at the bottom, and the cloak is placed on top of it with the dashed line being the boundary of the cloak. In this case, $\gamma = 0.96$, which

is sufficiently close to the case of isotropic materials with value one for γ. Therefore, it can be neglected and we can set isotropic $\mu = 1$ for the cloak as an approximation. When there is an incident beam shining on the cloak together with the bump, chosen at 45 degrees here as an example, the cloak cancels out the original scattering on the bump and it just looks like a flat mirror with specular reflection. Although the scattering is faithfully cancelled out, we note that due to the isotropic approximation, there is unavoidably a small phase aberration, or lateral shift, in the cloak.[13] However, it will diminish to zero as the size of the cloak increases for the width and the height. Such carpet cloak can be realized with varying density of a hole lattice[9] or nanometer-sized silicon structures.[10] As an example, Fig. 11.4(c) shows the first infrared carpet cloak fabricated by drilling holes on a silicon-on-insulator wafer.[9] The density profile of the hole lattice provides the inhomogeneous permittivity profile. It cloaks a small bump and when an infrared Gaussian beam illuminates on it, the reflected light profile is just like that reflected from a flat mirror when the cloak is on top of the bump, as shown in Fig. 11.4(d).

We have been using a Laplace equation to get one set of coordinate lines. It is interesting to point out that a Poisson equation (also an electrostatic analog) with a point source can also be used to generate a quasiconformal map and the corresponding material profile. Suppose we are transforming a circular disk of radius 1 in the primed space to a regular hexagon in the unprimed space to act as a 6-way collimator when a point source is put at the center. In this case, we can set to solve a Poisson-like equation on V in both spaces

$$\begin{aligned} \nabla'^2 V &= 2\pi\delta(\vec{r}\,'), \\ \nabla^2 V &= 2\pi q\delta(\vec{r}\,), \end{aligned} \tag{11.10}$$

with $V = 0$ set at the outer boundary of the device. It can then be proved that $V = \ln r'$ in the primed space and the induced transformation medium (from Eq. (11.6) and Eq. (11.5) with closed path integration around the point charge) has the form of

$$\epsilon = \gamma|\nabla r'|^2, \quad \gamma = 1/q. \tag{11.11}$$

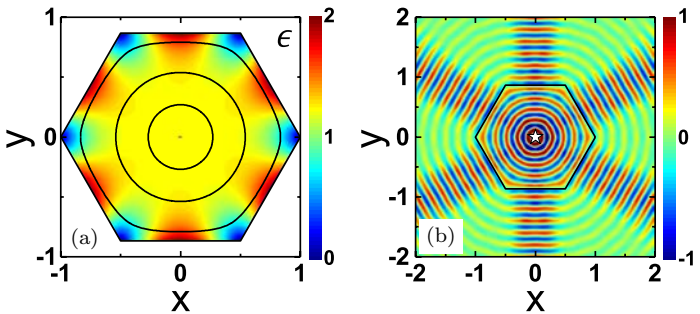


Fig. 11.5. Conformal TO medium with analog problem using a unit point charge within cavity. (a) Induced permittivity (color map) in the physical space. Solid lines are having constant $r' = 0.3$, 0.6, and 0.9. (b) A 6-way collimator using the TO medium with point source at origin as excitation, at a wavelength of 0.2.

The size of the point source q can be used to control the constant anisotropy of the generated medium. As a special case, we set $q = 1$ so that the quasiconformal map becomes strictly conformal. Figure 11.5(a) shows the numerically solved constant-r' lines in the unprimed (physical) space. Now, we put a point source at the origin, due to the transformation, the wavefront should arrive the outer boundary of the hexagon as if the point-source has its wavefront arriving the outer boundary of a circular disk at the same phase. As the boundary is flat, it becomes a beam travelling along the normal direction of each facet. Figure 11.5(b) shows the corresponding full-wave simulation, beams of 6 different directions are collimated from the point source at origin.

11.4. Linear Transformation

In contrast to quasi-conformal map which generates isotropic and inhomogeneous dielectric materials, linear transformation leads to materials which are homogeneous and anisotropic.[16,17] Under this transformation, a line in the initial coordinate is still transformed to a line in the physical space. Therefore, for an arbitrary triangular region, we can always find a unique linear transformation to map it to another one. This paves the way for the design of homogeneous cloaks and other transformation optical devices. Figures 11.6(a) and

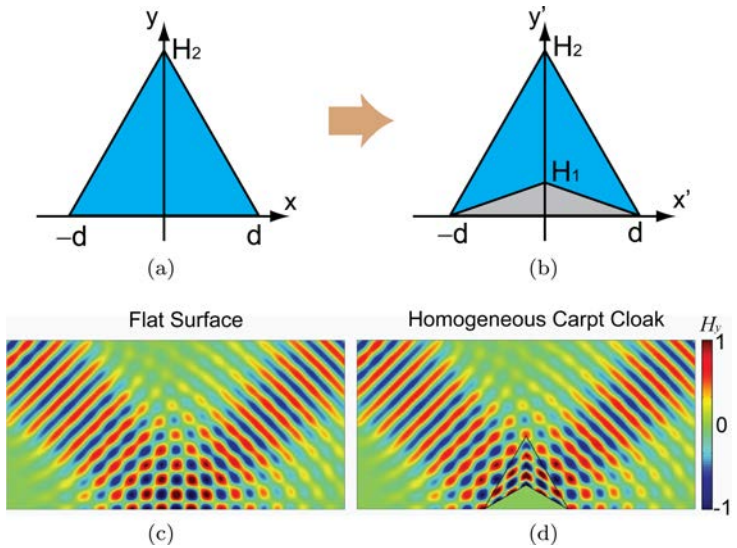


Fig. 11.6. (a, b) The bilinear transformation for the design of a carpet cloak. (c, d) Full-wave simulation of H_z-profiles with an incident beam at 45 degrees to the (c) flat ground and (d) the homogeneous carpet cloak.

(b) show the schematic of designing a homogeneous carpet cloak by applying a bilinear transformation which compresses the triangular region along the y' axis into a polygonal region (again, we restrict ourselves in two-dimensional cases with $z = z'$ assumed for simplicity):

$$x = x', \quad y = \frac{H_2 - H_1}{H_2} y' + \frac{d - |x'|}{d} H_1, \tag{11.12}$$

By using Eq. 11.2 (with the roles of ϵ and μ swapped) for the H_z-polarization to transform from vacuum, we obtain the permittivity and permeability profiles as

$$\epsilon = \begin{pmatrix} \dfrac{H_2}{H_2 - H_1} & \dfrac{H_1 H_2/d}{H_1 - H_2}\mathrm{sgn}(x) \\ \dfrac{H_1 H_2/d}{H_1 - H_2}\mathrm{sgn}(x) & \dfrac{H_2 - H_1}{H_2} + \dfrac{H_1^2 H_2/d^2}{H_2 - H_1} \end{pmatrix}$$
$$\mu = \frac{H_2}{H_2 - H_1} \tag{11.13}$$

We invoke the so called reduced-parameter approximation by lumping the permeability into the permittivity to design a non-magnetic ($\mu = 1$) carpet cloak, which now takes the form of

$$\epsilon = \begin{pmatrix} \dfrac{H_2^2}{(H_2 - H_1)^2} & -\dfrac{H_1 H_2^2/d}{(H_2 - H_1)^2}\mathrm{sgn}(x) \\ -\dfrac{H_1 H_2^2/d}{(H_2 - H_1)^2}\mathrm{sgn}(x) & 1 + \dfrac{H_1^2 H_2^2/d^2}{(H_2 - H_1)^2} \end{pmatrix} \tag{11.14}$$

When the background medium is changed to a dielectric medium (in an actual experiment), the above permittivity tensor is simply scaled by the background permittivity. Moreover, the reduced parameter approximation introduces a small mismatch of impedance or a small reflection on the interface between the cloak and the background medium, which is usually negligible.

Consider a transverse-magnetic (TM, magnetic field perpendicular to the cloak device) polarized Gaussian beam incident obliquely upon such a carpet cloak on top of a flat surface. The anisotropic cloak layer guides the beam around the bump, making the output beam propagating exactly the same way as that reflected from a flat surface, as shown in Figs. 11.6(c) and (d).

Note that the permittivity tensor in Equation (11.14) can be diagonalized by rotating the optical axis. Therefore, we may find natural birefringent crystals[18–20] or metamaterials with effective uniaxial material profiles, such as dielectric grating structures,[21–24] to realize such kind of cloaks. The latter offers more freedom as we can engineer the geometrical parameters, such as the filling factor of the gratings, to achieve anisotropy not found in natural materials. Figure 11.7 shows two examples of carpet cloaks designed with linear transformation. The top one is the first macroscopic carpet cloak for the whole visible spectrum,[18] made of two calcite crystals glued together. The bottom one is an infrared cloak made of silicon grating structures.[21]

The linear transformation can be further applied to design an omnidirectional cloak.[25,26] The basic principle is to expand a line segment into a cloaked region by stretching or compressing the space in two orthogonal directions, as illustrated in Fig. 11.8. The cloak

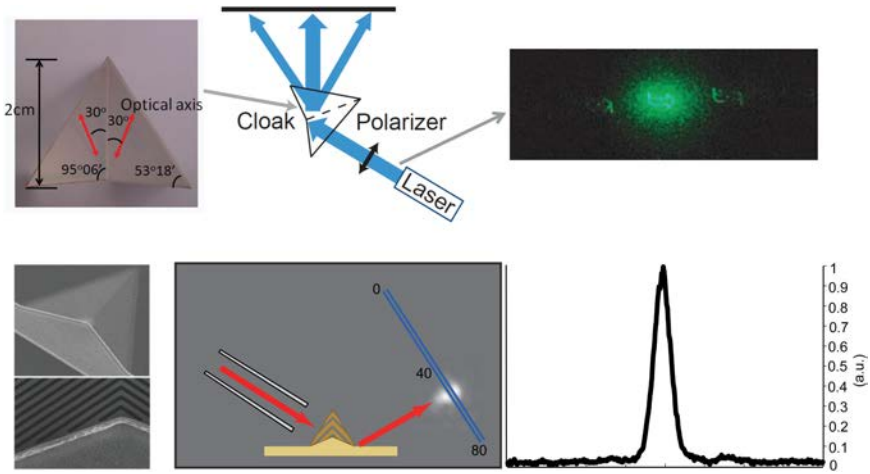


Fig. 11.7. (top) The experimental realization of the first macroscopic carpet cloak for the whole visible spectrum, made of two calcite crystals glued together.[18] (bottom) The experimental realization of a carpet cloak made of silicon gratings, working at infrared spectrum.[21]

shown is achieved by applying two different linear transformations. The gray regions are obtained by applying a space compression along $+y'$ and $-y'$ directions, while the pink regions are more complicated: The spaces are expanded along the y' axis and then compressed along the x' axis, as described by the following transformation:

$$\begin{aligned} x &= x', \\ y &= \operatorname{sgn}(y') \left(\frac{a_2 - a_1}{a_2 - a_0} |y'| + \frac{a_2(a_1 - a_0)}{b_2(a_2 - a_0)} (b_2 - |x'|) \right), \end{aligned} \tag{11.15}$$

for the gray regions and

$$\begin{aligned} x &= \operatorname{sgn}(x') \left(\frac{b_2 - b_1}{b_2 - b_0} |x'| + \frac{b_2(b_1 - b_0)}{a_0(b_2 - b_0)} (a_0 - |y'|) \right), \\ y &= \frac{a_1}{a_0} y' \end{aligned} \tag{11.16}$$

for the pink regions. As the white region in Fig. 11.8(a) is small enough, a nearly perfect cloak is obtained.

The resultant cloak consists of eight blocks of birefringent prisms and only two kinds of homogeneous materials are involved.

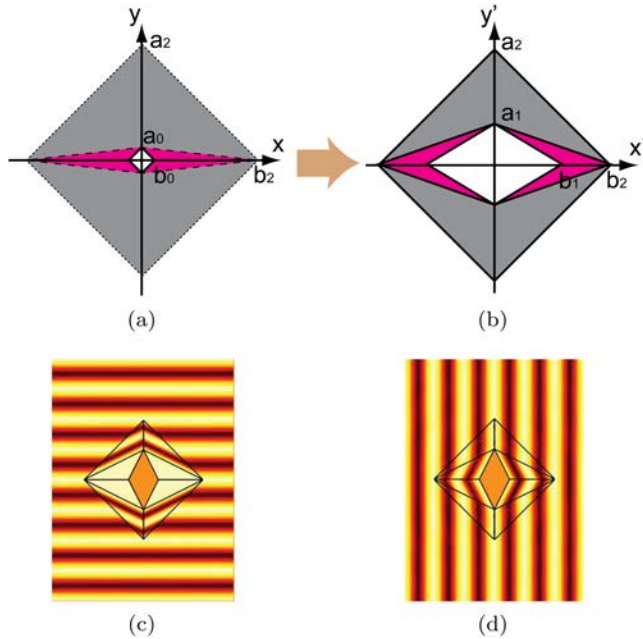


Fig. 11.8. (a, b) The linear transformation for the design of an omni-directional cloak. (c, d) Finite-element simulation of H_z-profiles with incident TM plane waves from two orthogonal directions.

The permeability and permittivity are

$$\epsilon = \begin{pmatrix} \dfrac{a_2 - a_0}{a_2 - a_1} & \dfrac{a_2(a_1 - a_0)}{b_2(a_1 - a_2)}\mathrm{sgn}(xy) \\ \dfrac{a_2(a_1 - a_0)}{b_2(a_1 - a_2)}\mathrm{sgn}(xy) & \dfrac{b_2^2(a_2 - a_1)^2 + a_2^2(a_1 - a_0)^2}{b_2^2(a_2 - a_0)(a_2 - a_1)} \end{pmatrix},$$
$$\mu = \frac{a_2 - a_0}{a_2 - a_1} \tag{11.17}$$

for the gray regions, and

$$\epsilon = \begin{pmatrix} \dfrac{a_0^2(b_2 - b_1)^2 + b_2^2(b_1 - b_0)^2}{a_0 a_1 (b_2 - b_0)(b_2 - b_1)} & \dfrac{b_2(b_1 - b_0)}{a_0(b_1 - b_2)}\mathrm{sgn}(xy) \\ \dfrac{b_2(b_1 - b_0)}{a_0(b_1 - b_2)}\mathrm{sgn}(xy) & \dfrac{a_1(b_2 - b_0)}{a_0(b_2 - b_1)} \end{pmatrix},$$
$$\mu = \frac{a_0(b_2 - b_0)}{a_1(b_2 - b_1)} \tag{11.18}$$

for the pink regions. Thus, without any further simplifications, a nearly perfect cloak composed of eight blocks of homogenous materials is obtained. Numerical simulations demonstrate that the cloak works effectively for light incident from any direction, as shown in Figs. 11.8 (c, d).

Apart from invisibility cloaks, the linear transformation can also be applied to design other photonic devices.[44–47] One example is a planar waveguide adapter which adapts electromagnetic waves for waveguides with different widths, specifically, to compress the energy from a planar waveguide with a width d_1 into another with a width d_2 as shown in Figure 11.9. The compression from the original waveguide in the virtual space with width d_1 is done in two steps. Each step corresponds to a transformation of a triangular region to another triangular region, in which the linear transformation can be uniquely determined. The linear map (with $z = z'$) can be written as

$$x = \alpha_x x' + \alpha_y y' + \alpha_0, \quad y = \beta_x x' + \beta_y y' + \beta_0, \tag{11.19}$$

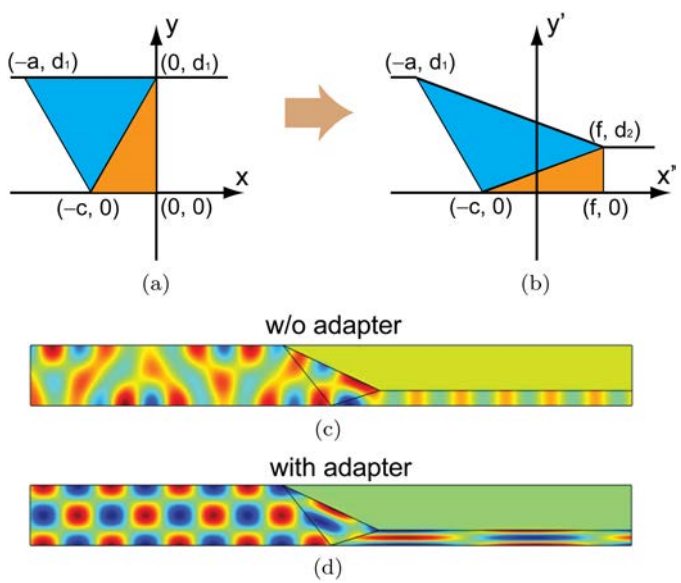


Fig. 11.9. (a, b) The linear transformation for the design of a planar waveguide adapter. (c, d) Finite element simulation of H_z-profiles with TM_{02} mode incident from the left.

where $\alpha_x = (a+f)/a, \alpha_y = f(a-c)/(ad_1), \alpha_0 = fc/a$, $\beta_x = (d_2 - d_1)/a, \beta_y = ((a-c)d_2 + cd_1)/(ad_1)$ and $\beta_0 = c(d_2 - d_1)/a$ for the upper (blue) triangular region, and

$$x = \frac{c+f}{c}x' + f, \quad y = \frac{d_2}{d_1}y', \tag{11.20}$$

for the bottom (brown) triangular region. The corresponding permeability and permittivity can then be obtained (from transforming vacuum region within the original waveguide) as

$$\epsilon = \frac{1}{\alpha_x\beta_y - \alpha_y\beta_x}\begin{pmatrix} \alpha_x^2 + \alpha_y^2 & \alpha_x\beta_x + \alpha_y\beta_y \\ \alpha_x\beta_x + \alpha_y\beta_y & \beta_x^2 + \beta_y^2 \end{pmatrix},$$
$$\mu = \frac{1}{\alpha_x\beta_y - \alpha_y\beta_x} = \frac{ad_1}{(a-c)d_2 + (c+f)d_1} \tag{11.21}$$

for the upper (blue) triangular region, and

$$\epsilon = \begin{pmatrix} \dfrac{(c+f)d_1}{cd_2} & 0 \\ 0 & \dfrac{cd_2}{(c+f)d_1} \end{pmatrix}, \quad \mu = \frac{cd_1}{(c+f)d_2}, \tag{11.22}$$

for the bottom (brown) triangular region.

Equations (11.21) and (11.22) indicate that the adapter can be realized with nonmagnetic materials provided that its geometrical parameters satisfy

$$c = \frac{ad_2}{d_1 + d_2}, \quad f = a\left(\frac{d_1 - d_2}{d_1 + d_2}\right). \tag{11.23}$$

Figures 11.9(c) and (d) depict the results when electromagnetic waves transmit through the two planar waveguides with different widths. Without the waveguide adapter, the mode profile was severely

disturbed, and only a small portion of electromagnetic energy was transmitted into the thinner waveguide. On the contrary, the homogeneous nonmagnetic adapter can smoothly guide electromagnetic waves through the waveguide without reflection. The similar concept can also be applied to the design of plasmonic waveguide adaptors (see Ref. 48 for more details).

11.5. Conclusion

Quasiconformal and linear mappings are two kinds of transformations that can be used to design carpet cloaks and a range of transformation optical devices. Both methods allow devices to be fabricated with dielectric materials, making transformation optics applications more straight-forward to implement and with large frequency bandwith. Moreover, they each have their own advantages and can be used for different purposes. Quasiconformal mapping can minimize the anisotropy in the transformed medium, yielding optical devices realized using isotropic permittivity profiles. Such an approach is suitable for both TE and TM polarizations of light. On the other hand, linear transformation gives rise to homogeneous material parameters, thereby allowing for the realization of macroscopic transformational optical devices for TM light polarization.

Acknowledgments

J. L. would like to acknowledge funding support from the European Union's Seventh Framework Programme under Grant Agreement No. 630979. Y. L. would like to acknowledge support from Singapore Ministry of Education (MOE) AcRF TIER 1 under Grant No RG72/15, AcRF TIER 2 under Grant No MOE2015-T2-1-145.

References

1. Stevens, M. and Merilaita, S. (2011), *Animal Camouflage: Mechanisms and Function* (Cambridge University press, UK).
2. Pendry, J. B., Schurig, D. and Smith, D. R. (2006). Controlling electromagnetic fields, *Science* **312**, pp. 1780–1782.

3. Leonhardt, U. (2006). Optical conformal mapping, *Science* **312**, pp. 1777–1780.
4. Schurig, D., Mock, J. J., Justice, B. J., Cummer, S. A., Pendry, J. B., Starr, A. F. and Smith, D. R. (2006). Metamaterial electromagnetic cloak at microwave frequencies, *Science* **314**, pp. 977–980.
5. Jiang, W. X., Cui, T. J., Yang, X. M., Cheng, Q., Liu, R. and Smith, D. R. (2008). Invisibility cloak without singularity, *Appl. Phys. Lett.* **93**, p. 194102.
6. Leonhardt, U. and Tyc, T. (2009). Broadband invisibility by non-Euclidean cloaking, *Science* **323**, pp. 110–112.
7. Li, J. and Pendry, J. B. (2008). Hiding under the carpet: a new strategy for cloaking, *Phys. Rev. Lett.* **101**, pp. 203901.
8. Liu, R., Ji, C., Mock, J. J., Chin, J. Y., Cui, T. J. and Smith, D. R. (2009). Broadband ground-plane cloak, *Science* **323**, pp. 366–369.
9. Valentine, J., Li, J., Zentgraf, T., Bartal, G. and Zhang, X. (2009). An optical cloak made of dielectrics, *Nat. Mater.* **8**, pp. 568–571.
10. Gabrielli, L. H., Cardenas, J., Poitras, C. B. and Lipson, M. (2009). Silicon nanostructure cloak operating at optical frequencies, *Nat. Photonics* **3**, pp. 461–463.
11. Ergin, T., Stenger, N., Brenner, P., Pendry, J. B. and Wegener, M. (2010). Three-dimensional invisibility cloak at optical wavelengths, *Science* **328**, pp. 337–339.
12. Ma, H. F. and Cui, T. J. (2010). Three-dimensional broadband ground-plane cloak made of metamaterials, *Nat. Commun.* **1**, p. 21.
13. Zhang, B., Chan, T. and Wu, B. I. (2010). Lateral shift makes a ground-plane cloak detectable, *Phys. Rev. Lett.* **104**, p. 233903.
14. Zhou, F., Bao, Y., Cao, W., Stuart, C. T., Gu, J., Zhang, W. and Sun, C. (2011). Hiding a realistic object using a broadband terahertz invisibility cloak, *Sci. Rep.* **1**, p. 78.
15. Gharghi, M., Gladden, C., Zentgraf, T., Liu, Y., Yin, X., Valentine, J. and Zhang, X. (2011). A carpet cloak for visible light, *Nano lett.* **11**, pp. 2825–2828.
16. Luo, Y., Zhang, J., Chen, H., Ran, L., Wu. B.-I. and Kong, J. A. (2009). A rigorous analysis of plane-transformed invisibility cloaks, *IEEE Trans. Antennas Propagat* **57**, pp. 3926–3933.
17. Xi S., Chen, H., Wu. B.-I. and Kong, J. A. (2009). One-directional perfect cloak created with homogeneous material, *IEEE Microw. Compon. Lett.* **19**, pp. 131–133.
18. Chen, X., Luo, Y., Zhang, J., Jiang K., Pendry, J. B. and Zhang, S. (2011). Macroscopic invisibility cloaking of visible light, *Nat. Commun* **2**, p. 176.
19. Zhang, B., Luo, Y., Liu, X. and Barbastathis, G. (2011). Macroscopic invisibility cloak for visible light, *Phys. Rev. Lett.* **106**, p. 033901.
20. Liang, D., Gu, J., Han, J., Yang, Y., Zhang, S. and Zhang, W. (2012). Robust large dimension terahertz cloaking, *Adv. Mater.* **24**, pp. 916–921.

21. Zhang, J., Liu, L., Luo, Y., Zhang, S. and Mortensen, N. A., (2011). Homogeneous optical cloak constructed with uniform layered structures, *Opt. Express* **9**, pp. 8625–8631.
22. Xu, X., Feng, Y., Xiong, S., Fan, J., Zhao, J. M. and Jiang, T. (2011). Broad band invisibility cloak made of normal dielectric multilayer, *Appl. Phys. Lett.* **99**, p. 154104.
23. Popa, B. I., Zigoneanu, L. and Cummer, S. A. (2011). Experimental acoustic ground cloak in air, *Phys. Rev. Lett.* **106**, p. 253901.
24. Zigoneanu, L., Popa, B. I. and Cummer, S. A. (2014). Three-dimensional broadband omnidirectional acoustic ground cloak, *Nat. Mater* **13**, pp. 352–355.
25. Li, W., Guan, J., Sun, Z., Wang, W. and Zhang, Q. (2009). A near-perfect invisibility cloak constructed with homogeneous materials, *Opt. Express* **17**, pp. 23410–23416.
26. Chen, H. and Zheng, B. (2012). Broadband polygonal invisibility cloak for visible light, *Sci. Rep.* **2**, p. 255.
27. Landy, N. and Smith, D. R. (2013). A full-parameter unidirectional metamaterial cloak for microwaves, *Nat. Mater.* **12**, pp. 25–28.
28. Rahm, M., Schurig, D., Roberts, D. A., Cummer, S. A., Smith, D. R. and Pendry, J. B. (2008). Design of electromagnetic cloaks and concentrators using form-invariant coordinate transformations of Maxwell's equations, *Photon. Nanostr. Fundam. Appl.* **6**, pp. 87–95.
29. Kildishev, A. V. and Shalaev, V. M. (2008). Engineering space for light via transformation optics, *Opt. Lett.* **33**, pp. 43–45.
30. Genov, D. A., Zhang, S. and Zhang, X. (2009). Mimicking celestial mechanics in metamaterials, *Nat. Phys.* **5**, pp. 687–692.
31. Chen, H., Hou, B., Chen, S., Ao, X., Wen, W. and Chan, C. T. (2009). Design and experimental realization of a broadband transformation media field rotator at microwave frequencies, *Phys. Rev. Lett.* **102**, p. 183903.
32. Lai, Y., Ng, J., Chen, H., Han, D., Xiao, J., Zhang, Z. Q. and Chan, C. T. (2009). Illusion optics: the optical transformation of an object into another object, *Phys. Rev. Lett.* **102**, p. 253902.
33. Landy, N. I. and Padilla, W. J. (2009). Guiding light with conformal transformations, *Opt. Express* **17**, pp. 14872–14879.
34. Chang, Z., Zhou, X., Hu, J. and Hu, G. (2010). Design method for quasi-isotropic transformation materials based on inverse Laplace's equation with sliding boundaries, *Opt. Express* **18**, pp. 6089–6096.
35. Zentgraf, T., Valentine, J., Tapia, N., Li, J. and Zhang, X. (2010). An optical "Janus" device for integrated photonics, *Advanced Materials* **22**, pp. 2561–2564.
36. Landy, N. I., Kundtz, N. and Smith, D. R. (2010). Designing three-dimensional transformation optical media using quasiconformal coordinate transformations, *Phys. Rev. Lett.* **105**, p. 193902.

37. Mei, Z. L., Bai, J. and Cui, T. J. (2011). Experimental verification of a broadband planar focusing antenna based on transformation optics, *New J. Phys.* **13**, p. 063028.
38. Yang, R., Tang, W. and Hao, Y. (2011). A broadband zone plate lens from transformation optics, *Opt. Express* **19**, pp. 12348–12355.
39. García-Meca, C., Martínez, A. and Leonhardt, U. (2011). Engineering antenna radiation patterns via quasi-conformal mappings, *Opt. Express* **19**, pp. 23743–23750.
40. Wu, Q., Turpin, J. P. and Werner, D. H. (2012). Integrated photonic systems based on transformation optics enabled gradient index devices, *Light Sci. Appl.* **1**, p. e38.
41. Gabrielli, L. H., Liu, D., Johnson, S. G. and Lipson, M. (2012). On-chip transformation optics for multimode waveguide bends, *Nat. Commun.*,**3**, p. 1217.
42. Liu, D., Gabrielli, L. H., Lipson, M. and Johnson, S. G. (2013). Transformation inverse design, *Opt. Express* **21**, pp. 14223–14243.
43. Weber, O., Myles, A. and Zorin, D. (2012). Computing extremal quasiconformal maps, *Computer Graphics Forum* **31**, pp. 1679–1689.
44. Zhang, J., Luo, Y. and Mortensen, N. A. (2010). Transmission of electromagnetic waves through sub-wavelength channels, *Opt. Express* **18**, pp. 3864–3870.
45. Wang, Y., Zhang, D. H., Wang, J., Yang, X., Li, D. and Xu, Z. (2011). Waveguide devices with homogeneous complementary media, *Opt. Lett.* **36**, pp. 3855–3857.
46. Xu, H., Zhang, B., Yu, T., Barbastathis and Sun, H. (2012). Dielectric waveguide bending adapter with ideal transmission: practical design strategy of area-preserving affine transformation optics, *JOSA B* **29**, pp. 1287–1290.
47. Wee, W. H., Ye, Y. J. and Luo, Y. (2016). Towards a practical compact magnifying superlens — a simple simplical design, *J. Opt.* **18**, p. 044011.
48. Zhang, J., Xiao, S., Wubs, M. and Mortensen, N. A. (2011). Surface plasmon wave adapter designed with transformation optics, *ACS Nano* **5**, pp. 4359–4364.

CHAPTER 12

Hyperelastic Cloaking: Transformation Elasticity Using Pre-stressed Solids

ANDREW N. NORRIS

Rutgers University, USA
norris@rutgers.edu

WILLIAM J. PARNELL

University of Manchester, UK
William.J.Parnell@manchester.ac.uk

12.1. Introduction

Various approaches to cloaking in the elastodynamic setting have been presented but these can conveniently be partitioned into *passive* and *active* methods. Passive cloaking is usually based on transformation elastodynamics to motivate a choice of inhomogeneous metamaterial [Milton *et al.* (2006); Brun *et al.* (2009); Norris and Shuvalov (2011)] whereas active cloaking uses active sources [Norris *et al.* (2013); O'Neill *et al.* (2014); Futhazar *et al.* (2015)]. Passive cloaking theories in electromagnetism and acoustics are underpinned by the *transformation* or *change-of-variables* method [Greenleaf *et al.* (2003); Pendry *et al.* (2006)]: material properties of the cloak are defined by spatial transformations. The general

principals involved for both applications are similar so that the first applications in electromagnetism (see e.g. [Schurig *et al.* (2006)]) were similarly applied to acoustics [Cummer and Schurig (2007); Chen and Chan (2007); Cummer *et al.* (2008)] but various aspects of the theory are quite different. The fundamental property underlying transformations in both the acoustics and electromagnetic cases is the observation that the Laplacian in the original coordinates maps to a differential operator in physical coordinates that involves a tensor that can be interpreted as having the new, transformed, material properties embedded within it [Greenleaf *et al.* (2007)]. The equivalence between the Laplacian in the original coordinates and the new operator involves an arbitrary divergence free tensor [Norris (2008)], implying that for the acoustic case the transformed material properties are not unique, unlike the case of electromagnetism. For a given transformation function one can achieve acoustic cloaking by a variety of materials, ranging from fluids with anisotropic inertia, to quasi-solids with isotropic density but anisotropic stiffness [Norris (2008, 2009)]. In the electromagnetic case, non-uniqueness of the material properties does not apply: the permittivity and permeability tensors must be proportionate for a transformation of the vacuum.

A crucial aspect of transformation optics and acoustics is that the mapped properties correspond to exotic material properties far removed from the realm of the original material. This aspect is accentuated in transformation elasticity where shear and compressional waves couple and so simultaneous control of both wave types is required. The first study of transformation elastodynamics was the work by Milton *et al.* [Milton *et al.* (2006)], which concluded that the transformed materials are described by the Willis equations. This constitutive theory for material response is dispersive, involving coupling between stress and velocity, in addition to anisotropic inertia [Strutt (1871); Willis (1997); Milton and Willis (2007)]. Brun *et al.* [Brun *et al.* (2009)] considered the transformation of isotropic elasticity in cylindrical coordinates and found transformed material properties with isotropic inertia and elastic behavior of Cosserat type. The governing equations for Cosserat elastic materials [Cosserat and

Cosserat (1909)] are the same as those of standard linear elasticity except that the elastic moduli do *not* satisfy the minor symmetry, i.e. $C^{\text{eff}}_{jik\ell} \neq C^{\text{eff}}_{ijk\ell}$ (although they do satisfy the major symmetry $C^{\text{eff}}_{k\ell ij} = C^{\text{eff}}_{ijk\ell}$). This implies a non-symmetric stress, $\mathbf{T}^T \neq \mathbf{T}$ (where the superscript T denotes transpose) which depends not only on the strain (the symmetric part of the displacement gradient) but also on the local rotation $\frac{1}{2}(\nabla\mathbf{u} - (\nabla\mathbf{u})^T)$ where $\mathbf{u}$ is the displacement.

Thorough analysis of transformation elastodynamics reveals that the range of mapped material properties is highly non-unique, as in the case of acoustics [Norris and Shuvalov (2011)]. This explains the divergence in previously obtained results [Milton *et al.* (2006); Brun *et al.* (2009)]. The transformed elastodynamic constitutive parameters may be characterized through their dependence on (i) the transformation (mapping function) and (ii) on the relation between the displacement fields in the two descriptions, represented by matrices: $\mathbf{F}$, the deformation matrix, and $\mathbf{A}$ (the so-called *gauge matrix*), respectively. It was shown that requiring stress to be symmetric implies $\mathbf{A} = \mathbf{F}$ [Norris and Shuvalov (2011)] and that the material must be of Willis form, as in Milton *et al.* [Milton *et al.* (2006)]. Alternatively, setting $\mathbf{A} = \mathbf{I}$ results in Cosserat materials with non-symmetric stress but isotropic density, as in Brun *et al.* [Brun *et al.* (2009)] and also by Vasquez *et al.* [Guevara Vasquez *et al.* (2013)]. A different approach to transformation elasticity has been proposed that employs inextensible fibers embedded in an elastic material [Olsson and Wall (2011); Olsson (2011)]. This has the advantage that the effective material has isotropic density and retains both the minor and major symmetries of the stiffness tensor. Micropolar materials have also been employed in order to render cloaking effects for Rayleigh surface waves [Khlopotin *et al.* (2015)] and experiments regarding the seismic case were considered by [Brulé *et al.* (2015)]. A variety of behaviour of *static* properties associated with elastic media have been predicted and measured, see the work of [Bückmann *et al.* (2014, 2015a,b); Kadic *et al.* (2014, 2015)]. For recent thorough reviews of elastic metamaterials see [Kadic *et al.* (2013)], [Christensen *et al.* (2015)] and [Srivastava (2015)].

Although significant progress has recently been made in the context of waves in thin plates, [Farhat *et al.* (2009); Stenger *et al.* (2012); Colquitt *et al.* (2014)], the major difficulty associated with all current theories of elastodynamic metamaterials in a general setting is that the required materials are, in all cases, beyond the realms of possibility practically speaking. One way to circumvent many of the difficulties regarding elastodynamic cloaking is to use pre-stressed nonlinear elastic materials. Correctly-chosen pre-stress coupled with specific strain energy functions yields incremental material properties that are exactly right in order to guide linear elastic waves around chosen regions of space, enabling a cloaking effect [Parnell (2012)]. What is more, this class of materials displays non-symmetric stress of the type necessary to achieve elastodynamic cloaking [Norris and Parnell (2012)]: moduli with the major symmetry ($C^{\text{eff}}_{ijk\ell} = C^{\text{eff}}_{k\ell ij}$) that do not display the minor symmetry ($C^{\text{eff}}_{ijk\ell} \neq C^{\text{eff}}_{jik\ell}$) are found in the theory of incremental motion superimposed on finite deformation [Ogden (2007)].

The starting point for the *theory of hyperelastic cloaking* is the formal equivalence of the constitutive parameters of transformation elasticity with the density and moduli for incremental motion after finite pre-stress. This offers the possibility for achieving elasticity of the desired form by proper selection of the finite (hyperelastic) strain energy function. Such an approach has been shown to be successful in the context of *antiplane* or *horizontally polarized shear* (SH) wave motion. By using the neo-Hookean strain energy function for incompressible solids and applying a radially symmetric cylindrical pre-strain, [Parnell (2012)] and [Parnell *et al.* (2012)] showed that the resulting small-on-large equations are identically those required for cloaking of the SH wave motion. Comparisons with layered metamaterials and other nonlinear media were made in [Parnell *et al.* (2013)]. All of this theory is detailed in Sec. 12.4. The general *compressible* two-dimensional case was considered in [Norris and Parnell (2012)] where it was shown that for the two dimensional coupled compressional/shear problem, the form of the finite strain energy is restricted in form for isotropic elasticity. The equivalence between the

transformation and the finite pre-strain also limits the type of transformation possible. This contrasts with the acoustic and electromagnetic problems for which the transformation is arbitrary. The elastic transformation is restricted in form because the pre-stress must be statically equilibrated, implying that the transformation must satisfy a partial differential equation. It was also shown that this constraint can be cast as $\mathrm{tr}\mathbf{F} =$ constant, subject to symmetry constraints. In order to accommodate hyperelastic invariance it was shown that a so-called *semi-linear* strain energy function is the appropriate material required for cloaking. The three dimensional setting is still open for discussion: technical difficulties arise that shall be mentioned briefly later on. Following on from this work, a theory for the disentanglement of compressional and shear waves using pre-stress of a neo-Hookean medium was developed in [Chang *et al.* (2015); Galich and Rudykh (2015)]. The SH wave problem was also considered by [Wu and Gao (2015)].

In this review of hyperelastic cloaking, we start with a discussion of transformation elastodynamics in Sec. 12.2 before describing the theory of nonlinear elasticity and associated details of *small-on-large* in Sec. 12.3. The antiplane elastic wave problem is summarized in Sec. 12.4. The general hyperelastic invariance theory is described in Sec. 12.5 and applied in the two-dimensional, isotropic scenario in Sec. 12.6. We close with a summary and conclusions in Sec. 12.7.

12.2. Transformation Elastodynamics

First introduce the equations of linear elasticity governing the infinitesimal particle displacement $\mathbf{u}^0(\mathbf{X}^0, t)$ for $\mathbf{X}^0 \in B_0$, the so-called *virtual configuration*, terminology that will be clarified shortly:

$$\mathrm{Div}_0\mathbf{T}^0 = \rho_0\ddot{\mathbf{U}}^0, \quad \mathbf{T}^0 = \mathbf{C}^0\mathrm{Grad}_0\mathbf{U}^0 \quad \text{in } B_0. \tag{12.1}$$

Here $\mathbf{T}^0(\mathbf{X}^0, t)$ is the linear Cauchy stress, ρ_0 is the (scalar) mass density and $\mathbf{C}^0$ is the elastic modulus tensor with full symmetry properties $C^0_{IJKL} = C^0_{JIKL}$, $C^0_{IJKL} = C^0_{KLIJ}$; the first identity

expresses the symmetry of the stress and the second is the consequence of an assumed strain energy density function. Note that we use Div_0 and Grad_0 to refer to the divergence and gradient operators respectively, with respect to $\mathbf{X}_0$ and $\dot{f}$ refers to the time derivative of the function f.

12.2.1. *Transformation notation*

Of interest is how the equations of linear elastodynamics (12.1) transform given a spatial transformation (or mapping) from $\mathbf{X}^0 \in B_0$ to $\mathbf{x} \in b$, which will be referred to as the current (or *physical*) configuration. Component subscripts in upper and lower case $(I, J, \ldots, i, j, \ldots)$ are used to distinguish between explicit dependence upon $\mathbf{X}_0$ and $\mathbf{x}$, and the summation convention on repeated subscripts is assumed. The transformation or mapping is assumed to be one-to-one and invertible. It shall be shown that perfect cloaking requires that the transformation is one-to-many at a single point in B_0. This can be avoided by considering near-cloaks, where, for instance, a small hole in B_0 is mapped to a larger hole in b. The transformation gradient associated with the transformation is defined as $\mathbf{F}^0 = \mathrm{Grad}_0\mathbf{x}$ with inverse $\mathbf{F}_0^{-1} = \mathrm{grad}\,\mathbf{X}^0$, or in component form $F^0_{iI} = \partial x_i/\partial X^0_I$, ${F^0_{iI}}^{-1} = \partial X^0_I/\partial x_i$. The Jacobian of the transformation is $J_0 = \det \mathbf{F}^0$ and grad refers to the gradient operator with respect to $\mathbf{x}$.

Particle displacement in the transformed domain, $\mathbf{u}^{\mathrm{eff}}(\mathbf{x}, t)$, is assumed to be related to the displacement in the virtual domain by the non-singular matrix $\mathbf{A}$ as

$$\mathbf{U}^0 = \mathbf{A}^T\mathbf{u}^{\mathrm{eff}} \qquad (U^0_I = A_{iI}u_i^{\mathrm{eff}}) \tag{12.2}$$

where the superscript T denotes transpose. The choice of the matrix, $\mathbf{A}^T$ in (12.2) means that the gauge $\mathbf{A}$ and the transformation gradient $\mathbf{F}^0$ are similar objects, although at this stage they are unrelated. Milton *et al.* [Milton *et al.* (2006)] specify $\mathbf{A} = \mathbf{F}^0$ since this is the only choice that guarantees a symmetric stress (see §12.2.2). Identifying (see Milton *et al.* 2006) $\mathrm{d}\mathbf{X}_0$ and $\mathrm{d}\mathbf{x}$ with $\boldsymbol{U}^0$ and $\boldsymbol{u}^{\mathrm{eff}}$, respectively, and using $\mathrm{d}\mathbf{X}_0 = {\mathbf{F}^0}^{-1}\,\mathrm{d}\mathbf{x}$ would lead one to expect $\mathbf{A}^T = {\mathbf{F}^0}^{-1}$.

However, the displacements are not associated with the coordinate transformation and $\mathbf{F}^0$ and $\mathbf{A}$ are independent quantities.

12.2.2. *The transformed equations of linear elastodynamics*

Under the transformation (or change of coordinates) $\mathbf{X}_0 \to \mathbf{x}$ the equations of motion and constitutive relations (12.1) transform to [Norris and Shuvalov (2011)]

$$\begin{aligned} T^{\text{eff}}_{ij,i} &= \dot{p}^{\text{eff}}_j, \quad T^{\text{eff}}_{ij} = C^{\text{eff}}_{ijk\ell} u^{\text{eff}}_{\ell,k} + S^{\text{eff}}_{ij\ell}\, \dot{u}^{\text{eff}}_\ell, \\ p^{\text{eff}}_\ell &= S^{\text{eff}}_{ij\ell}\, u^{\text{eff}}_{j,i} + \rho^{\text{eff}}_{j\ell} \dot{u}^{\text{eff}}_j, \quad \text{in } b, \end{aligned} \tag{12.3}$$

with parameters $\mathbf{C}^{\text{eff}}$, $\mathbf{S}^{\text{eff}}$ and $\boldsymbol{u}^{\text{eff}}$ defined as follows in the Fourier time domain (dependence $e^{-i\omega t}$ understood but omitted)

$$C^{\text{eff}}_{ijk\ell} = J_0 C^0_{IJKL} \mathcal{Q}_{ijIJ} \mathcal{Q}_{k\ell KL}, \tag{12.4}$$

$$S^{\text{eff}}_{ijl} = (-i\omega)^{-1} J_0 C^0_{IJKL} \mathcal{Q}_{ijIJ} \mathcal{Q}_{k\ell KL,k}, \tag{12.5}$$

$$\rho^{\text{eff}}_{jl} = \rho_0 J_0^{-1} A_{jK} A_{\ell K} + (-i\omega)^{-2} J_0 C^0_{IJKL} \mathcal{Q}_{ijIJ,i} \mathcal{Q}_{k\ell KL,k}, \tag{12.6}$$

where $\mathcal{Q}_{ijIJ} = J_0^{-1} F^0_{iI} A_{jJ}$. The elastic moduli and the density satisfy the symmetries

$$C^{\text{eff}}_{ijk\ell} = C^{\text{eff}}_{k\ell ij}, \quad \rho^{\text{eff}}_{j\ell} = \rho^{\text{eff}}_{\ell j}, \tag{12.7}$$

although these are not the full symmetries for the Willis constitutive model (which requires the additional "minor" symmetry $C^{\text{eff}}_{ijk\ell} = C^{\text{eff}}_{jik\ell}$). Equations (12.3)–(12.6) are the fundamental result of elastic transformation theory [Norris and Shuvalov (2011)].

The absence of the minor symmetries under the interchange of i and j in $C^{\text{eff}}_{ijk\ell}$ and $S^{\text{eff}}_{ij\ell}$ of (12.4) and (12.5) implies that the stress is generally asymmetric. Symmetric stress is guaranteed if $\mathcal{Q}_{ijIJ} = \mathcal{Q}_{jiIJ}$, and occurs if the gauge matrix is of the form $\mathbf{A} = \alpha \mathbf{F}^0$, for any scalar $\alpha \neq 0$, which may be set to unity with no loss in generality. This $\mathbf{A}$ recovers the results of [Milton *et al.* (2006)] that the transformed material is of the Willis form. As

noted in [Milton *et al.* (2006)], it is the only choice for $\mathbf{A}$ that yields symmetric stress.

The equations in the transformed domain, i.e. the physical scenario, clearly display a great deal of non-uniqueness, corresponding to a vast realm of possible material properties. Our preference is for non-dispersive (i.e. independent of frequency) materials, in particular, those with the least "unusual" properties, so that they can conceivably be related to actual materials. In this regard, isotropic density is achieved by taking the constant matrix $\mathbf{A}$ proportional to the identity, $\mathbf{A} = \alpha\mathbf{I}$, with $\alpha = 1$ without loss of generality. In this case $\boldsymbol{\rho}^{\text{eff}} = \rho^{\text{eff}}\mathbf{I}$, $\mathbf{S}^{\text{eff}} = 0$, with non-dispersive density and elastic moduli given by

$$\rho^{\text{eff}} = J_0^{-1}\rho_0, \quad C^{\text{eff}}_{ijk\ell} = J_0^{-1}\, F^0_{iI} F^0_{kK}\, C^0_{IjK\ell}. \tag{12.8}$$

The equations of motion in the current (physical) domain are then

$$(C^{\text{eff}}_{ijk\ell} u^{\text{eff}}_{\ell,k})_{,i} = \rho^{\text{eff}}\, \ddot{u}^{\text{eff}}_j. \tag{12.9}$$

The effective moduli of $(12.8)_2$ satisfy the major symmetry $(12.7)_1$ but $C^{\text{eff}}_{ijk\ell} \neq C^{\text{eff}}_{jik\ell}$, indicating a non-symmetric stress. Departure from symmetric stress is possible in continuum theories such as Cosserat elasticity and micropolar theories of elasticity. Another context admitting non-symmetric stress is the theory of small-on-large, i.e. wave propagation in pre-stressed nonlinear elastic media, which we shall now describe.

12.3. Finite Deformation Elasticity and Small-On-Large

Also known as the *theory of incremental deformation*, the *theory of small-on-large* analyses small perturbations about a (stable) nonlinearly pre-stressed equilibrium state. Here we are interested in perturbations that are time-harmonic ($e^{-i\omega t}$ for consistency with the analysis above). In order to derive the appropriate equations it is most useful to consider two finite deformations that are close to each other.

12.3.1. *Kinematics*

Consider an undeformed body of homogeneous elastic material in the state B, with surface S, and let it be subjected to the two different stable deformations

$$\mathbf{x} = \mathbf{x}(\mathbf{X}, t), \quad \overline{\mathbf{x}} = \overline{\mathbf{x}}(\mathbf{X}, t), \tag{12.10}$$

which take a point $\mathbf{X} \in B$ to the point $\mathbf{x} \in b$ or $\overline{\mathbf{x}} \in \overline{b}$ respectively. The position vectors $\mathbf{X}, \mathbf{x}$ and $\overline{\mathbf{x}}$ have common origin. Define the displacement of these deformations to be

$$\mathbf{U} = \mathbf{x} - \mathbf{X}, \quad \overline{\mathbf{U}} = \overline{\mathbf{x}} - \mathbf{X}. \tag{12.11}$$

The associated deformation gradient tensors are

$$\mathbf{F} = \text{Grad}\, \mathbf{x}, \quad \overline{\mathbf{F}} = \text{Grad}\, \overline{\mathbf{x}} \tag{12.12}$$

where Grad is the gradient operator with respect to $\mathbf{X}$. Note that $\mathbf{F} = \mathbf{R}\boldsymbol{\mathcal{U}} = \boldsymbol{\mathcal{V}}\mathbf{R}$ where $\boldsymbol{\mathcal{U}}$ and $\boldsymbol{\mathcal{V}}$ are the right and left stretch tensors. Analogous forms can be stated for $\overline{\mathbf{F}}$. The principal stretches associated with the deformations in (12.12) are λ_j and $\overline{\lambda}_j, j = 1, 2, 3$ respectively, which in general are inhomogeneous. We also note that in general, principal axes vary with position in the deformed medium.

Note at this point that $\mathbf{x} \in b$ in the previous section corresponded to the position vector in the "transformed" domain. This notation should not cause confusion since it will always be clear in which context we are describing the analysis (i.e. transformation theory or small-on-large) and indeed later on we shall equate the two analyses in order to describe the theory of hyperelastic cloaking. What matters more is that we distinguish between the dependent variables and material properties in the two scenarios: in the transformation setting these have the superscript "eff" as noted in the previous section.

Carrying on with the small-on-large analysis, define the difference between position vectors in the perturbed and deformed configurations as

$$\mathbf{u} = \overline{\mathbf{x}} - \mathbf{x} \tag{12.13}$$

and let us assume that (in some sense)

$$|\overline{\mathbf{x}} - \mathbf{x}| = |\mathbf{u}| \ll 1 \tag{12.14}$$

which amounts to saying that the configurations b and $\overline{b}$ are "close".

Next define the deformation gradient tensors

$$\mathbf{f} = \operatorname{grad} \overline{\mathbf{x}}, \quad \overline{\mathbf{F}} = \operatorname{Grad} \overline{\mathbf{x}} = \mathbf{f}\mathbf{F}. \tag{12.15}$$

where grad is the gradient operator with respect to $\mathbf{x}$. Define

$$\boldsymbol{\gamma} = \operatorname{grad} \mathbf{u} = \operatorname{grad} (\overline{\mathbf{x}} - \mathbf{x}) = \operatorname{grad} \overline{\mathbf{x}} - \operatorname{grad} \mathbf{x} = \mathbf{f} - \mathbf{I} \tag{12.16}$$

so that

$$\mathbf{f} = \mathbf{I} + \boldsymbol{\gamma} \tag{12.17}$$

and so

$$\boldsymbol{\Gamma} = \operatorname{Grad} \mathbf{u} = \operatorname{Grad} (\overline{\mathbf{x}} - \mathbf{x}) = \mathbf{f}\mathbf{F} - \mathbf{F} = (\mathbf{f} - \mathbf{I})\mathbf{F} = \boldsymbol{\gamma}\mathbf{F}. \tag{12.18}$$

We have $|\boldsymbol{\Gamma}| \ll 1$ and $|\boldsymbol{\gamma}| \ll 1$, where the notation $|\mathbf{B}|$ with $\mathbf{B}$ as a second order tensor is defined as

$$|\mathbf{B}| = (\mathbf{B} : \mathbf{B}^T)^{1/2}. \tag{12.19}$$

Note that

$$\boldsymbol{f}^{-1} \approx \mathbf{I} - \boldsymbol{\gamma} \tag{12.20}$$

and is correct to terms that are linear in $\boldsymbol{\gamma}$.

12.3.2. *Stress*

Defining $J = \operatorname{Det}(\mathbf{F})$, (which is unity for an incompressible medium) the Cauchy and nominal stress in the finitely deformed body b are given by [Ogden (1997)]

$$\mathbf{T} = J^{-1}\mathbf{F}\frac{\partial \mathcal{W}}{\partial \mathbf{F}}, \quad \mathbf{S} = \frac{\partial \mathcal{W}}{\partial \mathbf{F}}, \tag{12.21}$$

for a compressible body and for an incompressible body we have

$$\mathbf{T} = \mathbf{F}\frac{\partial \mathcal{W}}{\partial \mathbf{F}} + Q\mathbf{I}, \quad \mathbf{S} = \frac{\partial \mathcal{W}}{\partial \mathbf{F}} + Q\mathbf{F}^{-1}. \tag{12.22}$$

We have used $\mathbf{T}$ for both linear and nonlinear Cauchy stress above. It will always clear as to which we will refer. Note that we follow the convention introduced by Ogden, that component-wise differentiation follows the rule

$$S_{ij} = \frac{\partial \mathcal{W}}{\partial F_{ji}}. \tag{12.23}$$

Above, $\mathcal{W}$ is the strain energy function, defining the constitutive behaviour of the medium in question and Q is the Lagrange multiplier, accommodating the constraint of incompressibility.

For reference, the Piola-Kirchhoff stress is often known as $\mathbf{P} = \mathbf{S}^T$ and the *second* Piola-Kirchhoff stress is denoted as

$$\boldsymbol{\mathcal{S}} = \mathbf{F}^{-1}\mathbf{P}. \tag{12.24}$$

It is straightforward to show, by linearization that the corresponding stresses in $\overline{b}$ can be written

$$\overline{\mathbf{T}} = \mathbf{T} + \boldsymbol{\tau}, \quad \overline{\mathbf{S}} = \mathbf{S} + \mathbf{s} \tag{12.25}$$

where $\boldsymbol{\tau}$ and $\mathbf{s}$ are the incremental Cauchy and nominal stress, respectively. Upon linearizing about the deformed state b and writing $\overline{Q} = Q+q$, one can show that for the compressible and incompressible cases,

$$\mathbf{s} = \mathbf{L} : \boldsymbol{\gamma}\mathbf{F}, \quad \mathbf{s} = \mathbf{L} : \boldsymbol{\gamma}\mathbf{F} + q\mathbf{F}^{-1} - Q\mathbf{F}^{-1}\boldsymbol{\gamma}, \tag{12.26}$$

where $\mathbf{L} = \partial^2\mathcal{W}/\partial\mathbf{F}^2$ (equivalent to $\boldsymbol{\mathcal{A}}$ in [Ogden (1997)]) and its Cartesian components are defined via

$$L_{IjK\ell} = \partial^2\mathcal{W}/\partial F_{jI}\partial F_{\ell K}, \tag{12.27}$$

noting again the convention on subscripts here. Colon notation indicates $\mathbf{A} : \mathbf{b} = A_{ijk\ell}b_{\ell k}$. In (12.26) quadratic and higher terms (in $\boldsymbol{\gamma}$) have been neglected since we are linearizing. Defining the *push forward of the incremental nominal stress* as $\boldsymbol{\zeta} = J^{-1}\mathbf{F}\mathbf{s}$, we have for

a compressible medium

$$\boldsymbol{\zeta} = \mathbf{M} : \boldsymbol{\gamma}. \tag{12.28}$$

Here the Cartesian components of the tensor $\mathbf{M}$ are defined by

$$M_{ijk\ell} = J^{-1} \frac{\partial^2 \mathcal{W}}{\partial F_{jM} \partial F_{\ell N}} F_{iM} F_{kN}. \tag{12.29}$$

We note that $\mathbf{M}$ is equivalent to $\boldsymbol{\mathcal{A}}_0$ in [Ogden (1997)]. For an incompressible material, we have

$$\boldsymbol{\zeta} = \mathbf{M} : \boldsymbol{\gamma} + q\mathbf{I} - Q\boldsymbol{\gamma}. \tag{12.30}$$

Q is determined from the large deformation problem but q is an unknown in the incremental problem.

We note that the incremental Cauchy stress for a compressible medium is given by

$$\boldsymbol{\tau} = \boldsymbol{\zeta} - \mathrm{tr}(\boldsymbol{\gamma})\mathbf{T} + \boldsymbol{\gamma}\mathbf{T} \tag{12.31}$$

whereas for an incompressible material, since $\mathrm{tr}(\boldsymbol{\gamma}) = 0$ we have

$$\boldsymbol{\tau} = \boldsymbol{\zeta} + \boldsymbol{\gamma}\mathbf{T}. \tag{12.32}$$

12.3.3. *The tensors* $\mathbf{L}$ *and* $\mathbf{M}$

Note that in general the tensors $\mathbf{L}$ and $\mathbf{M}$ do *not* possess the minor symmetries that are usually possessed by tensors of elastic moduli. In general $\mathbf{M}$ is rather laborious to determine because the second derivatives of $\mathcal{W}$ must be evaluated. The strain energy function $\mathcal{W}$ depends on n invariants $I_1, I_2, \ldots, I_n$ and so

$$\frac{\partial \mathcal{W}}{\partial \mathbf{F}} = \sum_{i=1}^{n} W_i \frac{\partial I_i}{\partial \mathbf{F}} \tag{12.33}$$

and

$$\frac{\partial^2 \mathcal{W}}{\partial \mathbf{F} \partial \mathbf{F}} = \sum_{i=1}^{n} \sum_{j=1}^{n} \mathcal{W}_{ij} \frac{\partial I_i}{\partial \mathbf{F}} \otimes \frac{\partial I_j}{\partial \mathbf{F}} + \sum_{i=1}^{n} \mathcal{W}_i \frac{\partial^2 I_i}{\partial \mathbf{F} \partial \mathbf{F}} \tag{12.34}$$

where we have denoted $\mathcal{W}_i = \partial\mathcal{W}/\partial I_i, \mathcal{W}_{ij} = \partial^2\mathcal{W}/\partial I_i \partial I_j$ for $i, j \in \{1, 2, \ldots, n\}$ where the index 3 is not included in the case of an incompressible medium. We see that this reduces to determining first and second derivatives of the strain invariants with respect to $\mathbf{F}$.

When the medium is *isotropic*, the components of $\mathbf{L}$ reduce to [Ogden (1997)] (no summation over repeated indices)

$$L_{iijj} = \mathcal{W}_{ij}, \tag{12.35}$$

$$L_{ijij} - L_{ijji} = \frac{\mathcal{W}_i + \mathcal{W}_j}{\lambda_i + \lambda_j}, \quad i \neq j, \tag{12.36}$$

$$L_{ijij} + L_{ijji} = \begin{cases} \frac{\mathcal{W}_i - \mathcal{W}_j}{\lambda_i - \lambda_j}, & i \neq j,\ \lambda_i \neq \lambda_j, \\ \mathcal{W}_{ii} - \mathcal{W}_{ij}, & i \neq j,\ \lambda_i = \lambda_j. \end{cases} \tag{12.37}$$

where here we have introduced the short-hand notation $\mathcal{W}_i = \partial\mathcal{W}/\partial\lambda_i$ and $\mathcal{W}_{ij} = \partial^2\mathcal{W}/\partial\lambda_i\partial\lambda_j$, $i, j \in \{1, 2, 3\}$ and recall that $J = 1$ for an incompressible medium. Similarly the components of the tensor $\mathbf{M}$ can be stated as

$$JM_{iijj} = \lambda_i\lambda_j\mathcal{W}_{ij}, \tag{12.38}$$

$$J(M_{ijij} - M_{ijji}) = \lambda_i\mathcal{W}_i, \quad i \neq j, \tag{12.39}$$

$$J(M_{ijij} + M_{ijji}) = \begin{cases} \lambda_i\mathcal{W}_i + 2\lambda_i\lambda_j\dfrac{\lambda_j\mathcal{W}_i - \lambda_i\mathcal{W}_j}{\lambda_i^2 - \lambda_j^2}, & i \neq j, \lambda_i \neq \lambda_j, \\ \lambda_i^2(\mathcal{W}_{ii} - \mathcal{W}_{ij}), & i \neq j, \lambda_i = \lambda_j. \end{cases} \tag{12.40}$$

12.3.4. *Incremental equations of motion*

We have, in succession

$$\overline{\mathrm{div}}(\overline{\mathbf{T}}) = \overline{\rho}\frac{\partial^2\overline{\mathbf{U}}}{\partial t^2} \tag{12.41}$$

and $\overline{\mathbf{U}} = \mathbf{U} + \mathbf{u}$. But the deformation to b is static, and $\mathbf{U}$ is independent of t, so that

$$\frac{\partial^2\overline{\mathbf{U}}}{\partial t^2} = \frac{\partial^2\mathbf{u}}{\partial t^2}. \tag{12.42}$$

It can also be shown that

$$\overline{\rho} = \rho + O(|\mathbf{u}|) \tag{12.43}$$

and so

$$\overline{\mathrm{div}}(\overline{\mathbf{T}}) = \rho \frac{\partial^2 \mathbf{u}}{\partial t^2}. \tag{12.44}$$

Finally it can be shown that

$$\overline{\mathrm{div}}(\overline{\mathbf{T}}) = \mathrm{div}\, \boldsymbol{\zeta} \tag{12.45}$$

so that the incremental equations of motion are

$$\mathrm{div}\, \boldsymbol{\zeta} = \rho \frac{\partial^2 \mathbf{u}}{\partial t^2}, \tag{12.46}$$

noting that if the density if the reference configuration is ρ_r then $\rho = J^{-1}\rho_r$.

What is striking is that the incremental equations of small-on-large have equivalent form to those for transformation elastodynamics as stated in (12.9) and indeed the modulus tensor $\mathbf{M}$ *does not possess the minor symmetries*, so that it seems we have found an alternative mechanism for elastodynamic cloaking using pre-stressed solids. What needs to be determined is which strain energy functions yield invariance and this is achieved by equating the equations of transformation elastodynamics (12.9) with those of small-on-large (12.46). This shall be done via a general framework in Sec. 12.5 but before this we describe some results in the context of antiplane elastic waves.

12.4. Antiplane Elastic Waves

Let us first discuss some results associated with the scalar problem of antiplane elasticity or horizontally polarized shear (SH) waves. We first show how transformation elasticity yields required cloaking properties and illustrate how, in principal these may be obtained from layered media. We then go on to describe a special result in the context of scattering from a cavity in pre-stressed media and how this result informs hyperelastic cloaking.

12.4.1. *Transformation cloaking*

With reference to Sec. 12.2, consider an unbounded homogeneous linear elastic material with shear modulus μ_0 and density ρ_0 and introduce a Cartesian coordinate system (X_0, Y_0, Z_0) and cylindrical polar coordinate system (R_0, Θ_0, Z_0) with some common origin $\mathbf{O}$. This is the virtual configuration. Planar variables are related in the usual manner, $X_0 = R_0 \cos\Theta_0, Y_0 = R_0 \sin\Theta_0$. Suppose that there is a time-harmonic line source, polarized in the Z_0 direction and located at (R', Θ'), with circular frequency ω and amplitude C_0 (which is a force per unit length in the Z_0 direction). This generates antiplane elastic waves with the only non-zero displacement component in the Z_0 direction of the form $\mathbf{U}_0 = \mathbf{e}_{Z_0}\Re[W_0(X_0, Y_0)\exp(-i\omega t)]$ where $\mathbf{e}_{Z_0}$ is a unit basis vector in the Z_0 direction. The displacement W_0 is governed by

$$\mathrm{div}_0\,(\mu_0 \mathrm{grad}_0\, W_0) + \rho_0\omega^2 W_0 = \frac{C_0}{R'}\delta(R_0 - R')\delta(\Theta_0 - \Theta'). \quad (12.47)$$

The assumed mapping from the virtual configuration to the physical configuration for a cloak for antiplane waves (cf. two dimensional acoustics) expressed in plane cylindrical polar coordinates, takes the form of the identity mapping for all $R_0 > D$ and

$$r = g(R_0), \qquad \theta = \Theta_0, \qquad z = Z_0, \qquad \text{for } 0 \le R_0 \le D, \qquad (12.48)$$

for some chosen monotonically increasing function $g(R_0)$ with $g(0) \equiv a \in [0, D)$, $g(D) = D \in \mathbb{R}^+$ (the set of positive real numbers) such that $D < R'$, i.e. the line source remains *outside* the cloaking region. The cloaking region is thus defined by $r \in [a, d]$ where $d = D$. Note once more that upper and lower case variables are employed for the untransformed and transformed configurations respectively. Under this mapping the form of the governing equation (12.47) remains unchanged for $R_0 = r > D$, whereas for $0 \le R_0 \le D$, corresponding to the transformed domain $a \le r \le d$, the transformed equation takes the form (in transformed cylindrical polar coordinates

$r, \theta = \Theta$)

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\mu_r^{\text{eff}}(r)\frac{\partial w^{\text{eff}}}{\partial r}\right) + \frac{\mu_\theta^{\text{eff}}(r)}{r^2}\frac{\partial^2 w^{\text{eff}}}{\partial \theta^2} + \rho^{\text{eff}}(r)\omega^2 w^{\text{eff}} = 0 \tag{12.49}$$

where (see Eqs. (26), (27) in [Norris and Shuvalov (2011)])

$$\mu_r^{\text{eff}}(r) = \frac{\mu_0^2}{\mu_\theta^{\text{eff}}(r)} = \mu_0 \frac{R_0}{r}\frac{dg}{dR_0}, \qquad \rho^{\text{eff}}(r) = \rho_0 \frac{R_0}{r}\left(\frac{dg}{dR_0}\right)^{-1}. \tag{12.50}$$

Hence, *both* the shear modulus and density must be inhomogeneous and the shear modulus must be anisotropic. Material properties of this form cannot be constructed exactly since the shear modulus μ_θ^{eff} becomes unbounded as $r \to r_1$ (the inner boundary of the cloak). In this limit the density behaves as $\rho^{\text{eff}} = (pca)^{-1}\rho_0 R_0^{2-p} + \ldots$ where p, $c > 0$ define the mapping in the vicinity of the inner boundary according to $r = a + cR_0^p + \ldots$ as $R_0 \to 0$. In practice of course, approximations are required as described in e.g. [Schurig *et al.* (2006); Farhat *et al.* (2008); Zhang *et al.* (2011)]. Note that, as expected [Norris and Shuvalov (2011)], the total mass is conserved since, regardless of the mapping, the integral of the density $\rho^{\text{eff}}(r)$ over $r \in [a, d]$ is $\pi D^2 \rho_0$.

A frequently quoted example due to its simplicity is the mapping

$$g(R_0) = a + R_0\left(\frac{d-a}{d}\right) \tag{12.51}$$

so that the required cloak properties are

$$\mu_r^{\text{eff}}(r) = \mu_0\left(1 - \frac{a}{r}\right), \quad \mu_\theta^{\text{eff}}(r) = \mu_0\left(\frac{r}{r-a}\right),$$
$$\rho^{\text{eff}}(r) = \rho_0\left(1 - \frac{a}{r}\right)\left(\frac{d}{d-a}\right)^2. \tag{12.52}$$

We note that μ_r^{eff} and ρ^{eff} tend to zero and μ_θ^{eff} becomes unbounded as $r \to a$ on the inner boundary of the cloak.

12.4.2. *Cloak realisation via homogenization of layered media*

From the transformation method above it is clear that a perfect cloak cannot be produced (as should be expected). However if a material that has approximately the same properties can be constructed (we call this material a metamaterial) the cloak could be achieved approximately, i.e. some wave scattering will occur but its magnitude in comparison to the incident energy will be small. In order to achieve this we require a metamaterial that has a very small mass density *and* radial shear modulus near the inner cloak boundary but also a very large azimuthal shear modulus near the inner boundary. Additionally the material must be structured in such a way so that the material properties are inhomogeneous and vary radially in the manner described in (12.52) above. One can imagine various mechanisms for producing such anisotropy and inhomogeneity. One such scheme was proposed by Sanchez-Dehesa and Torrent [Sánchez-Dehesa and Torrent (2008)] in the acoustics context. They imagined using a layered medium as depicted in Figs. 12.1 and 12.2 with each layer itself possessing constant material properties. The radial variation in moduli

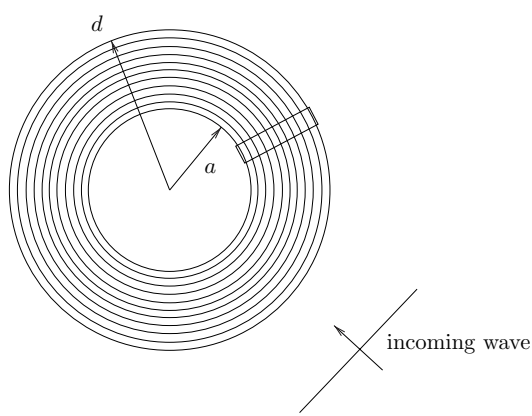


Fig. 12.1. Cloak with outer radius d and inner radius a. A zoom into the rectangular region highlighted is given in Fig. 12.2 where we see that locally the material is approximately a layered medium.

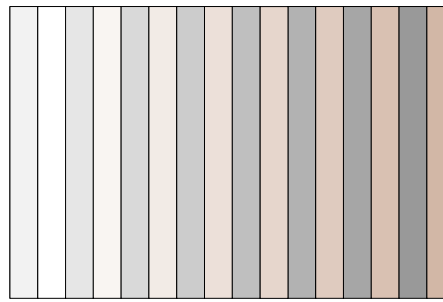


Fig. 12.2. Local geometry of the cloak — i.e. a layered medium. Each "cell" consists of two layers — here denoted with alternating shading. Local effective properties of the two layer medium are determined from classical expressions for effective properties of layered media.

is therefore piecewise constant and induces anisotropy. It becomes a better approximation to the required metamaterial properties as we increase the number of layers as we shall see below. The construction of a finely layered medium itself is a rather difficult engineering task and since we also require the properties of the layers to vary with radial distance (by using some additional microstructure within each layer) this increases the technological difficulty.

In order to induce the necessary anisotropy we introduce a cell that consists of two isotropic layers. The properties of each layer vary radially with the cell position as depicted by the different shadings in Fig. 12.2. The effective antiplane shear properties of such a two layered cell (with properties μ_1, ρ_1 and μ_2, ρ_2 with volume fraction ϕ of phase 1 material, are [Backus (1962)]

$$\mu_r^* = \left(\frac{\phi}{\mu_1} + \frac{1-\phi}{\mu_2}\right)^{-1}, \quad \mu_\theta^* = \phi\mu_1 + (1-\phi)\mu_2,$$
$$\rho^* = \phi\rho_1 + (1-\phi)\rho_2. \tag{12.53}$$

We note that we have assigned the property in the direction perpendicular to the layer surfaces as μ_r^* whereas that in the parallel direction is μ_θ^* neglecting the curvature effects here. As in [Sánchez-Dehesa and Torrent (2008)] let us choose $\phi = 1/2$ so that both layers

are of equal thickness. Therefore (12.53) become

$$\frac{\mu_r^*}{\mu_0} = 2\left(\frac{\mu_0}{\mu_1} + \frac{\mu_0}{\mu_2}\right)^{-1}, \quad \frac{\mu_\theta^*}{\mu_0} = \frac{1}{2\mu_0}(\mu_1 + \mu_2), \quad \frac{\rho^*}{\rho_0} = \frac{1}{2\rho_0}(\rho_1 + \rho_2) \tag{12.54}$$

where we have introduced the "background" material properties of the exterior region μ_0 and ρ_0 in order to non-dimensionalize.

These properties vary from cell to cell because we choose different material properties for the two individual layers in each cell. This therefore permits us to determine the necessary material properties within the cell by equating these homogenized properties (as functions of r) with the *required* cloaking properties. Clearly in practice the effective properties would be piecewise constant but as the layer thickness gets smaller this dependence becomes smoother.

The choice of approach is non-unique. As an example, let us first set $\mu_2 = \mu_0^2/\mu_1$ so that if we now consider the effective properties to be a (continuous) function of r we require the homogenized properties of the layered cloak to take the form

$$\frac{\mu_r^*(r)}{\mu_0} = 2\left(\frac{\mu_0}{\mu_1(r)} + \frac{\mu_1(r)}{\mu_0}\right)^{-1}, \quad \frac{\mu_\theta^*(r)}{\mu_0} = \frac{1}{2}\left(\frac{\mu_0}{\mu_1(r)} + \frac{\mu_1(r)}{\mu_0}\right). \tag{12.55}$$

Equating μ_θ^* with the necessary cloak property μ_θ^{eff} as given for the specific mapping in (12.52), and solving for μ_1 we obtain

$$\mu_1(r) = \mu_\theta^{\text{eff}}(r) + \sqrt{(\mu_\theta^{\text{eff}}(r))^2 - \mu_0^2} \tag{12.56}$$

so that

$$\frac{\mu_1(r)}{\mu_0} = \frac{r}{(r-a)} + \sqrt{\left(\frac{r}{r-a}\right)^2 - 1}. \tag{12.57}$$

This also fixes the function $\mu_2(r)$ via the relation $\mu_2 = \mu_0^2/\mu_1$ and ensures that the shear modulus cloaking properties are achieved.

To prescribe the appropriate density one option is to set the phase speeds of each layer to be equal $c_1(r) = c_2(r)$. We wish to equate the

effective density ρ^* with the cloaked density property $\rho^{\text{eff}}(r)$ so that

$$\rho^*(r) = \frac{1}{2}(\rho_1(r) + \rho_2(r)) = \rho_0 \left(1 - \frac{a}{r}\right) \alpha^2 \tag{12.58}$$

where we have introduced $\alpha = d/(d-a)$. Substituting $\rho_1(r) = \mu_1(r)c_1^2(r)$ and $\rho_2(r) = \mu_2(r)c_1^2(r)$ into (12.58) then yields

$$c_1^2(r) = \frac{2\rho_0\alpha^2(r-a)}{r(\mu_1(r) + \mu_2(r))}, \tag{12.59}$$

from which we can determine $\rho_1(r)$ and $\rho_2(r)$. Alternatively, one could insist that for example $\rho_1(r)$ has the same functional dependence on r as $\mu_1(r)$ which then fixes $\rho_2(r)$ via (12.58).

Suppose we have N cells in the cloak, i.e. we have a total of $2N$ layers in the region between $r = a$ and $r = d$. These are the alternating 1 and 2 phases but where the material properties of these phases vary with r. The layer thickness is

$$\delta = \frac{(d-a)}{2N}. \tag{12.60}$$

The ordering of the layers does not matter, so begin on the inner face of the cloak with phase 1, then 2. We assign the (constant) material properties of the nth layer as those of the inhomogeneous functions evaluated at their outer interface, i.e. on $r = a + n\delta$, $n = 1, 2, \ldots, 2N$ so that the outermost layer (a phase 2 layer) has properties evaluated at $r = d$. Proceeding in this manner also avoids the obvious singularity at $r = a$.

In Figs. 12.3 and 12.4 we illustrate the use of the above technique in the situation where there are 50 and 200 layers within the cloak region and a plane wave is incident from the right. The phase properties, varying as a function of r are chosen using the technique above. In this case we have chosen $D = d = 2a$ and $ka = 4\pi$ which are the same parameters as those used in [Sánchez-Dehesa and Torrent (2008)]. The point source is located at (R', Θ') as above. In order to solve this scattering problem we write down the solution to the governing (Helmholtz) equation in the mth layer analytically via the

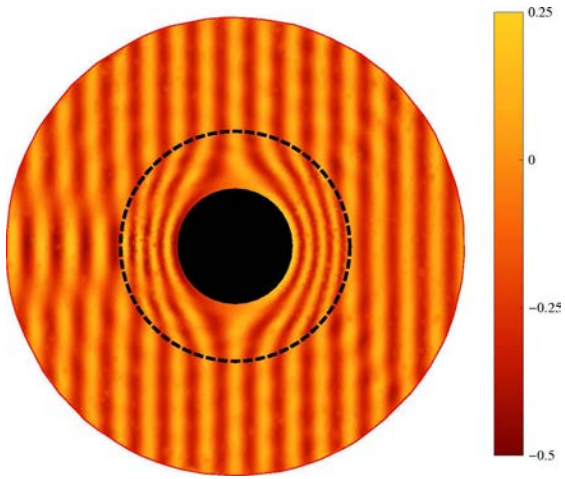


Fig. 12.3. A cloak consisting of 50 layers ($N = 25$) constructed via the homogenisation method. The incident wave is a plane wave of unit amplitude impinging from the right.

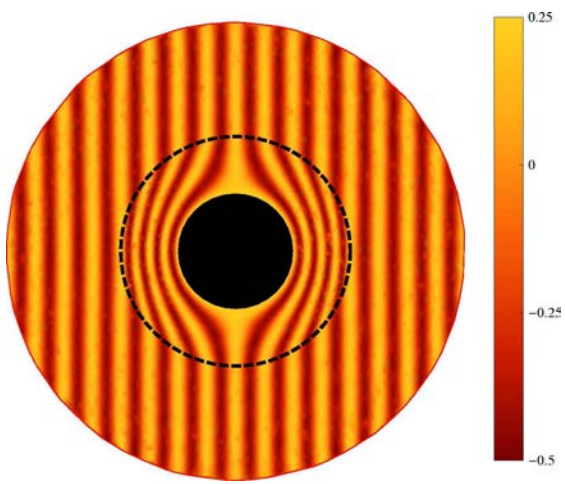


Fig. 12.4. A cloak consisting of 200 layers ($N = 100$) constructed via the homogenisation method. The incident wave is a plane wave of unit amplitude impinging from the right.

standard separation of variables expansion, i.e.

$$w_m(r,\theta) = \sum_{n=-\infty}^{\infty} (c_{mn}\mathrm{H}_n^{(1)}(k_m R) + d_{mn}\mathrm{H}_n^{(2)}(k_m r))e^{in(\theta-\Theta')} \tag{12.61}$$

where k_m is the wavenumber in the mth layer and $\mathrm{H}_n^{(j)}$ is the Hankel function of nth order and jth type ($j = 1, 2$ correspond to outgoing and incoming waves respectively). We satisfy continuity of displacement and traction on each of the interfaces between the layers with a traction free condition on $r = a$. On $r = D$ we impose continuity of traction and displacement with the solution in the external medium, i.e.

$$w(r,\theta) = w_i + w_s = \frac{C_0}{4i\mu_0}\mathrm{H}_0^{(1)}(k_0 S) + \sum_{n=-\infty}^{\infty} (-i)^n a_n \mathrm{H}_n^{(1)}(k_0 r)e^{in(\theta-\Theta')}, \tag{12.62}$$

where $S = \sqrt{(x-X')^2 + (y-Y')^2}$ and k_0 is the wavenumber in the host medium and we note the "incident" and "scattered" partition of the field. The coefficients a_n are the *scattering coefficients* which should be small if cloaking is successful. This solution scheme generates a large algebraic system which we invert numerically for the coefficients c_{mn}, d_{mn} and a_n. This approach is described in further detail in [Parnell *et al.* (2013)]. We plot the field in Figs. 12.3 and 12.4 for the cases of 50 layers ($N = 25$) and 200 layers ($N = 100$) respectively. We see the qualitative improvement in cloaking (reduction of the shadow to the left) when we increase the number of layers from 50 to 200. We have taken $\Theta' = 0$ and considered the limiting case of an incident plane wave of unit amplitude which is derived by taking the limit $R' \to \infty$ together with source strength

$$C_0 \to C_{\mathrm{pw}} = 2i\mu_0\sqrt{2\pi k_0 R'}e^{i(\frac{\pi}{4}-k_0 R')}. \tag{12.63}$$

Note that very few papers introduce a quantitative measure of the success of cloaking via metamaterials. What we will do here is to

introduce such a measure, i.e. the scattering cross section γ (see (12.103)) and calculate its magnitude. We do this later when comparing this transformation (layered metamaterial) approach and the hyperelastic method which we shall introduce shortly.

As noted earlier, although the above metamaterial construction is feasible it is certainly formidable. An alternative construction would be to begin with a cell which itself is transversely isotropic so that the initial homogenization step is not required. This strategy still requires a very large library of available material properties, now anisotropic. Another option is to continue to use homogenization for the local effective properties, but to limit the number of available isotropic materials. Variation in effective properties is then obtained by selecting the relative concentrations. Surprisingly, a considerable degree of cloaking can be achieved with this approach, even with as few as three distinct materials [Norris and Nagy (2010)].

12.4.3. *Hyperelastic cloaking using a neo-Hookean medium*

Before describing how nonlinear pre-stress of an elastic medium can be used to construct finite cloaks analogous to those just described using layered materials, we shall first describe the solution of a canonical scattering problem associated with antiplane waves in pre-stressed nonlinear elastic media of neo-Hookean type, summarizing the results of [Parnell (2012)].

12.4.3.1. *Initial large deformation*

With reference to Sec. 12.3 consider an isotropic *incompressible* material of infinite extent in all directions with a cylindrical cavity of initial radius A at the origin. Assume a finite deformation as that depicted on the left of Fig. 12.5, which arises due to increasing the pressure inside the cavity region, leading to an increase in the radius (to say a) of the cavity. In order to describe this deformation mathematically write

$$R = R(r), \quad \Theta = \theta, \quad Z = \frac{z}{\zeta}, \tag{12.64}$$

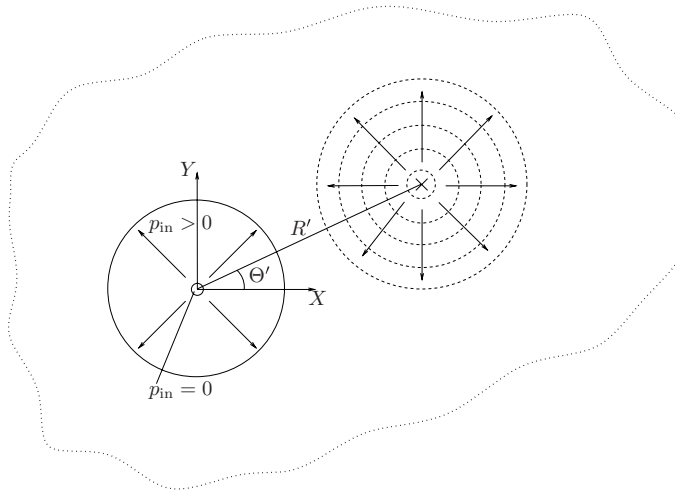


Fig. 12.5. In an unbounded medium, a time-harmonic line source, located at (R', Θ') is indicated by crossed lines and we show schematically the direction of outward propagating waves and wave-fronts. This wave is scattered from a small cavity of radius $A \ll \Lambda$ where Λ is the wavelength of the source, with internal pressure $p_{\rm in} = 0$ in the undeformed configuration. When the pressure is increased, $p_{\rm in} > 0$, the cavity increases its radius to a such that $a/\Lambda = O(1)$. A region of inhomogeneous deformation develops around the deformed cavity.

where (R, Θ, Z) and (r, θ, z) are cylindrical polar coordinates in the undeformed and deformed configurations, denoted B and b respectively. The function $R(r)$ will be determined from the radial equation of equilibrium and incompressibility condition below and $\zeta \in \mathbb{R}^+$ is an axial stretch.

The principal stretches for this deformation in the radial, azimuthal and longitudinal directions, respectively, are

$$\lambda_1 = \lambda_r = \frac{dr}{dR} = \frac{1}{R'(r)}, \qquad \lambda_2 = \lambda_\theta = \frac{r}{R(r)}, \qquad \lambda_z = \zeta, \tag{12.65}$$

and the deformation gradient tensor $\mathbf{F}$ (referred to cylindrical polar coordinates) is

$$\mathbf{F} = \begin{pmatrix} \lambda_r & 0 & 0 \\ 0 & \lambda_\theta & 0 \\ 0 & 0 & \lambda_z \end{pmatrix} = \begin{pmatrix} (R'(r))^{-1} & 0 & 0 \\ 0 & r/R(r) & 0 \\ 0 & 0 & \zeta \end{pmatrix}. \tag{12.66}$$

Since the medium is incompressible

$$\lambda_r \lambda_\theta \lambda_z = \frac{r\zeta}{R(r)R'(r)} = 1, \tag{12.67}$$

which is easily solved to yield

$$R(r) = \sqrt{\zeta(r^2 + M)}, \tag{12.68}$$

where $M = (A^2 - a^2)/\zeta$. In this section let us now consider the case when there is no axial stretch, i.e. $\zeta = 1$; the equation linking the internal pressure p_{in} and the radii a and A will be derived shortly. In order to make further progress consider a *neo-Hookean* host medium, having strain energy function [Ogden (1997)]

$$\mathcal{W}_{\text{NH}} = \frac{\mu}{2}(I_1 - 3) = \frac{\mu}{2}(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3). \tag{12.69}$$

Integrating the radial equation of equilibrium and applying the traction condition $\mathbf{t}|_{r=a} = \mathbf{T}\mathbf{n}|_{r=a} = p_{\text{in}}\mathbf{e}_r$ where $\mathbf{e}_r$ is the unit vector in the radial direction and $\mathbf{n} = -\mathbf{e}_r$ is the (outward) unit normal from the elastic material, so that $T_{rr}|_{r=a} = -p_{\text{in}}$, we find that the radial stress is determined by

$$T_{rr}(r) + p_{in} = \frac{\mu}{2}\left(M\left(\frac{1}{r^2} - \frac{1}{a^2}\right) + \log\left(\frac{r^2 + M}{a^2 + M}\right) - \log\left(\frac{r^2}{a^2}\right)\right). \tag{12.70}$$

Using the condition that $T_{rr} \to 0$ as $r \to \infty$ in (12.70) yields

$$\frac{p_{\text{in}}}{\mu} = \frac{1}{2}\left(1 - \frac{A^2}{a^2} + \log\left(\frac{a^2}{A^2}\right)\right), \tag{12.71}$$

which specifies the relationship of the deformed to undeformed radius ratio a/A in terms of the scaled pressure p_{in}/μ.

12.4.3.2. *Incremental wave equation*

As depicted in Fig. 12.5, SH waves are excited in the pre-stressed medium by virtue of a line source, originally at the location (R', Θ') of strength C_0 (per unit length) which is moved to the location (r', Θ')

after the initial (finite) deformation. The total incremental deformation is of the form

$$\mathbf{u} = \Re[w(x,y)\mathbf{e}_z \exp(-i\omega t)], \tag{12.72}$$

where $\mathbf{e}_z$ is the unit basis vector in the z direction. Using the theory as presented in Sec. 12.3 the incremental equation governing w takes the form (details of this derivation are also given in the Appendix of [Parnell (2012)])

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\mu_r(r)\frac{\partial w}{\partial r}\right) + \frac{\mu_\theta(r)}{r^2}\frac{\partial^2 w}{\partial \theta^2} + \rho\omega^2 w = \frac{C_0}{r_0}\delta(r-r')\delta(\theta-\theta') \tag{12.73}$$

where

$$\mu_r(r) = \mu\left(\frac{r^2+M}{r^2}\right), \qquad \mu_\theta(r) = \mu\left(\frac{r^2}{r^2+M}\right). \tag{12.74}$$

Note that as $r \to \infty$, (12.73) reduces to

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial w}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 w}{\partial \theta^2} + K^2 w = \frac{C_0}{\mu r_0}\delta(r-r')\delta(\theta-\theta') \tag{12.75}$$

where $K^2 = \rho\omega^2/\mu$ is the wavenumber associated with the undeformed material. Note that as $r \to \infty$ the deformation does not affect the wave; the governing equation is simply the scalar wave equation in the far field.

12.4.3.3. *Scattering in the pre-stressed medium*

In Appendix B of [Parnell (2012)] it is shown that in an *unstressed* medium, given an incident point source field located at $(X,Y) = (X',Y')$ of the form

$$W_i = \frac{C_0}{4i\mu}\mathrm{H}_0^{(1)}(K\sqrt{(X-X')^2+(Y-Y')^2}) \tag{12.76}$$

the scattered field that emanates from a cylindrical cavity of radius A takes the form

$$W_s = \sum_{n=-\infty}^{\infty} (-i)^n A_n \mathrm{H}_n^{(1)}(KR) e^{in(\Theta-\Theta')} \tag{12.77}$$

where $X = R\cos\Theta, Y = R\sin\Theta$,

$$A_n = \frac{C_0(-1)^n}{4\mu i^{n-1}} \frac{\mathrm{J}_n'(KA)}{\mathrm{H}_n^{(1)\prime}(KA)} \mathrm{H}_n^{(1)}(KR') \tag{12.78}$$

and the prime here on the Bessel functions indicates differentiation with respect to their argument.

In the knowledge that $\zeta = 1$, the equation governing the incremental waves in the configuration b is (12.73), or explicitly

$$\frac{1}{r}\frac{\partial}{\partial r}\left(\left(r + \frac{M}{r}\right)\frac{\partial w}{\partial r}\right) + \frac{1}{(r^2+M)}\frac{\partial^2 w}{\partial \theta^2} + K^2 w = \frac{C_0}{\mu r_0}\delta(r-r')\delta(\theta-\theta'), \tag{12.79}$$

noting again that K is the wavenumber associated with the undeformed material. In this specific case of a neo-Hookean elastic material we can, in fact, solve the problem analytically and we show, perhaps surprisingly, that the scattering coefficients are completely unaffected by the pre-stress and therefore even though the cavity has increased significantly in size, its scattering effect on incident waves remains weak, when the wavelength of the incident wave is long compared to the *undeformed* cavity radius A.

To illustrate the method of solution, introduce the following mapping into (12.79):

$$R^2 = r^2 + M, \quad \Theta = \theta \tag{12.80}$$

and we note that this mapping corresponds exactly to the initial finite deformation (12.64) and (12.68). Upon defining $W(R,\Theta) =$

$w(r(R), \theta(\Theta))$, we find

$$\frac{\partial^2 W}{\partial R^2} + \frac{1}{R}\frac{\partial W}{\partial R} + \frac{1}{R^2}\frac{\partial^2 W}{\partial \Theta^2} + K^2 W = \frac{C_0}{\mu}\frac{1}{r_0}\delta(r - r')\delta(\Theta - \Theta'). \tag{12.81}$$

It is straightforward to show that

$$\frac{1}{r'}\delta(r - r') = \frac{1}{R'}\delta(R - R') \tag{12.82}$$

and hence (12.81) becomes

$$\nabla^2 W + K^2 W = \frac{C_0}{\mu}\frac{1}{R'}\delta(R - R')\delta(\Theta - \Theta'). \tag{12.83}$$

This can therefore be straightforwardly solved. The incident field in the mapped domain is given by (12.76), which then gives rise to the scattered field as in (12.77) except that A_n are replaced by a_n, the scattering coefficients in the pre-stressed domain. Since the antiplane traction-free boundary condition on $R = A$ is not affected by the imposed pre-stress, the scattering coefficients a_n are given by

$$a_n = A_n = \frac{C_0(-1)^n}{4\mu i^{n-1}}\frac{\mathrm{J}_n'(KA)}{\mathrm{H}_n'(KA)}\mathrm{H}_n(KR'). \tag{12.84}$$

We now map back to the deformed configuration, using (12.80); the incident and scattered fields in the deformed configuration are therefore

$$w_i(r) = \frac{C_0}{4i\mu}\sum_{n=-\infty}^{\infty} e^{in(\theta-\theta')} \times \begin{cases} \mathrm{H}_n(K\sqrt{r'^2 + M})\mathrm{J}_n(K\sqrt{r^2 + M}), & r < r', \\ \mathrm{H}_n(K\sqrt{r^2 + M})\mathrm{J}_n(K\sqrt{r'^2 + M}), & r > r', \end{cases} \tag{12.85}$$

and

$$w_s(r) = \sum_{n=-\infty}^{\infty} (-i)^n a_n \mathrm{H}_n\left(K\sqrt{r^2 + M}\right) e^{in(\theta-\theta')} \tag{12.86}$$

where we used Graf's addition theorem to obtain the form (12.85) [Martin (2006)].

Perhaps most important to note is that the scattering coefficients a_n depend on the *initial* distance R' between the centre of the cavity and the source location, and the *undeformed* cavity radius A and they are unaffected by the pre-stress, $a_n = A_n$. The wave field close to the cavity is affected by the pre-stress and in fact waves are bent around the cavity, whereas in the far-field the wave that is seen will be identical to the field that would be seen in that region when scattering takes place from the undeformed cavity. Since scattering is weak in the latter case ($KA \ll 1$), it will also be weak from the deformed (large) cavity ($Ka = O(1)$).

The analytical solution to the scattering problem above was first derived in [Parnell and Abrahams (2011)] and used in [Parnell and Abrahams (2012)] to consider antiplane wave scattering in a pre-stressed microvoided composite with a neo-Hookean host phase. A general asymptotic methodology that determines the scattered field when the hyperelastic medium is more general than neo-Hookean was introduced in [Shearer *et al.* (2015)].

12.4.3.4. *Cloaking*

In order to illustrate the above theory for a specific case, consider the situation where $a/A = 20$ with $KA = 2\pi/20, Ka = 2\pi$. We set $KR' = 8\pi, \Theta' = 0$. On introducing the wavelength Λ via $\Lambda = 2\pi/K$, this choice ensures that $\Lambda/A = 20$ (weak scattering) whereas $\Lambda/a = 1$. Having the ratio $\Lambda/a = 1$ in an unstressed medium would normally result in strong scattering and the notable presence of a scattered field and shadow region. Using pre-stress however, the scattering effect of a cylindrical cavity of radius a such that $Ka = O(1)$ is equivalent to that of the undeformed cavity with $KA \ll 1$, i.e. very weak. Note that an inner pressure of around $p_{\text{in}}/\mu = 3.5$ ensures this deformation.

We illustrate this situation with three plots of the total wave field in Fig. 12.6. On the left we show the total field $W = W_i + W_s$ (defined in (12.76)–(12.78)) resulting from scattering from the small cylindrical cavity ($KA = 2\pi/20$) in an undeformed medium. Since $\Lambda \gg A$, scattering is very weak and the scattering effect cannot be

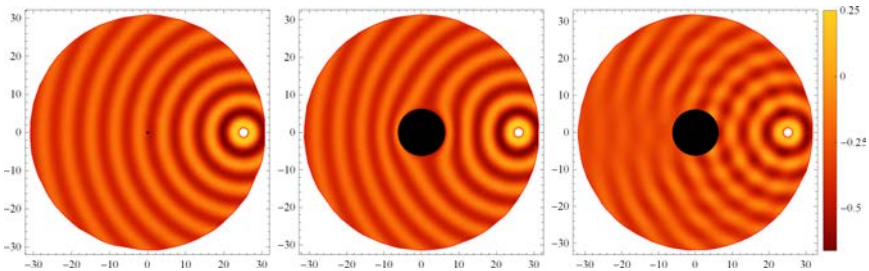


Fig. 12.6. Total antiplane wave-field due to a line source located at $R' = 8\pi$, $\Theta' = 0$ in an undeformed configuration. The source is shown as a small finite region to avoid the large values near its origin. We show the total field in the case of a small cavity with $KA = 2\pi/20$ (left), a deformed cavity in the pre-stressed medium ($A \to a$) with $Ka = 2\pi$ (middle) and a large cavity with $KA = 2\pi$ in an unstressed medium (right). This illustrates the possibility of cloaking via pre-stress.

seen on the plot. In the middle plot we show the equivalent problem but with the pre-stressed cavity inflated from A to a so that $Ka = 2\pi$ and the total displacement field plotted is $w = w_i + w_s$ as defined in (12.85) and (12.86). As shown theoretically above, the scattered field remains the same as for the undeformed problem and therefore even though $Ka = 2\pi = O(1)$, scattering is weak. This is in contrast to the plot on the right of the Fig. where we show the total field wave $W = W_i + W_s$ for an undeformed cavity with $KA = 2\pi$, illustrating the total wave field that *would* result in an unstressed medium in this scenario and the fact that it gives rise to strong scattering and a noticable shadow region.

The above invariance result shall now be used in a more general setting in order to illustrate how hyperelastic deformation can be used to generate a finite cloak

12.4.4. *Finite antiplane elastic cloaks using nonlinear pre-stress*

12.4.4.1. *Invariance using a neo-Hookean cloak*

A general objective is of course to develop a *finite* cloak, rather than an infinite (hyperelastic) cloak as described in the previous section.

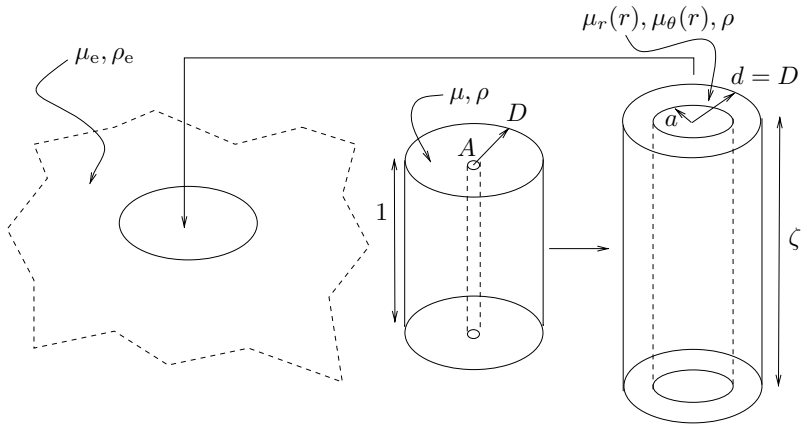


Fig. 12.7. The incompressible neo-Hookean cylindrical annulus is pre-stressed as depicted on the right. This annulus then creates a cloak when slotted into a cylindrical cavity in an unbounded elastic medium, as illustrated on the left.

One mechanism to achieve this in the SH case was described in [Parnell *et al.* (2013)]. Consider the configuration as depicted in Fig. 12.7: an elastic material within which is located a cylindrical cavity of radius D. Let us assume that the density of this medium is ρ_e and its axial shear modulus (corresponding to shearing on planes parallel to the axis of the cylindrical cavity) is μ_e. Additionally we take a cylindrical annulus of isotropic *incompressible* neo-Hookean material (with strain energy function defined by (12.69)) with associated shear modulus μ and density ρ and with inner and outer radii A and D respectively with $A \ll D$. The exact nature of this latter relationship will be described shortly. We shall consider deformations of the cylindrical annulus in order that it can act as an elastodynamic cloak to incoming antiplane elastic waves. We deform the material so that its inner radius is significantly increased (to a) but its outer radius $d = D$ remains unchanged. The deformed cylindrical annulus can then slot into the existing cylindrical cavity region within the unbounded (unstressed) domain. We choose μ and ρ so that subsequent waves satisfy the necessary continuity conditions on $r = d = D$. We consider the initial deformation of the cylindrical annulus domain as depicted in Fig. 12.7. Since the material is

incompressible, the deformation is induced *either* by applying a uniform axial stretch ζ *or* a radial pressure difference $p_{\text{out}} - p_{\text{in}}$ where p_{out} and p_{in} denote the pressures applied to the outer and inner face of the cylindrical annulus respectively. The ensuing deformation in the annulus is therefore described via (12.64) but now we take $\zeta \neq 1$ unlike the hyperelastic cloak case considered in the previous subsection.

In this case, ensuring that $d = D$ means that instead of (12.68) we find that $M = D^2(\zeta^{-1} - 1)$. Furthermore imposing the requirement that $R(a) = A$ we find that

$$\zeta = \frac{D^2 - A^2}{D^2 - a^2}. \tag{12.87}$$

One can easily determine the relation between applied pressures and deformation as with the unbounded case above, and it is deduced that

$$\frac{\zeta}{\mu}(p_{\text{in}} - p_{\text{out}}) = \frac{1}{2\zeta}\left(1 - \frac{A^2}{a^2}\right) + \log\left(\frac{a}{A}\right). \tag{12.88}$$

Now assume that the cylindrical annulus has been pre-stressed in an appropriate manner and slotted into the unbounded elastic material with perfect bonding at $r = d$. We consider wave propagation in this medium given a time-harmonic antiplane line source located at (R', Θ') with $R' > d$. In $r > d$ the antiplane wave with corresponding displacement which we shall denote by $w(r, \theta)$, is again governed by (12.75) but with wavenumber $k_e = \omega\sqrt{\rho_e/\mu_e}$ and of course no inhomogeneous term on the right hand side. In the region $a \leq r \leq d = D$, the wave satisfies a different equation since this annulus region has been pre-stressed according to the deformation (12.64) and (12.68). We can obtain the governing equation using the theory of small-on-large [Ogden (1997)]. It was shown in [Parnell *et al.* (2012)] that the wave in this region satisfies (12.73) but now where $\mu_r(r)$ and $\mu_\theta(r)$ are defined via

$$\frac{\zeta}{\mu}\mu_r(r) = \frac{\mu}{\zeta\mu_\theta(r)} = 1 + \frac{M}{r^2}. \tag{12.89}$$

The density remains as ρ.

Let us introduce the identity mapping for $r > D$ and

$$R^2 = \zeta(r^2 + M), \quad \Theta = \theta, \text{ for } a \leq r \leq D \tag{12.90}$$

which corresponds to the actual physical deformation (12.68). Finally define $W(R, \Theta) = w(r(R), \theta(\Theta))$. It is then straightforward to show that the equation governing wave propagation in the entire domain $R \geq A$ is (12.83), *provided that we choose* $\mu = \zeta\mu_{\mathrm{e}}$ *and* $\rho = \zeta\rho_{\mathrm{e}}$. These relations ensure that the wavenumbers in the exterior and cloak regions are the same and they also maintain continuity of traction on $R = D$. Furthermore since (12.90) corresponds to the actual deformation, the inner radius a maps back to A. Therefore with the appropriate choice of cloak material properties, the scattering problem in the undeformed and deformed configurations are equivalent. As with the unbounded cloak case considered in the previous chapter we can therefore solve the equation in the undeformed configuration and map back to the deformed configuration to obtain the physical solution. Decomposing the solution into incident and scattered parts $W = W_i + W_s$ where W_i is as defined in (12.76) and W_s is as defined in (12.77) with $A_n \to a_n$ which is determined by imposing the traction-free condition on $R = A$.

We want the wave field with respect to the *deformed* configuration, so we map back using (12.90) in order to find $w = w_i + w_s$. The incident wave is most conveniently determined by using Graf's addition theorem in order to distinguish between the regions $r < R'$ and $r > R'$, as was described in [Parnell (2012)]. The incident and scattered fields are then, respectively,

$$w_i(r) = \frac{C_0}{4i\mu_{\mathrm{e}}} \sum_{n=-\infty}^{\infty} e^{in(\theta-\theta')} \times \begin{cases} \mathrm{H}_n^{(1)}(KR')J_n(K\sqrt{\zeta(r^2+M)}), & a \leq r < D, \\ \mathrm{H}_n^{(1)}(KR')J_n(Kr), & D \leq r < R', \\ \mathrm{H}_n^{(1)}(Kr)J_n(KR'), & r > R', \end{cases} \tag{12.91}$$

$$w_s(r) = -\frac{C_0}{4i\mu_e}\sum_{n=-\infty}^{\infty} e^{in(\theta-\theta')}\,\frac{\mathrm{J}_n'(KA)}{\mathrm{H}_n^{(1)\prime}(KA)}\mathrm{H}_n^{(1)}(KR') \times \begin{cases} \mathrm{H}_n^{(1)}\left(K\sqrt{\zeta(r^2+M)}\right), & a \le r < D, \\ \mathrm{H}_n^{(1)}(Kr), & r \ge D. \end{cases} \tag{12.92}$$

The key to cloaking is to ensure that the scattered field is small compared with the incident field, i.e. $a_n \ll 1$. Note from (12.84) that a_n are solely dependent on the initial annulus inner radius A (and source distance R') but are *independent* of the deformed inner radius a. Therefore we must choose A such that $KA \ll 1$ which will ensure negligible scattering. We illustrate with some examples in Fig. 12.8, showing that the hyperelastic cloak appears to work relatively well. Quantitative analysis using the notion of a *scattering cross section* will be described shortly.

12.4.4.2. *Near-cloaks with other media*

Mathematically it has been shown that neo-Hookean material behaviour gives rise to the invariance principal and the idea of hyperelastic cloaking. However one could suppose that since many nonlinear materials behave in a neo-Hookean manner for relatively small deformations these alternative media may also be reasonable candidates for a cloak. Here we consider the Mooney-Rivlin strain energy function as a candidate for more realistic materials and compare this with layered media in the spirit of the work by [Sánchez-Dehesa and Torrent (2008)], summarizing the work of [Parnell *et al.* (2013)].

The neo-Hookean model for hyperelastic materials is an approximation which is acknowledged to be valid up to "moderate" strains. Clearly here, the larger the strain we induce the larger the deformed cavity and the more effective the cloaking is. Therefore we are certainly interested in the large strain regime. An improved model for hyperelastic materials of a rubber nature is

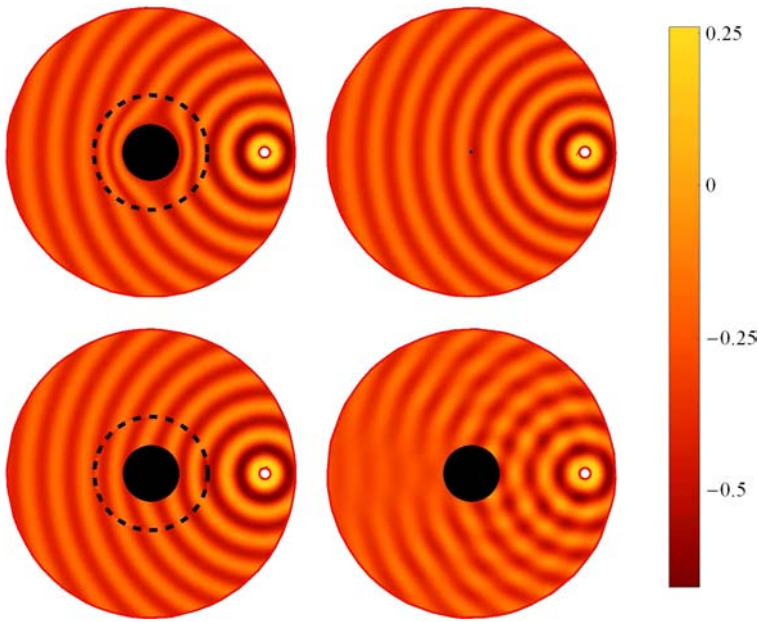


Fig. 12.8. Cloaking of antiplane shear waves. An antiplane line source is located at $Kr = KR' = 8\pi$, $\Theta' = 0$, shown as a white circle. **Upper left**: A region of (nondimensionalized) radius $Ka = 2\pi$ is cloaked using a classic linear elastic cloak $g(R) = a + R\left(\frac{d-a}{d}\right)$ in $2\pi \leq Kr \leq 4\pi$ as defined in (12.51). **Upper right**: Scattering from a cavity of radius $KA = 2\pi/20$ in an unstressed medium. **Lower left**: A hyperelastic cloak in $2\pi \leq Kr \leq 4\pi$ generated from an annulus with initial inner radius $KA = 2\pi/20$. **Lower right**: Scattering from a cavity with radius $KA = 2\pi$ in an unstressed medium. Scattering and the shadow region presence in the latter is significant, as compared with that for an equivalent sized cavity for the hyperelastic cloak.

the so-called *Mooney-Rivlin* model with associated strain energy function

$$\begin{aligned}\mathcal{W}_{\mathrm{MR}} = \frac{\mu}{2}(S_1(\lambda_r^2 + \lambda_\theta^2 + \lambda_z^2 - 3) \\ + (1 - S_1)(\lambda_r^2\lambda_\theta^2 + \lambda_r^2\lambda_z^2 + \lambda_\theta^2\lambda_z^2 - 3))\end{aligned} \tag{12.93}$$

where $S_1 \in [0, 1]$ is an additional material parameter.

As some background to the above, the neo-Hookean model was apparently first proposed by Treloar in his 1943 paper [Treloar (1943)]. It was shown by Rivlin, in 1948 [Rivlin (1948a)], that the deformation produced on a unit cube of such a material by the action of three equally and oppositely directed forces acting normally on its faces is uniquely determined provided that the forces per unit area, measured in the deformed state, are specified. The Mooney-Rivlin model was proposed by Mooney in his 1940 paper [Mooney (1940)] and the same uniqueness theorem as above was proved by Rivlin in [Rivlin (1948b)] for this model.

In the case of a pre-stressed Mooney-Rivlin annulus, the wave equation in the cloak region cannot be solved explicitly. It therefore has to be solved numerically inside this finite domain. *Outside* the cloak region we represent the field in the form $w = w_i + w_s$ where

$$w_i(r) = \frac{C}{4i\mu_{\rm e}} \sum_{n=-\infty}^{\infty} e^{in(\theta-\theta')} \times \begin{cases} \mathrm{H}_n^{(1)}(k_e R')\mathrm{J}_n(k_e r), & D \le r < R', \\ \mathrm{H}_n^{(1)}(k_e r)\mathrm{J}_n(k_e R'), & r > R', \end{cases} \tag{12.94}$$

and

$$w_s(r) = \sum_{n=-\infty}^{\infty} (-i)^n a_n^{\rm MR} \mathrm{H}_n^{(1)}\left(k_e r\right) e^{in(\theta-\theta')} \tag{12.95}$$

and we must understand how the scattering coefficients $a_n^{\rm MR}$ vary with pre-stress. From the analysis above we know that at fixed frequency, they are invariant under pre-stress in the neo-Hookean case ($S_1 = 1$ in (12.93)) which makes it ideal for cloaking.

In the cloak region we denote the wave field as w_c and we pose the total field in the modal expansion

$$w_c(r,\theta) = \sum_{n=-\infty}^{\infty} (-i)^n \mathrm{F}_n(r) e^{in(\theta-\theta')}. \tag{12.96}$$

This yields the governing ordinary differential equation

$$\frac{1}{r}\frac{d}{dr}\left(r\mu_r(r)\frac{d\mathrm{F}_n}{dr}\right) + \left(\rho\omega^2 - \frac{\mu_\theta(r)}{r^2}\right)\mathrm{F}_n = 0 \tag{12.97}$$

where

$$\mu_r(r) = \frac{T\mu}{\zeta^2}\left(1 + \frac{m}{r^2}\right), \quad \mu_\theta(r) = \frac{T\mu}{\zeta^2}\left(1 - \frac{m}{(r^2 + M)}\right) \tag{12.98}$$

where $m = MS_1\zeta/T$ and $T = 1 + (\zeta - 1)S_1$. Equation (12.97) is solved numerically in the cloak region subject to $d\mathrm{F}_n/dr = 0$ on the inner boundary (traction free) and continuity of displacement and traction with the solution in the exterior domain on $r = R_2$, the latter condition being defined by

$$\left[\mu_r(r)\frac{\partial w_c}{\partial r}\right]\Bigg|_{r=D} = \left[\mu_\mathrm{e}\left(\frac{\partial w_i}{\partial r} + \frac{\partial w_s}{\partial r}\right)\right]\Bigg|_{r=D}. \tag{12.99}$$

Note the simplification $\mu_r(D) = \mu/\zeta^2$ upon using $M = D^2(\zeta^{-1} - 1)$. We can eliminate the scattering coefficients a_n^{MR} between the two conditions at $R = D$ which leads to the single condition

$$\begin{aligned}\frac{1}{k_0\zeta}\mathrm{F}_n'(D) - \left(\frac{\mathrm{H}_n^{(1)'}(k_eD)}{\mathrm{H}_n^{(1)}(k_eD)}\right)\mathrm{F}_n(D) \\ = \frac{C_0}{4i\mu_\mathrm{e}(-i)^n}\mathrm{H}_n^{(1)}(k_eR')\left(\mathrm{J}_n'(k_eD) - \frac{\mathrm{J}_n(k_eD)}{\mathrm{H}_n^{(1)}(k_eD)}\right)\end{aligned} \tag{12.100}$$

on $r = D$. This approach is frequently known in the literature as a "Dirichlet to Neumann mapping". This condition together with the traction free condition, $d\mathrm{F}_n/dr = 0$ on $r = a$, closes the system and enables a numerical solution to be found once material properties are specified. The scattering coefficients are then determined by substituting the solution in either of the continuity conditions at $r = D$. Various aspects of this condition, including the behaviour of the derivative of the Hankel function, pertaining to non-reflecting boundary conditions for time domain solutions in two and three dimensions are discussed in [Alpert *et al.* (2000)].

We have not yet specified the shear modulus and density of the cloak region nor the additional parameter S_1. We note here again that in fact in the neo-Hookean case the only reason to choose $\mu = \zeta\mu_\mathrm{e}$ and

$\rho = \zeta\rho_{\rm e}$ was to ensure the invariance under mapping and hence the ability to determine analytically the scattering coefficients under pre-stress. Choosing different material properties would lead to different scattering coefficients determined by a numerical solution. To avoid an exhaustive material parameter investigation let us fix $S_1 \neq 1$ for the Mooney-Rivlin case and choose $\mu = \zeta\mu_{\rm e}$ and $\rho = \zeta\rho_{\rm e}$ as in the neo-Hookean case.

With these parameters the ODE in (12.97) was solved in Mathematica 7, using the command NDSolve. In particular the Gelfand-Lokutsiyevskii chasing method is used [Berezin and Zhidkov (1965)]. The AccuracyGoal was set to 10 digits and the WorkingPrecision was set to 50. Results will be provided in the next section.

Using the hyperelastic cloaking technique, the anisotropic, inhomogeneous material moduli in the cloaking region, defined by (12.98), are induced naturally by the pre-stress and therefore exotic metamaterials devised by the use of upscaling (homogenization) procedures, as described in the previous section are not required. Furthermore such metamaterials by their nature require a microstructure of finite size so that dispersive effects will become apparent at some frequency. This will not be the case in the context of hyperelastic cloaks since the naturally induced properties are continuously varying. Additionally the density of the cloak is homogeneous in the hyperelastic case.

12.4.4.3. *Scattering cross section*

We wish to ascertain quantitatively the effectiveness of cloaking using the techniques introduced above. In order to do this we need a measure of the magnitude of scattering. A convenient such quantity is the scattering cross section (SCS). It is convenient to measure this for an incident plane wave and therefore in the theory above we fix $\Theta' = 0$ and take the appropriate limit $R' \to \infty$ together with the source strength limit (12.63) in order to yield an incident left propagating plane wave of unit amplitude, i.e. $w_i = e^{-ik_e x}$. In this limit, we also note that on the right hand side of (12.100)

$$\frac{C_0}{4i\mu_{\rm e}(-i)^n}\mathrm{H}_n^{(1)}(k_e R') \to 1. \tag{12.101}$$

Given the general form of a scattered antiplane wave field as

$$w_s = \sum_{n=-\infty}^{\infty} (-i)^n a_n \mathrm{H}_n^{(1)}(k_e r) e^{in(\theta-\Theta')} \tag{12.102}$$

where a_n are the associated scattering coefficients, we define the scattering cross section for antiplane wave scattering from an inclusion of radius r_1 as [Lewis *et al.* (1976)]

$$\gamma_{\mathrm{SH}} = \frac{2}{k_e r_1} \left(|a_0|^2 + 2 \sum_{n=1}^{\infty} |a_n|^2 \right). \tag{12.103}$$

We note that for plane wave incidence on a traction free cavity of radius a in an otherwise uniform elastic medium we would have

$$a_n = -\frac{\mathrm{J}_n'(k_e a)}{\mathrm{H}_n^{(1)'}(k_e a)} \tag{12.104}$$

and this case therefore corresponds to having no cloak around the cavity. This is clearly a useful comparison case. We would of course expect that having a good cloak leads to a reduction in scattering as compared with this case.

In order to provide a fair comparison of results let us fix the inner and outer radii of the cloak as $a = 2\pi$ and $D = 4\pi$ respectively. We then plot γ_{SH} as a function of $k_e a$ by varying k_e which is equivalent to considering incident waves of different wavelengths (or frequencies). For the hyperelastic cloak we fix the initial inner radius as $A = \pi/10$. Reducing A further would further reduce the associated scattering cross section.

In Fig. 12.9 we plot γ_{SH} for the case of a layered cloak as introduced above and we plot this for an increasing number of layers $2^j, j = 1, 2, \ldots, 8$. The number of cells N as defined above is $N = 2^{j-1}$. We also plot the case of scattering with no cloak (the red curve, notable by its lack of oscillations). For low frequencies $k_e a < 1$ scattering is always small and so we pay attention mainly to higher frequencies when $k_e a > 1$. As the number of layers increases we see the significant reduction in scattering cross section, to the extent that for $k_e a > 1$ and in the plot range considered, by the

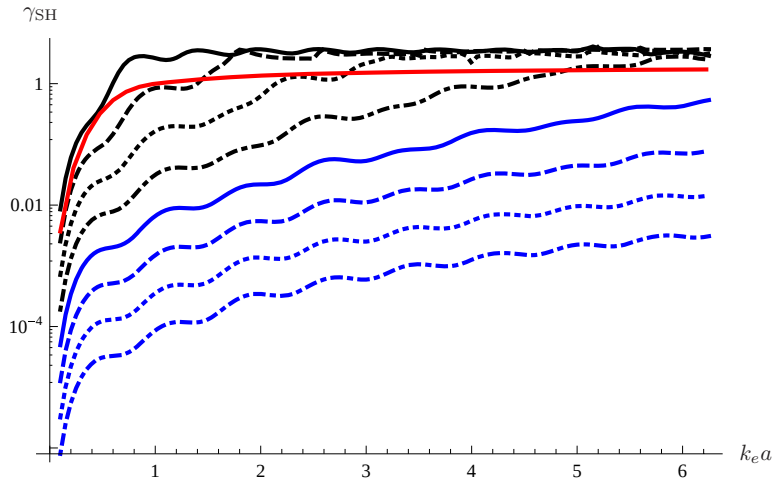


Fig. 12.9. Plot (on a logarithmic scale) of the scattering cross section for the layered metamaterial cloak with an increasing number of layers $2, 4, 8, 16, 32, 64, 128$ and 256 (black then blue curves, top to bottom). We also plot the case of no cloak to enable comparison (red curve with no oscillations). The improvement at larger N is clear. We note the oscillations for the cloak case are due to the layers being of finite width.

time we reach 128 layers $\gamma_{SH} \simeq 0.01$ and for 256 layers $\gamma_{SH} \simeq 0.001$. We note that we need at least 32 layers before a significant cloaking effect is achieved, although 16 layers achieves good cloaking up to around $k_e a = \pi$ with an approximate reduction in scattering of 90%. The reason for the poor performance for small N is clearly because by having a small number of layers we are introducing a number of large impedance mismatches and hence significant scattering. The magnitude of these mismatches reduce significantly as we increase the number of layers. We note the oscillations in the plotted curves, which are clearly associated with the fact that the cloak is constructed from layers of finite width.

In Fig. 12.10 we retain a number of plots from the layered case (the cases of $4, 16, 64$ and 256 layers) in order to compare these results with the hyperelastic case. The neo-Hookean cloak exhibits outstanding cloaking performance, the results being equivalent to the 256 layer metamaterial. Whilst the Mooney-Rivlin materials do not achieve the

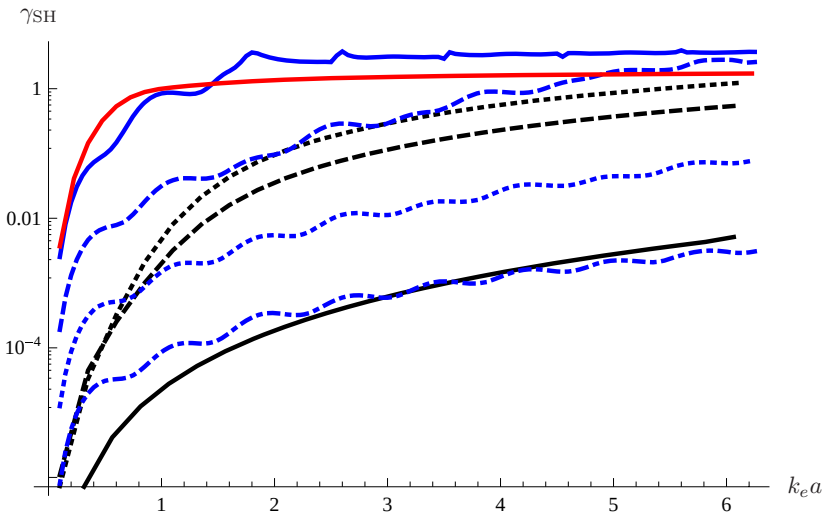


Fig. 12.10. Plot (on a logarithmic scale) of the scattering cross section for the layered metamaterial cloak with an increasing number of layers 4, 16, 64 and 256 (top to bottom — curves with oscillations) together with the corresponding plots for the hyperelastic cloaks (smooth monotonic curves) corresponding to neo-Hookean, Mooney-Rivlin $S_1 = 0.9$ and $S_1 = 0.8$ (bottom to top). The initially uppermost (red) smooth curve is associated with scattering when no cloak is present, which enables useful comparisons to be made.

same level of performance as the neo-Hookean cloak we see that for much of the parameter range considered $\gamma \simeq 0.1$. This amounts to a 90% reduction in scattering which in practical terms is extremely useful.

12.5. Transformational Invariance for Hyperelastic Cloaking

Let us now broaden our perspective and consider the general elastodynamics problem of coupled compressional and shear waves. The objective is to find possible hyperelastic solids, i.e. strain energy functions $\mathcal{W}$, such that the equations for small-on-large motion are equivalent to those required after transformation of a material assumed to be homogeneous with properties $\{\rho_0, C^0_{IjK\ell}\}$.

The connection between the transformation and the small-on-large theories is made by first identifying the displacement fields as equivalent, $\mathbf{u}^{\text{eff}}(\mathbf{x},t) = \mathbf{u}(\mathbf{x},t)$, and then requiring that the equations of motion (12.9) and (12.46) (coupled with (12.28)) are the same. This is the case if

$$\rho = \gamma\rho^{\text{eff}}, \quad M_{ijk\ell} = \gamma C^{\text{eff}}_{ijk\ell}, \tag{12.105}$$

for some positive constant γ. Hence,

$$J^{-1}\rho_r = \gamma J_0^{-1}\,\rho_0, \quad J^{-1}\,F_{iM}F_{kN}\,L_{MjN\ell} = \gamma J_0^{-1}\,F^0_{iI}F^0_{kK}\,C^0_{IjK\ell}. \tag{12.106}$$

The reference density ρ_r can then be chosen so that $\gamma = 1$, and Eq. (12.106) then implies that the hyperelastic material is defined by

$$\rho_r = \rho_0 J_0^{-1} J, \qquad L_{MjN\ell} = J_0^{-1} J\,F^{-1}_{Mi}F^0_{iI}\,F^{-1}_{Nk}F^0_{kK}\,C^0_{IjK\ell}. \tag{12.107}$$

Equation $(12.107)_1$ is automatically satisfied if the transformation and the finite deformation are related in the following manner:

$$\mathbf{F} = \left(g\rho_r/\rho_0\right)^{1/3}\mathbf{F}^0\mathbf{G}^{-1}, \quad g = \det\mathbf{G}, \tag{12.108}$$

for some non-singular $\mathbf{G}$. Equation (12.108) combined with the expression for $L_{MjN\ell}$ in Eq. (12.107) yields a second order differential equation for the strain energy function,

$$\frac{\partial^2\mathcal{W}}{\partial F_{jM}\partial F_{\ell N}} = \left(\frac{\rho_r}{g^2\rho_0}\right)^{1/3} G_{MI}G_{NK}C^0_{IjK\ell}. \tag{12.109}$$

Recall that ρ_0 and $C^0_{IjK\ell}$ are constant, but at this stage the remaining quantities in (12.109), i.e. ρ_r and $\mathbf{G}$, are not so constrained. The density in the reference configuration could be inhomogeneous, $\rho_r = \rho_r(\mathbf{X})$. In that case (12.109) would not have a general solution for $\mathcal{W}$ unless $\mathbf{G}$ also depends upon $\mathbf{X}$ in such a manner that the right hand side is independent of $\mathbf{X}$. This suggests that permitting ρ_r to be inhomogeneous does not provide any simplification, and we therefore take the reference density to be constant, although not necessarily the same as ρ_0. The quantity $\mathbf{G}$ could, in principle, be a matrix function

of $\mathbf{F}$, but this makes the integration of (12.109) difficult, if not impossible. We therefore restrict attention to constant $\mathbf{G}$. Consideration of the important case of isotropic elasticity in §12.5.1 indicates that the degrees of freedom embodied in $\mathbf{G}$ do not provide any significant additional properties, and therefore from hereon-in we take $\mathbf{G} = \mathbf{I}$, and set $\rho_r = \rho_0$ with no loss in generality. In this case the solution of (12.109) such that $\mathcal{W} = 0$ under zero deformation ($\mathbf{F} = \mathbf{I}$) is

$$\mathcal{W} = \frac{1}{2}(F_{jM} - \delta_{jM})(F_{\ell N} - \delta_{\ell N})\, C^0_{MjN\ell}. \tag{12.110}$$

Equation (12.110) provides a formal solution for $\mathcal{W}$ that is consistent with (12.109). However, the dependence of $\mathcal{W}$ in (12.110) upon $\mathbf{F}$ points to a fundamental difficulty, since the strain energy should be a function of the right stretch tensor $\boldsymbol{\mathcal{U}}$. The two are not equal in general, unless

$$\mathbf{R} = \mathbf{I} \quad \Leftrightarrow \quad \mathbf{F} = \boldsymbol{\mathcal{U}} = \boldsymbol{\mathcal{V}}. \tag{12.111}$$

We henceforth assume (12.111) to be the case: that is, *we restrict consideration to deformations that are everywhere rotation-free*. Equation (12.110) then suggests the following possible form of the finite strain energy

$$\mathcal{W} = \frac{1}{2}E_{jM}E_{\ell N}\, C^0_{MjN\ell} \quad \text{where} \quad \mathbf{E} \equiv \boldsymbol{\mathcal{U}} - \mathbf{I}. \tag{12.112}$$

Although this has realistic dependence on $\boldsymbol{\mathcal{U}}$, it will not in general satisfy Eq. (12.109), i.e. $\partial^2\mathcal{W}/\partial F_{jM}\partial F_{\ell N} \neq C^0_{MjN\ell}$. We return to this crucial point for isotropic elasticity in §12.5.3 where we demonstrate that Eq. (12.109) is satisfied by the isotropic form of (12.112) under additional conditions. Note that the strain measure $\mathbf{E}$, which is sometimes called the *extension tensor*, has as conjugate stress measure $\boldsymbol{\mathcal{S}}_a = \partial\mathcal{W}/\partial\mathbf{E} = \frac{1}{2}(\boldsymbol{\mathcal{S}}\boldsymbol{\mathcal{U}} + \boldsymbol{\mathcal{U}}\boldsymbol{\mathcal{S}})$ where $\boldsymbol{\mathcal{S}}$ is the second Piola-Kirchhoff tensor introduced in (12.24) [Dill (2007), §2.5].

Let us now summarize the two approaches of transformation elasticity and the theory of small on large and explore their connections.

Transformation elastodynamics

Untransformed (unphysical) medium, density ρ_0, moduli $C^0_{ijk\ell}$,

$$\frac{\partial}{\partial X_i}\left(C^0_{ijk\ell}\frac{\partial U^0_\ell}{\partial X_k}\right)+\rho_0\omega^2 U^0_j=0.$$

Mapping $\mathbf{x}=\boldsymbol{\chi}_0(\mathbf{X}),\mathbf{F}^0=\mathrm{Grad}\boldsymbol{\chi_0}$ to the transformed (physical) domain:

$$\frac{\partial}{\partial x_i}\left(C^{\mathrm{eff}}_{ijk\ell}\frac{\partial u^{\mathrm{eff}}_\ell}{\partial x_k}\right)+\rho^{\mathrm{eff}}\omega^2 u^{\mathrm{eff}}_j=0,$$

where $\rho^{\mathrm{eff}}=J_0^{-1}\rho_0$ and

$$C^{\mathrm{eff}}_{ijk\ell}=\frac{1}{J_0}F^0_{im}F^0_{kn}C^0_{mjn\ell}.$$

Small-on-large

Reference (undeformed) medium of density ρ_r, defined by strain energy function $\mathcal{W}$.

Finite deformation $\mathbf{x}=\boldsymbol{\chi}(\mathbf{X}),\mathbf{F}=\mathrm{Grad}\boldsymbol{\chi}$ and small-on-large equations:

$$\frac{\partial}{\partial x_i}\left(M_{ijk\ell}\frac{\partial u_\ell}{\partial x_k}\right)+\rho\omega^2 u_j=0,$$

where $\rho=J^{-1}\rho_r$ and

$$M_{ijk\ell}=J^{-1}F_{iM}F_{kN}\frac{\partial^2\mathcal{W}}{\partial F_{jM}\partial F_{\ell N}}.$$

Connecting the two theories

Properties are defined by equating the following quantities:

$$\mathbf{u}^{\mathrm{eff}}=\mathbf{u},\qquad \rho^{\mathrm{eff}}=\rho,\qquad C^{\mathrm{eff}}_{ijk\ell}=M_{ijk\ell}.$$

We restrict attention henceforth to the case of hyperelastic materials that are *isotropic* in the undeformed state.

12.5.1. *Isotropy — the semilinear strain energy function*

The initial moduli are $C^0_{\alpha j\beta l} = \lambda_0\delta_{\alpha j}\delta_{\beta l} + \mu_0(\delta_{\alpha\beta}\delta_{jl} + \delta_{l\alpha}\delta_{j\beta})$ with original Lamé moduli $\mu_0 > 0$, λ_0 and Poisson's ratio $\nu = \lambda_0/[2(\lambda_0 + \mu_0)] \in (-1, \frac{1}{2})$. We consider the isotropic version of the hyperelastic strain energy in (12.112),

$$\mathcal{W} = \frac{\lambda_0}{2}(\mathrm{tr}\mathbf{E})^2 + \mu_0\mathrm{tr}(\mathbf{E})^2 = \frac{\lambda_0}{2}(i_1 - 3)^2 + \mu_0\big((i_1 - 1)^2 - 2(i_2 - 1)\big), \tag{12.113}$$

with the latter expression in terms of two of the three invariants of $\boldsymbol{\mathcal{U}}$: $i_1 = \lambda_1 + \lambda_2 + \lambda_3$, $i_2 = \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1$ where λ_1, λ_2, λ_3 are the principal stretches of $\boldsymbol{\mathcal{U}}$. Materials with strain energy (12.113) have been called semilinear [Lur'e (1968)] because of its relative simplicity and the linear form of the Piola-Kirchhoff stress $\mathbf{P}$ which we recall is related to the Cauchy stress by $\boldsymbol{T} = J^{-1}\mathbf{F}\mathbf{P}^T$; thus $\mathbf{P} = 2\mu_0\mathbf{F} + \big(\lambda_0(\mathrm{tr}\mathbf{E}) - 2\mu_0\big)\mathbf{R}$. John [John (1960)] proposed the strain energy (12.113) based on the explicit form of its complementary energy density in terms of $\mathbf{P}$, a property also noted by others [Zubov (1970); Raasch (1975)]. The semilinear strain energy is a special case of the more general harmonic strain energy function [John (1960)]. Plane strain solutions for harmonic strain energy are reviewed in [Ogden (1997), §5.2]. Sensenig [Sensenig (1964)] examined the stability of circular tubes under internal pressure, while Jafari *et al.* [Jafari *et al.* (1984)] considered both internal and external pressure loading. The latter study has implications for the stability of the pre-strain developed here, see §12.6.2.3.

For semi-linear materials, the Cauchy finite pre-stress follows as

$$\boldsymbol{T} = J^{-1}\big[\lambda_0(i_1 - 3)\boldsymbol{\mathcal{V}} + 2\mu_0(\boldsymbol{\mathcal{V}}^2 - \boldsymbol{\mathcal{V}})\big]. \tag{12.114}$$

It is emphasized that we are restricting attention to deformations with $\boldsymbol{\mathcal{U}} = \mathbf{F} = \mathbf{F}^t$, so that the Piola-Kirchhoff stress is also symmetric

with $\mathbf{P} = \lambda_0(\mathrm{tr}\mathbf{E})\mathbf{I} + 2\mu_0\mathbf{E}$. Applying the equilibrium equation for the finite deformation,

$$\mathrm{Div}\mathbf{P} = 0 \quad \Rightarrow \quad \lambda_0 x_{\alpha,\alpha j} + 2\mu_0 x_{j,\alpha\alpha} = 0. \tag{12.115}$$

We seek solutions with symmetric deformation gradient, $x_{j,\alpha} = x_{\alpha,j}$, and hence $x_{\alpha,\alpha j} = x_{j,\alpha\alpha}$. Consequently Eq. (12.115) is satisfied by finite deformations satisfying either of the equivalent conditions $x_{j,\alpha\alpha} = 0$ and $x_{\alpha,\alpha j} = 0$. Thus:

$$\text{if } \; x_{j,\alpha} = x_{\alpha,j} \quad \text{then } x_{\alpha,\alpha j} = 0 \quad \Leftrightarrow \quad x_{j,\alpha\alpha} = 0. \tag{12.116}$$

Since the two partial differential equations in (12.116) are the same we need only seek solutions of one. Focusing on $x_{\alpha,\alpha j} = 0$ we conclude that the most general type of deformation $\mathbf{x}(\mathbf{X})$ is described by

$$\mathrm{Div}\ \mathbf{x} = c \ (= \text{constant} > 0), \quad \text{where Div } \mathbf{x} = \big(\mathrm{Div}\ \mathbf{x}\big)^T. \tag{12.117}$$

The appearance of the positive constant of integration in $(12.117)_1$ means that the sum of the principal stretches is fixed,

$$\lambda_1 + \lambda_2 + \lambda_3 = c, \quad (i_1 = c). \tag{12.118}$$

Further implications of the general solution (12.117) for a material that is isotropic in its undeformed state are explored in greater detail in the next section. For now we note that the pre-stress follows from (12.114) and (12.118) as

$$\boldsymbol{T} = 2\mu_0 J^{-1}\Big(\boldsymbol{\mathcal{V}}^2 - \boldsymbol{\mathcal{V}} + \frac{(c-3)\nu}{1-2\nu}\boldsymbol{\mathcal{V}}\Big). \tag{12.119}$$

12.5.2. *The limit of* $\nu = \frac{1}{2}$

It is of interest to consider the limit of the isotropic solution for $\nu = \frac{1}{2}$. By assumption the pre-stress must remain finite. Consequently, using Eq. (12.119), it becomes clear that in the limit as $\nu \to \frac{1}{2}$ the constant

of integration $c \equiv 3$, i.e.

$$\mathcal{W} = \mu_0 \,\mathrm{tr}(\boldsymbol{\mathcal{U}}^2 - \mathbf{I}), \qquad \mathbf{T} = pJ^{-1}\,\boldsymbol{\mathcal{V}} + 2\mu_0 J^{-1}\,(\boldsymbol{\mathcal{V}}^2 - \boldsymbol{\mathcal{V}}), \tag{12.120}$$

$$\mathrm{Div}\,\mathbf{x} = 3, \qquad \mathbf{F} = \mathbf{F}^T\,(= \boldsymbol{\mathcal{U}} = \boldsymbol{\mathcal{V}}), \tag{12.121}$$

where the scalar $p(\mathbf{X})$ defines the constraint reaction stress (the factor J^{-1} is included for later simplification). The latter arises from the limiting process of $\nu \to \frac{1}{2}$ in Eq. (12.119), and has also been shown to be the unique form of the reaction stress for the constraint $\mathrm{tr}\boldsymbol{\mathcal{V}} = 3$ [Beatty and Hayes (1992)]. Note that in writing $\mathbf{T}$ in (12.120) we maintain a term proportional to $\boldsymbol{\mathcal{V}}$ in the second term rather than incorporating it with the constraint term. This form is consistent with the requirement that $p = 0$ and hence $\mathbf{T} = 0$ in the undeformed state $\mathbf{x} \equiv \mathbf{X}$. The equilibrium equation for the pre-strain follows from Eq. (12.120) as $\mathrm{Grad}\,p + 2\mu_0 \mathrm{Div}\,\mathrm{Grad}\,\mathbf{x} = 0$, and since $\mathrm{Div}\,\mathrm{Grad}\,\mathbf{x} = 0$ (see Eq. (12.116)) it follows that $p =$ constant.

Several aspects of (12.120) are noteworthy. The limit of $\nu = \frac{1}{2}$ is usually associated with incompressibility, i.e. the constraint $J = 1$ or equivalently $i_3 \equiv \lambda_1\lambda_2\lambda_3 = 1$, although the reason underlying this identification originates in linear elasticity and is therefore by no means required. Strictly speaking, the isochoric constraint $i_3 = 1$ conserves volume under the deformation. Here we find that $\nu = \frac{1}{2}$ implies the kinematic constraint on the deformation that $i_1 = \lambda_1 + \lambda_2 + \lambda_3 = 3$. The latter is associated with the notion of incompressibility in linear elasticity in the form $\mathrm{tr}\boldsymbol{E} = 0$ and in the present context can be viewed as a "semilinear" feature, in keeping with the descriptor [Lur'e (1968)] for the strain energy function (12.113). The kinematic condition, $\mathrm{Div}\,\mathbf{x} = 3$ or equivalently

$$\lambda_1 + \lambda_2 + \lambda_3 = 3 \quad \big(\mathrm{tr}\mathbf{V} = 3\big), \tag{12.122}$$

has been called the *Bell constraint* [Beatty and Hayes (1992)] by virtue of the fact that Bell [Bell (1985, 1989)] showed it to be

consistent with numerous sets of data for metals in finite strain. Solids satisfying this constraint have been called *Bell materials* [Beatty and Hayes (1992)]. In contrast to the constraint $\lambda_1\lambda_2\lambda_3 = 1$ it can be shown that volume decreases for every deformation of a Bell material, and hence isochoric deformations are not possible [Beatty and Hayes (1992)].

Another feature of the $\nu = \frac{1}{2}$ limit is that the strain energy in (12.120) has the functional dependence $\mathcal{W} = \mu_0(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3)$. It is interesting to compare this with the strain energy for a neo-Hookean solid, $\mathcal{W}_{\rm NH} = \frac{\mu_0}{2}(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3)$, associated with incompressibility (i.e. $i_3 = 1$). Both strain energies reduce to the incompressible form for linear elasticity, and the factor $\frac{1}{2}$ appearing in $\mathcal{W}$ but not in $\mathcal{W}_{\rm NH}$ can be attributed to the different constraints in each case ($i_1 = 3$ or $i_3 = 1$). Parnell [Parnell (2012)] and Parnell *et al.* [Parnell *et al.* (2012)] considered neo-Hookean materials in the context of transformation elasticity for isochoric deformation. The present results indicates that the same form of the strain energy but with a different constraint yields a quite distinct class of volume decreasing deformations. This aspect will be examined further in the next section in terms of specific examples.

12.5.3. *Consistency condition*

It remains to show that the semilinear strain energy (12.113) satisfies

$$L_{\alpha j\beta l} = \lambda_0\delta_{\alpha j}\delta_{\beta l} + \mu_0(\delta_{\alpha\beta}\delta_{jl} + \delta_{l\alpha}\delta_{j\beta}) \tag{12.123}$$

where we recall that $L_{ijk\ell}$ was defined in (12.27). Since the modulus tensor $\mathbf{L}$ is isotropic, it it sufficient to show the equivalence in any orthogonal system of coordinates. We choose the principal coordinate system, in which the non-zero components satisfy those given in (12.35)–(12.37). Using $\mathcal{W}$ and c as defined in Eqs. (12.113) and (12.118) gives

$$\mathcal{W}_i = \lambda_0\,(c-3) + 2\mu_0\,(\lambda_i - 1), \tag{12.124}$$

$$\mathcal{W}_{ii} = \lambda_0 + 2\mu_0, \quad \mathcal{W}_{ij} = \lambda_0, \quad i \neq j. \tag{12.125}$$

These satisfy (12.35), and (12.37). The remaining conditions (12.36) become

$$\mathcal{W}_i + \mathcal{W}_j = 0 \quad \Rightarrow \quad (\lambda_0 + \mu_0)\,(c-3) - \mu_0\,(\lambda_k - 1) = 0, \qquad i \neq j \neq k \neq i. \quad (12.126)$$

Equation (12.126) constitutes three conditions, which taken together imply the unique but trivial solution $\lambda_i = 1$, $i \in \{1, 2, 3\}$, i.e. zero pre-strain. We avoid this by *restricting attention to two dimensional dynamic solutions only*, either in-plane (P/SV) or out-of-plane (SH) motion.

12.5.3.1. *In-plane (P/SV) motion*

The small-on-large displacements for in-plane motion are of the form $u_1(x_1, x_2, t)$, $u_2(x_1, x_2, t)$, $u_3 = 0$. The condition (12.126) then only needs to be satisfied in the single instance $i, j = 1, 2$, implying that the out-of-plane extension is related to the sum of the in-plane extensions by

$$\lambda_3 = 1 - \frac{1}{2\nu}\,(\lambda_1 + \lambda_2 - 2). \quad (12.127)$$

Since λ_3 is strictly positive, this places an upper limit on the sum of the in-plane extensions: $\lambda_1 + \lambda_2 < 2(1 + \nu)$.

12.5.3.2. *Out-of-plane (SH) motion*

The out-of-plane SH motion is of the form $u_1 = u_2 = 0$, $u_3(x_1, x_2, t)$. The requirement now is that L_{1313} and L_{2323} are both equal to μ_0 in order to recover the out-of-plane equation of motion and associated tractions. Using (12.36) and (12.37)

$$L_{1313} - \mu_0 = \left(\frac{\lambda_0 + \mu_0}{\lambda_1 + \lambda_3}\right)\big[c - 3 - (1 - 2\nu)(\lambda_2 - 1)\big], \quad (12.128)$$

$$L_{2323} - \mu_0 = \left(\frac{\lambda_0 + \mu_0}{\lambda_2 + \lambda_3}\right)\big[c - 3 - (1 - 2\nu)(\lambda_1 - 1)\big], \quad (12.129)$$

where c is the constant from Eq. (12.118). In this form it is clear that if $\nu \neq \frac{1}{2}$ then in-plane pre-stretches must be the same, $\lambda_1 = \lambda_2 =$

$1+(c-3)/(1-2\nu)$, and therefore all the stretches are constant (since c is a constant). This rules out the possibility of SH cloaking since we require that the in-plane pre-strain be inhomogeneous. However, *if both* $\nu = \frac{1}{2}$ *and* $c = 3$ simultaneously hold, then $L_{1313} = L_{2323} = \mu_0$ for inhomogeneous and unequal in-plane stretches λ_1 and λ_2. We are therefore led to the conclusion that SH cloaking requires a separate limit of the semilinear strain energy, one satisfying the Bell constraint (12.122) for which the strain energy and stress are given by (12.120). Note that we do not get the neo-Hookean strain energy in this limit.

12.6. Two Dimensional Hyperelastic Cloaking for Isotropic Elasticity

12.6.1. *Radially symmetric cylindrical deformations*

Consider deformations that are radially symmetric, $r = r(R)$, $\theta = \Theta$, in cylindrical coordinates (r, θ, x_3) and (R, Θ, X_3). The stretch in the x_3-direction is assumed fixed, $\lambda_3 =$ constant. The deformation gradient for $r = r(R)$ is irrotational with

$$\left(\mathbf{F}^T = \right) \quad \mathbf{F} = \lambda_r \mathbf{I}_r + \lambda_\theta \mathbf{I}_\theta + \lambda_3 \mathbf{I}_3, \quad \lambda_r = \frac{dr}{dR}, \quad \lambda_\theta = \frac{r}{R}, \tag{12.130}$$

where $\mathbf{I}_r = \boldsymbol{e}_r \otimes \boldsymbol{e}_r$, $\mathbf{I}_\theta = \boldsymbol{e}_\theta \otimes \boldsymbol{e}_\theta$ and $\mathbf{I}_3 = \boldsymbol{e}_3 \otimes \boldsymbol{e}_3$. The condition (12.118) implies that the sum of the in-plane principal stretches is constant, say c_0, and the constraint (12.127) relates this to c of Eq. (12.122),

$$(1-2\nu)c_0 + 2\nu c = 2(1+\nu) \quad \text{where} \quad c_0 = \lambda_r + \lambda_\theta. \tag{12.131}$$

Equation (12.117) for $\mathbf{x}$ reduces to an ordinary differential equation for $r(R)$,

$$\frac{dr}{dR} + \frac{r}{R} = c_0, \tag{12.132}$$

with general solution

$$r = \frac{c_0}{2} R + c_1 R^{-1}, \quad c_1 = \text{constant}. \tag{12.133}$$

Note that the free parameter c_0 may be expressed in terms of either c or λ_3, using Eqs. (12.127) and (12.131). Using Eqs. (12.119), (12.132) and (12.130) it follows that the principal stretches and stresses for the radially symmetric cylindrical configuration are

$$\lambda_r = 2 - \lambda_\theta + 2\nu(1 - \lambda_3), \quad \lambda_\theta = \frac{r}{R}, \quad \lambda_3, \tag{12.134}$$

and

$$T_{rr} = \frac{\mu_0}{\lambda_3\lambda_\theta}(\lambda_r - \lambda_\theta), \quad T_{\theta\theta} = \frac{\mu_0}{\lambda_3\lambda_r}(\lambda_\theta - \lambda_r),$$
$$T_{zz} = \frac{2\mu_0}{\lambda_r\lambda_\theta}(1 + \nu)(\lambda_3 - 1). \tag{12.135}$$

Note that

$$\frac{r}{R} \to 1, \quad T_{rr} \to 0 \quad \text{as } r \to \infty \quad \text{iff} \quad \lambda_3 = 1 \quad (\Leftrightarrow c_0 = 2). \tag{12.136}$$

12.6.2. *Two types of cloaking*

12.6.2.1. *Conventional cloaking (CC)*

The conventional concept of a cloaking material is that it occupies a finite region, in this case, the shell $R \in [A, D]$ which maps to an equivalent shell in physical space with the same outer surface and an inner surface of radius larger than the original, i.e. $r \in [a, D]$, $a \in (A, D)$. Applying (12.133) with the two constraints $r(A) = a$ and $r(D) = D$ yields

$$r = R + (a - A)\left[\frac{\left(\frac{D}{R}\right)^2 - 1}{\left(\frac{D}{A}\right)^2 - 1}\right]\frac{R}{A}, \quad R \in [A, D], \quad \text{(CC)} \tag{12.137}$$

which specifies the previously free parameter λ_3 (also c and c_0) as

$$\lambda_3 = 1 - \frac{1}{\nu} + \left[\frac{\frac{a}{A} - 1}{\left(\frac{D}{A}\right)^2 - 1}\right] > 1. \tag{12.138}$$

The constraint $\lambda_3 > 0$ therefore sets a lower limit in the permissible value of the outer radius as

$$D > A\left(1 + \frac{1}{\nu}\left(\frac{a}{A} - 1\right)\right)^{1/2}. \tag{12.139}$$

The mapping (12.137) must also be one-to-one within the shell with $\lambda_r = dr/dR > 0$. This means that there should be no zero of $dr/dR = 0$ for $R \in [A, D]$. The convex nature of the solution (12.137) implies there is only one zero, say at $R = \mathcal{R}$. Since $\operatorname{sgn} dr/dR = \operatorname{sgn}(R - \mathcal{R})$, it follows that $\mathcal{R} < A$ must hold. Noting from (12.133) that $\mathcal{R}^2 = 2c_1/c_0$, and using (12.137) to infer c_0 and c_1, the condition $\mathcal{R}^2 < A^2$ becomes

$$a < a_{\max} \equiv 2A \Big/ \left[1 + \left(\frac{A}{D}\right)^2\right]. \tag{12.140}$$

The magnification factor $\frac{a}{A} \geq 1$, which measures the ratio of the radius of the mapped hole to the radius of the original one, is therefore bounded according to

$$\frac{a}{A} < 2. \tag{12.141}$$

In order to achieve a reasonable degree of cloaking one expects that the magnification factor is large, so that the mapped hole corresponds to an original hole of small radius and hence small scattering cross section. The limitation expressed by (12.141) therefore places a severe restriction on the use of the hyperelastic material as a conventional cloak. Note that the upper limit on a in (12.140) is not strictly achievable because $a = a_{\max}$ implies $\lambda_r(A) = r'(A) = 0$ and hence the principal stresses $\sigma_\theta^{\text{pre}}, \sigma_{zz}^{\text{pre}}$ become infinite at $r = a$. Examples of the radial deformation are given in Fig. 12.11.

The hyperelastic mapped solid has other aspects that further diminish its attractiveness as a conventional cloaking material. Specifically, a non-zero traction must be imposed on both the outer and inner boundaries to maintain the state of prestress. Noting that

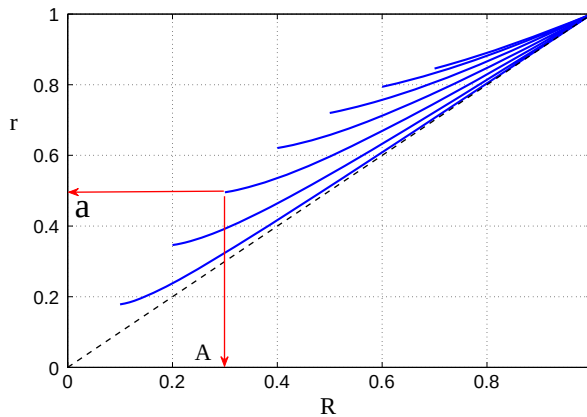


Fig. 12.11. The transformed radius r as a function of R for 2D radially symmetric pre-strain, from Eq. (12.137). The seven curves correspond to $A = 0.1, 0.2, \ldots, 0.7$ with $B = 1$ in each case. The value of the mapped inner radius for each curve is $a \equiv r(A) = 0.9\, a_{\max}$ where $a_{\max}$ is defined in Eq. (12.140). The dashed line indicates $r = R$. Mappings that lie above this line represent spatial compression, applicable to cloaking.

the radial stress is

$$T_{rr}(R) = -\frac{2\mu_0}{\lambda_3}\frac{R}{r}\left[\frac{r}{R} - 1 - \nu(1-\lambda_3)\right], \tag{12.142}$$

with r given in (12.137), yields

$$T_{rr}(A) = T_{rr}(D) - \frac{2\mu_0}{\lambda_3}\left(1 - \frac{A}{a}\right), \quad T_{rr}(D) = \frac{2\mu_0}{\lambda_3}\nu(1-\lambda_3). \tag{12.143}$$

The necessity of the inner traction at $r = a$ is a reasonable condition, but the requirement for an equilibrating traction at $r = D$ is physically difficult. One way to avoid this is to let $D \to \infty$, which will be considered next.

12.6.2.2. *Hyperelastic cloaking (HC)*

The hyperelastic material is now considered as infinite in extent. The cloaking effect is caused by allowing a radially symmetric hole in the unstressed configuration to be expanded under the action of an

internal pressure to become a larger hole. We therefore require that the traction at infinity is zero, and that r/R tends to unity, so that (12.136) applies. Then setting the mapped hole radius to $r(A) = a$ $(> A)$ implies the unique mapping

$$r = R + (a - A)\big(A/R\big), \quad R \in [A, \infty). \quad \text{(HC)} \tag{12.144}$$

This deformation is simply the limiting case of (12.137) for $D \to \infty$. Note that the restriction (12.141) still applies to the magnification factor $\frac{a}{A}$, in order to ensure $\lambda_r > 0$ for $r > a$. The traction at the inner surface is a pressure which follows from (12.143) in the limit $D \to \infty$, $\lambda_3 \to 1$, as

$$T_{rr}(A) = -p_{\text{in}}, \quad \text{where} \;\; p_{\text{in}} = 2\mu_0\Big(1 - \frac{A}{a}\Big). \tag{12.145}$$

It is interesting to note that the internal pressure is independent of the Poisson's ratio ν and it is therefore the same as p_{in} found by [Parnell (2012)] considering SH incremental motion.

12.6.2.3. *Stability of the pre-strain*

Jafari *et al.* [Jafari *et al.* (1984)] examined the stability of a finite thickness tube composed of material with harmonic strain energy, which includes semilinear strain energy as a special case. They showed that radially symmetric two dimensional finite deformations are stable under interior pressurization with zero exterior pressure. This implies that the finite pre-strain HC is stable. The stability of the CC deformation (12.137) does not appear to have been considered, and remains an open question. However, the stability of the HC, corresponding to $D \to \infty$, means there exists a minimum $D_{\min}$ for which CC stability is ensured for all $D > D_{\min}$.

12.6.3. *The limiting case when* $\nu = \frac{1}{2}$

In this limit the constraint (12.122) applies and the pre-stress for the radially symmetric deformation follows from (12.120) with constant

"pressure" p (see §12.5.2) as

$$T_{rr} = \frac{2\mu_0}{\lambda_3\lambda_\theta}(\lambda_r - \gamma_0), \qquad T_{\theta\theta} = \frac{2\mu_0}{\lambda_3\lambda_r}(\lambda_\theta - \gamma_0), \tag{12.146}$$

where the value of the constant $\gamma_0 = 1 + p/(2\mu_0)$ depends on the specified boundary conditions, and $\lambda_r = dr/dR$, $\lambda_\theta = r/R$, with $r(R)$ given by Eq. (12.133) for $c_0 \equiv 2$. For instance, in the case of hyperelastic cloaking as defined in §12.6.2.2 we find, noting the result (12.136), that $p = 0$, yielding the same interior pressure p_{in} as Eq. (12.145).

12.6.4. *Numerical examples*

We illustrate the above theory in the two dimensional setting where we consider wave scattering from a cylindrical cavity with and without a cloak, where the cloak is a conventional cloak created via pre-stress. We shall show that *partial cloaking* is achieved, in the sense that scattering is significantly reduced by the presence of the cloak. We are not able to achieve perfect cloaking since the cavity has to be of finite radius initially and furthermore, the hyperelastic theory above restricts the expansion to be at most twice the initial radius, i.e. $a < 2A$. We consider two cases: horizontally polarized shear (SH) waves and coupled compressional/in-plane shear (P/SV) waves. We take $D/a = 2$ which upon using (12.140) gives an initial inner to outer cloak radius ratio $D/A = 1/(2 - \sqrt{3}) \approx 3.732$ and $a/A = \beta = 1/(2(2 - \sqrt{3})) \approx 1.866$.

For the SH and P/SV wave examples considered, we need to formulate the elastic wave scattering problem in the context of the SH and P/SV problems. These are classical problems but it is convenient to summarize them here for completeness.

12.6.4.1. *Elastic wave scattering from cylindrical cavities*

Brief summaries of the two wave scattering problems are now given; for further details see e.g. Eringen and Suhubi [Eringen and Suhubi (1975)]. Scattering is considered from a cylindrical cavity of radius

A, located at the origin of a Cartesian coordinate system $\mathbf{X} = (X_1, X_2, X_3)$, related to a cylindrical polar coordinate system via $\mathbf{X} = (R\cos\Theta, R\sin\Theta, Z)$. An incident wave is generated by a line source of amplitude C (a force per unit length) located at the point (R', Θ'). We take $\Theta' \in [0, 2\pi)$, subtended from the positive X axis.

SH wave scattering

In this case the line source is polarized in the Z direction thus creating incident horizontally polarized shear (SH) waves which are then scattered from the cavity without mode conversion. The total wave field in this domain will therefore be $\mathbf{U} = (0, 0, W(X, Y))$ where W satisfies

$$(\nabla^2 + K_s^2)W = \frac{C}{R'}\delta(R - R')\delta(\Theta - \Theta') \tag{12.147}$$

with $K_s^2 = \rho\omega^2/\mu_0$ and $C = C_0/\mu_0$. We seek W in the form $W = W_i + W_s$ where $W_i = (C/4i)\mathrm{H}_0(K_s S)$ is the incident field and $S = \sqrt{(X - X')^2 + (Y - Y')^2}$ with $X' = R'\cos\Theta', Y' = R'\sin\Theta'$. We have defined $\mathrm{H}_0(K_s S) = \mathrm{H}_0^{(1)}(K_s S) = \mathrm{J}_0(K_s S) + i\mathrm{Y}_0(K_s S)$, the Hankel function of the first kind, noting that J_0 and Y_0 are Bessel functions of the first and second kind respectively, of order zero. Together with the $\exp(-i\omega t)$ time dependence in the problem, this ensures an outgoing field from the source. Graf's addition theorem allows us to write this field relative to the coordinate system (R, Θ) centred at the origin of the cavity [Martin (2006)] and we can use the form appropriate on $R = A$ in order to enforce the traction free boundary condition $\mu_0 \partial W/\partial R = 0$ on $R = A$, yielding the scattered field in the form

$$W_s = \sum_{n=0}^{\infty} \varepsilon_n D_n \mathrm{H}_n(K_s R)\cos(n(\Theta - \Theta')) \quad \text{with}$$

$$D_n = C\frac{i}{4}\frac{\mathrm{J}_n'(K_s A)}{\mathrm{H}_n'(K_s A)}\mathrm{H}_n(K_s R'). \tag{12.148}$$

where H_n and J_n are respectively Hankel and Bessel functions of the first kind of order n. We have also defined $\varepsilon_0 = 1$, $\varepsilon_n = 2$, $n \geq 1$. If we take $R' \to \infty$ and $C_0 = 2i\mu_0\sqrt{2\pi K_s R'}e^{i(\pi/4 - K_s R')}$, the incident wave of unit amplitude takes the (plane-wave) form $W_i = \exp\{iK_s(X\cos\Theta_{inc} + Y\sin\Theta_{inc})\}$ where $\Theta_{inc} = \Theta' - \pi \in [-\pi, \pi)$ is the angle of incidence subtended from the *negative* X axis. The scattered wave W_s takes the form $(12.148)_1$ with $D_n \to D_n^{\mathrm{pw}} \equiv -i^n \mathrm{J}'(K_s A)/\mathrm{H}_n'(K_s A)$. The scattering cross section of the cylindrical cavity for plane SH wave incidence from a cavity of radius r_1 was defined in (12.103).

P/SV wave scattering

In this case the line source at (R', Θ') with amplitude C_0 is a compressional source. Thus the incident field consists purely of in-plane compressional waves. Due to mode conversion, the scattered field consists of coupled in-plane compressional (P) and vertically polarized shear (SV) waves. The total wave field will therefore be $\mathbf{U} = (U(X,Y), V(X,Y), 0)$ and using the Helmholtz decomposition $\mathbf{U} = \nabla\Phi + \nabla \times (\Psi\mathbf{k})$, we deduce that

$$\nabla^2\Phi + K_p^2\Phi = \frac{C}{R'}\delta(R - R')\delta(\Theta - \Theta'), \quad \nabla^2\Psi + K_s^2\Psi = 0 \tag{12.149}$$

where $K_p^2 = \omega^2\rho/(\lambda_0 + 2\mu_0)$, $K_s^2 = \omega^2\rho/\mu_0$ and $C = C_0/(\lambda_0 + 2\mu_0)$. We seek the wave field in the form $\Phi = \Phi_i + \Phi_s, \Psi = \Psi_s$ where $\Phi_i = (C/4i)\mathrm{H}_0(K_p S)$ is the incident compressional wave with S defined above in the SH scattering problem. The traction free ($T_{RR} = 0$, $T_{R\Theta} = 0$) boundary condition on $R = A$ is satisfied by using Graf's addition theorem, and the scattered field is

$$\Phi_s = \sum_{n=0}^{\infty} \varepsilon_n A_n H_n(K_p R)\cos(n(\Theta - \Theta')),$$

$$\Psi_s = \sum_{n=0}^{\infty} \varepsilon_n B_n H_n(K_s R)\sin(n(\Theta - \Theta')). \tag{12.150}$$

The scattering coefficients are

$$A_n = \frac{i}{4} C\mathrm{H}_n(K_pR')[\mathcal{I}_n^1(K_pA)M_n^{22}(K_sA) - \mathcal{I}_n^2(K_pA)M_n^{12}(K_sA)]/\Delta_n, \tag{12.151}$$

$$B_n = \frac{i}{4} C\mathrm{H}_n(K_pR')[\mathcal{I}_n^2(K_pA)M_n^{11}(K_pA) - \mathcal{I}_n^1(K_pA)M_n^{21}(K_pA)]/\Delta_n, \tag{12.152}$$

where

$$\mathcal{I}_n^1(x) = \left(n^2 + n - \tfrac{1}{2}(K_sA)^2\right)J_n(x) - xJ_{n-1}(x), \tag{12.153}$$

$$\mathcal{I}_n^2(x) = n(n+1)J_n(x) - nxJ_{n-1}(x), \tag{12.154}$$

$$M_n^{11}(x) = -M_n^{22}(x) = \left(n^2 + n - \tfrac{1}{2}(K_sA)^2\right)H_n(x) - xH_{n-1}(x), \tag{12.155}$$

$$M_n^{12}(x) = -M_n^{21}(x) = -n(n+1)H_n(x) + nxH_{n-1}(x), \tag{12.156}$$

$$\Delta_n = M_n^{11}(K_pA)M_n^{22}(K_sA) - M_n^{21}(K_pA)M_n^{12}(K_sA). \tag{12.157}$$

If we take $R' \to \infty$ together with $C_0 = 2i(\lambda_0 + 2\mu_0)\sqrt{2\pi K_pR'}$ $e^{i(\pi/4 - K_pR')}$ the incident wave of unit amplitude takes the (plane-wave) form $\Phi_i = \exp\{iK_p(X\cos\Theta_{inc} + Y\sin\Theta_{inc})\}$ where Θ_{inc} was introduced above for the SH wave case. The plane wave scattered fields take the form in (12.150) with A_n, $B_n \to A_n^{\mathrm{pw}}$, B_n^{pw} where the latter are given by (12.151)–(12.152) under the replacement $\frac{i}{4}C\mathrm{H}_n(K_pR') \to -i^n$. The scattering cross section γ_{P} of the cylindrical cavity for plane compressional wave incidence (subscript P indicating this fact) from a cavity of radius r_1 is [Lewis *et al.* (1976)]

$$\gamma_{\mathrm{P}} = \frac{2}{K_pr_1}\sum_{n=0}^{\infty}\varepsilon_n(|A_n^{\mathrm{pw}}|^2 + |B_n^{\mathrm{pw}}|^2). \tag{12.158}$$

12.6.4.2. *Hyperelastic cloaking for SH waves*

Consider first the case when the line source is located at a distance R' from the centre of the cavity with $R'/D = 2$ and the source is of unit amplitude, $C = 1$. In this case the shear wavenumber K_s of the medium is defined by $K_s^2 = \omega^2/c_s^2 = \rho_0\omega^2/\mu_0$ where ρ_0 is the density

of the medium in the undeformed configuration. We use the solution to the SH wave scattering problem above to solve the corresponding (conventional, pre-stress) cloak problem, the difference arising merely due to the modified argument due to the hyperelastic deformation (and invariance of equations). We shall always consider the case when $R' > D$, the outer cloak boundary. Thus in $R > D$ the solution can be written as (12.148), noting that the scattering coefficients are equivalent to scattering coefficients for a cavity of radius A. Therein resides the reduction in scattering. In $a < R < D$, the total field is given by $W_i + W_s$ but with an argument given by

$$R(r) = c_0^{-1}(r + \sqrt{r^2 - 2c_0c_1}) \tag{12.159}$$

i.e. that corresponding to the hyperelastic deformation described above (see equation (12.133)).

We take 30 terms in the modal sum (12.148) for the wave field, sufficient for convergence of the solution. Figure 12.12 shows both the total (top) and scattered (bottom) fields corresponding to the following problems: scattering from a cavity of radius A with $K_sA = 2\pi$ in an undeformed medium (left) and scattering from a cavity with the presence of a hyperelastic cloak (right) with undeformed (A) and deformed (a) inner radii defined via $a = \beta A$ where we recall that $\beta = 1.866$ was defined at the start of section 12.6.4. The outer cloak boundary D is defined by $K_sD = 4\pi$ (right). This demonstrates significantly reduced scattering due to the presence of the hyperelastic cloak as compared with the non-cloaked case. Indeed we are able to quantify this by determining the reduction in scattering cross section, defined in (12.103) for plane wave incidence. Without the cloak we have $\gamma_{\mathrm{SH}} = 5.39$ whereas with the cloak $\gamma_{\mathrm{SH}} = 2.61$ resulting in a 51.5% reduction in scattering. Figure 12.13 illustrates the scattering cross section γ_{SH} (left) together with the percentage reduction in scattering (right).

12.6.4.3. *Hyperelastic cloaking for P/SV waves*

The incident wave is now taken to be a pure compressional wave; a line source is located at a distance R' from the centre of the cavity

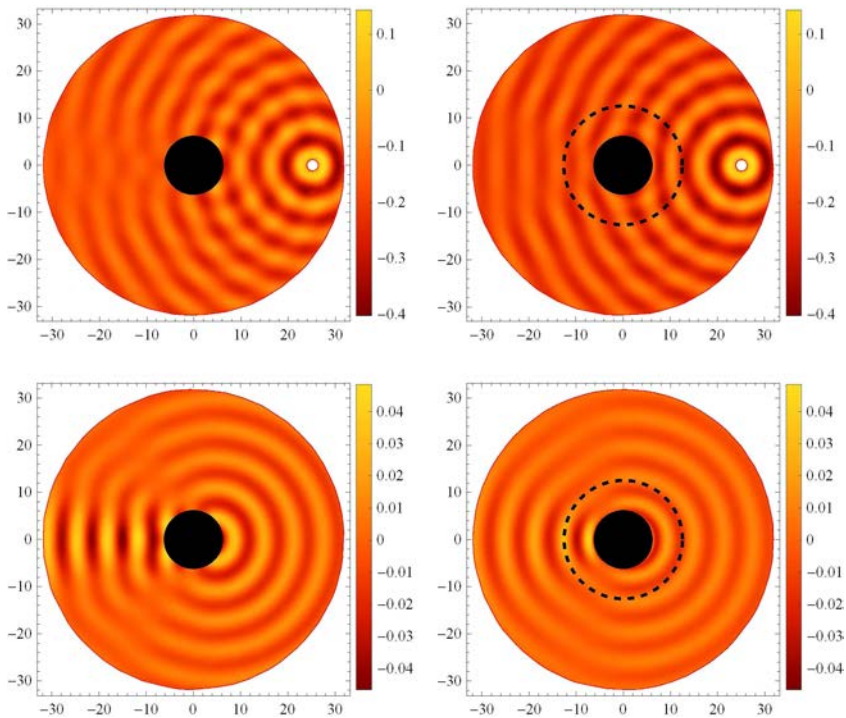


Fig. 12.12. SH wave field. Left: Total (top) and scattered (bottom) fields corresponding to an undeformed cavity with scaled radius $K_s A = 2\pi$. Right: Total (top) and scattered (bottom) fields corresponding to a conventional cloak generated via pre-stress where the scaled deformed inner radius is $K_s a = 2\pi$ and *initial* inner cavity radius defined by $a = \beta A$ where $\beta \approx 1.863$.

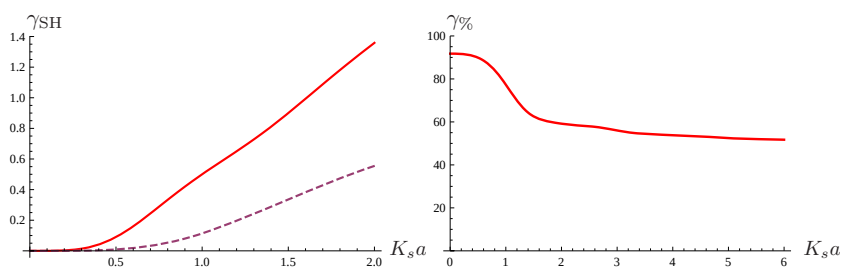


Fig. 12.13. Scattering cross section γ_{SH} (left) and percentage reduction in scattering cross section $\gamma_{\%}$ by using a hyperelastic cloak (right) (with $a = \beta A$ where $\beta \approx 1.863$.), both plotted against scaled cavity radius $K_s a$ for the SH wave case. The cross section is plotted without (solid) and with (dashed) a hyperelastic cloak. A significant reduction in scattering is achieved by using a hyperelastic cloak.

with $R'/D = 2$ and it is of unit amplitude $C = 1$. In the P/SV case, in addition to the shear wavenumber K_s we also introduce the compressional wavenumber K_p. We employ the P/SV scattering problem summarized above as a means of determining the solution for the cloak problem. This solution is employed in the exterior region together with the same solution but with modified argument (due to the hyperelastic deformation) in the cloak region. Thus in $R > D$ the solution can be written as (12.150) with scattering coefficients A_n and B_n given by (12.151) and (12.152) respectively, noting that they are equivalent to scattering coefficients for a cavity of radius A and therefore as in the SH case, a reduction in scattering is present. Note that here a different effect is introduced as compared with the SH case: shear waves are produced as a result of mode conversion on the boundary of the cavity. In $a < R < D$, the total field is given by the sum of the scattered and incident fields but with the argument as given in (12.159) due to the hyperelastic deformation.

We take 30 terms in the modal sums (12.150), which is sufficient for convergence of the solution. Figure 12.14 shows the scattered fields corresponding to the P-wave (top) and S-wave (bottom) fields associated with $\nu = 1/3$ and for the following problems: scattering from a cavity of radius A with $K_s A = 2\pi$ in an undeformed medium (left) and scattering from a cavity with the presence of a hyperelastic cloak (right) with undeformed (A) and deformed (a) inner radii defined via $a/A = \beta$. The outer cloak boundary D is defined by $K_s D = 4\pi$ (right). Scattering is significantly reduced due to the presence of the hyperelastic cloak as compared with the non-cloaked case although it is relatively difficult to see this directly with the plots. As with the SH case, let us quantify this by determining the reduction in scattering cross section, defined in (12.158) for plane wave incidence. Without the cloak $\gamma_{\mathrm{P}} = 13.564$ whereas with the cloak $\gamma_{\mathrm{P}} = 7.258$ resulting in a 46.48% reduction in scattering. Figure 12.15 illustrates the scattering cross section γ_{P} (left) together with the percentage reduction in scattering (right) compared for three different Poisson ratios: $\nu = 1/3, 7/15$ and $49/99$. Note that for very low frequencies there is a huge reduction in scattering, close to 100%. This

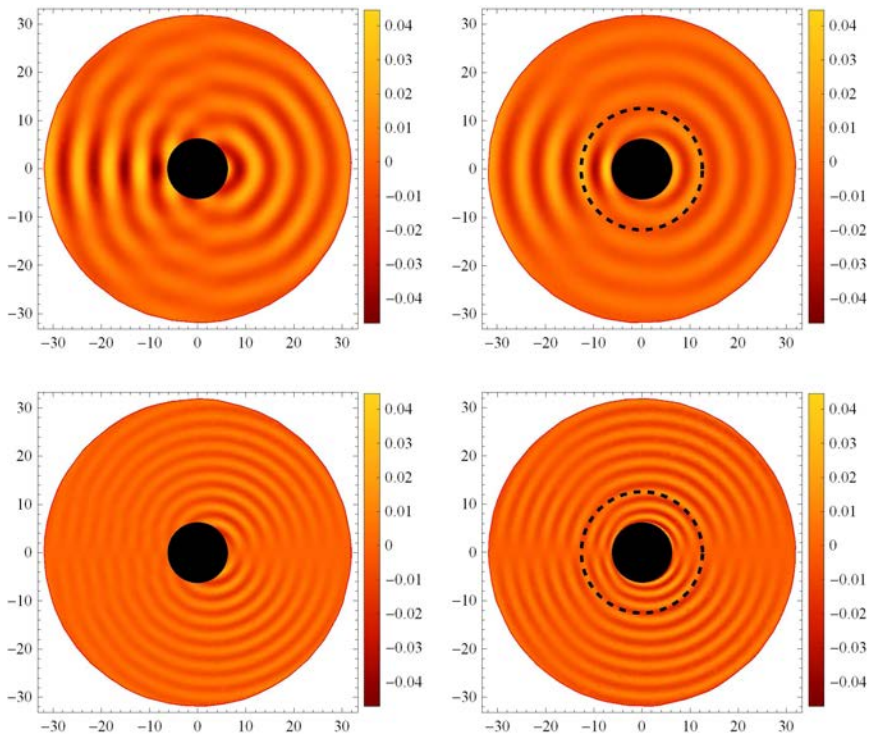


Fig. 12.14. Scattered fields for the in-plane P/SV problem for an incident field generated by a compressional source at $R' = 8\pi, \Theta' = 0$. Left: Compressional (top) and shear (bottom) fields corresponding to an undeformed cavity with scaled radius $K_pA = 2\pi$. Right: Compressional (top) and shear (bottom) fields corresponding to a conventional cloak generated via pre-stress where the scaled deformed inner radius is $K_pa = 2\pi$ and *initial* inner cavity radius is $K_pA = \pi$ so that $a = \beta A$ where $\beta \approx 1.863$.

tails off at higher frequencies but still remains at around 50% reduction in scattering which is clearly significant. Reduction is larger for smaller Poisson ratios. We also note the rather interesting result that the peak in the cross section actually induces an increase in scattering at some values of K_pa as compared with the case without the cloak although this is only for a narrow range of such values. This can be associated with the increasing disparity in the P and and SV wave numbers as ν tends to $\frac{1}{2}$, noting that $K_s^2/K_p^2 = 2(1-\nu)/(1-2\nu)$.

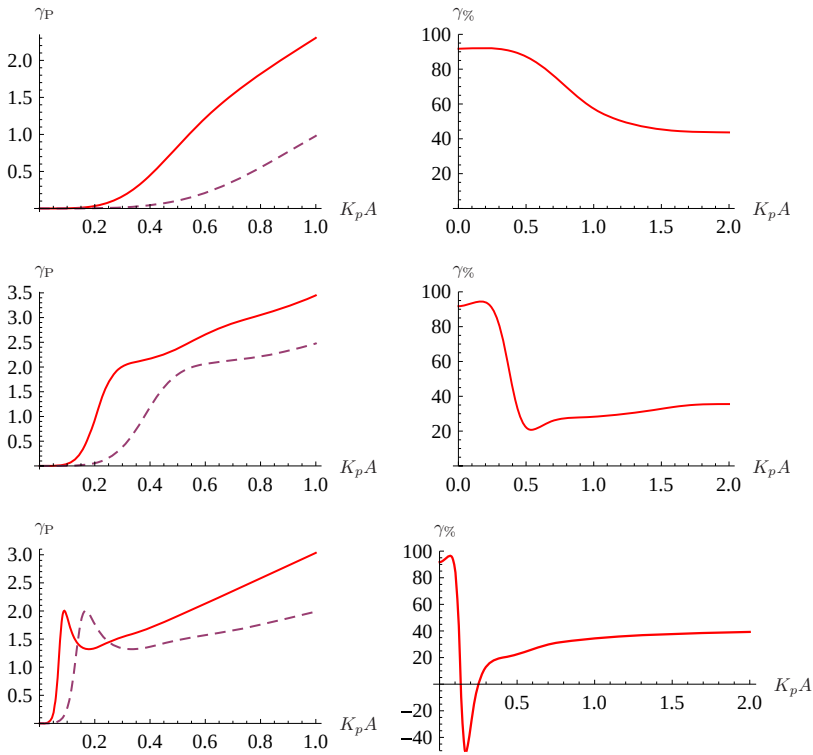


Fig. 12.15. Scattering of P/SV waves from a cylindrical cavity. Left: Scattering cross section from the undeformed cavity (solid) with radius K_pa and from a deformed cavity with initial scaled radius K_pA such that $a = \beta A$ where $\beta \approx 1.863$ (dashed). Right: Percentage *reduction* in scattering cross section $\gamma_{\%}$ due to pre-stress. We have $\nu = 1/3$ (top), $\nu = 7/15$ (middle) and $\nu = 49/99$ (bottom). Note that for the latter case the peak in scattering cross section results in a narrow range of values of K_pa where the cloak *increases* scattering. For other values, there is significant reduction in scattering, especially at very low frequencies.

12.7. Summary and Discussion

In this chapter, we have indicated the close connection between transformation elasticity and small-on-large theory. In particular this means that the latter provides a mechanism for realizing materials that can be employed for cloaking problems, in principle. In particular, for elastodynamic cloaking it has been shown that an alternative

to complex inhomogeneous metamaterials are pre-stressed hyperelastic media. Here we have reviewed this field and provided the associated theoretical basis for the theory of hyperelastic cloaking. A number of results have been presented in the context of SH waves for pre-stressed incompressible hyperelastic media where it has been shown that the neo-Hookean cloak achieves perfect invariance in the sense that the transformation is equal to the finite deformation (that yields the pre-stress). In this case there is no theoretical limit on the deformation, except for the practical ability of the material in question to be deformed, in a stable manner to the required deformation. Clearly the larger the deformation, the more useful the associated "cloak" would be. More realistic strain energy functions have also been described. Although they do not provide perfect invariance as in the neo-Hookean case they appear to perform well (theoretically) in terms of their cloaking potential in pre-stressed states.

In the general compressible case it was shown that the *semilinear* strain energy function of Eq. (12.113) yields the correct incremental moduli required for transformation of isotropic elasticity. As for the incompressible case for SH waves the connection between transformation elastodynamics and the theory of small-on-large is that the transformation equals the finite deformation. The fact that the pre-stress must be in a state of equilibrium places a constraint on the type of transformations allowed. Specifically, they are limited by the condition (12.118), or equivalently, $\mathrm{tr}\boldsymbol{\mathcal{V}}$ =constant, which yields stable radially symmetric pre-strain [Jafari *et al.* (1984)] for hyperelastic cloaking. This implies that the actual size of a cylindrical target can be increased in area by a factor of 4, its radius by factor of two, without any change to the scattering cross section. The restricted form of the transformation is not surprising considering the fact that the theory can simultaneously control more than one wave type, in contrast to acoustics. It should also be noted that the invariance property also requires that $\mathbf{R} = \mathbf{I}$ so that not all deformations of semi-linear materials will have invariant elastodynamics. Having said that one could expect that provided local rotations are small this non-invariant effect would be weak.

In the two-dimensional invariant problems for which results were provided, it was shown that the presence of a conventional cloak generated by the use of pre-stress leads to a significant reduction in the scattering cross section from the cavity, as compared with scattering from a cavity without a cloak. This effect is particularly striking at low frequencies and for small Poisson ratios. We should note that in general one has to consider stability of nonlinear elastic solids in the large deformation regime. While we have not undertaken a full stability analysis, we have noted that the deformation for what we have termed hyperelastic cloaking (HC) is automatically stable (see §12.6.2.3).

This work sheds some light on transformation methods in other wave problems. In acoustics and electromagnetism there is no constraint on the transformation; any one-to-one mapping is permitted. In principle, there is no constraint for transformation elasticity either, although the transformed materials are quite difficult, if not impossible, to obtain, especially since they are required to lose the minor symmetry in their corresponding elastic modulus tensor. The equivalence of transformation elasticity and small-on-large theory provides a unique and potentially realizable solution, although with a limited range of transformations allowed. It would be desirable to relax this constraint, which as already explained, does not appear for the related problem of SH wave motion in *incompressible* hyperelastic solids. The limit of incompressibility offers a clue to a possible resolution for solids with Poisson's ratio close to one-half.

It is also of interest to consider the overlap between the present approach of equating small-on-large with transformation elastodynamics on the one hand and the use of solids with Poisson's ratio close to -1 on the other. Such auxetic materials have been found useful in transformation optics [Shin *et al.* (2014, 2015)] because of their intrinsic property that any deformation is close to conformal, which leads to simplifications in transformation optics [Li and Pendry (2008)], and by analogy, in transformation acoustics [Norris (2012)]. Conformal mappings, by definition, have right and left stretch tensors $\boldsymbol{\mathcal{U}}$ and $\boldsymbol{\mathcal{V}}$ proportional to the identity. This makes them unique

among transformations in that they preserve isotropy in optics and acoustics and are of great interest for this simplification in material properties. Conformal mapping can be achieved to a good degree of approximation by considering the deformation of a solid with Poisson's ratio close to -1. In this limit the ratio of the bulk to shear modulus tends to zero so that volumetric deformation is much easier (less energetic) than shearing. Applications in 2D [Shin *et al.* (2014)] and 3D [Shin *et al.* (2015)] transformation optics have shown that the mechanically defined deformations are close to conformal. Conformal mappings do not provide the same simplification for transformation elastodynamics since it can be easily verified that the transformed moduli still yield non-symmetric stress. However, the kinematic connection between auxeticity (ν close to -1) and transformation optics and acoustics suggests possibly interesting small-on-large effects for elastic strain energy functions with Poisson's ratio close to -1 in the linear regime.

A number of problems are stimulated by the present theory. In particular, taking the *incompressible limit*, $\nu \to 1/2$ of the compressible problem would be of significant interest. We also note that manufacturing nonlinear elastic solids with specific strain energy functions can be difficult to achieve in practice, although this is certainly no harder than generating complex metamaterials, which appears to be the current alternative. The three-dimensional setting for hyperelastic cloaking has not yet been tackled, since it is restricted by the condition of $\mathbf{R} = \mathbf{I}$ for both the initial deformation and subsequent linear elastodynamics. Achieving this would be of interest, but perhaps it is only possible approximately, i.e. small rotations are required. This remains to be seen via further study but could still be useful in the context of "near cloaks". Near-invariance is also of interest for the P/SV problem. If the constraint on the expansion is relaxed, then perhaps materials with strain energy functions can be found that provide near-invariance in the context of larger cavity expansions. Finally it should be noted that the concept of invariance and "near-invariance" offers potential in other

small-on-large problems and could provide interesting results for elastodynamic metamaterials. Such effects have already been illustrated with the concept of elastic wave disentanglement in [Chang *et al.* (2015); Galich and Rudykh (2015)] phononic cloaks in [Barnwell *et al.* (2016)] and deformation-independent band-gaps in [Zhang and Parnell (2017)].

Bibliography

[1] Alpert, B., Greengard, L. and Hagstrom, T. (2000). Rapid evaluation of nonreflecting boundary kernels for time-domain wave propagation, *J. Applied Phys.* **37**, pp. 1138–1164.

[2] Backus, G. (1962). Long-wave elastic anisotropy produced by horizontal layering, *J. Geophys. Research* **67**, pp. 4427–4440.

[3] Barnwell, E. G., Parnell, W. J. and Abrahams, I. D. (2016). Antiplane elastic wave propagation in pre-stressed periodic structures; tuning, band-gap switching and invariance, *Wave Motion* **63**, pp. 98–110.

[4] Beatty, M. F. and Hayes, M. A. (1992). Deformations of an elastic, internally constrained material. Part 1: Homogeneous deformations, *J. Elasticity* **29**, 1, pp. 1–84, doi:10.1007/BF00043445.

[5] Bell, J. F. (1985). Contemporary perspectives in finite strain plasticity, *Int. J. Plasticity* **1**, pp. 3–27.

[6] Bell, J. F. (1989). Experiments on the kinematics of large plastic strain in ordered solids, *Int. J. Solids Struct.* **25**, pp. 267–277.

[7] Berezin, I. and Zhidkov, N. (1965). *Computing Methods*, Vol. 2 (Pergamon).

[8] Brulé, S., Javelaud, E., Enoch, S. and Guenneau, S. (2015). Experiments on seismic metamaterials: Molding surface waves, *Wave Motion* **58**, pp. 53–67.

[9] Brun, M., Guenneau, S. and Movchan, A. B. (2009). Achieving control of in-plane elastic waves, *Appl. Phys. Lett.* **94**, pp. 061903+.

[10] Bückmann, T., Kadic, M., Schittny, R. and Wegener, M. (2015a). Mechanical cloak design by direct lattice transformation, *Proc. Nat. Acad. Sciences* **112**(16), pp. 4930–4934.

[11] Bückmann, T., Kadic, M., Schittny, R. and Wegener, M. (2015b). Mechanical metamaterials with anisotropic and negative effective mass-density tensor made from one constituent material, *Physica Status Solidi* (*b*).

[12] Bückmann, T., Thiel, M., Kadic, M., Schittny, R. and Wegener, M. (2014). An elasto-mechanical unfeelability cloak made of pentamode metamaterials, *Nature Communications* **5**.

[13] Chang, Z., Guo, H.-Y., Li, B. and Feng, X.-Q. (2015). Disentangling longitudinal and shear elastic waves by neo-hookean soft devices, *Appl. Phys. Lett.* **106**(16), p. 161903, doi:10.1063/1.4918787.

[14] Chen, H. and Chan, C. T. (2007). Acoustic cloaking in three dimensions using acoustic metamaterials, *Appl. Phys. Lett.* **91**(18), p. 183518+, doi: 10.1063/1.2803315.

[15] Christensen, J., Kadic, M., Kraft, O. and Wegener, M. (2015). Vibrant times for mechanical metamaterials, *MRS Communications* **5**(03), pp. 453–462.

[16] Colquitt, D., Brun, M., Gei, M., Movchan, A., Movchan, N. and Jones, I. (2014). Transformation elastodynamics and cloaking for flexural waves, *J. Mech. Phys. Solids* **72**, pp. 131–143.

[17] Cosserat, E. and Cosserat, F. (1909). *Théorie des Corps Déformables* (Librairie Scientifique A. Hermann et fils, Paris. ISBN 978-1429704847, reprinted by Cornell University Library).

[18] Cummer, S. A., Popa, B. I., Schurig, D., Smith, D. R., Pendry, J., Rahm, M. and Starr, A. (2008). Scattering theory derivation of a 3D acoustic cloaking shell, *Phys. Rev. Lett.* **100**(2), p. 024301+, doi:10.1103/PhysRevLett.100.024301.

[19] Cummer, S. A. and Schurig, D. (2007). One path to acoustic cloaking, *New J. Phys.* **9**(3), p. 45+, doi:10.1088/1367-2630/9/3/045.

[20] Dill, E. (2007). *Continuum Mechanics: Elasticity, Plasticity, Viscoelasticity* (CRC: Boca Raton, FL).

[21] Eringen, A. and Suhubi, E. (1975). *Elastodynamics, Vol. II* (Academic Press, New York).

[22] Farhat, M., Enoch, S., Guenneau, S. and Movchan, A. B. (2008). Broadband cylindrical acoustic cloak for linear surface waves in a fluid, *Phys. Rev. Lett.* **101**(13), p. 134501+, doi:10.1103/PhysRevLett.101.134501.

[23] Farhat, M., Guenneau, S. and Enoch, S. (2009). Ultrabroadband elastic cloaking in thin plates, *Phys. Rev. Lett.* **103**(2), p. 024301+, doi: 10.1103/physrevlett.103.024301.

[24] Futhazar, G., Parnell, W. J. and Norris, A. N. (2015). Active cloaking for flexural waves in thin plates, *J. Sound. Vib.* **356**, pp. 1–19.

[25] Galich, P. and Rudykh, S. (2015). Comment on disentangling longitudinal and shear elastic waves by neo-Hookean soft devices [*Appl. Phys. Lett.* **106**, 161903 (2015)], *Appl. Phys. Lett.* **107**(5), p. 056101.

[26] Greenleaf, A., Kurylev, Y., Lassas, M. and Uhlmann, G. (2007). Full-wave invisibility of active devices at all frequencies, *Comm. Math. Phys.* **275**(3), pp. 749–789, doi:10.1007/s00220-007-0311-6.

[27] Greenleaf, A., Lassas, M. and Uhlmann, G. (2003). On nonuniqueness for Calderon's inverse problem, *Math. Res. Lett.* **10**, pp. 685–693.

[28] Guevara Vasquez, F., Milton, G. W., Onofrei, D. and Seppecher, P. (2013). Transformation elastodynamics and active exterior acoustic cloaking, in R. V. Craster and S. Guenneau (eds.), *Acoustic Metamaterials*, *Springer Series in Materials Science*, Vol. 166 (Springer), pp. 297–325.

[29] Jafari, A. H., Abeyaratne, R. and Horgan, C. O. (1984). The finite deformation of a pressurized circular tube for a class of compressible materials, *Z. Angew. Math. Phys.* **35**(2), pp. 227–246, doi:10.1007/BF00947935.

[30] John, F. (1960). Plane strain problems for a perfectly elastic material of harmonic type, *Comm. Pure Appl. Math.* **13**(2), pp. 239–296, doi:10.1002/cpa.3160130206.

[31] Kadic, M., Bückmann, T., Schittny, R., Gumbsch, P. and Wegener, M. (2014). Pentamode metamaterials with independently tailored bulk modulus and mass density, *Physical Review Applied* **2**(5), p. 054007.

[32] Kadic, M., Bückmann, T., Schittny, R. and Wegener, M. (2013). Metamaterials beyond electromagnetism, *Reports on Progress in Physics* **76**(12), p. 126501.

[33] Kadic, M., Bückmann, T., Schittny, R. and Wegener, M. (2015). Experiments on cloaking in optics, thermodynamics and mechanics, *Phil. Trans. R. Soc. A* **373**(2049), p. 20140357.

[34] Khlopotin, A., Olsson, P. and Larsson, F. (2015). Transformational cloaking from seismic surface waves by micropolar metamaterials with finite couple stiffness, *Wave Motion* **58**, pp. 53–67.

[35] Lewis, T., Kraft, D. and Hom, N. (1976). Scattering of elastic waves by a cylindrical cavity in a solid, *J. Appl. Phys.* **47**(5), pp. 1795–1798.

[36] Li, J. and Pendry, J. B. (2008). Hiding under the carpet: A new strategy for cloaking, *Phys. Rev. Lett.* **101**(20), p. 203901+.

[37] Lur'e, A. (1968). Theory of elasticity for a semilinear material, *Prikl. Mat. Mekh* (*J. Appl. Math. Mech.*) **32**, pp. 1053–1069 (1068–1085), doi:10.1016/0021-8928(68)90034-8.

[38] Martin, P. A. (2006). *Multiple Scattering: Interaction of Time-harmonic Waves with N Obstacles* (Cambridge University Press, New York).

[39] Milton, G. W., Briane, M. and Willis, J. R. (2006). On cloaking for elasticity and physical equations with a transformation invariant form, *New J. Phys.* **8**, pp. 248–267, doi:10.1088/1367-2630/8/10/248.

[40] Milton, G. W. and Willis, J. R. (2007). On modifications of Newton's second law and linear continuum elastodynamics, *Proc. R. Soc. A* **463**(2079), pp. 855–880, doi:10.1098/rspa.2006.1795.

[41] Mooney, M. (1940). A theory of large elastic deformation, *J. Applied Phys.* **11**, pp. 582–592.

[42] Norris, A. N. (2008). Acoustic cloaking theory, *Proc. R. Soc. A* **464**, pp. 2411–2434, doi:10.1098/rspa.2008.0076.

[43] Norris, A. N. (2009). Acoustic metafluids, *J. Acoust. Soc. Am.* **125**(2), pp. 839–849, doi:10.1121/1.2817359.

[44] Norris, A. N. (2012). Comment on "Design of acoutstic devices with isotropic material via conformal transformation" [Appl. Phys. Lett. 97 044101 (2010)], *Appl. Phys. Lett.* **100**, p. 066101.

[45] Norris, A. N., Amirkulova, F. A. and Parnell, W. J. (2013). Active elastodynamic cloaking, *Math. Mech. Solids* **19**(6), pp. 603–625, doi:10.1177/1081286513479962.

[46] Norris, A. N. and Nagy, A. J. (2010). Acoustic metafluids made from three acoustic fluids, *J. Acoust. Soc. Am.* **128**, 4, pp. 1606–1616, doi:10.1121/1.2817359.

[47] Norris, A. N. and Parnell, W. J. (2012). Hyperelastic cloaking theory: Transformation elasticity with prestressed solids, *Proc. R. Soc. A* **468**, pp. 2881–2903, doi:10.1098/rspa.2012.0123.

[48] Norris, A. N. and Shuvalov, A. L. (2011). Elastic cloaking theory, *Wave Motion* **49**, pp. 525–538, doi:10.1016/j.wavemoti.2011.03.002.

[49] Ogden, R. W. (1997). *Non-Linear Elastic Deformations* (Dover Publications).

[50] Ogden, R. W. (2007). Incremental statics and dynamics of pre-stressed elastic materials, in M. Destrade and G. Saccomandi (eds.), *Waves in Nonlinear Pre-Stressed Materials: CISM International Centre for Mechanical Sciences*, Vol. 495, chap. 1 (Springer), pp. 1–26.

[51] Olsson, P. (2011). Non-uniqueness of time-domain reflection from 3D planar elastic layers, *Wave Motion* **48**, pp. 539–549, doi:10.1016/j.wavemoti.2011.04.009.

[52] Olsson, P. and Wall, D. J. N. (2011). Partial elastodynamic cloaking by means of fiber-reinforced composites, *Inverse Problems* **27**(4), p. 045010+, doi:10.1088/0266-5611/27/4/045010.

[53] O'Neill, J., Selsil, O., McPhedran, R. C., Movchan, A. B. and Movchan, N. V. (2014). Active cloaking of finite defects for flexural waves in elastic plates, *ArXiv e-prints* eprint 1403.0816.

[54] Parnell, W. J. (2012). Nonlinear pre-stress for cloaking from antiplane elastic waves, *Proc. R. Soc. A* **468**, pp. 563–580, doi:10.1098/rspa.2011.0477.

[55] Parnell, W. J. and Abrahams, I. (2011). The effective wavenumber of a pre-stressed nonlinear microvoided composite, *J. Physics: Conference series* **269**, p. 012007.

[56] Parnell, W. J. and Abrahams, I. (2012). Antiplane wave scattering from a cylindrical void in a pre-stressed incompressible neo-Hookean material, *Comm. Comp. Physics* **11**, pp. 367–382.

[57] Parnell, W. J., Norris, A. N. and Shearer, T. (2012). Employing pre-stress to generate finite cloaks for antiplane elastic waves, *Appl. Phys. Lett.* **100**, p. 171907.

[58] Parnell, W. J., Shearer, T. and Norris, A. (2013). Antiplane elastic wave cloaking using metamaterials, homogenization and nonlinear hyperelasticity, *Wave Motion* **50**, pp. 1140–1152.

[59] Pendry, J. B., Schurig, D. and Smith, D. R. (2006). Controlling electromagnetic fields, *Science* **312**(5781), pp. 1780–1782, doi:10.1126/science.1125907.

[60] Raasch, I. (1975). A variational principle for non-linear elastodynamics and its application to the hybrid stress model, *Int. J. Nonlinear Mech.* **10**(5), pp. 215–222, doi:10.1016/0020-7462(75)90013-X.

[61] Rivlin, R. (1948a). Large elastic deformations of isotropic materials 1. Fundamental concepts, *Philos. Trans. A* **240**, pp. 459–508.

[62] Rivlin, R. (1948b). A uniqueness theorem in the theory of highly-elastic materials, *Proc. Cambridge Philos. Soc.* **44**, pp. 595–597.

[63] Sánchez-Dehesa, J. and Torrent, D. (2008). Acoustic cloaking in two dimensions: A feasible approach, *New J. Phys.* **10**, p. 063015.

[64] Schurig, D., Mock, J. J., Justice, B. J., Cummer, S. A., Pendry, J. B., Starr, A. F. and Smith, D. R. (2006). Metamaterial electromagnetic cloak at microwave frequencies, *Science* **314**(5801), pp. 977–980, doi:10.1126/science.1133628.

[65] Sensenig, C. B. (1964). Instability of thick elastic solids, *Comm. Pure Appl. math.* **17**(4), pp. 451–491, doi:10.1002/cpa.3160170406.

[66] Shearer, T., Parnell, W. J. and Abrahams, I. D. (2015). Antiplane wave scattering from a cylindrical cavity in pre-stressed nonlinear elastic media, *Proc. R. Soc. A* **471**(2182), p. 20150450, doi:10.1098/rspa.2015.0450.

[67] Shin, D., Kim, J., Yoo, D.-S., and Kim, K. (2015). Design of 3D isotropic metamaterial device using smart transformation optics, *Optics Express* **23**, 17, p. 21892, doi: 10.1364/oe.23.021892.

[68] Shin, D., Urzhumov, Y., Lim, D., Kim, K., and Smith, D. R. (2014). A versatile smart transformation optics device with auxetic elasto-electromagnetic metamaterials, *Sci. Rep.* **4**, doi:10.1038/srep04084.

[69] Srivastava, A. (2015). Elastic metamaterials and dynamic homogenization: A review, *International Journal of Smart and Nano Materials* **6**(1), pp. 41–60.

[70] Stenger, N., Wilhelm, M. and Wegener, M. (2012). Experiments on elastic cloaking in thin plates, *Phys. Rev. Lett.* **108**(1), p. 014301+, doi:10.1103/physrevlett.108.014301.

[71] Strutt, J. W. (1871). On double refraction, *Philosophical Magazine Series 4* **41**, pp. 519–528.

[72] Treloar, L. (1943). The elasticity of a network of long-chain molecules-ii, *Trans. Faraday Soc.* **39**, pp. 241–246.

[73] Willis, J. R. (1997). *Dynamics of composites*, *Continuum Micromechanics: CISM Lecture Notes*, Vol. 495, chap. 1 (Springer, Wien/New York), pp. 265–290.

[74] Wu, L. and Gao, P. (2015). Manipulation of the propagation of out-of-plane shear waves, *Int. J. Solids Struct.* **69–70**, pp. 383–391, doi:10.1016/j.ijsolstr.2015.05.012.

[75] Zhang, S., Xia, C. and Fang, N. (2011). Broadband acoustic cloak for ultrasound waves, *Phys. Rev. Lett.* **106**, p. 024301+, doi:10.1103/PhysRevLett.106.024301.

[76] Zubov, L. M. (1970). The stationary principle of complementary work in non-linear theory of elasticity, *Prikl. Mat. Mekh* (*J. Appl. Math. Mech.*) **34**, pp. 241–245 (228–232).

[77] Zhang, P. and Parnell, W. J. (2017). Soft phononic crystals with deformation-independent band gaps. *Proc. R. Soc. A.* **473**. doi:10.1098/rspa.2016.0865.

CHAPTER 13

Acoustic Metamaterials with Conical Dispersions

XUEQIN HUANG, MENG XIAO and CHE TING CHAN*

Hong Kong University of Science and Technology,
Hong Kong, P. R. China

FENGMING LIU

China University of Geosciences, P. R. China

13.1. Introduction

Acoustic metamaterial is an artificial material consisting of subwavelength resonant structures designed to control acoustic waves in order to achieve unusual functionalities and to realize novel phenomena such as negative refraction,[1–7] subwavelength imaging[8–16] and invisible cloaking.[17–33] As in the case of electromagnetics, the initial effort on acoustic metamaterial was focused on realizing a negative refractive index, which in the case of acoustics requires effective negative bulk modulus and negative mass density. Recently, zero-index and near zero-index metamaterials have also attracted much attention in electromagnetic wave systems.[34–51]

*Corresponding author: phchan@ust.hk

This class of artificial material has been studied theoretically[34–49] and experimentally.[50–53] They exhibit many intriguing properties such as reflectionless transmission through an arbitrary cross-section waveguide,[34–40,51,52] tailoring wave front,[41–43] and cloaking objects embedded in channels.[44–47,50] The natural question to ask is whether one can also realize zero-index material in acoustic waves. The answer is affirmative. One-dimensional (1D) acoustic transmission line metamaterials with negative/zero/positive index has been reported in the literature.[54] Two-dimensional (2D) acoustic metamaterials have been realized in 2D acoustic crystals.[55] Subsequently, three-dimensional (3D) zero-index material has also been realized.[49]

In this chapter, we will see that effective zero refractive index can be realized in systems with Dirac cone dispersions at the zone center.[48–50,55] Dirac cones describe the linear dispersion relationships between energy and momentum. Such dispersions can be found at the Brillouin zone boundary (K or K' point) in graphene with a honeycomb lattice.[56–69] As a consequence of the conical dispersion, graphene exhibits many interesting transport phenomena such as Klein tunneling,[62] quantum Hall effect[63–65] and Zitterbewegung.[66–69] The existence of conical dispersions at the Brillouin zone boundary can be traced to the lattice symmetry and time reversal symmetry. In fact, independent of the nature of the wave, the band structure of a periodic honeycomb lattice will most likely possess Dirac cones at the K and K' points. Conical dispersions have been found at the Brillouin zone boundary in electromagnetic wave[70–78] and acoustic wave.[79–82] The Dirac spectra are also found in plasmonic honeycomb lattices.[83] Many novel wave transport properties such as Klein tunneling and Zitterbewegung are due to the conical dispersion and the nature of wave, be it electronic, electromagnetic or acoustic, is of secondary importance. For example, while Zitterbewegung is commonly associated with graphene, it had also been observed in acoustic waves systems in simulations and experiments.[80] Conical dispersion at the center of the Brillouin zone is a different story. While lattice symmetry protects the linear dispersion at K and K' points, time reversal symmetry makes the

formation of cones difficult at the zone center. It is well known that time reversal symmetry generally implies quadratic dispersions at $\mathbf{k} = 0$, while conical dispersions need linear dispersions. As will be demonstrated below, conical dispersions at $\mathbf{k} = 0$ can be realized in 2D photonic[48–50] and acoustic[55] systems using accidental degeneracy. The conical dispersions consist of a triply degenerate state with two cones touching at one point and a flat branch intersecting at the same frequency. In the following, we shall call this three-fold degenerate point at the zone center a "Dirac-like point" in order to distinguish this point from the Dirac point in the zone boundary of graphene which has two degrees of freedom. In acoustic waves, we will see that if the triply degenerate state is derived from the monopolar and dipolar excitations, the system can be described as a zero-index metamaterial with effective density and reciprocal of bulk modulus equal to zero simultaneously.[55] This is hence a subtle relation between the zero-index material and conical dispersions.

One interesting implication of Dirac-like cones in 2D acoustic wave system is the existence of interface states. We will see that the interface states can be always found at the boundary of two semi-infinite acoustic crystals (ACs) with system parameters slightly perturbed from the Dirac-like cone formation condition. The assured existence of interface states can be traced to the geometric phase of the bulk bands.[84–86]

While the concept of Dirac point is intrinsically 2D, we will see that using accidental degeneracy, the notion of a Dirac-like point can be generalized from 2D to 3D in acoustic systems.[49] If the accidental degenerate state is derived from monopolar and dipolar acoustic excitations, effective medium theory[87] can be applied to describe the system as an isotropic zero-index material with mass density and reciprocal of bulk modulus equal to zero simultaneously in all directions.

Elastic wave is also a classic wave but with more degrees of freedom and hence the physics is more complicated than acoustic waves. These additional degrees of freedom make the properties of the conical dispersion at $\mathbf{k} = 0$ are rather different from those in acoustic systems. We will see that accidental degeneracy can also

generate conical dispersion at $\mathbf{k} = 0$ in elastic crystals.[88] The subtle relationship between the conical dispersions and zero-index materials is also applicable to elastic wave system, but in this case the triply degenerate state should be derived from the dipolar and quadrupolar excitations.[88] The conical dispersion has the novel consequence that the equi-frequency contours become circular, which is unusual in elastic crystals with a square lattice.[89,90] In addition, the eigen-modes in the conical dispersions exhibit the so-called "super-anisotropy",[91] as they support purely transverse mode in one direction but purely longitudinal mode in another.

We will also show that 3D acoustic meta-crystals consisting of coupled acoustic cavities with properly designed interlayer coupling can exhibit Weyl points in the 3D band structure.[92] Weyl points are 3D conical dispersions and are formed as intersecting points of 3D linear bands.[93–95] Weyl points are more robust than 2D Dirac points. They are topologically protected in the sense that pairs of Weyl points of opposite chirality can only be created or destroyed together[94] and hence destroying Weyl points generally require a significant change of system parameters while a 2D Dirac points can be gapped by either breaking parity or time reversal symmetry.

This chapter is organized as follows. In the second section, we consider the conical dispersions in acoustic crystals and the realization of zero-index materials. We extend the realization of zero-index materials using conical dispersions from 2D to 3D in the third section. In the fourth section, we show that the conical dispersion in phononic crystals can be used to achieve super-anisotropy and we discuss its related transport properties. In the fifth section, we briefly introduce the realization of Weyl points in 3D acoustic metacrystals.

13.2. Conical Dispersions in Two-dimensional Acoustic Crystals

Let us first see how we can obtain conical dispersions at the zone center in ACs.[55] The calculated band structure of a 2D AC consisting of a square lattice of rubber cylinders embedded in water is shown in

Fig. 13.1(a). Here, the lattice constant is a, the radii of the cylinders are $R = 0.224a$. The mass densities of rubber and water are respectively $\rho = 1.3 \times 10^3 kg/m^3$, $\rho_0 = 1.0 \times 10^3 kg/m^3$; and the velocities of waves in rubber and water are $v = 490\, m/s$ and $v_0 = 1490\, m/s$, respectively. For simplicity, we treat rubber as an acoustic medium, and we ignore the consideration of transverse waves inside. From Fig. 13.1(a), we see that there is a triply degenerate state at $\mathbf{k} = 0$ at $f = 0.561 v_0/a$. Two linear bands touch at the degenerate frequency and these two degree of freedoms generate a conical dispersion with two cones touching at a Dirac-like point. There is an additional flat band intersecting at the same frequency. We note that time reversal symmetry requires quadratic dispersion bands at $\mathbf{k} = 0$. So, the typical dispersions are like those shown in Fig. 13.1(c) and Fig. 13.1(d). Linear bands at $\mathbf{k} = 0$, required to form conical dispersions, can be obtained only if we have accidental degeneracy. To see how the accidental degeneracy is related to the general case of quadratic dispersions, we fix the mass densities and velocities of waves of the rubber and water, and vary the radii of the rubber cylinders from $R = 0.224a$. The triply degenerate state at $\mathbf{k} = 0$ shown in Fig. 13.1(a) immediately breaks into a single state and a doubly degenerate state, both showing the expected quadratic dispersion. For a smaller radius $R = 0.2a$, the frequency of the single state at $\mathbf{k} = 0$ is lower than that of the doubly degenerate state as shown in Fig. 13.1(c). While for a larger radius $R = 0.25a$, the frequency of the single state is higher than that of the doubly degenerate state as shown in Fig. 13.1(d). Therefore, there should be an intermediate radius ($R = 0.224a$) where the single and doubly states are at the same frequency at $\mathbf{k} = 0$, which is the result shown in Fig. 13.1(a).

In order to better understand the properties of the triply degenerate states, we plot the pressure distributions of the eigen-modes near the Dirac-like point with a small $\mathbf{k}$ along ΓX direction ($k_x = 0.04\pi/a, k_y = 0$) as shown in Figs. 13.1(e)–13.1(g). Figures 13.1(e) and 13.1(f) show the real and imaginary parts of the pressure distributions of the eigen-mode on the upper Dirac-like cone at

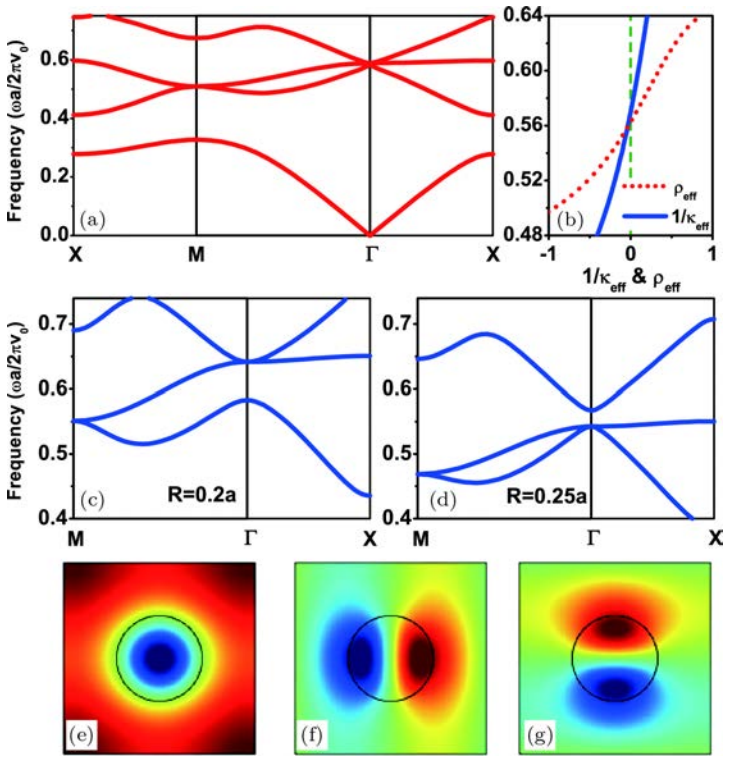


Fig. 13.1. (a) The band structure of a 2D square lattice acoustic crystal consisting of rubber cylinders embedded in water. The lattice constant is a, the radii of the cylinders are $R = 0.224a$. The mass densities of rubber and water are $\rho = 1.3 \times 10^3 kg/m^3$ and $\rho_0 = 1.0 \times 10^3 kg/m^3$, respectively. The velocities of waves in rubber and water are $v = 490m/s$ and $v_0 = 1490m/s$, respectively. (b) The effective mass density ρ_{eff} and reciprocal of bulk modulus $1/\kappa_{eff}$ of the acoustic crystal as a function of frequency. (c)–(d) The band structures of two acoustic crystals with different radii of the rubber cylinders in water. (c) is for $R = 0.2a$, and (d) is for $R = 0.25a$. (e)–(g) The pressure distributions of the eigen-modes near the Dirac-like point with $k_x = 0.04\pi/a$. (e) and (f) represent the real and imaginary parts of the pressure distributions at $f = 0.592v_0/a$ in the upper cone. (g) shows the real part of the pressure distribution of the flat band at $f = 0.589v_0/a$.

$f = 0.592v_0/a$. It is clear from the figure that they are the combinations of the monopolar and dipolar excitations. The dipole is a longitudinal one with displacement parallel to the **k** direction. The results for the lower Dirac-like cone are the same but not shown

here. Figure 13.1(g) shows the real part of the pressure distributions of the eigen-mode at $f = 0.589v_0/a$ on the flat band. The imaginary part is almost zero. We see that this is a quasi-transverse dipole with displacement perpendicular to the **k** direction. Due to its quasi-transverse nature, this mode cannot be excited by an incident longitudinal acoustic plane wave in the effective medium limit and hence this very flat mode is sometimes called a deaf mode.

As the Dirac-like cone is located at $\mathbf{k} = 0$ and derived from the lowest order monopolar and dipolar excitations, we can apply the standard effective medium theory to extract effective medium parameters. In 2D, there is a one-to-one correspondence between one polarization of electromagnetic waves and acoustic waves, i.e. $\rho \leftrightarrow \mu$, $1/\kappa \leftrightarrow \varepsilon$ and hence effective medium theory developed for electromagnetic waves[87] can be adapted to calculate the effective parameters of the AC with a simple change of variables. The expressions for the effective mass density ρ_{eff} and reciprocal of bulk modulus $1/\kappa_{eff}$ are respectively:

$$\frac{\rho_{eff}(\omega) - \rho_0 \frac{J_1(k_0 r_0)}{k_0 r_0 J_1'(k_0 r_0)}}{\rho_{eff}(\omega) - \rho_0 \frac{Y_1(k_0 r_0)}{k_0 r_0 Y_1'(k_0 r_0)}} = \frac{Y_1'(k_0 r_0)}{iJ_1'(k_0 r_0)} \left(\frac{D_1(\omega)}{1 + D_1(\omega)} \right), \tag{13.1}$$

$$\frac{1/\kappa_{eff}(\omega) + \frac{2}{\kappa_0} \frac{J_0'(k_0 r_0)}{k_0 r_0 J_0(k_0 r_0)}}{1/\kappa_{eff}(\omega) + \frac{2}{\kappa_0} \frac{Y_0'(k_0 r_0)}{k_0 r_0 Y_0(k_0 r_0)}} = \frac{Y_0(k_0 r_0)}{iJ_0(k_0 r_0)} \left(\frac{D_0(\omega)}{1 + D_0(\omega)} \right), \tag{13.2}$$

Here, ρ_0, κ_0, k_0 are the mass density, bulk modulus and wave vector of the background material, r_0 is the outer radius in the single-site coherent potential approximation type effective medium theory ($r_0 = a/\sqrt{\pi}$). $J_m(x)$, $Y_m(x)$ and $J_m'(x)$, $Y_m'(x)$ are the Bessel and Neumann functions and their first-order derivatives, respectively. $D_m(\omega)$ are the Mie scattering coefficients. Using Eqs. (13.1) and (13.2), ρ_{eff} and $1/\kappa_{eff}$ as a function of frequency are shown in Fig. 13.1(b). At the Dirac-like point frequency, we found that ρ_{eff} and $1/\kappa_{eff}$ are equal to zero simultaneously.

From the above discussion, we can realize a Dirac-like point at $\mathbf{k} = 0$ using accidental degeneracy and we found that at the Dirac-like point frequency, effective medium theory shows that the

AC should behave as a double-zero-index material at the Dirac-like point frequency with ρ_{eff} and $1/\kappa_{eff}$ equal to zero simultaneously. We note that in electromagnetic wave system, photonic crystals with C_{4v} symmetry can also be tuned so that the frequencies of the single state and the doubly degenerate state are at the same frequency so that a Dirac-like point at $\mathbf{k} = 0$ can be formed and the photonic crystal can also be mapped by effective medium theory to a zero index medium.[48–50] We note however that effective medium theory can be legitimately applied only in the small $\mathbf{k}$ and low frequency regime and for the lowest order (in this case monopole and dipole) excitations.[87] As such, only when the Dirac-like point is derived from the monopolar and dipolar excitations can the system be faithfully described as a zero-index acoustic material by the effective medium theory. A double-zero-index material must have a Dirac-like cone at $\mathbf{k} = 0$, but not all the Dirac-like point at $\mathbf{k} = 0$ (e.g. those formed with higher multipoles) can be described as a zero-index material.[48,55]

Figure 13.2(a) shows the band structure of a 2D AC consisting of steel cylinders with a square lattice embedded in water. Here, the mass density and Lamb constants of the steer cylinders are $\rho_{steel} = 7.9 \times 10^3 kg/m^3$, $\lambda_{steel} = 1.15 \times 10^{11} N/m^2$ and $\mu_{steel} = 8.28 \times 10^{10} N/m^2$. The radius of the steel cylinder is $R = 0.194a$. There is also a triply degenerate state forming a Dirac-like point at $\mathbf{k} = 0$. In this system, the wave velocity of steel is higher than that of waver, and as a consequence, the displacement fields mainly localize in water. This is very different from the case shown in Fig. 13.1(a), in which the displacement fields mainly localize in rubber cylinders with slow wave speed. The modes forming the conical dispersion are not dominated by monopolar and dipolar excitations but have significant projections in higher multipoles. In addition, the Dirac-like point is also high in frequency. It is not expected that the effective medium theory can be applied to describe the system. If one still insists to use the effective medium formulas to obtain ρ_{eff} and $1/\kappa_{eff}$ as a function of frequency as shown in Fig. 13.2(b), we will see that both ρ_{eff} and $1/\kappa_{eff}$ are not equal to zero at the Dirac-like point frequency. That is to say, this system cannot be mapped to a zero-index material at

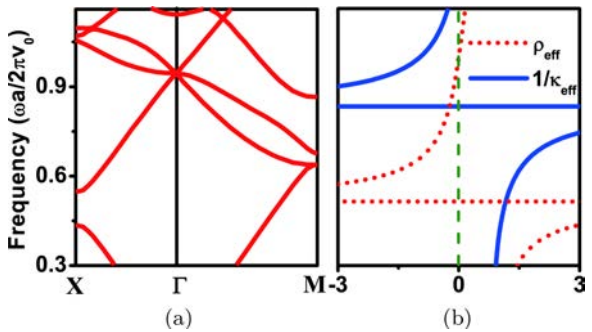


Fig. 13.2. (a) The band structure of a 2D acoustic crystal consisting of steel cylinders embedded in water with a square lattice. The radius of the steel cylinder is $R = 0.194a$. (b) The effective mass density ρ_{eff} and reciprocal of bulk modulus $1/\kappa_{eff}$ of the acoustic crystal as a function of frequency.

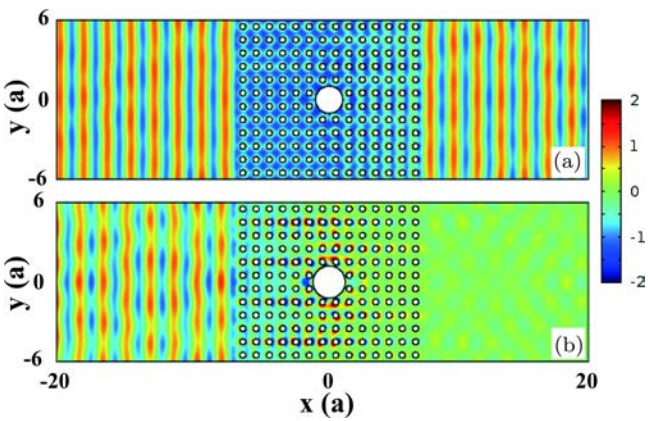


Fig. 13.3. (a) The simulated pressure field distributions for a 2D acoustic crystal with a silicone rubber defect (a), and a PMMA defect (b).

the Dirac-point frequency. We also note that the standard effective medium description is good only near the zone center and is expected to fail at large k-vectors. For example, Dirac-like points can also be found at the M point periodic systems with C_{4v} symmetry but the physics near those Dirac-like point cannot be faithfully described by a zero-index medium.[48]

It has been shown in the literature that an electromagnetic zero index medium can "cloak" an object inside a waveguide.[44–47,49,50] We have to clarify here that the cloaking effect inside a zero index channel is different from the invisibility cloak enabled by transformation optics in which the form invariance of the Maxwell equation under a coordinate transformation establishes a correspondence between the material constitutive parameters and coordinate transformation.[96–102] Although transformation optics was developed for electromagnetic waves, the same idea can be applied to acoustics to design acoustic invisibility cloaks.[17–33] As those transformation optics based cloaks are not directly related to conical dispersions,[33] we will not discuss them in detail here. We will focus on discussing the channel cloaking effect enabled by the existence of Dirac-like cones, which is demonstrated numerically in Fig. 13.3. In Fig. 13.3(a), a silicon rubber defect is embedded in an AC with parameters shown in Fig. 13.1(a). A plane wave illuminates from the left, and we see that there is almost no reflection and the exit wave on the right side preserves its plane wave front. The defect is apparently cloaked by the effective zero-index AC. Apart from cloaking objects embedded in it, zero-index materials can also realize total reflection dependent on the parameters of the object embedded in it. Figure 13.3(b) shows that a PMMA defect is embedded in it, the appearance of total reflection is observed. For those ACs with Dirac-like cones that can be mapped to a zero-index material, ρ_{eff} and $1/\kappa_{eff}$ are equal to zero simultaneous, there is no phase change when waves propagate through the material. Such zero phase change has many interesting implications including tunneling through an arbitrary waveguide and tailoring the wave front.[55]

In the above discussion, we revealed the relationship between the Dirac-like point physics and zero-index material. Zero-refractive-index is related to a triply-degenerate state at $\mathbf{k} = 0$ forming a Dirac-like point with two linear dispersions and an additional flat band intersecting at the same frequency. The existence of the flat band is consistent with the constitutive parameters approaching zero at some particular frequency. At first glance, the quasi-transverse flat band

with a zero group velocity (in homogeneous material limit) cannot be excited by incident plane waves and hence its existence has no role in the Dirac-like point physics. This types of modes are sometimes called "deaf modes" and are usually ignored in the wave propagation. However, its existence does have some subtle consequences. For instance, if this flat band is ignored, the dispersion near the Dirac-like point can be described by a 2×2 Hamiltonian which can be mapped to that of a spinor, and such systems can potentially carry a non-zero Berry phase.[103,104] On the other hand, the dispersion near the triply degenerate state requires a 3×3 matrix to describe and is similar to a (pseudo) spin-1 system[103,104] and the Berry phase should be zero. Using the multiple scattering theory, we found that the Berry phase of the Dirac-like cone is indeed equal to zero.[48] The Berry phase can have qualitative implications when we consider effects such as disorder induced localization/anti-localization.[56] Therefore, the existence of the quasi-transverse state cannot be ignored when we consider the Dirac-like point.

One implication of the Dirac-like cone at $\mathbf{k} = 0$ and in particular the existence of the flat band is the existence of interface states at the boundary between two semi-infinite ACs. The schematic picture of an interface along the y-direction is shown in Fig. 13.4(a). The mass densities, wave velocities and radii of the cylinders in the left and right ACs are ρ_1, v_1, R_1 and ρ_2, v_2, R_2, respectively. The mass density and wave velocity of the background material are ρ_0, v_0. For simplicity, we consider a system in which both the left and right ACs consist of rubber cylinders embedded in water but with different radii. The mass densities and velocities of rubber are $\rho_1 = \rho_2 = 1.3 \times 10^3 kg/m^3$ and $v_1 = v_2 = 490\, m/s$. The radii of rubber cylinders in the left and right ACs are $R_1 = 0.2a$, $R_2 = 0.25a$, respectively. The bulk band structures of the left and right ACs are shown in Figs. 13.1(c) and 13.1(d), respectively. The corresponding projected band structures are shown in Fig. 13.4(b). The blue color denotes the projected bands for the left AC and the red color marks those for the right AC. There are three common band gaps in the projected band structures marked with white colors. A common band gap is a

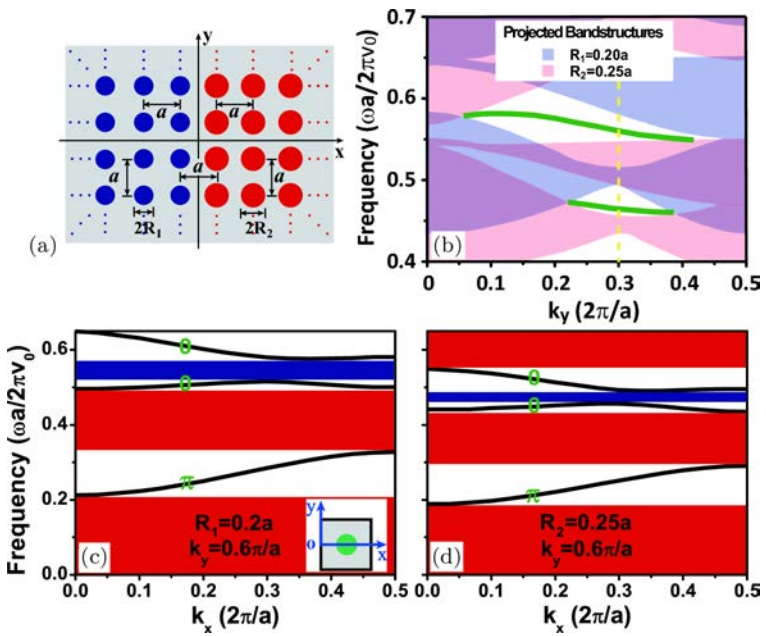


Fig. 13.4. (a) Schematic picture of an interface along the y-direction separating two semi-infinite 2D acoustic crystals with a square lattice of cylinders embedded in water. The mass densities, wave velocities and radii of the cylinders in the left and right acoustic crystals are ρ_1, v_1, R_1 and ρ_2, v_2, R_2, respectively. The background is water with ρ_0, v_0. The lattice constants for both acoustic crystals are a. (b) The projected band structures of these two acoustic crystals along the k_y direction with blue color for the left acoustic crystal and red color for the right acoustic crystal. Here, $\rho_1 = \rho_2 = 1.3 \times 10^3 kg/m^3$, $v_1 = v_2 = 490\, m/s$, $\rho_0 = 1.0 \times 10^3 kg/m^3$, $v_0 = 1490\, m/s$, $R_1 = 0.2a$, $R_2 = 0.25a$. The green lines represent the band dispersions of the interface states. The bulk band structures with a fixed $k_y = 0.6\pi/a$ for the acoustic crystals with (c) $R_1 = 0.2a$ and (d) $R_2 = 0.25a$. The inset in (c) shows the coordinate for calculating the Zak phase. The origin is located on the left boundary of the unit cell. The Zak phases of the bulk bands are labeled with green colors. The characters of the band gaps are labelled by the sign of $\mathrm{Im}(Z)$ with red color for $\mathrm{Im}(Z) < 0$ and blue for $\mathrm{Im}(Z) > 0$.

necessary condition for the existence of interface states but it is not a sufficient condition. In these three common gaps, there are only two branches of interface states shown as the green lines in Fig. 13.4(b). One point we want to emphasize is that the interface states of the

ACs are robust in the sense that they always exist, no matter how small the size of the common gap.

It is well known that the interface state formation condition at the interface between two homogeneous materials can be formulated analytically. If the effective medium theory can be applied, the existence of interface states in discrete systems can also be explained by the same formalism. However, the interface states shown in Fig. 13.4(b) extend to large **k** where conventional effective medium theory is invalid. We note that the condition for the formation of an interface state can be expressed as the surface impedance condition: $Z_L(\omega, k_y) + Z_R(\omega, k_y) = 0$,[105] where $Z_{L(R)}(\omega, k_y)$ is the surface impedance of the semi-infinite AC on the left (right) side for a given k_y. We can obtain the surface impedance of the AC using a layer-by-layer multiple scattering theory (MST) which considers the 2D AC as a stack of 1D ACs. The mathematical details can be found in the literature[106] and will not be given here. It can be shown that the surface impedance is uniquely defined as long as inter-layer scattering is dominated by the zero-order scattering. In a lossless material, $Z(\omega, k_y)$ is purely imaginary for frequencies inside a band gap, and $\mathrm{Im}(Z(\omega, k_y)) > 0$ or $\mathrm{Im}(Z(\omega, k_y)) < 0$. For a given k_y, $\mathrm{Im}(Z(\omega, k_y))$ decreases monotonically from $+\infty$ to 0 with increasing frequency in a region with $\mathrm{Im}(Z(\omega, k_y)) > 0$, while $\mathrm{Im}(Z(\omega, k_y))$ decreases monotonically from 0 to $-\infty$ with increasing frequency in a region with $\mathrm{Im}(Z(\omega, k_y)) < 0$. Therefore, knowing the surface impedances and their monotonic behavior for frequencies inside a band gap, the interface state formation condition shows that there must exist one and only one interface state inside the common gap if $\mathrm{Im}(Z_L(\omega, k_y))$ and $\mathrm{Im}(Z_R(\omega, k_y))$ have different signs and interface states cannot exist if $\mathrm{Im}(Z_L(\omega, k_y))$ and $\mathrm{Im}(Z_R(\omega, k_y))$ have the same sign.

In the following, we will apply a bulk-interface correspondence relationship that relates the surface impedance to bulk band geometric phases to interpret the existence of interface states.[84–86] It is known that the guaranteed interface states are frequently related to the topological properties of the bulk bands.[107–126] To give a "geometric" interpretation of the formation of interface states, we

unfold the projected band structures to reduced 1D bulk bands. For example, the projected bands in Fig. 13.4(b) come from the reduced 1D bulk bands with k_x varying from $-\pi/a$ to π/a for a fixed k_y. The reduced 1D bands of two ACs at $k_y = 0.6\pi/a$ are plotted in Figs. 13.4(c) and 13.4(d). The geometric Zak phases[84–86,127] of the reduced 1D bands can be calculated using the formula:

$$\varphi_n = i \int \langle u_{nk_x,k_y} | \frac{1}{\kappa(\vec{r})} \partial_{k_x} | u_{nk_x,k_y} \rangle dk_x, \tag{13.3}$$

where $u_{n\vec{k}}$ is the cell periodic part of the Bloch function of the pressure field for the n^{th} band at a particular $\vec{k}$, $\kappa(\vec{r})$ is position dependent bulk modulus. The geometric Zak phase is calculated using the periodic gauge and the origin is chosen at the left boundary of the unit cell as shown in the inset of Fig. 13.4(c). The Zak phases for the three lowest bands are $\pi, 0, 0$, respectively. In the 1D system with inversion symmetry, the bulk-interface correspondence relates the ratio of the signs of $\mathrm{Im}(Z(\omega, k_y))$ in two adjacent gaps, say the n^{th} and $(n-1)^{th}$ gaps, to the Zak phase of the band sandwiched in between,[84–86] and the relationship can be written as:

$$\frac{\mathrm{Sgn}[\mathrm{Im}(Z_n(\omega, k_y))]}{\mathrm{Sgn}[\mathrm{Im}(Z_{n-1}(\omega, k_y))]} = e^{i(\varphi_{n-1}+\pi)}. \tag{13.4}$$

It can be proved that $\mathrm{Im}(Z(\omega, k_y)) < 0$ in the lowest gap. Equation (13.4) then gives the signs of $\mathrm{Im}(Z(\omega, k_y))$ for the band gaps at higher frequencies. In Figs. 13.4(c) and 13.4(d), the red color denotes the gaps with $\mathrm{Im}(Z(\omega, k_y)) < 0$ and blue marks band gaps with $\mathrm{Im}(Z(\omega, k_y)) > 0$. These results are consistent with that calculated by MST (not show here). Thus, using the band structure information of one unit cell, we can calculate the Zak phase of the bulk bands and then determine the sign of $\mathrm{Im}(Z(\omega, k_y))$ without doing any further calculations. In Figs. 13.4(c) and 13.4(d), there are two overlapping gaps that $\mathrm{Im}(Z(\omega, k_y))$ have different signs. The overlap of the second gap in Fig. 13.4(c) and the third gap in Fig. 13.4(d) gives rise to the lower branch of the interface states near the frequency $0.45v_0/a$ and the overlap of the third gap in Fig. 13.4(c) and the fourth gap in Fig. 13.4(d) gives rise to the higher frequency branch near frequency

$0.53v_0/a$. In the system shown in Fig. 13.4, we find two branches of interface states. We emphasize that the higher branch of interface states is guaranteed independent of the system parameters. The assured existence of interface states between the frequencies of the two flat bands of the ACs is shown in the following. Firstly, the quasi-transverse flat band supplies two band gaps, one above and one below its projected band structure. Secondly, the zero Zak phase of the flat band (required by symmetry) of the system dictates that the gaps above and below the flat band have opposite signs of $\mathrm{Im}(Z(\omega, k_y))$. The common gap is above the projected band of the flat band of one AC, while it is below that of the other AC. Together with the monotonic behaviors of $\mathrm{Im}(Z(\omega, k_y))$ inside the band gap, the AC on the left side must have an opposite sign of $\mathrm{Im}(Z(\omega, k_y))$ to the AC on the right side in the common gap. As a consequence, one interface state must exist in this gap. We note again that this behavior is specific to a Dirac-like dispersion at $\mathbf{k} = 0$.

13.3. Extension of Two-dimensional Conical Dispersions to Three Dimension

In this section, we will extend the concept of a Dirac-like point from 2D to 3D.[49] The aim is to realize an isotropic zero-index acoustic metamaterial. For electromagnetic waves, it is known that the 3D Dirac-like point at $\mathbf{k} = 0$ can be realized in a simple cubic photonic crystal through the accidental degeneracy of the electric and magnetic dipoles resulting in a six-fold degenerate state. For acoustic systems, we can use the accidental degeneracy of monopole and dipole as suggested in Ref. 128. Since there is no monopole in 3D in electromagnetic waves, the scheme proposed in Ref. 128 cannot be straightforwardly realized in a photonic system, but can be implemented in 3D acoustic systems. Using the accidental degeneracy of the A_{1g} and T_{1u} modes, a 3D Dirac-like point at $\mathbf{k} = 0$ with a four-fold degenerate state can be realized in a simple cubic AC. The bulk band structure of a 3D AC consisting of rubber spheres in water is shown in Fig. 13.5(a). The mass density and radii of rubber spheres are $\rho = 1.3 \times 10^3$ kg/m^3, $R = 0.255a$, respectively. a is the

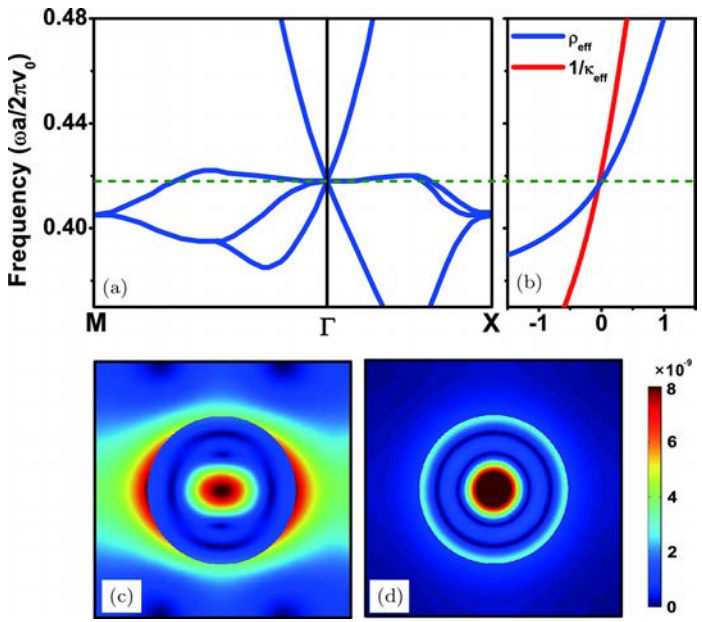


Fig. 13.5. (a) The band structure of a 3D simple cubic acoustic crystal consisting of rubber spheres in water. The mass density and radii of rubber spheres are $\rho = 1.3 \times 10^3 kg/m^3$, $R = 0.255a$. Here, a is the lattice constant. The mass density of water is $\rho_0 = 1.0 \times 10^3 kg/m^3$. The Lame constants of rubber and water are $\kappa = 1.17 \times 10^8 N/m^2$, $\kappa_0 = 2.22 \times 10^9 N/m^2$. (b) The effective mass density (ρ_{eff}) and reciprocal of bulk modulus ($1/\kappa_{eff}$) as a function of frequency obtained by effective medium theory. The horizontal green dash line marks the Dirac-like point frequency ($f_D = 0.418 v_0/a$) in the band structure which coincides with the frequency at which $\rho_{eff} = 1/\kappa_{eff} = 0$. The displacement field distributions of the eigen-modes at the Dirac-like point frequency are shown in (c) and (d).

lattice constant. The mass density of water is $\rho_0 = 1.0 \times 10^3$ kg/m^3. The Lame constants of rubber and water are $\kappa = 1.17 \times 10^8$ N/m^2, $\kappa_0 = 2.22 \times 10^9$ N/m^2. For simplicity, we ignore the shear wave within the rubber spheres due to the high velocity contrast between the rubber and water, and the main features will be the same if we also include the shear wave within the spheres.[4] Figure 13.5(a) shows a four-fold degenerate state at $\mathbf{k} = 0$ at frequency $f_D = 0.418 v_0/a$ consisting of two linear bands and another two quadratic bands intersecting at the same frequency. Here, v_0 is the acoustic velocity in water.

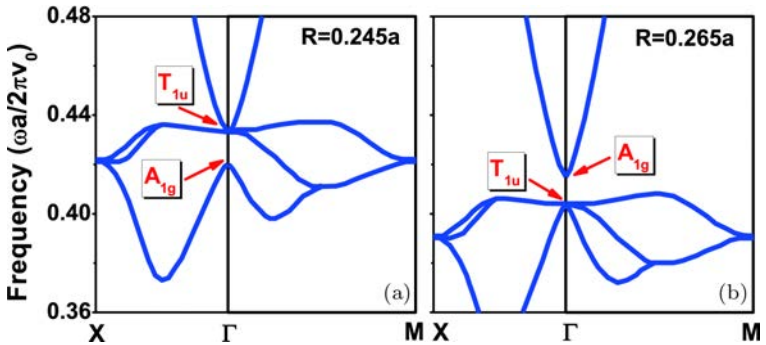


Fig. 13.6. The band structure of 3D acoustic crystals in a simple cubic lattice for different radii. $R = 0.245a$ is for (a), and $R = 0.265a$ for (b). Here, a is the lattice constant. The parameters of rubber and water are the same as Fig. 13.5.

The linear bands generate equi-frequency surfaces that are spheres with radii proportional to $(\omega - \omega_D)$. To see that the degeneracy is accidental, we change the radii of rubber spheres from $R = 0.255a$. The four-fold degenerate state immediately breaks into a single state and a triply degenerate state. The band structures for different radii are shown in Fig. 13.6. For smaller radius ($R = 0.245a$), the frequency of the single state (A_{1g} state) is lower than that of the triply degenerate state (T_{1u} state). For a larger radius ($R = 0.265a$), the frequency of the single state (A_{1g} state) is higher than that of the triply degenerate T_{1u} state. The inversion of the single and triply degenerate state as a function of sphere radius imply that there exists a certain radius at which these two states have the same frequency. We plot the displacement fields of the eigen-modes at the Dirac-like point frequency in Figs. 13.5(c) and 13.5(d). They represent the monopolar and dipolar excitations, respectively. More technical details concerning the existence of the 3D Dirac-like point can be found in Ref. 49.

Using effective medium theory,[87] we can calculate the effective mass density (ρ_{eff}) and reciprocal of bulk modulus ($1/\kappa_{eff}$) of this AC as a function of frequency and the results are shown in Fig. 13.5(b). Both ρ_{eff} and $1/\kappa_{eff}$ are equal to zero at the Dirac-like point frequency. We note here that the zero-index enabled by the 3D Dirac-like point is isotropic, while the 2D systems are inherently anisotropic

with the Dirac-like point giving a zero-index only for in-plane transport of acoustic waves.

13.4. Conical Dispersions in Two-dimensional Elastic Wave System

Dirac-like points have been found in 2D and 3D acoustic waves. Can the concept of Dirac-like point at $\mathbf{k} = 0$ be extended to elastic waves and if so, what are the consequences? To answer these questions, we consider a 2D phononic crystal (PC) consisting of rubber cylinders in a square lattice embedded in epoxy. The radius and mass density of rubber cylinder are respectively $R = 0.266a$, $\rho = 1.3 \times 10^3$ kg/m^3. The mass density of epoxy is $\rho_0 = 1.18 \times 10^3$ kg/m^3. The velocities of longitudinal and transverse waves in rubber (epoxy) are $v_l = 817$ m/s ($v_{l0} = 2605$ m/s) and $v_t = 335$ m/s ($v_{t0} = 1068$ m/s), respectively. The band structure of the PC for the in-plane modes is shown in Fig. 13.7(a). A triply degenerate state forms a Dirac-like point at $\mathbf{k} = 0$ at the frequency $f_D = 0.721 v_{t0}/a$. Similar to the case of acoustic waves, the elastic wave Dirac-like cone also has two linear bands and an additional flat band intersecting at the same frequency.[88] To better visualize the Dirac-like cone, we plot the 3D band structure near the Dirac-like point frequency in Fig. 13.7(b). The equi-frequency surfaces have circular contours at frequencies slightly above the Dirac-like point as shown in Fig. 13.7(c). At a first glance, this Dirac-like point is the same as that in the acoustic system but there are actually interesting implications unique to elastic waves. For example, the equi-frequency contours are generally anisotropic in a PC with a square lattice even in the low frequency limit.[89,90] Therefore, the circular isotropic equi-frequency contour is a specific property of the Dirac-like cone elastic wave system.

Applying multiple scattering theory, we can obtain the formation condition of the accidental degeneracy Dirac-like point. The band structure of a PC is determined by the secular equation:

$$\det \left| \begin{pmatrix} T^{ll} G^{l} & T^{lt} G^{t} \\ T^{tl} G^{l} & T^{tt} G^{t} \end{pmatrix} - I \right| = 0, \tag{13.5}$$

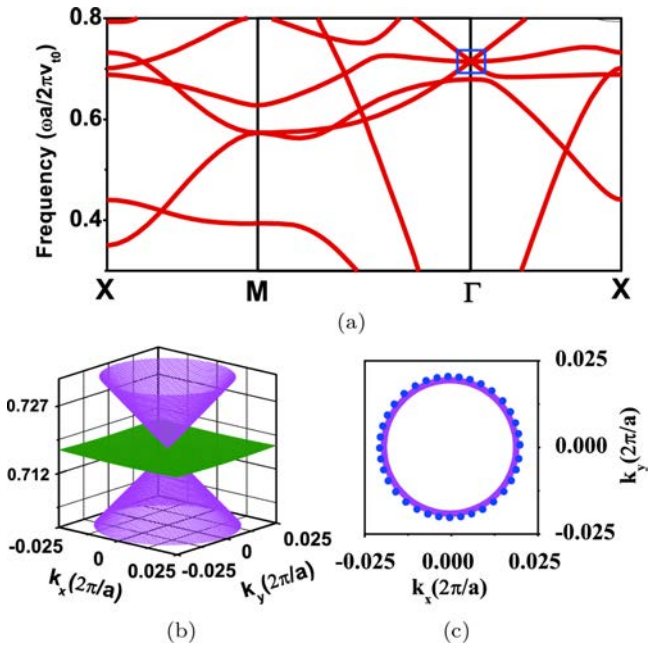


Fig. 13.7. (a) The band structure of a phononic crystal in a square lattice for the in-plane modes. A triply degenerate state at $\mathbf{k} = 0$ consists of two linear bands touching at the Dirac-like point frequency $f = 0.721 v_{t0}/a$ and a flat band intersecting at the same frequency. v_{t0} is the transverse wave velocity of epoxy. (b) Three-dimensional band structure near the Dirac-like point frequency. (c) The equi-frequency surfaces of the phononic crystal calculated by full-wave simulation (COMSOL, magenta circle) and by effective medium theory (blue dot) at frequency ($f = 0.73 v_{t0}/a$) which is above the Dirac-like point.

where $T^{\alpha\beta}$ is the T-matrix with matrix elements $T^{\alpha\beta}_{mm'} = D^{\alpha\beta}_m \delta_{mm'}$ and $D^{\alpha\beta}_m$ are the elastic Mie scattering coefficients of angular momentum m of the scatter. The indices $\alpha, \beta = l$ or t denote the longitudinal and transverse waves, respectively. The matrices $G^l$and $G^t$are given by the lattice sums $G^{\beta}_{m'm"} = \sum_{q \neq p} g^{\beta}_{m'm"} e^{i\vec{K}\cdot\vec{R}_q} = S(\beta, m' - m")$ with $(\beta = l, t)$. Explicit expressions of $D^{\alpha\beta}_m$ and $S(\beta, m)$ can be found in the literature.[129]

The eigen-modes near the Dirac-like point are derived from the dipolar ($m = \pm 1$) and quadrupolar ($m = \pm 2$) excitations. Hence, we only consider the dipolar and quadrupolar modes. In a square

lattice at $\mathbf{k} = 0$, the lattice symmetry requires that $S(\beta, 0) \neq 0$, $S(\beta, \pm 1) = S(\beta, \pm 2) = S(\beta, \pm 3) = 0$and $S(\beta, \pm 4) \neq 0$.[90] Thus, Eq. (13.5) reduces to:

$$\begin{vmatrix} D_2^{ll}S(l,0)-1 & 0 & 0 & D_2^{ll}S(l,4) \\ 0 & D_1^{ll}S(l,0)-1 & 0 & 0 \\ 0 & 0 & D_{-1}^{ll}S(l,0)-1 & 0 \\ D_{-2}^{ll}S(l,-4) & 0 & 0 & D_{-2}^{ll}S(l,0)-1 \\ D_2^{tl}S(l,0) & 0 & 0 & D_2^{tl}S(l,4) \\ 0 & D_1^{tl}S(l,0) & 0 & 0 \\ 0 & 0 & D_{-1}^{tl}S(l,0) & 0 \\ D_{-2}^{tl}S(l,-4) & 0 & 0 & D_{-2}^{tl}S(l,0) \end{vmatrix}$$

$$\begin{vmatrix} D_2^{lt}S(t,0) & 0 & 0 & D_2^{lt}S(t,4) \\ 0 & D_1^{lt}S(t,0) & 0 & 0 \\ 0 & 0 & D_{-1}^{lt}S(t,0) & 0 \\ D_{-2}^{ll}S(t,-4) & 0 & 0 & D_{-2}^{lt}S(t,0) \\ D_2^{tt}S(t,0)-1 & 0 & 0 & D_2^{tt}S(t,4) \\ 0 & D_1^{tt}S(t,0)-1 & 0 & 0 \\ 0 & 0 & D_{-1}^{tt}S(t,0)-1 & 0 \\ D_{-2}^{tt}S(t,-4) & 0 & 0 & D_{-2}^{tt}S(t,0)-1 \end{vmatrix} = 0. \tag{13.6}$$

Then, Eq. (13.6) can be rewritten as:

$$\begin{vmatrix} D_2^{ll}S(l,0)-1 & D_2^{lt}S(t,0) & D_2^{lt}S(t,4) & D_2^{ll}S(l,4) \\ D_2^{tl}S(l,0) & D_2^{tt}S(t,0)-1 & D_2^{tt}S(t,4) & D_2^{tl}S(l,4) \\ D_{-2}^{tl}S(l,-4) & D_{-2}^{tt}S(t,-4) & D_{-2}^{tt}S(t,0)-1 & D_{-2}^{tl}S(l,0) \\ D_{-2}^{ll}S(l,-4) & D_{-2}^{lt}S(t,-4) & D_{-2}^{lt}S(t,0) & D_{-2}^{ll}S(l,0)-1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{vmatrix}$$

$$\begin{vmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ D_1^{ll}S(l,0)-1 & D_1^{lt}S(t,0) & 0 & 0 \\ D_1^{tl}S(l,0) & D_1^{tt}S(t,0)-1 & 0 & 0 \\ 0 & 0 & D_{-1}^{ll}S(l,0)-1 & D_{-1}^{lt}S(t,0) \\ 0 & 0 & D_{-1}^{tl}S(l,0) & D_{-1}^{tt}S(t,0)-1 \end{vmatrix} = 0. \tag{13.7}$$

The cylindrical symmetry of the cylinders ensures that $D_1^{ll} = D_{-1}^{ll}, D_1^{tt} = D_{-1}^{tt}, D_1^{lt} = D_{-1}^{lt}$ and $D_1^{tl} = D_{-1}^{tl}$. In Eq. (13.7), the lower two block matrices satisfy

$$\begin{vmatrix} D_1^{ll}S(l,0) - 1 & D_1^{lt}S(t,0) \\ D_1^{tl}S(l,0) & D_1^{tt}S(t,0) - 1 \end{vmatrix} = \begin{vmatrix} D_{-1}^{ll}S(l,0) - 1 & D_{-1}^{lt}S(t,0) \\ D_{-1}^{tl}S(l,0) & D_{-1}^{tt}S(t,0) - 1 \end{vmatrix} = 0. \tag{13.8}$$

This shows that the dipolar modes are doubly degeneracy at the frequency ω_d. The upper block matrix in Eq. (13.7) satisfies:

$$\begin{vmatrix} D_2^{ll}S(l,0) - 1 & D_2^{lt}S(t,0) & D_2^{lt}S(t,4) & D_2^{ll}S(l,4) \\ D_2^{tl}S(l,0) & D_2^{tt}S(t,0) - 1 & D_2^{tt}S(t,4) & D_2^{tl}S(l,4) \\ D_{-2}^{tl}S(l,-4) & D_{-2}^{tt}S(t,-4) & D_{-2}^{tt}S(t,0) - 1 & D_{-2}^{tl}S(l,0) \\ D_{-2}^{ll}S(l,-4) & D_{-2}^{lt}S(t,-4) & D_{-2}^{lt}S(t,0) & D_{-2}^{ll}S(l,0) - 1 \end{vmatrix} = 0, \tag{13.9}$$

As $S(l, \pm 4) \neq 0$ and $S(t, \pm 4) \neq 0$, the quadrupolar modes at $\mathbf{k} = 0$ will interact with each other and split into two nondegenerate modes at different frequencies ω_{q1} and ω_{q2}. In general, there is no requirement for the quadrupolar modes and the dipolar modes at the same frequency, i.e. $\omega_{q1} \neq \omega_d$ or $\omega_{q2} \neq \omega_d$. However, through tuning the system parameters (such as the radius of the cylinder), we can achieve accidental degeneracy by making one of the quadrupolar modes and the dipolar modes to have the same frequency ($\omega_{q1} = \omega_d$ or $\omega_{q2} = \omega_d$). The triply degenerate state at $\mathbf{k} = 0$ shown in Fig. 13.7(a) is realized in this manner.

Figure 13.7(c) shows that the equi-frequency contours are circular near the elastic wave Dirac-like point. The dispersions are hence "super-isotropic"[91] as elastic wave equi-frequency contours are generally not isotropic for a square lattice.[89,90] However, we will see that the eigenvectors exhibit the so-called "super-anisotropic" behaviors. Figure 13.8 shows the displacement field distributions of the eigen-modes near the Dirac-like point with a small $\mathbf{k}$ along the ΓX ($k_x = 0.04\pi/a, k_y = 0$) and ΓM ($k_x = k_y = 0.04\pi/a$) directions.

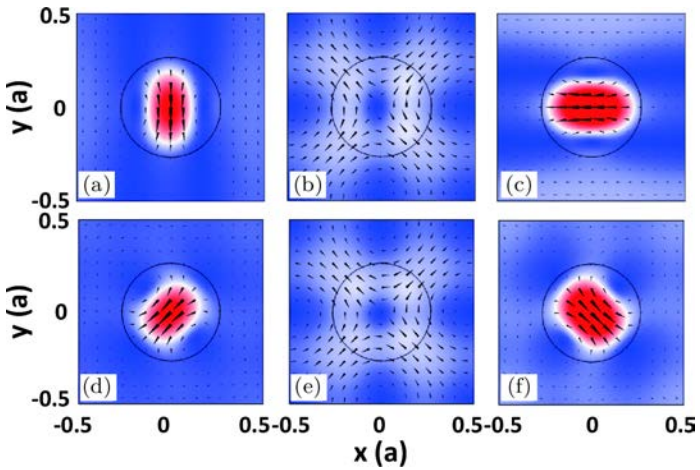


Fig. 13.8. The displacement field distributions of the eigen-modes near the Dirac-like point with a small **k** along the ΓX (a–c) and ΓM directions (d–f). (a) the real part and (b) the imaginary part of the displacement fields of the linear band at $f = 0.724v_{t0}/a$; (c) the real part of the displacement fields of the flat band at $f = 0.721v_{t0}/a$. (d) the real part and (e) the imaginary part of the displacement fields of the linear band at $f = 0.728v_{t0}/a$; (f) the real part of the displacement fields of the flat band at frequency $f = 0.721v_{t0}/a$. The arrows denote the directions of the displacements.

The real and imaginary parts of the displacement fields of the linear band state along ΓX direction on the higher Dirac-like cone (at $f = 0.724v_{t0}/a$) are shown in Figs. 13.8(a) and 13.8(b). We see that the eigen-mode is a linear combination of dipolar and quadrupolar excitations. The dipole displacement is perpendicular to **k** and the quadrupolar mode couples only to transversely polarized incident waves. The linear band is hence a band of transverse modes along ΓX direction. The eigen-mode for the lower Dirac-like cone is also a linear combination of a quadrupole and a transverse dipole (not shown here). The real part of the eigen-mode of the flat band is plotted in Fig. 13.8(c) and the imaginary part is almost zero. It is a dipolar mode with displacement parallel to **k**. Thus, the flat band is a longitudinal band near the Dirac-like point along ΓX direction.

In contrast, the eigen-modes along ΓM direction is entirely different. Figures 13.8(d) and 13.8(e) show the real and imaginary parts

of the displacement fields of the eigen-mode on the higher cone at $f = 0.728v_{t0}/a$. The eigen-mode along ΓM direction is also a combination of quadrupolar and dipolar excitations. The dipolar mode displacement is parallel to **k** and the quadrupolar mode can only be excited by longitudinal waves. This implies that the linear band is a longitudinal band along ΓM direction. The eigen-mode for the lower Dirac-like cone is similar to the higher one and not shown here. The real part of the eigen-mode of the flat band is plotted in Fig. 13.8(f). It is a dipolar mode with displacement perpendicular to **k**. The nature of the flat band is also different to that along ΓX direction. The Dirac-like cone modes are hence "super-anisotropic", in the sense that they are purely transverse in one direction but purely longitudinal in another. We note that "super-anisotropic" elastic wave metamaterials have been realized using a rather complex structural design[91] and the Dirac-like cone can offer a much simpler alternative.[88]

As the wave velocity of the rubber is lower than that of epoxy, the displacement fields are mainly localized in the rubber. Effective medium theory can be applied to such systems to describe the physics near the Dirac-like point. It is known that dipolar and quadrupolar resonances are associated with mass density and certain components of elastic moduli, respectively.[7,129] Here, we employ the boundary effective medium theory[91] to obtain the effective moduli C_{11}^{eff}, C_{12}^{eff} and C_{44}^{eff} of the PC as shown in Fig. 13.9(a). It is well known that both C_{11} and C_{12} are positive for natural materials. The strains S_{xx} and S_{yy} compress or extend simultaneously under external stress. Effective medium theory found that C_{11}^{eff} and C_{12}^{eff} have opposite signs near the Dirac-like point, i.e. $C_{11}^{eff} \approx -C_{12}^{eff}$. This is because of the quadrupolar mode, which induce the displacement compression in one direction and simultaneous extension in the orthogonal direction. Moreover, C_{11}^{eff},C_{12}^{eff} and C_{44}^{eff} show divergence behavior near the Dirac-like point frequency due to the quadrupolar resonance. In Fig. 13.9(b), ρ_{eff} is plotted as a function of frequency, showing that ρ_{eff} is equal to zero at the Dirac-like point frequency. Knowing the effective parameters of the PC, we can use the Christoffel's

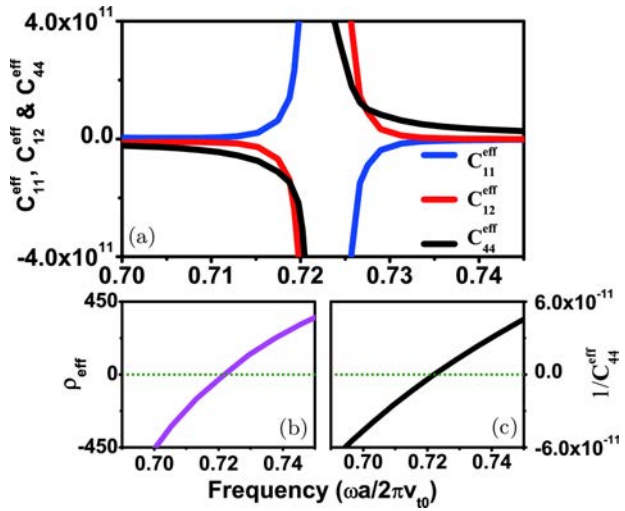


Fig. 13.9. (a) The effective moduli C_{11}^{eff}, C_{12}^{eff} and C_{44}^{eff}; (b) the effective mass density ρ_{eff}; and (c) the reciprocal of effective modulus, $1/C_{44}^{eff}$, of the PC as a function of frequency near the Dirac-like point.

equation to obtain the phase velocities v_1, v_2 of the two in-plane modes.

The Christoffel's equation is:

$$\rho_{eff} v^2 u_i^0 = \Gamma_{il} u_l^0, \tag{13.10}$$

where $\Gamma_{il} = c_{ijkl} n_j n_k$ is the Christoffel tensor, u_i^0 is the direction of the displacement, and n_i denotes the propagation direction. In 2D PC with a square lattice, Γ_{il} has the form:

$$\Gamma_{il} = \begin{vmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{12} & \Gamma_{22} \end{vmatrix}, \tag{13.11}$$

where $\Gamma_{11} = C_{11}^{eff} \cos^2\varphi + C_{44}^{eff} \sin^2\varphi$, $\Gamma_{12} = (C_{12}^{eff} + C_{44}^{eff}) \sin\varphi \cos\varphi$ and $\Gamma_{22} = C_{11}^{eff} \sin^2\varphi + C_{44}^{eff} \cos^2\varphi$. Here, φ denotes the angle between the propagation direction and the x axis. Solving Eq. (13.10), we obtain two eigenvalues $\gamma_m = \rho_{eff} v_m^2 (m = 1, 2)$ as functions of φ,

$$2\rho_{eff} v_1^2 = C_{11}^{eff} + C_{44}^{eff}$$

$$+\sqrt{\left(C_{11}^{eff}-C_{44}^{eff}\right)^2\cos^2 2\varphi+\left(C_{12}^{eff}+C_{44}^{eff}\right)^2\sin^2 2\varphi},$$

$$2\rho_{eff}v_2^2=C_{11}^{eff}+C_{44}^{eff}$$

$$-\sqrt{\left(C_{11}^{eff}-C_{44}^{eff}\right)^2\cos^2 2\varphi+\left(C_{12}^{eff}+C_{44}^{eff}\right)^2\sin^2 2\varphi}. \tag{13.12}$$

Above the Dirac-like point frequency, for $C_{11}^{eff}<0$, $C_{44}^{eff}>0$ and $C_{12}^{eff}\approx -C_{11}^{eff}$, the eigenvalues γ_1 and γ_2 reduce to $\gamma_1=\rho_{eff}v_1^2=C_{44}^{eff}$ and $\gamma_2=\rho_{eff}v_2^2=C_{11}^{eff}$ for any propagation direction φ. For $\rho_{eff}>0$, $v_1=\sqrt{C_{44}^{eff}/\rho_{eff}}$ is real and $v_2=\sqrt{C_{11}^{eff}/\rho_{eff}}$ is imaginary. Thus, only the branch with eigenvalue γ_1 is allowed for the upper cone.

Along the ΓX direction with $\varphi=0°$, $\Gamma_{11}=C_{11}^{eff}$, $\Gamma_{22}=C_{44}^{eff}$ and $\Gamma_{12}=0$, through Eq. (13.10), the eigenvectors of the eigenvalues γ_1 satisfy the following expressions:

$$\begin{aligned}(C_{11}^{eff}-C_{44}^{eff})u_{1x}^0=0\\(C_{44}^{eff}-C_{44}^{eff})u_{1y}^0=0\end{aligned}. \tag{13.13}$$

It can be seen that $u_{1x}^0=0$ and u_{1y}^0 is arbitrary. As the polarization of plane wave is defined by the angle θ with $\tan\theta=u_{1y}^0/u_{1x}^0$, the plane wave along the ΓX direction is a transverse wave with $\theta=90°$.

Along the ΓM direction with $\varphi=45°$, $\Gamma_{11}=(C_{11}^{eff}+C_{44}^{eff})/2$, $\Gamma_{22}=(C_{11}^{eff}+C_{44}^{eff})/2$ and $\Gamma_{12}=(C_{44}^{eff}-C_{11}^{eff})/2$, the eigenvectors of the eigenvalues γ_1 satisfy:

$$\begin{aligned}\frac{(C_{11}^{eff}-C_{44}^{eff})}{2}u_{1x}^0+\frac{(C_{44}^{eff}-C_{11}^{eff})}{2}u_{1y}^0=0\\\frac{(C_{44}^{eff}-C_{11}^{eff})}{2}u_{1x}^0+\frac{(C_{11}^{eff}-C_{44}^{eff})}{2}u_{1y}^0=0\end{aligned}. \tag{13.14}$$

Thus, the plane wave is a longitudinal wave with $\theta=45°$.

The results for frequencies below the Dirac-like point are similar to those above the Dirac-like point frequency. At the Dirac-like point frequency, both the eigenvalues γ_1 and γ_2 are allowed simultaneously, the

super-anisotropy vanishes. Considering all these cases above, near the Dirac-like point frequency, the velocities of the in-plane modes of the PCs only depend on the modulus C_{44}^{eff} for any propagation direction φ meaning that the equi-frequency surface is "circular". Hence, the isotropy of the equi-frequency contours and the "super-anisotropy" of the eigen-modes near the Dirac-like point frequency can be demonstrated analytically using the effective medium parameters.

Figure 13.9(c) shows the reciprocal of effective modulus $1/C_{44}^{eff}$ as a function of frequency which shows that $1/C_{44}^{eff}$ also approaches zero at the Dirac-like point frequency. Using ρ_{eff} and C_{44}^{eff}, the equi-frequency surface near the Dirac-like point can be given by $k = \omega\sqrt{\rho_{eff}/C_{44}^{eff}}$ as shown in Fig. 13.7(c) (blue dot). The result is consistent with that calculated by the full-wave simulation (magenta line shown in Fig. 13.7(c)), which indicates that the effective medium theory is applicable near the Dirac-like point. As ρ_{eff} and $1/C_{44}^{eff}$ are proportional to $\Delta\omega = \omega - \omega_D$ when approaching ω_D, where ω_D is the Dirac-like point frequency, it is easily seen that k is also proportional to $\Delta\omega$ implying that the dispersion relations are linear.

The effective medium theory relates the Dirac-like cone elastic wave system to an effective zero-index material with ρ_{eff} and $1/C_{44}^{eff}$ equal to zero simultaneously. It is known that homogeneous zero-index elastic wave materials can exhibit some interesting wave transport phenomena. We want to see whether our discrete system has the similar phenomena as the homogeneous case. In the above discussion, we know that the Dirac-like cone is super-anisotropic, thus it should support different transporting properties for differently polarized incident waves with different incident directions. Figures 13.10(a)–13.10(d) shows the incident plane wave along the ΓX direction. The transverse waves have total transmission independent of whether there is a defect (a steel bar) embedded in it or not as shown in Figs. 13.10(c) and 13.10(a). On the other hand, for longitudinal incident waves, total reflections appear as shown in Figs. 13.10(d) and 13.10(b) for the same configuration. This is because of the Dirac-like cone only supports transverse wave along ΓX direction. By contrast, the Dirac-like cone supports only longitudinal waves along the ΓM

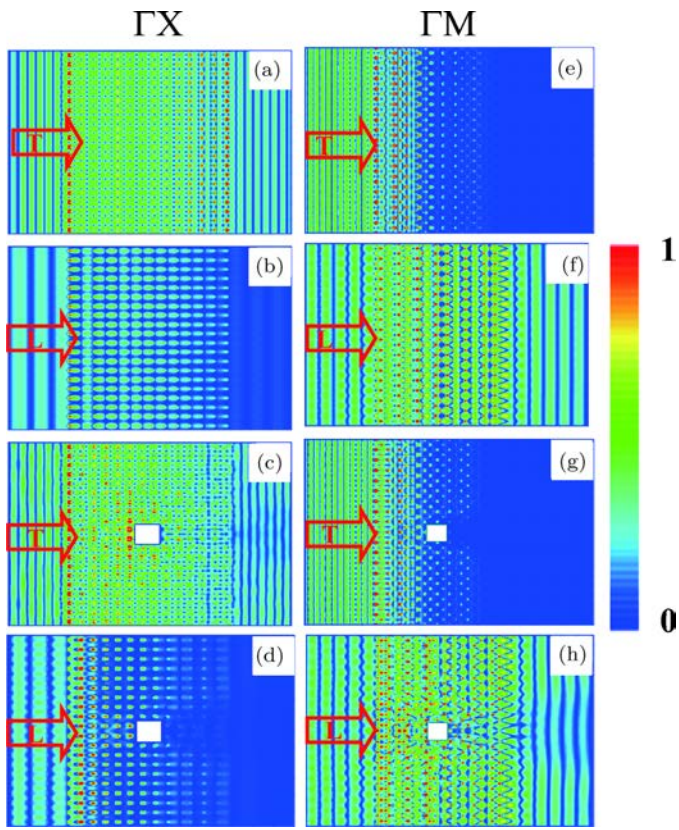


Fig. 13.10. Different polarized incident plane waves impinging the phononic crystal slab along ΓX (a)–(d) and ΓM (e)–(h) directions with and without imbedded steel defects. Along ΓX direction, (a) and (c) show that transverse waves can go through the slab without or with steel defect embedded in it, respectively; (b) and (d) show that longitudinal waves are totally reflected without or with steel defect embedded in the PC slab, respectively. By contrast, along ΓM direction, (e) and (g) shows that transverse waves are totally reflected; (f) and (h) show that longitudinal waves can pass through. The working frequency is slightly above the Dirac-like point at $f = 0.745v_{t0}/a$. T and L denote transverse and longitudinal incident waves, respectively.

direction. As a consequence, the transverse incident plane waves are rejected as shown in Figs. 13.10(e) and 13.10(g) and the longitudinal incident waves can pass through the slab without distortion as shown in Figs. 13.10(f) and 13.10(h). The simulations shown in

Figs. 13.10(c, d, g, h) give the impression that the steel defects are cloaked by the PC slab similar to the results in acoustic and electromagnetic waves, but the physics is more complex here due to the super-anisotropic character of the elastic wave Dirac-like cone. We note here that in order to avoid the influence of the flat band, we select a working frequency slightly above the Dirac-like point frequency at $f = 0.745 v_{t0}/a$.

13.5. Weyl Point in Acoustic Systems

In this section, we will first use a simple model to illustrate how Weyl points[130] can emerge in acoustic systems which possess linear dispersions along all the directions in the 3D moment space. The dispersion around the Weyl points can be approximated by the Weyl Hamiltonian[93–95]

$$H(k) = \sum k_i v_{ij} \sigma_j, \tag{13.15}$$

where $i, j \in \{x, y, z\}$, k_i is the wave vector around the Weyl points and σ_j is the Pauli matrix. Weyl points can be regarded as monopoles of Berry flux, and we can define the charges or chirality of these monopoles by $c = sgn[\det(v_{ij})] = \pm 1$. As an intuitive picture, we can envision that Berry flux emerges from the Weyl point that possesses a positive charge, and goes into the Weyl point with negative charge just like the flux coming out of the a positive monopole charge and ending with the negative monopole. When a Weyl point with positive charge encounters a Weyl point with negative charge, they will annihilate with each other. Except from the pair annihilation which typically requires strong perturbation, the Weyl points are otherwise stable and this explains why the Weyl points are considered as "topologically stable".

Although both Weyl points[130] and accidental degeneracy induced Dirac points[48–51,88] process linear dispersions around the degeneracy points along all directions, there are many differences between them.

(1) Time reversal symmetry (TRS) ensures that a Weyl point at wavevector $\mathbf{k}$ must have a Weyl point companion at $-\mathbf{k}$ with

the *same* charge.[93,94] Meanwhile the total charge of all Weyl points inside the Brillouin zone must be zero.[93,94] Hence, TRS requires that the minimum number of Weyl point should be 4. On the other hand, inversion symmetry ensures that a Weyl point at one $\mathbf{k}$ point must have a Weyl point companion at $-\mathbf{k}$ with the *different* charge.[93,94] Because the chirality required by the TRS and inversion symmetry is opposite, there must be no Weyl point in a system with both TRS and inversion symmetry. Hence, we must break at least one of these two symmetries to obtain Weyl points. However, both TRS and inversion symmetry can be kept when we construct accidental degeneracy induced Dirac-like points at the zone center. TRS and inversion symmetry also co-exist for Dirac points at zone boundaries.

(2) Weyl Hamiltonian involves only two linear dispersive bands, while besides linear dispersive band, we must have additional quadratic bands for accidental degeneracy induced Dirac-like points at the zone center. The minimum number of additional flat bands is 1 for 2D[50,131] and 2 for 3D.[49]

(3) Weyl points are stable in the parameter space and can only be annihilated with another Weyl point with opposite charge. On the contrary, we must introduce degeneracy between modes with appropriate symmetries to construct accidental degeneracy induced Dirac-like points.[131] Any parameter variation which lifts this accidental degeneracy will eliminate the Dirac-like points.

As we have discussed above, one must break either TRS or inversion symmetry to generate Weyl points. The TRS is difficult to break for acoustic systems which requires dynamic modulation[113] or moving fluid.[121,122] Here we only consider the effect of inversion symmetry breaking. To construct linear dispersions along all the directions in 3D momentum space, we start with linear dispersion in 2D momentum space. The simplest Hamiltonian which processes this property is

$$H(q_x, q_y) = q_x\sigma_x + q_y\sigma_y, \tag{13.16}$$

where $\mathbf{q}$ represents the wave vector with the origin at the degeneracy point. The physical system which can be described by this Hamiltonian is graphene using simple tight binding model with the nearest neighbor approximation. To construct a Weyl point, we still need to introduce an additional term $q_z\sigma_z$. In this case, this term means we need to introduce an odd function (relative to q_z) of the onsite energy difference along the third direction. This happens when the onsite energy on different sublattice sites as functions of k_z have some linear crossing points. Onsite energy difference can be introduced by adding a constant energy gating difference or interlayer coupling difference between difference sublattice sites. Here we consider the latter case and we work with an AA stacked honeycomb lattice. Other kinds of interactions which break the inversion symmetry may also introduce similar effects which we will not discuss here.

A unit cell of our structure is shown in Fig. 13.11(a). We use different colored bonds to represent different coupling coefficients and coupling is non-zero only between connected sites. The Hamiltonian which describes this nearest-neighbor tight-binding model can be written as

$$\begin{aligned} H = & \sum_{i,k} \varepsilon (a_{i,k} a_{i,k}^{\dagger} - b_{i,k} b_{i,k}^{\dagger}) \\ & + \sum_{\langle a_{i,k}, b_{j,k} \rangle} (t_n a_{i,k} b_{j,k}^{\dagger} + \text{H.c.}) \\ & + \sum_{i,k} (t_a a_{i,k} a_{i,k+1}^{\dagger} + t_b b_{i,k} b_{i,k+1}^{\dagger} + \text{H.c.}), \end{aligned} \tag{13.17}$$

where t_n, t_a and t_b are hopping coefficients and are all real, a (b) and $a^{\dagger}$ ($b^{\dagger}$) are the annihilation and creation operators on the sublattice site, respectively. Each lattice is specified by subscripts (i, k), where the first labels the position in each layer and the second labels the number of layers. For completeness, we also include the onsite energy difference ε. The second term is identical to graphene described by the nearest neighbor approximation where $\langle \cdot \rangle$ means hopping is only

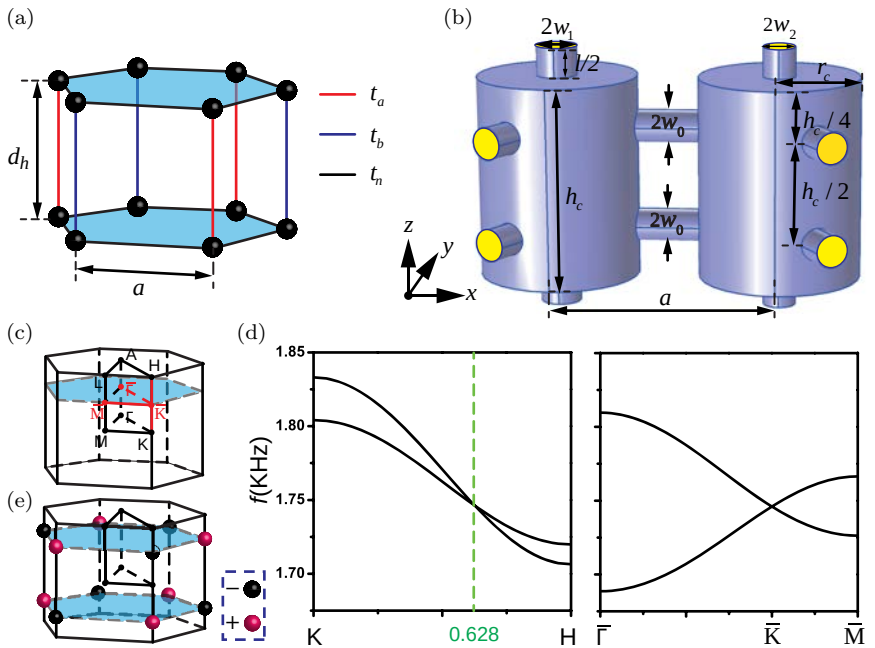


Fig. 13.11. Weyl points can emerge if interlayer coupling coefficients are different in a stacked lattice. (a) A unit cell of an AA stacked honeycomb lattice. Coupling is only nonzero between connected sites and different colored bonds represent different coupling coefficients. (b) A unit cell of the acoustic metacrystal under study. Light blue and yellow represent respectively the area where hard boundary and Bloch periodic boundary condition apply. (c) The reciprocal space with red lines marking the position where band dispersions are calculated in (d). (d) The band dispersion of the acoustic metacrystal under study, where the parameters used were $r_c = 4\text{cm}$, $a = 11\text{cm}$, $w_0 = w_2 = 0.8\text{cm}$, $w_1 = 1\text{cm}$, $h_c = 10\text{cm}$ and $l = 3\text{cm}$. The systems were filled with air (density $\rho = 1.3kg/m^3$, and the speed of sound $v = 343m/s$). A Weyl point is located at the $\bar{\text{K}}$ point with $k_z = 0.628\pi/(h_c + l)$. (e) Weyl points in the reciprocal space, where black and red represent Weyl points with negative and positive charges, respectively.

between nearest neighbor on the same layer. The third term represents interlayer hopping.

The corresponding Bloch Hamiltonian $H(\boldsymbol{k})$ is obtained using Fourier transform as

$$H(k) = \begin{pmatrix} \varepsilon + 2t_a \cos(k_z d_h) & t_n\beta \\ (t_n\beta)^* & -\varepsilon + 2t_b \cos(k_z d_h) \end{pmatrix}, \tag{13.18}$$

where $\beta = \exp(-ik_x a) + 2\cos(\sqrt{3}k_y a/2)\exp(ik_x a/2)$, a is the distance between the two sublattices and d_h is the interlayer distance. When $|t_a - t_b| > \varepsilon$, there exists some special values $k_z = \pm\arccos[\varepsilon/(t_b - t_a)]/d_h$ where the diagonal term are the same. Around these special k_z values, the difference between diagonal terms is an odd function. As β is independent of k_z and also has linear dispersion inside the $k_x - k_y$ plane along the KH direction (see Fig. 13.11(c)), the band dispersion around those points on the KH line and with $k_z = \pm\arccos[\varepsilon/(t_b - t_a)]/d_h$ should hence be linear along all the directions. Those points in the momentum space are Weyl points according to the definition of Eq. (13.15).

Now let us consider how to realize the above tight-binding model using real acoustic systems. In Fig. 13.11(b), we show a unit cell of the real acoustic metacrystal, where light blue and yellow show respectively the area where hard boundary and periodic boundary condition is applied. We use sound resonance cavities connected by small tubes. Such a coupled acoustic resonator metacrystal can be described quite well by the tight-binding model. The resonance cavities filled with air can be viewed as "meta-atoms", the hopping between atoms is realized with small connection tubes and the hopping strength can be tuned with the radius of the connection tubes. When the radius of the connection tube is not too large, the hopping strength is roughly proportional to the cross-sectional area of the connection tubes. Hence if we choose $w_1 \neq w_2$, the physics can be described by the Hamiltonian in Eq. (13.17). Indeed, as we calculated the band dispersion along the KH line as shown in Fig. 13.11(d), we found a linear crossing point at $k_z = 0.628\pi/(h_c + l)$. The crossing point is not at the middle point of KH which means $\varepsilon \neq 0$ if we map our acoustic system to a tight binding model. We note that when $\varepsilon = 0$, the crossing point should be at exactly $k_z = 0.5\pi/(h_c + l)$. This is because the connection tubes not only introduce coupling between resonance cavities but will also change the boundary condition of the resonance cavities. The change of the boundary condition from hard boundary to free boundary shifts the resonance frequency, which is equivalent to an onsite energy term in the Hamiltonian.

The right panel of Fig. 13.11(d) shows the band dispersion in the xy plane at the linear crossing point. It is clear that the dispersions are also linear along all the remaining directions. The metacrystal hence exhibits Weyl points in the momentum space.

As we have discussed above, Weyl points have associated topological charges. In Fig. 13.11(e), we show the signs of charge distribution in the reciprocal space, where red and black spheres represent positive and negative charges, respectively. As TRS is maintained in our system, Weyl points at $\mathbf{k}$ and $-\mathbf{k}$ have the same sign of charge. There are two additional mirror planes besides the TRS, i.e., the xz plane and the xy plane. Mirror symmetry relates two $\mathbf{k}$ vectors with only one component opposite, and hence the charges of Weyl points at those two $\mathbf{k}$ points should also be different as can be seen in Fig. 13.11(e).

We have given one example of constructing Weyl points in the acoustic metacrystals. Actually, there is another interesting view point of understanding this. As the unit cell in Fig. 13.11(a) is periodic along the z direction, the wave vector along the z direction, k_z, is a good quantum number. For each fixed k_z, the system is reduced to an effective 2D honeycomb lattice. Now the different interlayer coupling at different sublattice sites corresponds to different k_z-dependent onsite energy, which means that the effective in plane inversion symmetry is now broken by introducing the interlayer coupling difference. At these special k_z points where two band cross with each other, the effective in plane inversion symmetry is recovered, i.e., $H(x, y, k_z) = H(-x, -y, k_z)$. We know that breaking of inversion symmetry in honeycomb lattice lifts the degeneracy at the $\bar{K}$ points and the band gap opens. The gap opening through inversion symmetry breaking is a topologically trivial band gap,[132] i.e., the Chern numbers of the lower and upper band is zero (Here we only have two bands for this tight-binding model). This can also be seen from the distribution of Weyl points in Fig. 13.11(e). The Weyl points are located on the plane with $k_z = \pm 0.628\pi/(h_c + l)$ and the net charge on each plane vanishes. Weyl points are the sources or sinks of Berry flux, and hence the total flux through an arbitrary

2D band with fixed k_z vanishes. Chern number of a 2D band is proportional to the total Berry flux through it and hence it is also zero.

To have nontrivial topology for reduced 2D bands with fixed k_z, one has to break the effective time reversal symmetry for the reduced 2D system. While it is difficult, if not impossible, to break TRS in acoustic systems, we can introduce effective TRS breaking in the reduced 2D system with a fixed k_z. To achieve this goal, from the point view of Weyl points, we need to relocate the position of Weyl points such that the positive charges and negative charges are located at different k_z planes. We can then find some k_z values that the net Berry flux through a 2D band with these k_z does not vanish. Indeed, this goal can be realized if we introduce chiral interlayer coupling which breaks all the possible mirror symmetries with mirror plane parallel to the z-axis.[92]

Figure 13.12(a) shows an example of chiral coupling and we assume the hopping strength is real and denoted by t_c. The corresponding Bloch Hamiltonian is

$$H(k) = \begin{pmatrix} t_c f(k_z d_h) & t_n \beta \\ (t_n \beta)^* & t_c f(-k_z d_h) \end{pmatrix}, \tag{13.19}$$

where $f(k_z d_h) = 2\cos(\sqrt{3}k_y a - k_z d_h) + 4\cos(3k_x a/2)\cos(\sqrt{3}k_y a/2 + k_z d_h)$. Now for a fixed k_z, the interlayer coupling becomes next nearest neighbor coupling with a complex hoping coefficient $t_c e^{i\phi}$ with $\phi = \pm k_z d_h$, where the sign of ϕ depends on the hopping direction. On the right panel of Fig. 13.12(a), we use red arrows to denote the direction along which $\phi = -k_z d_h$. After a complete loop along the direction indicated by the red arrows, the total phase accumulated is $-3k_z d_h$. According to the rule of Peierls substitution,[133,134] we know there is nonzero gauge flux $(-3k_z d_h)$ enclosed by this loop. In the right panel of Fig. 13.12(a), we use dotted and crossed circles to denote the direction of local flux. The local flux distribution shows that our metacrystal at a fixed k_z is an acoustic analogue of the topological Haldane model.[132] The existence of gauge flux can also be understood in terms of the Weyl points. In Fig. 13.12(b), we show the distribution of the Weyl points and their associated charges.

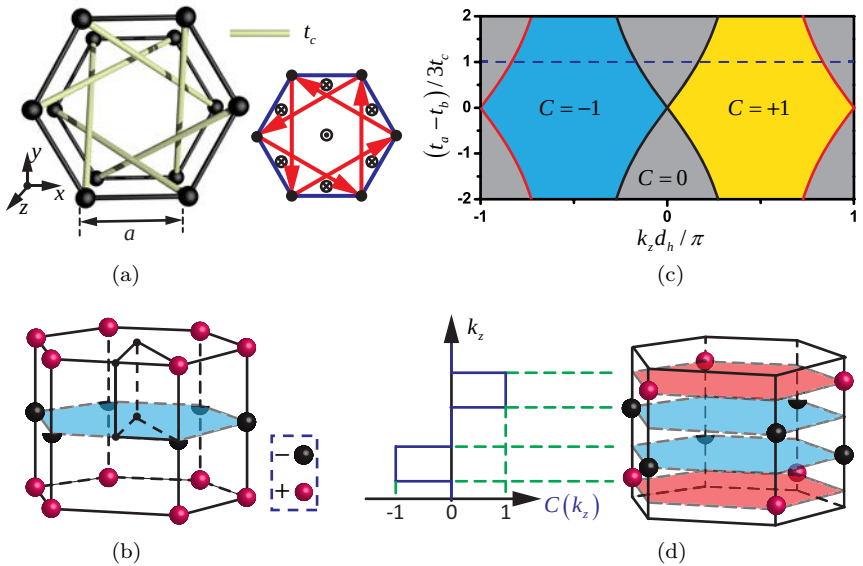


Fig. 13.12. (a) Left panel shows a unit cell of an AA stacked honeycomb lattice with chiral coupling. The chiral hopping is realized with apple-green bonds with strength denoted by t_c. Right panel shows that this chiral coupling is an analogue of Haldane model for a fixed nonzero k_z, where dotted and crossed circle represents the local gauge flux and the red arrows represent the hopping direction with positive phase. (b) The distribution of Weyl points with this chiral kind coupling and the signs of their associated charge. (c) Phase diagram of the parameter space when considering both the different onsite couplings at different sublattice sites and the chiral coupling, where cyan, gray and yellow represent region with Chern numbers -1, 0 and $+1$ for the lower band, respectively. Red/black curves show the trajectories of the Weyl points with positive/negative charges. (d) Right panel shows the distribution of Weyl points with the parameters marked by the blue dashed line in (c). Left panel shows the corresponding value of Chern numbers of the lower band as a function of k_z.

Weyl points with positive charges and negative charges are located at $k_z = \pm\pi/d_h$ and $k_z = 0$ plane, respectively. As the Weyl points are sources or sinks of Berry flux, the total flux through a 2D band with an arbitrary k_z is nonzero.

Next, we consider the effect of incorporating both the different onsite couplings at different sublattice sites and the chiral coupling,

and the Hamiltonian becomes

$$H(k) = \begin{pmatrix} 2t_a \cos(k_z d_h) + t_c f(k_z d_h) & t_n \beta \\ (t_n \beta)^* & 2t_b \cos(k_z d_h) + t_c f(-k_z d_h) \end{pmatrix}. \tag{13.20}$$

As β only depends on k_x and k_y, and is zero along the KH (see Fig. 13.11(c)) direction, the Weyl points should also be located along the KH direction and are given by the equality of the diagonal terms, which is

$$2(t_a - t_b)\cos(k_z d_h) + t_c[f(k_z d_h) - f(-k_z d_h)] = 0. \tag{13.21}$$

In Fig. 13.12(c), red and black curves show the trajectories of the Weyl points with positive and negative charges, respectively. These red and black curves divide the parameter space into regions with different topological properties. Consider a specific set of coupling coefficients as marked by the blue dashed line in Fig. 13.12(c), and the corresponding distribution of Weyl points are shown in the right panel of Fig. 13.12(d). To calculate the Chern number of the lower band, we start with $k_z = \pi/d_h$ and gradually decrease the value of k_z. At $k_z = \pi/d_h$, the Chern number is 0. With the decreasing of k_z, we first encounter a plane with net charge $+1$, and the Chern number goes to $+1$. Then we encounter a plane with net charge -1 and the Chern number decrease by 1 and jumps back to zero, and so on so forth. As a result, we obtain the Chern number as a function of k_z as shown in the left panel of Fig. 13.12(d) with solid blue line. We also label the different topological regions with the Chern number of the lower band in Fig. 13.12(c), where cyan, gray and yellow represent region with Chern numbers -1, 0 and $+1$, respectively.

The realization of this kind of chiral coupling in acoustic metacrystals with the connected resonance cavities (like Fig. 13.11(b)) is conceptually straight forward. We can in principal realize all the regions in the phase diagram of Fig. 13.12(c) as we tune the coupling coefficients and vary k_z. Actually, we can easily design metacrystals that go beyond the tight-binding description.[92]

The crucial part of breaking the effective TRS in a reduced 2D system with a fixed k_z is the chiral coupling which breaks the inversion symmetry as well as all the possible mirror symmetry with mirror plane parallel to the z-axis.

To recap, we discussed in this section the realization of Weyl points in acoustic metacrystals, and discussed the similarity and difference between the Weyl points and accidental degeneracy induced Dirac points. We find that acoustic Weyl points can be realized by engineering the interlayer coupling of a stacked honeycomb lattice of coupled acoustic cavities. In particular, we discussed the effects of different coupling at different sublattice sites and chiral coupling. These two kinds of couplings correspond respectively to the effective inversion symmetry breaking and effective time reversal symmetry breaking if we consider a reduced 2D system at a fixed k_z. If both types of coupling are incorporated in the metacrystal, we can realize an acoustic analogue of the topological Haldane model. The above systems can be realized using acoustic metacrystals comprising connected resonance cavities.

13.6. Conclusion

In this chapter, we showed that conical dispersions at the Brillouin zone center can be realized in acoustic and phononic crystals using accidental degeneracy. We then focus on building the relationship between zero-index metamaterials and conical dispersions at $\mathbf{k} = 0$, and reveal the subtle relationship between these two seemingly unrelated topics. We then see that the Dirac-like point concept can be generalized from 2D to 3D, and thereby achieving an isotropic zero-index acoustic metamaterial. For elastic waves, Dirac-like cones give additional interesting physics such as super-anisotropy. Last but not the least, we discussed the physics associated with an acoustic Weyl point and proposed coupled-cavity structures to realize such 3D conical dispersions and their implications in acoustic metacrystals.

Acknowledgement

This work was supported by Hong Kong Research Grant Council grant AoE/P-02/12.

References

1. Liu, Z., Zhang, X., Mao, Y., Zhu, Y. Y., Yang, Z., Chan, C. T. and Sheng, P. (2000). Locally resonant sonic materials, *Science* **289**, pp. 1734–1736.
2. Fang, N., Xi, D., Xu, J., Ambati, M., Srituravanich, W., Sun, C. and Zhang, X. (2006). Ultrasonic metamaterials with negative modulus, *Nat. Mater.* **5**, pp. 452–456.
3. Christensen, J. and Abajo, F. J. G. (2012). Anisotropic metamaterials for full control of acoustic waves, *Phys. Rev. Lett.* **108**, p. 124301.
4. Li, J. and Chan, C. T. (2004). Double-negative acoustic metamaterial, *Phys. Rev. E* **70**, p. 055602.
5. Zhang, X. and Liu, Z. (2004). Negative refraction of acoustic waves in two-dimensional phononic crystals, *Appl. Phys. Lett.* **85**, pp. 341–343.
6. Feng, L., Liu, X. P., Chen, Y. B., Huang, Z. P., Mao, Y. W., Chen, Y. F., Zi, J. and Zhu, Y. Y. (2005). Negative refraction of acoustic waves in two-dimensional sonic crystals, *Phys. Rev. B* **72**, p. 033108.
7. Ding, Y., Liu, Z., Qiu, C. and Shi, J. (2007). Metamaterial with simultaneously negative modulus and mass density, *Phys. Rev. Lett.* **99**, p. 093904.
8. He, Z., Cai, F., Ding, Y. and Liu, Z. (2008). Subwavelength imaging of acoustic waves by a canalization mechanism in a two-dimensional phononic crystal, *Appl. Phys. Lett.* **93**, p. 233503.
9. Sukhovich, A., Merheb, B., Muralidharan, K., Vasseur, J. O., Pennec, Y., Deymier, P. A. and Page, J. H. (2009). Experimental and theoretical evidence for subwavelength imaging in phononic crystals, *Phys. Rev. Lett.* **102**, p. 154301.
10. Zhu, J., Christensen, J., Jung, J., Martin-Moreno, L., Yin, X., Fok, L., Zhang, X. and Garcia-Vidal, F. J. (2011). A holey-structured metamaterial for acoustic deep-subwavelength imaging, *Nat. Phys.* **7**, pp. 52–55.
11. Park, C. M., Park, J. J., Lee, S. H., Seo, Y. M., Kim, C. K. and Lee, S. H. (2011). Amplification of acoustic evanescent waves using metamaterial slabs, *Phys. Rev. Lett.* **107**, p. 194301.
12. Zhou, X. and Hu, G. (2011). Superlensing effect of an anisotropic metamaterial slab with near-zero dynamic mass, *Appl. Phys. Lett.* **98**, p. 263510.
13. Ao, X. and Chan, C. T. (2008). Far-field image magnification for acoustic waves using anisotropic acoustic metamaterials, *Phys. Rev. E* **77**, p. 025601.
14. Li, J., Fok, L., Yin, X., Bartal, G. and Zhang, X. (2009). Experimental demonstration of an acoustic magnifying hyperlens, *Nat. Mater.* **8**, pp. 931–934.

15. Rosny, J. and Fink, M. (2002). Overcoming the diffraction limit in wave physics using a time-reversal mirror and a novel acoustic sink, *Phys. Rev. Lett.* **89**, p. 124301.
16. Lemoult, F., Fink, M. and Lerosey, G. (2011). Acoustic resonators for far-field control of sound on a subwavelength scale, *Phys. Rev. Lett.* **107**, p. 064301.
17. Cummer, S. A. and Schurig, D. (2007). One path to acoustic cloaking, *New J. Phys.* **9**, p. 45.
18. Cai, L. W. and Sanchez-Dehesa, J. (2007). Analysis of cummer-schurig acoustic cloaking, *New J. Phys.* **9**, p. 450.
19. Cheng, Y. and Liu, X. J. (2008). Resonance effects in broadband acoustic cloak with multilayered homogeneous isotropic materials, *Appl. Phys. Lett.* **93**, p. 071903.
20. Liu, H. (2009). Virtual reshaping and invisibility in obstacle scattering, *Inverse Problems* **25**, p. 045006.
21. Miller, D. A. B. (2006). On perfect cloaking, *Opt. Express* **14**, pp. 12457–12466.
22. Vasquez, F. G., Milton, G. W. and Onofrei, D. (2009). Active exterior cloaking for the 2D Laplace and Helmholtz equations, *Phys. Rev. Lett.* **103**, p. 073901.
23. Vasquez, F. G., Milton, G. W. and Onofrei, D. (2009). Broadband exterior cloaking, *Opt. Express* **17**, pp. 14800–14805.
24. Schoenberg, M. and Sen, P. N. (1983). Properties of a periodically stratified acoustic half-space and its relation to a Blot fluid, *J. Acoust. Soc. Am.* **73**, pp. 61–67.
25. Cheng, Y., Yang, F., Xu, J. Y. and Liu, X. J. (2008). A multilayer structured acoustic cloak with homogeneous isotropic materials, *Appl. Phys. Lett.* **92**, p. 151913.
26. Torrent, D. and Sanchez-Dehesa, J. (2008). Acoustic cloaking in two dimensions: a feasible approach, *New J. Phys.* **10**, p. 063015.
27. Cheng, Y. and Liu, X. J. (2009). Three dimensional multilayered acoustic cloak with homogeneous isotropic materials, *Appl. Phys. A* **94**, pp. 25–30.
28. Greenleaf, A., Kurylev, Y., Lassas, M. and Uhlmann, G. (2008). Isotropic transformation optics: approximate acoustic and quantum cloaking, *New J. Phys.* **10**, p. 115024.
29. Norris, A. N. (2008). Acoustic cloaking theory, *Proc. R. Soc. A* **464**, pp. 2411–2434.
30. Pendry, J. B. and Li, J. (2008). An acoustic metafluid: realizing a broadband acoustic cloak, *New J. Phys.* **10**, p. 115032.
31. Farhat, M., Enoch, S., Guenneau, S. and Movchan, A. B. (2008). Broadband cylindrical acoustic cloak for linear surface waves in a fluid, *Phys. Rev. Lett.* **101**, p. 134501.
32. Farhat, M., Guenneau, S., Enoch, S., Movchan, A., Zolla, F. and Nicolet, A. (2008). A homogenization route towards square cylindrical acoustic cloaks, *New J. Phys.* **10**, p. 115030.

33. Chen, H. and Chan, C. T. (2010). Acoustic cloaking and transformation acoustics, *J. Phys. D: Appl. Phys.* **43**, p. 113001.
34. Silveirinha, M. and Engheta, N. (2006). Tunneling of electromagnetic energy through subwavelength channels and bends using ε-near-zero materials, *Phys. Rev. Lett.* **97**, p. 157403.
35. Silveirinha, M. and Engheta, N. (2007). Design of matched zero-index metamaterials using nonmagnetic inclusions in epsilon-near-zero media, *Phys. Rev. B* **75**, p. 075119.
36. Silveirinha, M. G. and Engheta, N. (2007). Theory of supercoupling, squeezing wave energy and field confinement in narrow channels and tight bends using ε near-zero metamaterials, *Phys. Rev. B* **76**, p. 245109.
37. Alu, A. and Engheta, N. (2008). Dielectric sensing in ε-near-zero narrow waveguide channels, *Phys. Rev. B* **78**, p. 045102.
38. Alu, A., Silveirinha, M. G. and Engheta, N. (2008). Transmission-line analysis of ε-near-zero-filled narrow channels, *Phys. Rev. E* **78**, p. 016604.
39. Edwards, B., Alu, A., Silveirinha, M. G. and Engheta, N. (2009). Reflectionless sharp bends and corners in waveguides using epsilon-near-zero effects, *J. Appl. Phys.* **105**, p. 044905.
40. Halterman, K. and Feng, S. (2008). Resonant transmission of electromagnetic fields through subwavelength zero-ε slits, *Phys. Rev. A* **78**, p. 021805.
41. Ziolkowski, R. W. (2004). Propagation in and scattering from a matched metamaterial having a zero index of refraction, *Phys. Rev. E* **70**, p. 046608.
42. Enoch, S., Tayeb, G., Sabouroux, P., Guerin, N. and Vincent, P. (2002). A metamaterial for directive emission, *Phys. Rev. Lett.* **89**, p. 213902.
43. Alu, A., Silveirinha, M. G., Salandrino, A. and Engheta, N. (2007). Epsilon-near-zero metamaterials and electromagnetic sources: tailoring the radiation phase pattern, *Phys. Rev. B* **75**, p. 155410.
44. Hao, J., Yan, W. and Qiu, M. (2010). Super-reflection and cloaking based on zero index metamaterial, *Appl. Phys. Lett.* **96**, p. 101109.
45. Jin, Y. and He, S. (2010). Enhancing and suppressing radiation with some permeability-near-zero structures, *Opt. Express* **18**, pp. 16587–16593.
46. Nguyen, V. C., Chen, L. and Halterman, K. (2010). Total transmission and total reflection by zero index metamaterials with defects, *Phys. Rev. Lett.* **105**, p. 233908.
47. Xu, Y. and Chen, H. (2011). Total reflection and transmission by epsilon-near-zero metamaterials with defects, *Appl. Phys. Lett.* **98**, p. 113501.
48. Chan, C. T., Hang, Z. H. and Huang, X. (2012). Dirac dispersion in two-dimensional photonic crystals, *Advances in Opto Electronics* **12**, p. 313984.
49. Chan, C. T., Huang, X., Liu, F. and Hang, Z. H. (2012). Dirac dispersion and zero-index in two dimensional and three dimensional photonic and phononic systems, *PIER B* **44**, pp. 163–190.
50. Huang, X., Lai, Y., Hang, Z. H., Zheng, H. and Chan, C. T. (2011). Dirac cones induced by accidental degeneracy in photonic crystals and zero-refractive-index materials, *Nat. Mater.* **10**, pp. 582–586.

51. Liu, R., Cheng, Q., Hand, T., Mock, J. J., Cui, T. J., Cummer, S. A. and Smith, D. R. (2008). Experimental demonstration of electromagnetic tunneling through an epsilon-near-zero metamaterial at microwave frequencies, *Phys. Rev. Lett.* **100**, p. 023903.
52. Edwards, B., Alu, A., Young, M. E., Silveirinha, M. and Engheta, N. (2008). Experimental verification of epsilon-near-zero metamaterial coupling and energy squeezing using a microwave waveguide, *Phys. Rev. Lett.* **100**, p. 033903.
53. Moitra, P., Yang, Y., Anderson, Z., Kravchenko, I. I., Briggs, D. P., Valentine, J. (2013). Realization of an all-dielectric zero-index optical metamaterial, *Nat. Photon.* **7**, pp. 791–795.
54. Bongard, F., Lissek, H. and Mosig, J. R. (2010). Acoustic transmission line metamaterial with negative/zero/positive refractive index, *Phys. Rev. B* **82**, p. 094306.
55. Liu, F., Huang, X. and Chan, C. T. (2012). Dirac cones at $k = 0$ in acoustic crystals and zero refractive index acoustic materials, *Appl. Phys. Lett.* **100**, p. 071911.
56. Neto, A. H. C., Guinea, F., Peres, N. M. R., Novoselov, K. S. and Geim, A. K. (2009). The electronic properties of grapheme, *Rev. Mod. Phys.* **81**, pp. 109–162.
57. Geim, A. K. and MacDonald, A. H. (2007). Graphene: exploring carbon flatland, *Phys. Today* **60**, pp. 35–41.
58. Geim, A. K. and Novoselov, K. S. (2007). The rise of graphene, *Nat. Mater.* **6**, pp. 183–191.
59. Novoselov, K. S., Geim, A. K., Morozov, S. V., Jiang, D., Zhang, Y., Dubonos, S. V., Gregorieva, I. V. and Firsov, A. A. (2004). Electric field effect in atomically thin carbon films, *Science* **306**, pp. 666–669.
60. Novoselov, K. S., Geim, A. K., Morozov, S. V., Jiang, D., Katsnelson, M. I., Grigorieva, I. V., Dubonos, S. V. and Firsov, A. A. (2005). Two-dimensional gas of massless Dirac fermions in grapheme, *Nature* **438**, pp. 197–200.
61. Morozov, S. V., Novoselov, K. S., Katsnelson, M. I., Schedin, F., Ponomarenko, L. A., Jiang, D. and Geim, A. K. (2006). Strong suppression of weak localization in graphene, *Phys. Rev. Lett.* **97**, p. 016801.
62. Katsnelson, M. I., Novoselov, K. S. and Geim, A. K. (2006). Chiral tunnelling and the Klein paradox in graphene, *Nat. Phys.* **2**, pp. 620–625.
63. Zhang, Y., Tan, Y. W., Stormer, H. L. and Kim, P. (2005). Experimental observation of the quantum Hall effect and Berry's phase in grapheme, *Nature* **438**, pp. 201–204.
64. Gusynin, V. P. and Sharapov, S. G. (2005). Unconventional integer quantum hall effect in graphene, *Phys. Rev. Lett.* **95**, p. 146801.
65. Novoselov, K. S., Jiang, Z., Zhang, Y., Morozov, S. V., Stormer, H. L., Zeitler, U., Maan, J. C., Boebinger, G. G., Kim, P. and Geim, A. K. (2007). Room-temperature quantum hall effect in graphene, *Science* **315**, p. 1379.
66. Katsnelsona, M. I. (2006). Zitterbewegung, chirality and minimal conductivity in graphene, *Eur. Phys. J. B* **51**, pp. 157–160.

67. Cserti, J. and David, G. (2006). Unified description of Zitterbewegung for spintronic, graphene and superconducting systems, *Phys. Rev. B* **74**, p. 172305.
68. Rusin, T. M. and Zawadzki, W. (2007). Transient Zitterbewegung of charge carriers in mono and bilayer graphene and carbon nanotubes, *Phys. Rev. B* **76**, p. 195439.
69. David, G. and Cserti, J. (2010). General theory of the Zitterbewegung, *Phys. Rev. B* **81**, p. 121417.
70. Plihal, M. and Maradudin, A. A. (1991). Photonic band structure of a two-dimensional system: The triangular lattice, *Phys. Rev. B* **44**, pp. 8565–8571.
71. Haldane, F. D. M. and Raghu, S. (2008). Possible realization of directional optical waveguides in photonic crystals with broken time-reversal symmetry, *Phys. Rev. Lett.* **100**, p. 013904.
72. Raghu, S. and Haldane, F. D. M. (2008). Analogs of quantum-Hall-effect edge states in photonic crystals, *Phys. Rev. A* **78**, p. 033834.
73. Ochiai, T. and Onoda, M. (2009). Photonic analog of graphene model and its extension: Dirac cone, symmetry and edge states, *Phys. Rev. B* **80**, p. 155103.
74. Ochiai, T. (2010). Topological properties of bulk and edge states in honeycomb lattice photonic crystals: the case of TE polarization, *J. Phys.: Condens. Matter* **22**, p. 225502.
75. Sepkhanov, R. A., Bazaliy, Y. B. and Beenakker, C. W. J. (2007). Extremal transmission at the Dirac point of a photonic band structure, *Phys. Rev. A* **75**, p. 063813.
76. Sepkhanov, R. A., Nilsson, J. and Beenakker, C. W. J. (2008). Proposed method for detection of the pseudospin-1/2 Berry phase in a photonic crystal with a Dirac spectrum, *Phys. Rev. B* **78**, p. 045122.
77. Diem, M., Koschny, T. and Soukoulis, C. M. (2010). Transmission in the vicinity of the Dirac point in hexagonal photonic crystals, *Physica B* **405**, pp. 2990–2995.
78. Zhang, X. (2008). Observing Zitterbewegung for photons near the Dirac point of a two-dimensional photonic crystal, *Phys. Rev. Lett.* **100**, p. 113903.
79. Mei, J., Wu, Y., Chan, C. T. and Zhang, Z. Q. (2012). First-principles study of Dirac and Dirac-like cones in phononic and photonic crystals, *Phys. Rev. B* **86**, p. 035141.
80. Zhang, X. and Liu, Z. (2008). Extremal transmission and beating effect of acoustic waves in two-dimensional sonic crystals, *Phys. Rev. Lett.* **101**, p. 264303.
81. Torrent, D. and Sanchez-Dehesa, J. (2012). Acoustic analogue of graphene: observation of Dirac cones in acoustic surface waves, *Phys. Rev. Lett.* **108**, p. 174301.
82. Lu, J., Qiu, C., Xu, S., Ye, Y., Ke, M. and Liu, Z. (2014). Dirac cones in two-dimensional artificial crystals for classical waves, *Phys. Rev. B* **89**, p. 134302.

83. Han, D., Lai, Y., Zi, J., Zhang, Z. Q. and Chan, C. T. (2009). Dirac spectra and edge states in honeycomb plasmonic lattices, *Phys. Rev. Lett.* **102**, p. 123904.
84. Xiao, M., Zhang, Z. Q. and Chan, C. T. (2014). Surface impedance and bulk band geometric phases in one-dimensional systems, *Phys. Rev. X* **4**, p. 021017.
85. Xiao, M., Ma, G., Yang, Z., Sheng, P., Zhang, Z. Q. and Chan, C. T. (2014). Geometric phase and band inversion in periodic acoustic systems, *Nat. Phys.* **11**, pp. 240–244.
86. Huang, X., Xiao, M., Zhang, Z. Q. and Chan, C. T. (2014). Sufficient condition for the existence of interface states in some two-dimensional photonic crystals, *Phys. Rev. B.* **90**, p. 075423.
87. Wu, Y., Li, J., Zhang, Z. Q. and Chan, C. T. (2006). Effective medium theory for magnetodielectric composites: beyond the long-wavelength limit, *Phys. Rev. B* **74**, p. 085111.
88. Liu, F., Lai, Y., Huang, X. and Chan, C. T. (2011). Dirac cones at $k = 0$ in phononic crystals, *Phys. Rev. B* **84**, p. 224113.
89. Ni, Q. and Cheng, J. (2005). Anisotropy of effective velocity for elastic wave propagation in two-dimensional phononic crystals at low frequencies, *Phys. Rev. B* **72**, p. 014305.
90. Wu, Y. and Zhang, Z. Q. (2009). Dispersion relations and their symmetry properties of electromagnetic and elastic metamaterials in two dimensions, *Phys. Rev. B* **79**, p. 195111.
91. Lai, Y., Wu, Y., Sheng, P. and Zhang, Z. Q. (2011). Hybrid elastic solids, *Nat. Mater.* **10**, pp. 620–624.
92. Xiao, M., Chen, W. J., He, W. Y. and Chan, C. T. (2015). Synthetic gauge flux and Weyl points in acoustic systems, *Nat. Phys.* **11**, p. 3458.
93. Lu, L., Fu, L., Joannopoulos, J. D. and Soljacic, M. (2013). Weyl points and line nodes in gyroid photonic crystals, *Nat. Photon.* **7**, pp. 294–299.
94. Wan, X., Turner, A. M., Vishwanath, A. and Savrasov, S. Y. (2011). Topological semimetal and Fermi-arc surface states in the electronic structure of pyrochlore iridates, *Phys. Rev. B* **83**, p. 205101.
95. Fang, Z., Nagaosa, N., Takahashi, K. S., Asamitsu, A., Mathieu, R., Ogasawara, T., Yamada, H., Kawasaki, M., Tokura, Y. and Terakura, K. (2003). The anomalous Hall effect and magnetic monopoles in momentum space, *Science* **302**, pp. 92–95.
96. Pendry, J. B., Schurig, D. and Smith, D. R. (2006). Controlling electromagnetic fields, *Science* **312**, pp. 1780–1782.
97. Leonhardt, U. (2006). Optical conformal mapping, *Science* **312**, pp. 1777–1780.
98. Schurig, D., Mock, J. J., Justice, B. J., Cummer, S. A., Pendry, J. B., Starr, A. F. and Smith, D. R. (2006). Metamaterial electromagnetic cloak at microwave frequencies, *Science* **314**, pp. 977–980.
99. Leonhardt, U. and Philbin, T. G. (2006). General relativity in electrical engineering, *New J. Phys.* **8**, p. 247.

100. Milton, G. W., Briane, M. and Willis, J. R. (2006). On cloaking for elasticity and physical equations with a transformation invariant form, *New J. Phys.* **8**, p. 248.
101. Leonhardt, U. and Philbin, T. G. (2009). Transformation optics and the geometry of light, *Prog. Opt.* **53**, pp. 69–152.
102. Chen, H., Chan, C. T. and Sheng, P. (2010). Transformation optics and metamaterials, *Nat. Mater.* **9**, pp. 387–396.
103. Lan, Z., Goldman, N., Bermudez, A., Lu, W. and Ohberg, P. (2011). Dirac-Weyl fermions with arbitrary spin in two-dimensional optical superlattices, *Phys. Rev. B* **84**, p. 165115.
104. Dora, B., Kailasvuori, J. and Moessner, R. (2011). Lattice generalization of the Dirac equation to general spin and the role of the flat band, *Phys. Rev. B* **84**, p. 195422.
105. Lawrence, F. J., Botten, L. C., Dossou, K. B., McPhedran, R. C. and Sterke, C. M. (2010). Photonic-crystal surface modes found from impedances, *Phys. Rev. A* **82**, p. 053840.
106. Lawrence, F. J., Botten, L. C., Dossou, K. B., Sterke, C. M. and McPhedran, R. C. (2009). Impedance of square and triangular lattice photonic crystals, *Phys. Rev. A* **80**, p. 023826.
107. Klitzing, K. (1986). The quantized Hall effect, *Rev. Mod. Phys.* **58**, pp. 519–531.
108. Hasan, M. Z. and Kane, C. L. (2010). Colloquium: topological insulators, *Rev. Mod. Phys.* **82**, pp. 3045–3067.
109. Qi, X. L. and Zhang, S. C. (2011). Topological insulators and superconductors, *Rev. Mod. Phys.* **83**, pp. 1057–1110.
110. Wang, Z., Chong, Y., Joannopoulos, J. D. and Soljacic, M. (2009). Observation of unidirectional backscattering-immune topological electromagnetic states, *Nature* **461**, pp. 772–775.
111. Rechtsman, M. C., Zeuner, J. M., Plotnik, Y., Lumer, Y., Podolsky, D., Dreisow, F., Nolte, S., Segev, M. and Szameit, A. (2013). Photonic Floquet topological insulators, *Nature* **496**, pp. 196–200.
112. Khanikaev, A. B., Mousavi, S. H., Tse, W. K., Kargarian, M., MacDonald, A. H. and Shvets, G. (2013). Photonic topological insulators, *Nat. Mater.* **12**, pp. 233–239.
113. Fang, K., Yu, Z. and Fan, S. (2012). Realizing effective magnetic field for photons by controlling the phase of dynamic modulation, *Nat. Photon.* **6**, pp. 782–787.
114. Hafezi, M., Demler, E. A., Lukin, M. D. and Taylor, J. M. (2011). Robust optical delay lines with topological protection, *Nat. Phys.* **7**, pp. 907–912.
115. Wang, Z., Chong, Y. D., Joannopoulos, J. D. and Soljacic, M. (2008). Reflection-free one-way edge modes in a gyromagnetic photonic crystal, *Phys. Rev. Lett.* **100**, p. 013905.
116. Yu, Z., Veronis, G., Wang, Z. and Fan, S. (2008). One-way electromagnetic waveguide formed at the interface between a plasmonic metal under a static magnetic field and a photonic crystal, *Phys. Rev. Lett.* **100**, p. 023902.

117. Kraus, Y. E., Lahini, Y., Ringel, Z., Verbin, M. and Zilberberg, O. (2012). Topological states and adiabatic pumping in quasicrystals, *Phys. Rev. Lett.* **109**, p. 106402.
118. Plotnik, Y., Rechtsman, M. C., Song, D., Heinrich, M., Zeuner, J. M., Nolte, S., Lumer, Y., Malkova, N., Xu, J., Szameit, A., Chen, Z. and Segev, M. (2014). Observation of unconventional edge states in 'photonic graphene', *Nature Mater.* **13**, pp. 57–62.
119. Rechtsman, M. C., Plotnik, Y., Zeuner, J. M., Song, D., Chen, Z., Szameit, A. and Segev, M. (2013). Topological creation and destruction of edge states in photonic graphene, *Phys. Rev. Lett.* **111**, p. 103901.
120. Lu, L., Joannopoulos, J. D. and Soljacic, M. (2014). Topological photonics, *Nat. Photon.* **8**, pp. 821–829.
121. Fleury, R., Sounas, D. L., Sieck, C. F., Haberman, M. R. and Alu, A. (2014). Sound isolation and giant linear nonreciprocity in a compact acoustic circulator, *Science* **343**, pp. 516–519.
122. Yang, Z., Gao, F., Shi, X., Lin, X., Gao, Z., Chong, Y. and Zhang, B. (2015). Topological acoustics, *Phys. Rev. Lett.* **114**, p. 114301.
123. Prodan, E. and Prodan, C. (2009). Topological phonon modes and their role in dynamic instability of microtubules, *Phys. Rev. Lett.* **103**, p. 248101.
124. Kane, C. L. and Lubensky, T. C. (2013). Topological boundary modes in isostatic lattices, *Nat. Phys.* **10**, pp. 39–45.
125. Paulose, J., Chen, B. G. and Vitelli, V. (2015). Topological modes bound to dislocations in mechanical metamaterials, *Nat. Phys.* **11**, pp. 153–156.
126. Ni, X., He, C, Sun, X. C., Liu, X. P., Lu, M. H., Feng, L. and Chen, Y. F. (2015). Topologically protected one-way edge mode in networks of acoustic resonators with circulating air flow, *New J. Phys.* **17**, p. 053016.
127. Zak, J. (1989). Berry's Phase for energy bands in solids, *Phys. Rev. Lett.* **62**, pp. 2747–2750.
128. Sakoda, K. (2012). Dirac cone in two- and three-dimensional metamaterials, *Opt. Express* **20**, pp. 3898–3917.
129. Wu, Y., Lai, Y. and Zhang, Z. Q. (2007). Effective medium theory for elastic metamaterials in two dimensions, *Phys. Re. B* **76**, p. 205313.
130. Hermann, W. (1929). Elektron und Gravitation. I, *Zeitschrift für Physik* **56**, pp. 330–352.
131. Sakoda, K. (2012). Universality of mode symmetries in creating photonic Dirac cones, *J. Opt. Soc. Am. B* **29**, pp. 2770–2778.
132. Haldane, F. D. M. (1988). Model for a quantum Hall effect without Landau levels: condensed-matter realization of the "parity anomaly", *Phys. Rev. Lett.* **61**, pp. 2015–2018.
133. Jimenez-Garcia, K., LeBlanc, L. J., Williams, R. A., Beeler, B. C., Perry, A. R. and Spielman, I. B. (2012). Peierls substitution in an engineered lattice potential, *Phys. Rev. Lett.* **108**, p. 225303.
134. Peierls, R. E. (1933). On the theory of diamagnetism of conduction electrons. *Z. Phys.* **80**, p. 763.

CHAPTER 14

An Interface Model for Homogenization of Acoustic Metafilms

JEAN-JACQUES MARIGO

École Polytechnique, France

AGNÈS MAUREL*

Institut Langevin, France

14.1. Introduction

Acoustic metamaterials consisting in massive materials perforated by periodic subwavelength holes [Pendry *et al.* (2004); Garcia-Vidal *et al.* (2005); Zhu *et al.* (2011)] or more sparse structures involving periodic arrangement of wires [Lemoult *et al.* (2013)] have been shown to be able to control the wave propagation with high flexibility. In comparison, phononic crystals have *a priori* higher dimensions because of their wavelength-scale period. However, if the metamaterial has subwavelength period, many of the observed phenomena are attributable to Fabry-Perot type resonances, resonances in the hole or resonances of the wires. Therefore, these structures have a

*Corresponding author: agnes.maurel@espci.fr

limitation in their thickness, which has to be at wavelength scale to produce such resonances (and thickness refers to the size in the direction perpendicular to the plane containing the periodic cells). In order to reduce the size of the devices, structures with subwavelength thickness have been developed. They are known as metasurfaces and metafilms. Despite the vanishing thickness in comparison to the incident wavelength, the capability of these ultrathin devices to control the wave propagation has been evidenced. This is because they are based on a resonance which is not related to their thickness. It can be a thin elastic membrane within the unit cell [Ma *et al.* (2014); Zhao *et al.* (2014)], or resonances of labyrinthine or curled elements squeezed in the unit cell [Peng *et al.* (2014); Li *et al.* (2014); Xie *et al.* (2014)].

Because of their subwavelength unit cell, homogenization techniques are natural tools to describe the effective properties of metamaterials. Standard homogenization methods [Zhu *et al.* (2013); Maurel *et al.* (2013)] or more empirical methods, as the retrieval techniques [Arslanagic *et al.* (2013); Castanié *et al.* (2014)] (often referred as Nicolson-Ross-Weir technique) and approximated modal approaches [Pendry *et al.* (2004); Garcia-Vidal *et al.* (2005); Kelders *et al.* (1998)] have been largely used for massive structures leading to effective bulk modulus and effective mass density, being the acoustic counterparts of the effective permittivity and permeability in electromagnetism. An enlightening review can be found in Ref. [Simovski (2011)]. When metafilms or metasurfaces are concerned, first attempts have been proposed which use the same retrieval parameters as for bulk materials and affecting a thickness value more or less relevant (see the discussions in § 3 of [Simovski (2011)] and in §2 of [Holloway *et al.* (2009)]). It is now admitted that these approaches are not pertinent and rather, jump conditions or transmission conditions of the fields across a zero thickness surface are thought. In the context of electromagnetism, effective surface parameters are expressed in term of effective surface admittances [Zhao *et al.* (2011)] or surface susceptibilities [Holloway

et al. (2009); Dimitriadis *et al.* (2012); Kuester *et al.* (2003); Holloway *et al.* (2005)]. In addition to be closely related to the notion of electric polarization density and surface magnetization, which are not easily transposable to the case of acoustic or elastic waves, these techniques require an inversion procedure to get the surface susceptibilities, from the measurement of the scattering coefficients. Thus, although the problem of the artificial thickness is avoided, the question of wether or not the obtained parameters depend on the scattering problem that has been considered remains.

In this paper, we present a homogenization method for vanishing thickness metamaterials which explicitly accounts for the microstructure of the unit cell. The problem ends with jump conditions for the acoustic pressure and for the normal acoustic velocity involving parameters being wave independent, by construction. This is because, as in the standard homogenization, the parameters are determined by solving (analytically or numerically) elementary cell problems in the static case (that is for zero frequency). The model relies on a separation of scales, a micro scale associated to the small scatterer size and a macro scale associated to the acoustic wavelength, and ε is the small parameter that measures the ratio of the two scales. Each scale is associated to a system of coordinates which is relevant or not to describe the variation of the acoustic fields wether we are close to the film or far from it. Thus, a separation of the space is used, into an outer region, typically the far field, where only the macro scale makes sense and an inner region, the near field, where both the micro and macro scales are needed. Expansions of the fields in power of ε are performed in both regions and finally, matching conditions are used between the two regions. This approach has been developed in the context of static elasticity, see [Marigo and Pideri (2011); David *et al.* (2012)] for a complete description. The case of wave propagation has been less regarded. We mention the works of Capdeville and Marigo for seismic waves in the time domain [Capdeville and Marigo (2007); Capdeville *et al.* (2010a); Guillot *et al.* (2010);

Capdeville *et al.* (2010b); Capdeville and Marigo (2013)], and the Refs. [Bonnet-Bendhia *et al.* (2004); Delourme (2010); Delourme *et al.* (2012)] in the frequency domain. Note also works using alternative forms of the matched asymptotic expansions [Sanchez-Hubert and Sanchez-Palencia (1982); Martin and Dalrymple (1988); Kakuno *et al.* (1992)].

The paper is organized as follows. In Section 14.2, we apply the interface model for the problem of acoustic wave propagation through a thin film made of a periodic row of sound hard inclusions. The jump conditions are derived, Eqs. (14.27), with the interface parameters $(\mathbf{B}, \mathsf{C})$ being associated to the elementary problems, Eqs. (14.17). Section 14.3 presents results in the case of rectangular inclusions in two dimensions, with thickness e. This choice is not casual. First, it allows for approximate, but explicit, solutions of the elementary problems. Thus, the interface parameters $(\mathbf{B}, \mathsf{C})$ are given, and associated transmission and reflection coefficients of a plane wave at oblique incidence, Eqs. (14.33), are obtained. Next, a comparison with the classical homogenization of layered media is possible since again, explicit solutions are available in this case [Garcia-Vidal *et al.* (2005); Oleinik *et al.* (2009)], Eq. (14.34). The validity of our interface model is inspected by comparison with full wave simulations, and discussed in the light of the comparison with the classical homogenization. The results show that the interface homogenization is accurate for $kh, ke < 1$, while classical homogenization can be used for $kh < 1$ and $e > h$ and it largely underestimates the scattering properties of thin structured layers. The small parameter ε can be forced to values of order unity; typically for ε and ke of order unity, the discrepancy between the exact solution and the solution of the homogenized problem do not exceed few per percent. We end with an illustrative example of an ultra thin structured film obstructing a waveguide with rigid walls. Again, an explicit solution of the homogenized problem is possible and a good agreement with direct numerics is obtained. Technical calculations are collected in the Appendices.

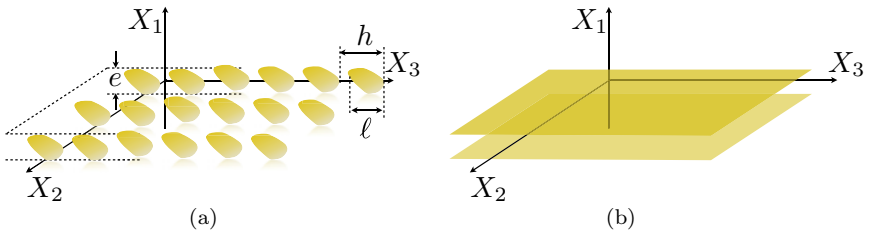


Fig. 14.1. Scattering problem for (a) the microstructured film (metafilm) and (b) the equivalent interface. $\mathbf{X} = (X_1, X_2, X_3)$ is the coordinate system in the real space.

14.2. The Homogenized Interface Model

The acoustic pressure $P(\mathbf{X})$ satisfies the Helmholtz equation outside the inclusions

$$\left(\Delta + k^2\right) P(\mathbf{X}) = 0, \tag{14.1}$$

with k the wavenumber (the time dependance is $e^{-i\omega t}$, with ω the frequency, and it is omitted in the following). Here, we restricted our study to sound hard inclusions, with Neumann boundary condition on the boundary of the inclusions $\partial_{\mathbf{n}} P = 0$, where $\partial_{\mathbf{n}}$ denotes the normal derivative. Also, the wave propagation is associated to a wave source and to a proper radiation condition that we do not need to specify at this point. The inclusions form a periodic arrangement located near the (X_2, X_3) plane (with typical thickness e along X_1, Fig. 14.1(a)). For simplicity, we assume that h is the period of the cell along X_2 and X_3 (we could assume two different periods along X_2 and X_3 without additional complexity, except in the notations). We inspect the possibility to replace this problem by a simpler one, where the structured film is replaced by an interface associated to jump conditions for the pressure P and its normal derivative, or normal velocity, $\partial_{X_1} P$ (the nature of the equivalent interface, in terms of its thickness, will be discussed).

In our acoustic problem, the natural small parameter is $\varepsilon = kh \ll 1$ and to be consistent, we need to write the Helmholtz equation in a dimensionless form, with $\mathbf{x} \equiv k\mathbf{X}$, and $p(\mathbf{x}) \equiv P(\mathbf{X})$. In

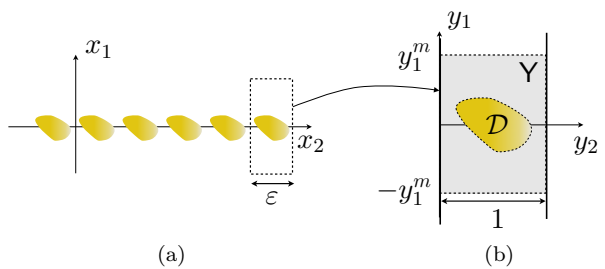


Fig. 14.2. (a) Geometry of the structured film in non dimensional coordinates $\mathbf{x} = (x_1, x_2, x_3)$; (b) inner problem in coordinates $\mathbf{y} = (y_1, y_2, y_3)$. $\mathcal{D}$ is the region occupied by the rigid inclusion. We define $\mathsf{Y} = (-y_1^m, y_1^m) \times (0,1)^2$, $\mathsf{Y}_\infty = \lim_{y_1^m \to +\infty} \mathsf{Y}$, and $\mathbf{x}' = (x_2, x_3)$, $\mathbf{y}' = (y_2, y_3)$.

$\mathbf{x} = (x_1, x_2, x_3)$, the period is now ε and the film thickness is ke (Fig. 14.2(a)); it is assumed that e and h are of the same order of magnitude. We get $(\Delta_{\mathbf{x}} + 1)\, p(\mathbf{x}) = 0$, which is written, for convenience using $\mathbf{u}$,

$$\begin{cases} \mathrm{div}_{\mathbf{x}} \mathbf{u}(\mathbf{x}) + p(\mathbf{x}) = 0, \quad \text{with} \;\; \mathbf{u} \equiv \boldsymbol{\nabla}_{\mathbf{x}} p, \\ \\ \mathbf{u}.\mathbf{n}_{|\partial\mathcal{D}} = 0, \end{cases} \tag{14.2}$$

and $\mathbf{u}$ is the acoustic velocity, in dimensionless form.

14.2.1. *The inner/outer expansions and associated matching conditions*

The idea is now to expand the solution with respect to the small parameter ε, namely

$$\begin{cases} p = p^0(\mathbf{x}) + \varepsilon p^1(\mathbf{x}) + \varepsilon^2 p^2(\mathbf{x}) + \ldots, \\ \\ \mathbf{u} = \mathbf{u}^0(\mathbf{x}) + \varepsilon \mathbf{u}^1(\mathbf{x}) + \varepsilon^2 \mathbf{u}^2(\mathbf{x}) + \ldots \end{cases} \tag{14.3}$$

In principle, this expansion can be used in the whole space (see e.g. [Martin and Dalrymple (1988)]). Nevertheless, if the spatial derivatives in Eq. (14.2) make ε to appear, the resolution may become tricky. Such complications are avoided if a displacement in $\mathbf{x}$ of order unity produces a variation in p of order unity, namely $\partial_{x_i} p \sim p$. This

is ensured in the far field of the film, where $\partial_{X_i} P \sim kP \to \partial_{x_i} p \sim p$. The story is different in the near field: there, the film generates evanescent waves whose strongest variations are associated to the smallest scale of the microstructure, say $\partial_{X_i} P \sim P/h \to \partial_{x_i} p \sim p/\varepsilon$. Then, the wavefield satisfies pseudo-periodic conditions, for instance along X_2, $P(X_1, X_2 + h, X_3) = P(X_1, X_2, X_3)e^{ik_2 h}$, with k_2 the component of the incident wavenumber along X_2 (the same applies along X_3). This behavior is associated to low variations of P ($\partial_{x_2} p \sim p$) for non normal incidence $k_2 \neq 0$. Thus, we have to deal with different scales in the far and near fields, and two scales in the near field.

The presented approach solves this annoying problem with two ingredients: first, a separation of the space into an inner and an outer regions, which correspond to the near and far fields, respectively. In the outer region, the natural coordinates $\mathbf{x} \equiv (x_1, x_2, x_3)$ are adapted and the expansion, Eq. (14.3), applies. In the inner region, the rapid variations of p and $\mathbf{u}$ are accounted introducing a new system of coordinates $\mathbf{y} = \mathbf{x}/\varepsilon$ such that $\partial_{y_i} p \sim p$ for the rapid variations of p; next, the slow variations along x_2 and x_3 are accounted for by keeping $\mathbf{x}' \equiv (x_2, x_3)$ as additional coordinates. Note that in the $\mathbf{y}$ coordinates, the cell has a thickness unity (Fig. 14.2(b)). Owing to this modification, the expansions read

$$\text{Outer expansion} \begin{cases} p = p^0(\mathbf{x}) + \varepsilon p^1(\mathbf{x}) + \varepsilon^2 p^2(\mathbf{x}) + \dots, \\ \mathbf{u} = \mathbf{u}^0(\mathbf{x}) + \varepsilon \mathbf{u}^1(\mathbf{x}) + \varepsilon^2 \mathbf{u}^2(\mathbf{x}) + \dots \end{cases}$$

$$\text{Inner expansion} \begin{cases} p = q^0(\mathbf{y}, \mathbf{x}') + \varepsilon q^1(\mathbf{y}, \mathbf{x}') + \varepsilon^2 q^2(\mathbf{y}, \mathbf{x}') + \dots, \\ \mathbf{u} = \mathbf{v}^0(\mathbf{y}, \mathbf{x}') + \varepsilon \mathbf{v}^1(\mathbf{y}, \mathbf{x}') + \varepsilon^2 \mathbf{v}^2(\mathbf{y}, \mathbf{x}') + \dots \end{cases} \tag{14.4}$$

Finally, both regions are connected in some boundary region, where the evanescent field is vanishing at small x_1 values corresponding to $y_1 = x_1/\varepsilon \to \pm\infty$. These matching conditions are written using Taylor expansions of $p^0(x_1, \mathbf{x}') = p^0(0, \mathbf{x}') + x_1 \partial_{x_1} p^0(0, \mathbf{x}') + \dots = p^0(0, \mathbf{x}') + \varepsilon y_1 \partial_{x_1} p^0(0, \mathbf{x}') + \dots$, same for $\mathbf{u}^0$, and identifying the terms

in ε^n in the inner and outer expansions, Eqs. (14.4). We get

$$\begin{cases} p^0(0^\pm, \mathbf{x}') = \lim\limits_{y_1 \to \pm\infty} q^0(\mathbf{y}, \mathbf{x}'), \\ \mathbf{u}^0(0^\pm, \mathbf{x}') = \lim\limits_{y_1 \to \pm\infty} \mathbf{v}^0(\mathbf{y}, \mathbf{x}'), \end{cases} \tag{14.5}$$

at leading order, and

$$\begin{cases} p^1(0^\pm, \mathbf{x}') = \lim\limits_{y_1 \to \pm\infty} \left[q^1(\mathbf{y}, \mathbf{x}') - y_1 \dfrac{\partial p^0}{\partial x_1}(0^\pm, \mathbf{x}') \right], \\ \mathbf{u}^1(0^\pm, \mathbf{x}') = \lim\limits_{y_1 \to \pm\infty} \left[\mathbf{v}^1(\mathbf{y}, \mathbf{x}') - y_1 \dfrac{\partial \mathbf{u}^0}{\partial x_1}(0^\pm, \mathbf{x}') \right], \end{cases} \tag{14.6}$$

at first order, and fortunately, we do not need to go at higher orders in the expansions. A last subtlety of the interface model is to assume that the functions q^n and $\mathbf{v}^n$ are periodic with respect to y_2 and y_3. Again, this is not meaningless in the present context if we have in mind the condition of pseudo periodicity. This condition is handled by the variables x_2 and x_3, for instance $q^n(\mathbf{y}, x_2 + \varepsilon, x_3) = e^{ik_2 h} q^n(\mathbf{y}, x_2, x_3)$ (note that the condition of pseudo periodicity applies for the outer solution $(p^n, \mathbf{u}^n)$, and from (14.5)–(14.6), it applies for $(q^n, \mathbf{v}^n)$ also). If one thinks to the q^n in terms of separable functions $q^n(\mathbf{y}, \mathbf{x}') = f(\mathbf{y})g(\mathbf{x}')$ (and this will be the case in the quasi-static limit), we recover the form of a Floquet solution, with $g(\mathbf{x}') = e^{i\mathbf{k}'.\mathbf{x}'/k}$ and f periodic with respect to $\mathbf{y}' \equiv (y_2, y_3)$.

Now, the matching conditions, Eqs. (14.5)–(14.6) can be written in an alternative form, that will be used later

$$\begin{cases} \left[\!\left[p^0 \right]\!\right] = \delta q^0, \qquad \left[\!\left[p^1 \right]\!\right] = \lim\limits_{y_1^m \to +\infty} \left[\delta q^1(y_1^m) - 2y_1^m \, \dfrac{\partial p^0}{\partial x_1}(0, \mathbf{x}') \right], \\ \left[\!\left[\mathbf{u}^0 \right]\!\right] .\mathbf{e}_1 = \delta v_1^0, \; \left[\!\left[\mathbf{u}^1 \right]\!\right] .\mathbf{e}_1 = \lim\limits_{y_1^m \to +\infty} \left[\delta v_1^1(y_1^m) - 2y_1^m \, \dfrac{\partial u_1^0}{\partial x_1}(0, \mathbf{x}') \right], \end{cases} \tag{14.7}$$

(Fig. 14.2), where we have defined

$$\delta f(y_1^m) \equiv f(+y_1^m) - f(-y_1^m), \quad \text{and} \quad \delta f \equiv \delta f(+\infty). \tag{14.8}$$

In the above expressions, the dependence of f on $\mathbf{y}'$ and $\mathbf{x}'$ are omitted.

14.2.2. *Equations governing the outer and inner terms at orders 0 and 1*

The equations in the outer and inner problems read, from Eqs. (14.2)

$$\text{Outer problem: } \operatorname{div}_{\mathbf{x}}\mathbf{u}^0 + p^0 = 0, \quad \mathbf{u}^0 = \boldsymbol{\nabla}_{\mathbf{x}} p^0,$$

$$\text{Inner problem: } \begin{cases} \operatorname{div}_{\mathbf{y}}\mathbf{v}^0 = 0, \ \boldsymbol{\nabla}_{\mathbf{y}} q^0 = \mathbf{0}, \\ \mathbf{v}^0 = \boldsymbol{\nabla}_{\mathbf{y}} q^1 + \boldsymbol{\nabla}_{\mathbf{x}'} q^0, \\ \operatorname{div}_{\mathbf{x}'}\mathbf{v}^0 + \operatorname{div}_{\mathbf{y}}\mathbf{v}^1 + q^0 = 0, \\ \mathbf{v}^0.\mathbf{n}_{|\partial\mathcal{D}} = \mathbf{v}^1.\mathbf{n}_{|\partial\mathcal{D}} = 0. \end{cases} \tag{14.9}$$

For the inner problem, we must add, at orders 0 and 1, the periodic conditions of q^n and $\mathbf{v}^n$, $n = 0, 1$, with respect to $\mathbf{y}'$.

These are the equations that will yield the elementary problems, and, together with the matching conditions, Eqs. (14.5)–(14.6), the jump conditions in the outer problem.

14.2.3. *The jump conditions for the outer problem of order 0*

The first two equations of the inner problem in (14.9) show that q^0 does not depend on $\mathbf{y}$ and that $\int \mathrm{d}\mathbf{y}' \ v_1^0(y_1 = -\infty, \mathbf{y}', \mathbf{x}') = \int \mathrm{d}\mathbf{y}' \ v_1^0(y_1 = +\infty, \mathbf{y}', \mathbf{x}')$. This latter relation is obtained by integrating $\operatorname{div}_{\mathbf{y}}\mathbf{v}^0 = 0$ over $\mathsf{Y}_\infty \backslash \mathcal{D}$ and using the boundary condition $\mathbf{v}^0.\mathbf{n} = 0$ on $\partial\mathcal{D}$ and the periodicity of $\mathbf{v}^0$ with respect to $\mathbf{y}'$. It follows from Eqs. (14.5) that (i) $q^0(\mathbf{x}') = p^0(0^{\pm}, \mathbf{x}')$ and (ii) $\int \mathrm{d}\mathbf{y}' \ v_1^0(\pm\infty, \mathbf{y}', \mathbf{x}') = u_1^0(0^{\pm}, \mathbf{x}')$, so that the jump conditions at leading order read

$$\left[\!\left[p^0 \right]\!\right] = \left[\!\left[\mathbf{u}^0 \right]\!\right] .\mathbf{e}_1 = 0. \tag{14.10}$$

The structured film is transparent at leading order, with the pressure and the normal velocity being continuous across the interface. In the context of static elasticity, for microcraks or microholes in an elastic

body, this has been referred as the "*the principle of the dressmaker*" ([Marigo and Pideri (2011)]): "*It is not necessary to sew entirely two pieces of fabrics in order to render invisible their relative opening, it is sufficient to sew them at a great number of points regularly spaced.*" This is not trivial *a priori*, and it fact, it is not expected if we have the limiting case of a rigid wall $\ell = h$ in mind, and this will be discussed below. For now, the solution $(p^0, \mathbf{u}^0)$ at this leading order is the solution of the Helmholtz equation in the absence of interface, which means that $(p^0, \mathbf{u}^0)$ correspond to the pressure and velocity associated to a given incident wave. At this stage, the appropriate radiation condition has to be accounted for to determine the unique solution of the problem at leading order.

To capture the effect of the interface, we need to go at the next order to find the jump conditions.

14.2.4. *The jump conditions for the outer problem of order 1 and the associated elementary problems*

At order 1, we have to consider a problem on the inner terms $(q^1, \mathbf{v}^0)$. To do so, it is convenient to introduce $(q, \mathbf{v})$ defined by

$$\begin{cases} q(\mathbf{y}, \mathbf{x}') \equiv q^1(\mathbf{y}, \mathbf{x}') - y_1 \dfrac{\partial p^0}{\partial x_1}(0, \mathbf{x}'), \\ \mathbf{v}(\mathbf{y}, \mathbf{x}') \equiv \mathbf{v}^0(\mathbf{y}, \mathbf{x}') - \mathbf{u}^0(0, \mathbf{x}'). \end{cases} \tag{14.11}$$

From the equations (14.9) in the inner problem, $(q, \mathbf{v})$ satisfy

$$\begin{cases} \operatorname{div}_{\mathbf{y}} \mathbf{v} = 0, \quad \mathbf{v} = \boldsymbol{\nabla}_{\mathbf{y}} q, \\ \lim\limits_{y_1 \to \pm\infty} \mathbf{v} = 0, \\ \mathbf{v}.\mathbf{n}_{|\partial\mathcal{D}} = -\mathbf{u}^0(0, \mathbf{x}').\mathbf{n}_{|\partial\mathcal{D}}, \\ q \text{ and } \mathbf{v}.\mathbf{n} \text{ being } \mathbf{y}' - \text{periodic}, \end{cases} \tag{14.12}$$

where we have used that $q^0(\mathbf{x}') = p^0(0, \mathbf{x}')$ and $\mathbf{u}^0 = \boldsymbol{\nabla}_{\mathbf{x}} p^0$, from the outer problem. For the boundary conditions at $y_1 \to \pm\infty$, we

also used the matching condition, Eq. (14.5), $\mathbf{u}^0(0,\mathbf{x}') = \mathbf{v}^0(y_1 = \pm\infty, \mathbf{y}', \mathbf{x}')$.

If this problem is solved, the jump conditions can be determined.

Inspecting the form of $[\![p^1]\!]$ in Eqs. (14.7), we see that the jump condition on p is simply determined using the definition of q in (14.11), yielding

$$\delta q^1(y_1^m) - 2y_1^m \frac{\partial p^0}{\partial x_1}(0,\mathbf{x}') = \delta q(y_1^m), \tag{14.13}$$

with the notations (14.8).

For the jump condition on $\mathbf{u}^1$, we integrate the third equation of the inner problem in (14.9) over $\mathsf{Y}\backslash\mathcal{D}$. This leads to

$$\delta v_1^1(y_1^m) - \frac{\partial u_1^0}{\partial x_1}(0,\mathbf{x}')\,(2y_1^m - S_{\mathcal{D}}) + \frac{\partial}{\partial x_\alpha}\int_{\mathsf{Y}\backslash\mathcal{D}} \mathrm{d}\mathbf{y}\; v_\alpha = 0, \tag{14.14}$$

where $\alpha = (2,3)$, and with $S_{\mathcal{D}}$ the volume of $\mathcal{D}$ in the $\mathbf{y}$ coordinates (we use the summation convention on repeated indices). We have used that (i) $\mathbf{v}^1$ is periodic with respect to $\mathbf{y}'$, (ii) $q^0(\mathbf{x}') = p^0(0,\mathbf{x}')$ from the previous section and (iii) $\mathrm{div}_{\mathbf{x}'}\mathbf{u}^0 + p^0 = -\partial_{x_1}u_1^0$ (from the outer problem). Note that $(2y_1^m - S_{\mathcal{D}})$ is simply the volume of $\mathsf{Y}\backslash\mathcal{D}$.

It is now sufficient to take the limit $y_1^m \to +\infty$ in Eqs. (14.13)–(14.14) and using (14.7) to get

$$\begin{cases} [\![p^1]\!] = \delta q, \\ [\![\mathbf{u}^1]\!]\,.\mathbf{e}_1 = -\dfrac{\partial u_1^0}{\partial x_1}(0,\mathbf{x}')\; S_{\mathcal{D}} - \dfrac{\partial}{\partial x_\alpha}\displaystyle\int_{\mathsf{Y}_\infty\backslash\mathcal{D}} \mathrm{d}\mathbf{y}\; v_\alpha(\mathbf{y},\mathbf{x}'), \end{cases} \tag{14.15}$$

with $\mathsf{Y}_\infty = (-\infty,\infty)\times(0,1)^2$.

As written above, the jump conditions are not satisfactory in the sense that both q and v_α depend on the value of $\mathbf{u}^0$ on $\partial\mathcal{D}$ (in (14.12)). However, it is possible to split the problem (14.12) on $(\mathbf{v},q)$ into simple problems on $q^{(i)}(\mathbf{y})$, $i = 1,2,3$, called elementary problems. These problems are "elementary" in the sense that they can be solved once and for all and this is because these problems are

independent of $\mathbf{x}'$ and of the value of $\mathbf{u}^0$ on $\partial\mathcal{D}$. Specifically, $(q, \mathbf{v})$ can be decomposed into

$$\begin{cases} q(\mathbf{y}, \mathbf{x}') = \dfrac{\partial p^0}{\partial x_i}(0, \mathbf{x}')\ q^{(i)}(\mathbf{y}) + \hat{q}(\mathbf{x}'), \\ \mathbf{v}(\mathbf{y}, \mathbf{x}') = \dfrac{\partial p^0}{\partial x_i}(0, \mathbf{x}')\ \boldsymbol{\nabla} q^{(i)}(\mathbf{y}), \end{cases} \tag{14.16}$$

with $q^{(i)}(\mathbf{y})$ satisfying

$$\text{the elementary problems:} \begin{cases} \Delta q^{(i)} = 0, \\ \boldsymbol{\nabla}\left[q^{(i)} + y_i\right].\mathbf{n} = 0, \text{ on } \partial\mathcal{D}, \\ \lim\limits_{y_1 \to \pm\infty} \boldsymbol{\nabla} q^{(i)} = 0, \\ q^{(i)} \text{ being } \mathbf{y}' - \text{periodic.} \end{cases} \tag{14.17}$$

It is easy to see that such decomposition ensures that $(q, \mathbf{v})$ is solution of (14.12) (using $\mathbf{u}^0 = \boldsymbol{\nabla}_{\mathbf{x}} p^0$). Note that $\hat{q}(\mathbf{x}')$ has been introduced since $q(\mathbf{y}, \mathbf{x}')$ in (14.12) is defined up to a $\mathbf{x}'$-dependent field, but its determination is not needed.

We can now express δq and v_α in (14.15) in terms of the elementary solutions $q^{(i)}$

$$\begin{cases} \delta q = \dfrac{\partial p^0}{\partial x_i}(0, \mathbf{x}')\ \delta q^{(i)}, \\ v_\alpha(\mathbf{y}, \mathbf{x}') = u_i^0(0, \mathbf{x}')\ \dfrac{\partial q^{(i)}}{\partial y_\alpha}(\mathbf{y}), \end{cases} \tag{14.18}$$

and the jump conditions (14.15) become

$$\begin{cases} \left[\!\left[p^1\right]\!\right] = \dfrac{\partial p^0}{\partial x_i}(0, \mathbf{x}')\ \delta q^{(i)}, \\ \left[\!\left[\mathbf{u}^1\right]\!\right].\mathbf{e}_1 = -\dfrac{\partial u_1^0}{\partial x_1}(0, \mathbf{x}')\ S_{\mathcal{D}} - \dfrac{\partial u_i^0}{\partial x_\alpha}(0, \mathbf{x}') \displaystyle\int_{\mathsf{Y}_\infty \setminus \mathcal{D}} \mathrm{d}\mathbf{y}\ \dfrac{\partial q^{(i)}}{\partial y_\alpha}(\mathbf{y}), \end{cases} \tag{14.19}$$

which can be written introducing the parameters $\mathbf{b}$ and c

$$\begin{cases} [\![p^1]\!] = b_i \, \dfrac{\partial p^0}{\partial x_i}(0, \mathbf{x}'), \\ [\![\mathbf{u}^1]\!] \, . \mathbf{e}_1 = \mathsf{c}_{ij} \, \dfrac{\partial u_i^0}{\partial x_j}(0, \mathbf{x}'), \end{cases} \tag{14.20}$$

with $b_i = \delta q^{(i)}$ and

$$\begin{cases} \mathsf{c}_{11} = -S_{\mathcal{D}}, \, \mathsf{c}_{1\beta} = -\displaystyle\int_{\mathsf{Y}_\infty \setminus \mathcal{D}} \mathrm{d}\mathbf{y} \, \frac{\partial q^{(1)}}{\partial y_\beta}, \\ \mathsf{c}_{\alpha 1} = 0, \qquad \mathsf{c}_{\alpha\beta} = -\displaystyle\int_{\mathsf{Y}_\infty \setminus \mathcal{D}} \mathrm{d}\mathbf{y} \, \frac{\partial q^{(\alpha)}}{\partial y_\beta}, \end{cases} \tag{14.21}$$

where α, β take the values 2 and 3.

It is essential that the elementary problems, Eqs. (14.17), do not depend on the incident wave (as the former problem Eqs. (14.12), does, through $\mathbf{u}^0$ in the boundary condition). This ensures that the parameters $\mathbf{b}$ and c are characteristic of the interface independently of the particular scattering problem that will be considered.

14.2.5. *Construction of the up to order 1 effective problem*

The final up to order 1 effective problem requires two additional steps. First, we have obtained jump conditions at order 0 and at order 1. Strictly, an iterative resolution of the problem is possible which would involve first the resolution of the order 0 outer problem (which does not see the structured layer) yielding $(\mathbf{u}^0, p^0)$, and then the resolution of the order 1 outer problem associated to the jump conditions (14.20) (the effective parameters entering in (14.20) being known after the elementary problems (14.17) have been solved). However, it is stressed in [David *et al.* (2012)] that such procedure is not suitable in practice, notably if we have numerical resolutions in mind, and rather, a unique outer problem involving $p^0 + \varepsilon p^1$ (and $u^0 + \varepsilon u^1$) has to be considered. To do so, we use $f = f^0 + \varepsilon f^1 + O(\varepsilon^2)$ and $[\![f]\!] = \varepsilon \, [\![f^1]\!] + O(\varepsilon^2)$ for any field f with $[\![f^0]\!] = 0$. From Eqs. (14.10)

and (14.20), we get

$$\begin{cases} [\![p]\!] = \varepsilon \, \dfrac{b_i}{2} \left[\dfrac{\partial p}{\partial x_i}(0^-, \mathbf{x}') + \dfrac{\partial p}{\partial x_i}(0^+, \mathbf{x}') \right] + O(\varepsilon^2), \\ [\![\mathbf{u}]\!] \cdot \mathbf{e}_1 = \varepsilon \, \dfrac{\mathsf{c}_{ij}}{2} \left[\dfrac{\partial u_i}{\partial x_j}(0^-, \mathbf{x}') + \dfrac{\partial u_i}{\partial x_j}(0^+, \mathbf{x}') \right] + O(\varepsilon^2), \end{cases} \tag{14.22}$$

where we have introduced for fields being discontinuous at $x_1 = 0$ the mean value of their limits $x_1 \to 0^{\pm}$.

Next, the above conditions will be written in an alternative form, equivalent at order $O(\varepsilon^2)$. Loosely speaking, the new jump conditions have to restore the initial thickness of the structured layer, and this is required following the intuitive argument that the equivalent interface has to support the acoustic energy in the actual layer in order to ensure that it supports a positive energy, see [Marigo and Maurel (2016)] for more details. In [Abdelmoula *et al.* (1998)], this is done by considering that the problem is written in a domain of fixed size in which the film of size ε has been inserted, resulting in a shift of the boundaries. In [Delourme (2010); Delourme *et al.* (2012)], this is done to overcome troubles in the homogenized problem due to the signs of the (b_i, c_{ij}), specifically some of these parameters have to be positive or negative to ensure existence and uniqueness of variational solutions in the homogenized problem. In this latter case, a family of equivalent jump conditions are proposed, which correspond to interfaces "enlarged" of an arbitrary thickness; here, we consider a shift of e. A shift of e in the real space corresponds to a shift of ke in the $\mathbf{x}$-space and we assume that $e/h = O(1)$, from which $ke = \varepsilon e/h = O(\varepsilon)$. To do so, we use the Taylor expansion of p^n, $n = 0, 1$,

$$\begin{cases} p^0 \left(\pm \dfrac{\varepsilon e}{2h}, \mathbf{x}' \right) = p^0(0, \mathbf{x}') \pm \dfrac{\varepsilon e}{2h} \dfrac{\partial p^0}{\partial x_1} \left(\pm \dfrac{\varepsilon e}{2h}, \mathbf{x}' \right) + O(\varepsilon^2), \\ p^1 \left(\pm \dfrac{\varepsilon e}{2h}, \mathbf{x}' \right) = p^1(0^{\pm}, \mathbf{x}') + O(\varepsilon), \end{cases} \tag{14.23}$$

to get

$$p\left(\frac{\varepsilon e}{2h},\mathbf{x}'\right)-p\left(-\frac{\varepsilon e}{2h},\mathbf{x}'\right)=\frac{\varepsilon e}{2h}\left[\frac{\partial p^0}{\partial x_1}\left(-\frac{\varepsilon e}{2h},\mathbf{x}'\right)+\frac{\partial p^0}{\partial x_1}\left(\frac{\varepsilon e}{2h},\mathbf{x}'\right)\right]+\varepsilon\left[\!\left[p^1\right]\!\right]+O(\varepsilon^2). \tag{14.24}$$

Finally, defining for any f, the mean value of f across the interface of thickness ke

$$\overline{f}=\frac{1}{2}\left[f\left(-\frac{\varepsilon e}{2h},\mathbf{x}'\right)+f\left(\frac{\varepsilon e}{2h},\mathbf{x}'\right)\right], \tag{14.25}$$

and coming back to (14.26), the new jump conditions read

$$\begin{cases}\left[\!\left[p\right]\!\right]_e=\varepsilon\ \left(b_i+\dfrac{e}{h}\,\delta_{i1}\right)\overline{\dfrac{\partial p}{\partial x_i}}+O(\varepsilon^2),\\ \left[\!\left[\mathbf{u}\right]\!\right]_e.\mathbf{e}_1=\varepsilon\ \left(\mathsf{c}_{ij}+\dfrac{e}{h}\,\delta_{i1}\delta_{j1}\right)\overline{\dfrac{\partial u_i}{\partial x_j}}+O(\varepsilon^2).\end{cases} \tag{14.26}$$

Now, the homogenized model is written in the real space for the pressure field $P(\mathbf{X})$ (Fig. 14.3): P satisfies the Helmholtz equation $\Delta P+k^2P=0$ except in $-e/2<X_1<e/2$ (this region is disregarded) and it experiences jump conditions which link its values at $X_1=\pm e/2$, with $\left[\!\left[P\right]\!\right]_e\equiv P(e/2,\mathbf{X}')-(-e/2,\mathbf{X}')$, to the mean values of its spatial derivatives $\overline{\partial_{X_i}P}\equiv\frac{1}{2}[\partial_{X_i}P(-e/2,\mathbf{X}')+\partial_{X_i}P(e/2,\mathbf{X}')]$

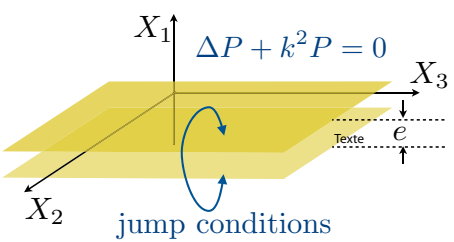


Fig. 14.3. In the homogenized problem, the Helmholtz equation applies in the whole-considered-space except in the region $-e/2<X_1<e/2$ which is disregarded; the jump conditions (14.27) apply between $X_1=-e/2$ and $e/2$.

(the same for $\partial_{X_1} P$):

$$\boxed{\text{Jump conditions} \begin{cases} [\![P]\!]_e = h\, B_i\, \overline{\dfrac{\partial P}{\partial X_i}}, \\ \left[\!\!\left[\dfrac{\partial P}{\partial X_1} \right]\!\!\right]_e = h\, \mathsf{C}_{ij}\, \overline{\dfrac{\partial^2 P}{\partial X_i \partial X_j}}, \end{cases}} \tag{14.27}$$

with

$$\begin{cases} B_1 = \dfrac{e}{h} + \delta q^{(1)},\ B_\alpha = \delta q^{(\alpha)}, \\ \mathsf{C}_{11} = \dfrac{e}{h} - S_{\mathcal{D}}, \quad \mathsf{C}_{1\beta} = -\displaystyle\int_{\mathsf{Y}_\infty \setminus \mathcal{D}} \mathrm{d}\mathbf{y}\ \frac{\partial q^{(1)}}{\partial y_\beta}, \\ \mathsf{C}_{\alpha 1} = 0, \qquad\quad \mathsf{C}_{\alpha\beta} = -\displaystyle\int_{\mathsf{Y}_\infty \setminus \mathcal{D}} \mathrm{d}\mathbf{y}\ \frac{\partial q^{(\alpha)}}{\partial y_\beta}, \end{cases} \tag{14.28}$$

where α, β take the values 2 and 3.

14.2.6. *Remarks*

As previously said, it is essential that the elementary problems, Eqs. (14.17), do not depend on the incident wave. This ensures that the interface parameters are characteristic of the structured film only. However, one can think to introduce a dependance on the frequency in the elementary problems, as it has been done for classical homogenization (high frequency homogenization or resonant homogenization) [Bouchitté and Felbacq (2004); Felbacq and Bouchitté (2005); Craster *et al.* (2011); Antonakakis *et al.* (2013)]. For acoustic waves, this could be done for metafilms containing long hard scatterers wrapped in the unit cell, thus allowing for the resonance of the enrolled element. For elastic or electromagnetic waves, this could be done for metafilms containing inclusions with a contrast in its material properties such that the wavelength inside the inclusions is of the same order than the size of the inclusions.

The following remarks can be done on the elementary problems, Eqs. (14.17)

- The problem for $i = 1$ is nothing but the problem of a uniform potential flow past a rigid obstacle in a duct. With velocity potential $\varphi(\mathbf{y}) = q^{(1)} + y_1$, we get $\Delta\varphi = 0, \partial_n\varphi_{|\partial\mathcal{D}} = 0$ and $\lim_{y_1\to\pm\infty}\varphi = y_1 + b^{\pm}$. The coefficient $b_1 = b^+ - b^-$ is called the blockage coefficient and it has been calculated for several shapes of obstacles (see [Martin and Dalrymple (1988)] and references therein). This coefficient can be non zero even for zero surface obstacle $S_{\mathcal{D}} = 0$, reducing to a plate of length ℓ.
- For inclusion shape being symmetric with respect to y_2, $q^{(1)}$ is symmetric also. It results that $\int_{\mathsf{Y}\setminus\mathcal{D}} \mathrm{d}\mathbf{y}\, \partial_{y_\alpha} q^{(1)} = 0$.
- The cell problem for $i = 2, 3$ is needed only for oblique incidence of the wave. Besides, for inclusion shape being symmetric with respect to y_2, $q^{(\alpha)}$, $\alpha = 2, 3$ are antisymmetric; in this case, $\delta q^{(\alpha)} = 0$.

Finally, we have said that the present derivation of the jump conditions follows from previous works [Marigo and Pideri (2011); David *et al.* (2012)] in the static case ($k = 0$). For the sake of completeness, we report in the Appendix 14.5 a short discussion on the links with this work.

14.3. The Case of Sound Hard Rectangular Inclusions in 2D

In the following, we consider a 2 dimensional array made of rectangular inclusions infinite along X_3 (otherwise of size ℓ along X_2 and thickness e along X_1). This inclusion shape being symmetric with respect to y_2, the preceding remarks on the elementary problems apply and the jump conditions (14.27) *simplify to*

$$\text{Jump conditions} \begin{cases} [\![P]\!]_e = hB\, \overline{\dfrac{\partial P}{\partial X_1}}, \\ \left[\!\!\left[\dfrac{\partial P}{\partial X_1}\right]\!\!\right]_e = hS\, \overline{\dfrac{\partial^2 P}{\partial X_1^2}} + hC\, \overline{\dfrac{\partial^2 P}{\partial X_2^2}}, \end{cases} \tag{14.29}$$

with $S \equiv (e/h - S_{\mathcal{D}})$, $B \equiv e/h + \delta q^{(1)}$ and $C \equiv -\int_{\mathsf{Y}_\infty \setminus \mathcal{D}} \mathrm{d}\mathbf{y}\ \partial_{y_2} q^{(2)}$. As previously, $\overline{f}$ means the mean value of f at $X_1 = \pm e/2$.

14.3.1. *The interface parameters*

An obvious interface parameter is $S = e/h(1 - \ell/h)$. Next, (B, C) can be calculated by solving numerically the elementary problems, Eqs. (14.17). A simple procedure to calculate (B, C) (based on mode matching) and the associated scripts are provided in the Appendice 14.6. Alternatively, we can use approximate but explicit expressions. An estimate of C is, as a rule of thumb,

$$\begin{cases} C \simeq e\left(1 - \dfrac{\ell}{h}\right) - \dfrac{\pi}{8}\left(1 - \dfrac{\ell}{h}\right)^2, & \text{if this leads to a positive value,} \\ C \simeq 0, & \text{otherwise,} \end{cases} \tag{14.30}$$

which means that C (which is always positive) can be neglected for small e/h values. A more precise estimate of B can be found in [Flagg and Newman (1971)] (see also [Suh *et al.* (2011)]). It is written as the sum of two terms, one of which does not vanish for $e = 0$ (blockage coefficient of a flat plate)

$$\begin{cases} B = \dfrac{e}{h - \ell} + B_0, \\ B_0 = \dfrac{2}{\pi} \log\left(\dfrac{d}{2} + \dfrac{1}{2d}\right), \quad \text{with } d \equiv \tan\left(\dfrac{\pi(h - \ell)}{4h}\right). \end{cases} \tag{14.31}$$

In fact, Ref. [Flagg and Newman (1971)] provides an approximation of $B_0 \simeq \frac{2}{\pi}\left[1 - \log\left(4(1 - \ell/h)\right) + \frac{1}{3}(1 - \ell/h)^2 + \frac{281}{180}(1 - \ell/h)\right]$ valid for ℓ/h close to 1. The expression of B_0 in Eq. (14.31) above can be found in [Morse and Ingard (1968)] (the calculation is done for flat plates using techniques of complex variables). The parameters (B, C) are reported in Fig. 14.4.

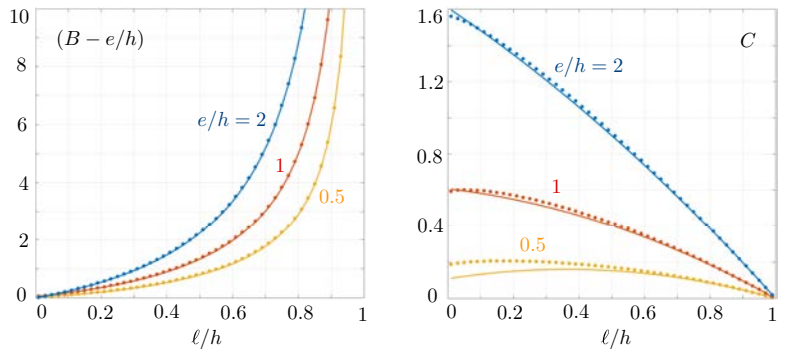


Fig. 14.4. Interface parameters B and C computed numerically by solving the elementary problems (14.17) (symbols), see Appendix 14.6. (a) $(B - e/h)$ as a function of ℓ/h for $e/h = 0.5$, 1 and 2; plain lines show the values given by (14.31). Same representation for C, with plain lines given by (14.30).

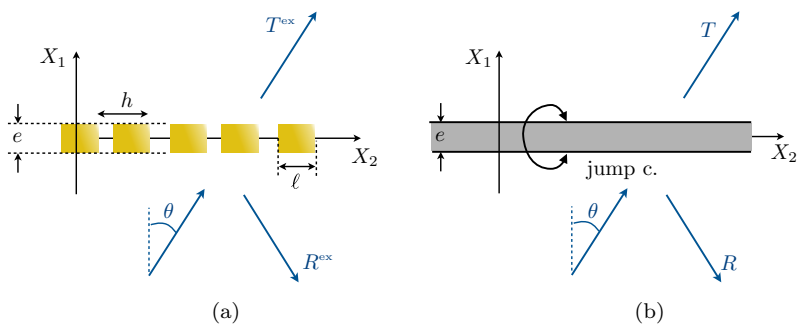


Fig. 14.5. Scattering of a plane wave at oblique incidence θ (a) in the real problem, (R^{ex}, T^{ex}) refer to the reflection and transmission coefficients of the incident plane wave, characteristic of the propagating waves and an evanescent near field is generated in the vicinity of the film, (b) in the homogenized problem, the jump conditions (14.29) do not produce evanescent field and (R, T) refers simply to the reflection and transmission coefficients of the incident wave.

14.3.2. *Validity of the interface model — The scattering of an incident plane wave*

We consider the problem of the scattering of an incident plane wave at oblique incidence θ on the structured film (Fig. 14.5). When replaced by the equivalent interface associated to the jump conditions Eqs. (14.29), the problem reduces to a one dimensional problem

whose solution reads

$$\begin{cases} P(X_1 < -e/2, X_2) = e^{\mathrm{i}k\sin\theta X_2}[e^{\mathrm{i}k\cos\theta(X_1-e/2)} + Re^{-\mathrm{i}k\cos\theta(X_1-e/2)}], \\ P(X_1 > e/2, X_2) = e^{\mathrm{i}k\sin\theta X_2}\ Te^{\mathrm{i}k\cos\theta(X_1+e/2)}, \end{cases} \tag{14.32}$$

with

$$\begin{cases} R = \mathrm{i}\,\dfrac{a-b}{(1-\mathrm{i}a)(1-\mathrm{i}b)}, \quad T = \dfrac{1+ab}{(1-\mathrm{i}a)(1-\mathrm{i}b)}, \\ \text{with } a \equiv \dfrac{kh}{2}\cos\theta\ \left(S + C\tan^2\theta\right),\ b \equiv \dfrac{kh}{2}\cos\theta\ B. \end{cases} \tag{14.33}$$

Note that for $\ell \to h$, we have $B \to +\infty$ (thus $b \to +\infty$) and $C \to 0$. We obtain $R \to -1/(1-\mathrm{i}a)$ and $T \to \mathrm{i}a/(1-\mathrm{i}a)$, with $a \propto S \neq 0$ except for a flat array $e = 0$. Obviously, we would expect $R = 1$, $T = 0$ in this case since the array becomes a sound hard wall. It is a classical problem in homogenization theories when a new small parameter is introduced, here $\varepsilon' = 1 - \ell/h$. In the present case, the problem $\varepsilon' \to 0$ appears in the jump conditions, Eq. (14.26); it has been assumed that $\varepsilon B \ll 1$ (as it has been assumed that $\varepsilon C \ll 1$ and $\varepsilon S \ll 1$). Inspecting B in Eq. (14.31), we see that this assumption can fail when $(h - \ell) \to 0$, depending on how fast ε and ε' go to zero: (i) if $e \neq 0$, it fails if $\varepsilon/\varepsilon' \to \infty$ and (ii) if $e = 0$, it fails if $\varepsilon \log \varepsilon' \to \infty$. This latter limit corresponds to the one established in [Sanchez-Hubert and Sanchez-Palencia (1982)] for perforated thin plate.

In the actual problem, the incident wave generates an evanescent field in the vicinity of the structured film, which may have a strong influence on the reflection and transmission of the plane wave. Below, we inspect the validity of our homogenized solution, namely its ability to incorporate the effect of the evanescent field in the interface parameters.

To begin with, we report in Figs. 14.6 and 14.7 the pressure fields for a structured interface of thickness e, with $ke = 0.5$ and 1 and, in the unit cell, different aspect ratio ℓ/h. For the considered geometries, the interface parameters are given in the table below. The notation X [Y] is used, with X the value of B or C given by the estimates

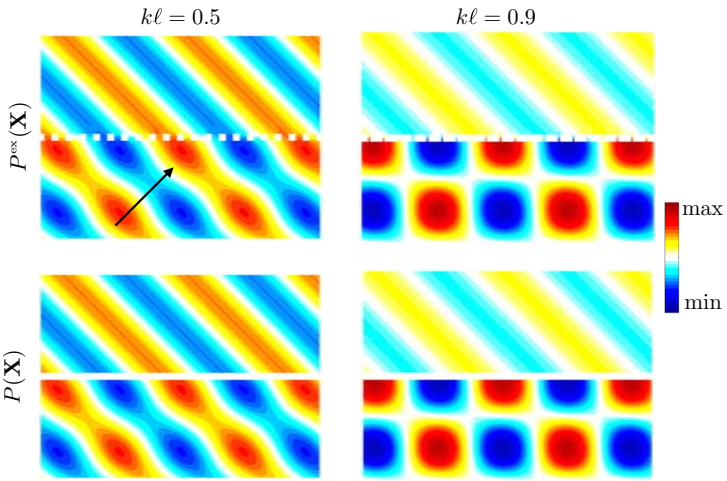


Fig. 14.6. Wavefields in the (X_1, X_2) plane for $ke = 0.5$, $kh = 1$, and $k\ell = 0.5$ or 0.9 (a) $P^{\text{ex}}(\mathbf{X})$ and (b) $P(\mathbf{X})$ coming from the interface model, Eqs. (14.32)–(14.33). The white arrow indicates the wavenumber of the incident plane wave.

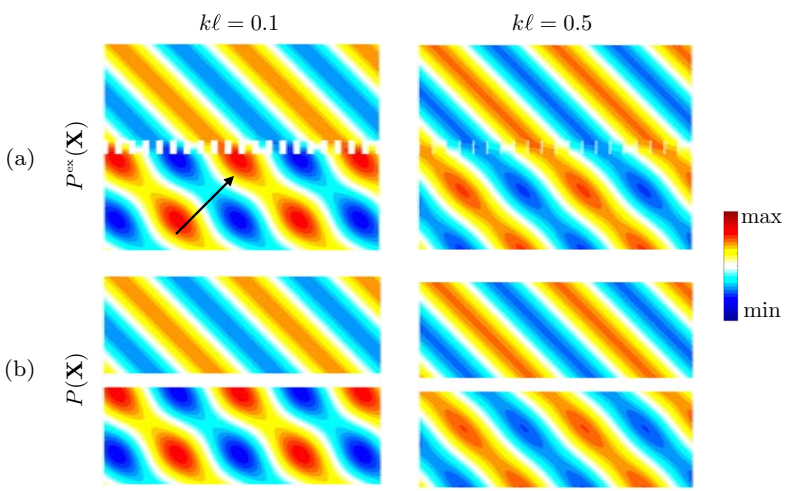


Fig. 14.7. Same representation as in Fig. 14.6 for $ke = 1$ $kh = 1$, and $k\ell = 0.1$ or 0.5.

(14.30)–(14.31), and Y the value obtained by solving the associated elementary problem (see Fig. 14.4 and Appendix 14.6).

$e/h = 0.5$	$\ell/h = 0.5$	$\ell/h = 0.9$
B	1.22 [1.25]	6.18 [6.22]
C	0.15 [0.17]	0.05 [0.05]

$e/h = 1$	$\ell/h = 0.1$	$\ell/h = 0.5$
B	1.12 [1.12]	2.22 [2.25]
C	0.58 [0.60]	0.40 [0.41]

The field $P^{\text{ex}}(\mathbf{X})$ refers to the solution of the actual problem calculated numerically (a multimodal method is used, see [Maurel *et al.* (2014)]) and $P(\mathbf{X})$ refers to the prediction coming from the interface model, Eq. (14.32), with Eqs. (14.33). We have considered $\theta = \pi/4$, $kh = 1$. This $\varepsilon = kh = 1$ value is obviously a limiting case for the validity of any homogenization model but the fields P^{ex} and P are found in good agreement, with $||P - P^{\text{ex}}||/||P^{\text{ex}}|| \sim 1\%$ for $ke = 0.5$ (in both cases reported in Figs. 14.6) and $||P - P^{\text{ex}}||/||P^{\text{ex}}|| \sim 5\%$ for $ke = 1$ (in both cases reported in Figs. 14.7); $||.||$ is the norm L_2, considering the field in $|X_1| > e/2$.

To go further, we inspect the variations of the reflection coefficient as a function of k and e (say for a fixed h value). This is done by comparing R in Eq. (14.33) (blue dotted lines in Figs. 14.8) with R^{ex} computed numerically (plain orange lines). We also report the behavior of $\tilde{R}$ (grey dotted lines) given by the classical homogenization of layered media (see Appendix 14.7.1). In this approach, the propagation inside the structured layer is explicitly accounted for (while the boundary layer effects are disregarded) and the structured film is replaced by an equivalent film of thickness e filled with an homogeneous anisotropic medium. Classical homogenization has been shown efficient to predict the extraordinary optical transmission at the Brewster angle [Maurel *et al.* (2013); Akarid *et al.* (2014)] and the behavior of "spoof plasmons", which are the waves guided

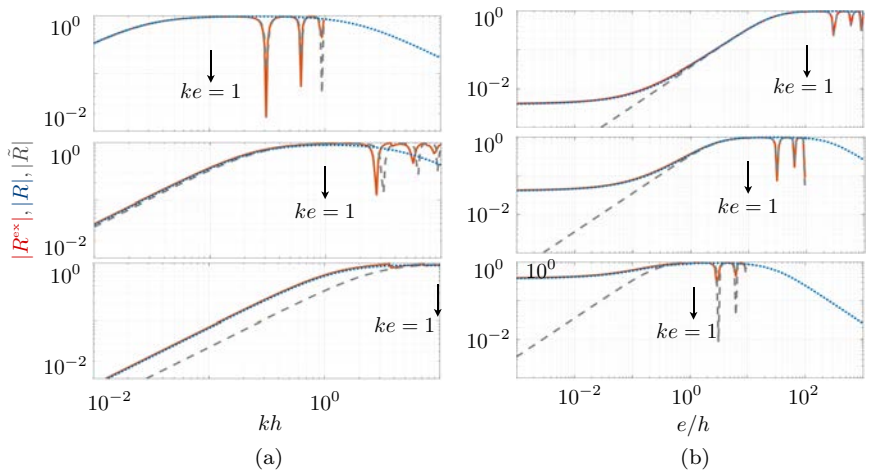


Fig. 14.8. Variations of the reflection coefficients R^{ex} calculated numerically (orange plain lines), R given by the interface model, Eq. (14.33) (blue dotted lines), and $\tilde{R}$ given by classical homogenization, Eq. (14.34) (grey dotted lines). (a) as a function of kh for $e/h = 10,\ 1,\ 10^{-1}$ (from the top to the bottom), and (b) as a function of e/h for $kh = 10^{-2},\ 10^{-1},\ 1$ (from the top to the bottom). In all plots, $\ell = 0.9h$ and $\theta = \pi/4$.

along the boundaries of a thick structured array [Garcia-Vidal *et al.* (2005); Mercier *et al.* (2015)]. It has also been applied for the design of a structure producing a rainbow type effect [Zhu *et al.* (2013)] (see the supplementary material in this reference). In the case of sound hard layers, this leads to a reflection coefficient $\tilde{R}$

$$\tilde{R} = \frac{(c^2 - 1)e^{\mathrm{i}ke\cos\theta/2}}{[c - \mathrm{i}\tan(ke/2)]\,[c + \mathrm{i}/\tan(ke/2)]}, \tag{14.34}$$

where $c = h\cos\theta/(h - \ell)$ (the derivation is given in the Appendix 14.7.1).

From Figs. 14.8, the ranges of validity of both homogenizations can be defined: the interface homogenization is valid for $ke, kh < 1$, while the classical homogenization is valid for $e/h > 1$, specifically $ke > 1 > kh$ (inspecting higher values of kh would reveal that both homogenizations fail for $kh > 1$, as expected). The error in $\tilde{R}$ for

small e/h is attributable to the behavior of B when $ke \to 0$, namely

$$\lim_{e\to 0} R \simeq -\mathrm{i}\frac{kh}{2}\cos\theta B_0, \tag{14.35}$$

and $B_0 \neq 0$ while $\tilde{R} \to 0$ (a discussion on the jump conditions obtained from the classical homogenization in the limit $ke \to 0$ is presented in the Appendix 14.7.2). Thus, classical homogenization underestimates the scattering properties of thin films.

14.3.3. *Application to the wave propagation in a guide obstructed by a hard perforated screen*

In addition to yield explicit solutions of several scattering problems, homogenization approaches are of particular interest if we have numerical resolutions in mind. This is particularly true if the problem under consideration involves very different scales between the largest one and the smallest one, and this is what we exemplify in this section. We consider the high frequency propagation of an acoustic wave in a rigid guide of height H, with $kH = 100$ (Fig. 14.9). The guide is obstructed, from wall to wall, by a rigid perforated screen with $h = H/100$ and $e = H/2000$ (thus, $kh = 1$ and $ke = 0.05$). Thus, the smallest scale is e and the largest scale is H, with $H = 2000e$ and this is demanding in terms of any computational method.

In general, the solution of guided wave problems can be expanded in modes, associated to transverse functions $\varphi_n(X_2)$ with vanishing

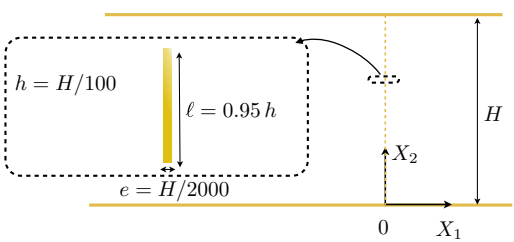


Fig. 14.9. Configuration of the rigid waveguide obstructed by a perforated screen at $-e/2 < X_1 < e/2$. The screen has a periodic spacing $h = H/100$ and thickness $e = H/2000$. The height of the rigid material in the screen is $\ell/h = 0.95$. The guide is studied in the high frequency regime, with $kH = 100$.

first derivative at $X_2 = 0, H$, namely

$$\begin{cases} P(X_1, X_2) = \displaystyle\sum_{n=0}^{N} P_n(X_1)\ \varphi_n(X_2), \\ \varphi_n(X_2) = \sqrt{2 - \delta_{n0}} \cos\left(\dfrac{n\pi X_2}{H}\right), \end{cases} \tag{14.36}$$

(the $\sqrt{2}$ stands for proper normalization of the transverse functions). At this stage, P stands for the solution of the actual problem or for the solution of the homogenized one. For $N \to +\infty$, (14.36) is the modal expansion of the solution, the P_n being the projections of P on the basis φ_n; in the numerics, a finite and sufficiently large N-value is used within a multimodal approach which accounts for the mode coupling due the presence of the screen, see [Maurel and Mercier (2012); Maurel *et al.* (2014)].

However, in the homogenized problem, no coupling is expected. This is because the jump conditions, Eqs. (14.29), apply for each mode owing to the orthogonality of the transverse functions, and they take the form

$$\begin{cases} [\![P_n]\!] = hB\overline{P_n'}, \\ [\![P_n']\!] = hS\overline{P_n''} - hC\left(\dfrac{n\pi}{H}\right)^2 \overline{P_n}, \end{cases} \tag{14.37}$$

where we used that $\varphi_n'' = -(n\pi/H)^2\varphi_n$. Next, in the parts of the waveguide free of scatterer, each mode satisfies $P_n'' + k_n^2 P_n = 0$, with

$$k_n \equiv \sqrt{k^2 - \left(\frac{n\pi}{H}\right)^2} \tag{14.38}$$

the component of the wavenumber along X_1. Except at the interface $-e/2 < X_1 < e/2$, this is the case and we can write for each P_n

$$P_n(X_1) = \begin{cases} P_n^{\text{inc}} \left[e^{\mathrm{i}k_n(X_1+e/2)} + r_n e^{-\mathrm{i}k_n(X_1+e/2)}\right], & \text{for } X_1 < -e/2, \\ P_n^{\text{inc}}\ t_n\ e^{\mathrm{i}k_n(X_1-e/2)}, & \text{for } \ X_1 > e/2, \end{cases} \tag{14.39}$$

with (r_n, t_n) the reflection and transmission coefficients of the mode n. Now, it is sufficient to apply the jump conditions, Eqs. (14.37), to get (r_n, t_n)

$$\begin{cases} r_n = \mathrm{i}\,\dfrac{a_n - b_n}{(1-\mathrm{i}a_n)(1-\mathrm{i}b_n)}, \quad t_n = \dfrac{1+a_n b_n}{(1-\mathrm{i}a_n)(1-\mathrm{i}b_n)}, \\ \text{with } a_n \equiv \dfrac{k_n h}{2}\left[S + C\left(\dfrac{n\pi}{k_n H}\right)^2\right],\ b_n \equiv \dfrac{k_n h}{2} B. \end{cases} \tag{14.40}$$

The interface parameters for this structuration ($e/h = 0.05$, $\ell/h = 0.95$) are : $B = 2.62[2.64]$, $C = 1.5\ 10^{-3}[1.7\ 10^{-3}]$ and $S = 2.5\ 10^{-3}$.

Figures 14.10 and 14.11 report the wavefields $P^{\mathrm{ex}}(\mathbf{X})$ calculated numerically and $P(\mathbf{X})$ given by the explicit expressions, Eqs. (14.36) to (14.40). The wavefield in the absence of the film for the same incident wave $P^{\mathrm{inc}}(X_1, X_2) = P_n^{\mathrm{inc}} e^{ik_n X_1} \varphi_n(X_2)$ is reported for comparison (in both cases, P^{inc} is a combination of the 100 first modes). The fields $P(\mathbf{X})$ and $P^{\mathrm{ex}}(\mathbf{X})$ are found in good agreement (with a relative discrepancy of about 1% in both cases), which confirms that the jump conditions correctly account for the effect of the structured screen.

Let us now inspect the reflection and transmission coefficients mode by mode. In the numerics, the scattering properties are encapsulated in matrices T and R being $N \times N$ matrices. We calculated R and T at the interfaces of the film, namely $P_m^{\mathrm{ex}}(-e/2) = [\delta_{mn} + \mathsf{R}_{mn}] P_n^{\mathrm{inc}}(-e/2)$, and $P_m^{\mathrm{ex}}(e/2) = \mathsf{T}_{mn} P_n^{\mathrm{inc}}(-e/2)$. We have checked that the scattering matrices are essentially diagonal (the relative weight of the off diagonal terms is about 10^{-4}), which confirms that the mode coupling is negligible. Figures 14.12 show the transmission and reflection coefficients (r_n, t_n) from Eqs. (14.40) and the diagonal terms $(\mathsf{R}_{nn}, \mathsf{T}_{nn})$ (the scattering properties do not depend on the source, so they are the same in the cases of the Figs. 14.10 and 14.11). The agreement is relatively good, with 1% and 5% respectively for the reflection and transmission coefficients, averaged for $n = 1, \ldots 100$.

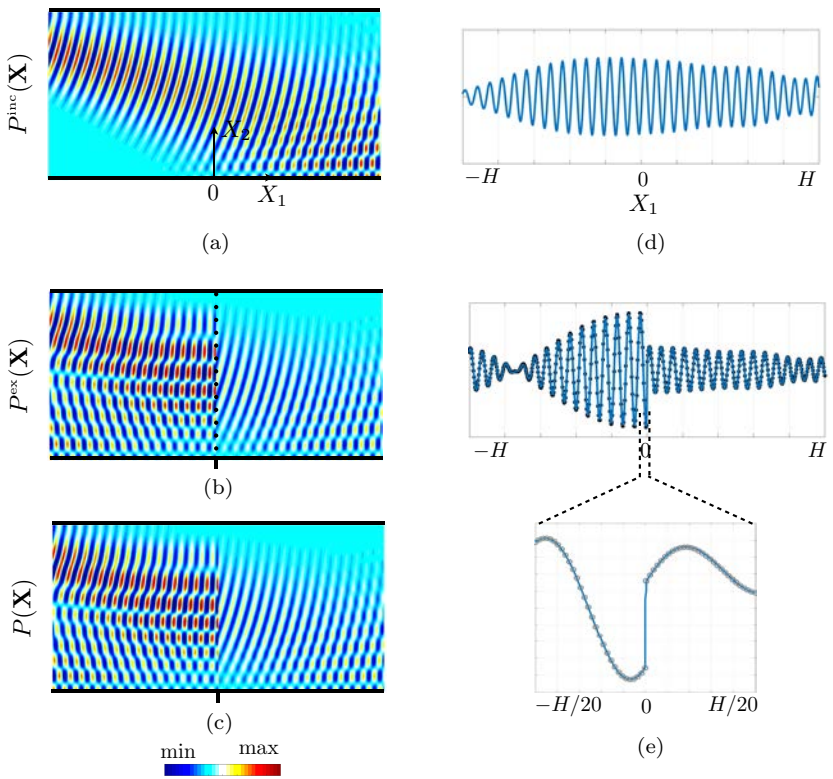


Fig. 14.10. Scattering by the perforated screen in the guide (geometry of Fig. 14.9). (a) Wavefield $P^{\rm inc}$ in the absence of the film (b) and (c) Wavefields $P^{\rm ex}$ calculated numerically and P in the homogenized problem, Eqs. (14.36) to (14.40). (d) shows the profile $P^{\rm inc}(X_1, H/2)$ (in arbitrary unit) and (e) show the profiles $P^{\rm ex}(X_1, H/2)$ (blue line) and $P(X_1, H/2)$ (black symbols); the discontinuity of the field across the structured film is visible in the zoom of the profile for $-H/20 < X_1 < H/20$.

As can be expected, the full wave calculations have been demanding. Indeed, it requires 3 different scales to be resolved, the largest one H, the intermediate one, given by the incident wavenumbers with k_n, $n = 0$ to 100, from which the typical wavelength is about $H/100$, and the smallest one, given by $e = H/2000$. This latter small scale imposes the truncation N in the numerics, which means that $|k_N|e \sim 1$. Thus $N\pi e/H \sim 1$, leading to $N \sim H/e = 2000$. In general,

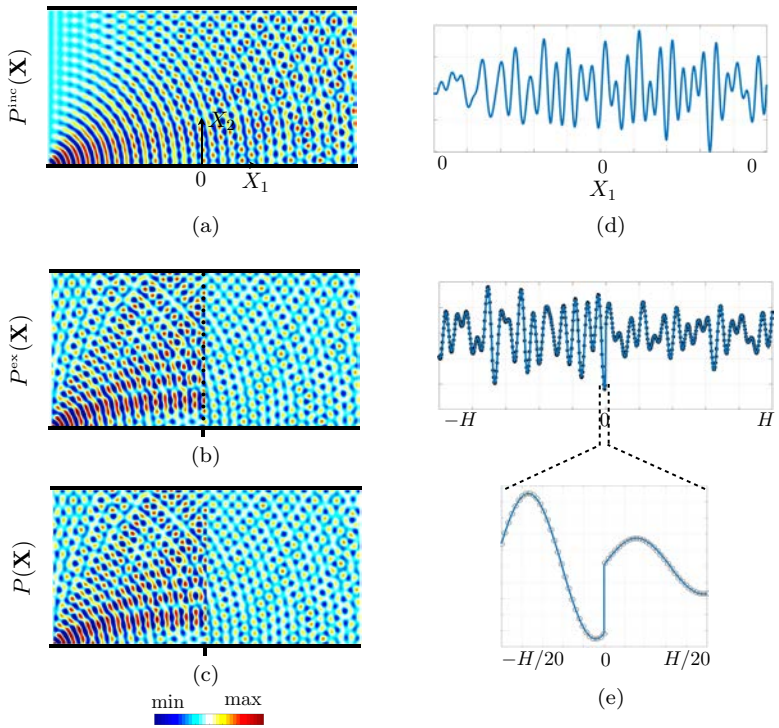


Fig. 14.11. Same representation as in Fig. 14.10 for an incident wave of beam type.

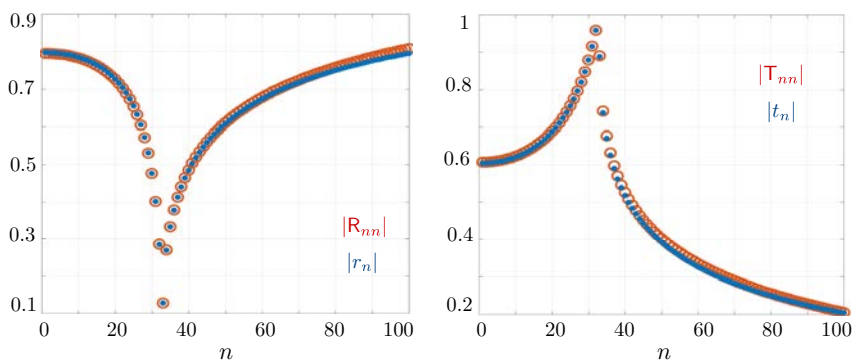


Fig. 14.12. Comparison of the reflection and transmission coefficients of the modes, $(\mathsf{R}_{nn}, \mathsf{T}_{nn})$ for $n = 1, \ldots, 100$ from direct numerics (computed with a truncation $N = 3000$), and (r_n, t_n) given by the interface model, Eqs. (14.40).

the smallest scale corresponds to the highest evanescent mode excited in the near field of the structured film, and indeed, $N = 2000 - 3000$ has been needed to reach a reasonably converged solution P^{ex} (for instance, the variation in P^{ex}, Eq. (14.36), from $N = 2000$ to 3000 is about 0.8 %). This need in a high truncation is expected and not expected. Indeed, it has been said, and it has been checked, that no mode coupling occurs in our configuration (see Appendix 14.9). Thus, in the present cases, the modes at $n > 100$ are not expected to contribute significantly to the total field since we have considered an incident wave being a combination of the first 100 modes only. The reason why high truncation is needed is in fact related to the convergence of R_{nn} and T_{nn} with N, a convergence which is often disregarded; the Appendix 14.9 presents a discussion on this aspect.

14.4. Concluding Remarks

We have presented an interface model to identify the effective behavior of a thin structured film composed of sound hard inclusions. Parameters characteristic of the interface enter in jump conditions for the acoustic pressure field and for the acoustic normal velocity. They are given by the resolution of elementary problems written in the static limit, and they are therefore wave independent by construction. We have validated this model in the case of a layer made of rectangular inclusions and for a plane wave at oblique incidence on the structured layer. Explicit expressions of the scattering coefficients deduced from the interface parameters have been shown to be accurate with a range of validity being $kh < 1$ and $ke < 1$. Results have been compared to the scattering coefficients given by classical homogenization of bulk material; these latter have a range of validity $kh < 1$ and $e/h > 1$, and largely underestimates the scattering properties of thin structured layers. Finally, the validity and the interest of such jump condition has been exemplified in the more involved case of an ultrathin structured layer obstructing a guide from wall to wall (there, due to the high frequency propagation in the guide, the numerical cost is heavy).

The present model is applicable to more involved geometries of the inclusions and to inclusions being penetrable for the wave as well as for three dimensional cases. More generally, it can be extended to a large class of wave problems, in acoustics, in elasticity and in electromagnetism.

14.5. Appendix A — The Present Derivation as a Simplified Case of a Previous One

In [Marigo and Pideri (2011)], the problem of an elastic body containing microcracks or micro holes periodically located on a surface is considered. The elastic body is submitted to external forces (or boundary conditions) on its boundaries, which impose variations of the external stresses on the scale which is typically the size of the body; this latter is assumed to be much larger than the spacing between the holes. The size of the body in the elastic problem is the equivalent of the incident wavelength in our acoustic problem.

The increased complexity in this work is that it requires to deal with vectors instead of scalars $p, q \rightarrow \mathbf{U}, \mathbf{V}$ (the elastic displacement vector) and a tensor instead of a vector $\mathbf{u}, \mathbf{v} \rightarrow \boldsymbol{\sigma}, \boldsymbol{\tau}$ (the stress tensor). Otherwise, the derivation of the interface parameters is the same.

One noticeable difference is that the calculation of the jump condition are performed after $\boldsymbol{\tau}^0$ and $\boldsymbol{\tau}^1$ have been extended by 0 in $\mathcal{D}$. This is possible in the static case since the equilibrium translates in $\mathrm{div}_{\mathbf{y}}\boldsymbol{\tau}^0 = 0$ and $\mathrm{div}_{\mathbf{y}}\boldsymbol{\tau}^1 + \mathrm{div}_{\mathbf{x}'}\boldsymbol{\tau}^0 = 0$ with boundary conditions $\boldsymbol{\tau}^0.\mathbf{n}_{|\partial\mathcal{D}} = \boldsymbol{\tau}^1.\mathbf{n}_{|\partial\mathcal{D}} = 0$. Thus, extending $\boldsymbol{\tau}^n$ by 0 in $\mathcal{D}$ is compatible with the equilibrium and makes the boundary conditions on $\partial\mathcal{D}$ to be automatically satisfied. This is not possible in our case, and this is why the integrations are performed in $\mathsf{Y}\backslash\mathcal{D}$ instead of in Y in [Marigo and Pideri (2011)]. The two results are of course identical. Namely, the Eq. (40) in this reference reads

$$[\![\sigma^1_{i1}]\!] = -\frac{\partial}{\partial x_\alpha}\int_{\mathsf{Y}} d\mathbf{y}\ \tau_{i\alpha}(\mathbf{x}', \mathbf{y}), \tag{14.41}$$

and $\boldsymbol{\tau} \equiv \boldsymbol{\tau}^0 - \boldsymbol{\sigma}^0(0, \mathbf{x}')$ has been extended implicitly by $-\boldsymbol{\sigma}^0(0, \mathbf{x}')$ in $\mathcal{D}$. It follows that

$$[\![\sigma^1_{i1}]\!] = \int_{\mathcal{D}} d\mathbf{y}\ \frac{\partial \boldsymbol{\sigma}^0_{i\alpha}}{\partial x_\alpha}(0, \mathbf{x}') - \frac{\partial}{\partial x_\alpha} \int_{\mathsf{Y}\setminus\mathcal{D}} d\mathbf{y}\ \tau_{i\alpha}(\mathbf{x}', \mathbf{y}), \tag{14.42}$$

and owing to $\frac{\partial \sigma^0_{i\alpha}}{\partial x_\alpha} = -\frac{\partial \sigma^0_{i1}}{\partial x_1}$ given by the outer problem, we get

$$[\![\sigma^1_{i1}]\!] = -\frac{\partial \boldsymbol{\sigma}^0_{i1}}{\partial x_1}(0, \mathbf{x}')S - \frac{\partial}{\partial x_\alpha} \int_{\mathsf{Y}\setminus\mathcal{D}} d\mathbf{y}\ \tau_{i\alpha}(\mathbf{x}', \mathbf{y}), \tag{14.43}$$

which is equivalent to our jump condition Eq. (14.15). Finally, the elementary problems satisfied in the static elastic case are given by Eqs. (28) are the equivalent of our Eqs. (14.17).

14.6. Appendix C — Multimodal Approach to Compute (B, C) for Rectangular Rigid Structuration

For piecewise constant geometry in the elementary problems, mode matching is a simple way to get B and C. In both elementary problems (14.17), the idea is to project the solution $q^{(i)}$ onto bases of transverse functions of $\mathbf{y}'$ being adapted in the different regions along y_1. Below, we give a simple procedure to compute (B, C) for rectangular rigid inclusions in two dimensions (in this case, only two bases of transverse functions $(f_n^-(\mathbf{y}'), f_n^+(\mathbf{y}')$ are needed). The simplicity lies in the fact that the procedure does not require more than the inversion of a matrix. The size N of the matrix fixes the spatial resolution : with a cell of size unity along y_2, the spatial resolution is typically $1/N$. As we assumed that all the dimensions of the inclusions is of order unity, large N is not required (otherwise, the calculation may become tricky since mode matching has low convergence, typically $1/N$).

In the following, we consider $h = 1$ for simplicity.

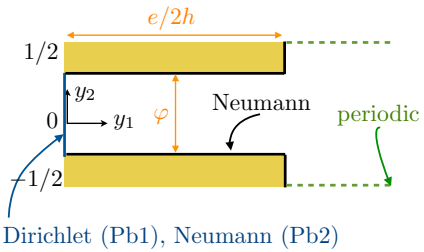


Fig. 14.13. Mode matching configuration. The solutions Q^- and Q^+ are written for $0 \le y_1 \le e/2h$ and $y_1 \ge e/2h$ respectively (with proper conditions at $y_1 = 0$), and the resolution involves only matching conditions at $y_1 = e/2h$.

14.6.1. *Calculation of B in the elementary problem i = 1*

Because $q^{(1)}$ is defined up to a constant and because the symmetry of the inclusion *w.r.t.* y_1, we can consider a function $q^{(1)}$ odd *w.r.t.* y_1, whence applying Dirichlet boundary condition at $y_1 = 0$ (namely $q^{(1)}(0, y_2) = 0$). It follows that $q^{(1)} \to b_1/2$ for $y_1 \to +\infty$ (and $b_1 = \delta q^{(1)}$). We use $Q(\mathbf{y}) = y_1 + q^{(1)}(\mathbf{y}) - b_1/2$ which satisfies $\Delta Q = 0$, $\boldsymbol{\nabla} Q.\mathbf{n}_{|\partial\mathcal{D}} = 0$, and $(Q - y_1) \to 0$ for $y_1 \to +\infty$. Finally, we use that $\int_{-\varphi/2}^{\varphi/2} \mathrm{d}y_2 \; \partial_{y_1} Q(0, y_2) = \int_{-1/2}^{1/2} \mathrm{d}y_2 \; \partial_{y_1} Q(+\infty, y_2) = 1$ (by integration of $\Delta Q = 0$). Thus, Q can be expanded as

$$Q(\mathbf{y}) = \begin{cases} Q^-(\mathbf{y}) = \dfrac{y_1}{\varphi} + q_0^- f_0^- & 0 \le y_1 \le e/2 \\ \qquad + \displaystyle\sum_{n=1}^{N^-} q_n^- \frac{\sinh a_n y_1}{\sinh a_n e/2} \; f_n^-(y_2), & \\ Q^+(\mathbf{y}) = y_1 & e/2 \le y_1, \\ \qquad + \displaystyle\sum_{n=-N^+, n\neq 0}^{N^+} q_n^+ e^{-|b_n|(y_1 - e/2)} f_n^+(y_2), & \end{cases} \tag{14.44}$$

with $a_n = n\pi/\varphi$, $b_n = 2n\pi$, and where

$$f_n^+(y_2) = e^{\mathrm{i}b_n y_2}, \quad f_n^-(y_2) = \sqrt{\frac{2 - \delta_{n0}}{\varphi}} \cos\left(a_n y_2 + \frac{n\pi}{2}\right), \tag{14.45}$$

are the transverse functions (forming a basis) adapted for solutions being respectively periodic and for solutions with zero derivatives at $y_2 = \pm\varphi/2$.

In the above projection, b_1 being unknown has ben encapsulated in the first coefficient q_0^- of the expansion, and it is easy to see that $Q^-(0, y_2) = q_0^- f_0^- = -b_1/2$ (thus $b_1 = -2q_0^-/\sqrt{\varphi}$), by construction.

Now, we will ask to $Q^\pm$ to match (on average) their values and their first derivatives at $y_1 = e/2$, and this latter matching on the derivatives will be done accounting for the Neumann boundary conditions at $y_1 = e/2$ and $|y_2| > \varphi/2$ (note that Q^- satisfies by construction the Neumann boundary condition on at $y_2 = \pm\varphi/2$, because of the choice of the f_m^-). To that aim, we use the following relations

$$\begin{cases} \displaystyle\int_{-\varphi/2}^{\varphi/2} \mathrm{d}y_2\; Q^-(e/2, y_2) f_m^-(y_2) = \int_{-\varphi/2}^{\varphi/2} \mathrm{d}y_2\; Q^+(e/2, y_2) f_m^-(y_2), \\ \displaystyle\int_{-\varphi/2}^{\varphi/2} \mathrm{d}y_2\; \frac{\partial Q^-}{\partial y_1}(e/2, y_2) f_m^{+*}(y_2) = \int_{-1/2}^{1/2} \mathrm{d}y_2\; \frac{\partial Q^+}{\partial y_1}(e/2, y_2) f_m^{+*}(y_2), \end{cases} \tag{14.46}$$

with f_m^{+*} the conjugate of f_m^+ (f_m^- is real). The first relation is the matching of the values in the region $y_2 \in [-\varphi/2, \varphi/2]$ where Q^- is defined. The second relation has more information: we have used that the $\partial_{y_1} Q^+ = 0$ at $y_1 = e/2$ for $|y_2| > \varphi/2$, from which

$$\int_{-1/2}^{1/2} \mathrm{d}y_2\; \frac{\partial Q^+}{\partial y_1}(e/2, y_2) f_n^{+*}(y_2) = \int_{-\varphi/2}^{\varphi/2} \mathrm{d}y_2\; \frac{\partial Q^+}{\partial y_1}(e/2, y_2) f_m^{+*}(y_2), \tag{14.47}$$

afterwards we ask, on average, $\partial_{y_1} Q^+ = \partial_{y_1} Q^-$ for $|y_2| < \varphi/2$. We get a matrix system for the two vectors $\mathbf{q}^- = (q_n^-)_{n=0,\dots,N^-}$ and $\mathbf{q}^+ = (q_n^+)_{n=0,\dots,N^+}$ of the form

$$\begin{pmatrix} \mathsf{I} & -{}^t\mathsf{F}^* \\ \mathsf{FA}_1 & \mathsf{B} \end{pmatrix} \begin{pmatrix} \mathbf{q}^- \\ \mathbf{q}^+ \end{pmatrix} = \begin{pmatrix} \mathbf{S}^s \\ \mathbf{S}^d \end{pmatrix}, \tag{14.48}$$

with I the $N^- \times N^-$ identity matrix, $\mathsf{A}_1 = \mathrm{diag}(a_n \mathrm{cotanh}(a_n e/2))$ with $\mathsf{A}_1(1,1) = 0$, $\mathsf{B} = \mathrm{diag}(|b_n|)$, and $F_{mn} = \int_{-\varphi/2}^{\varphi/2} \mathrm{d}y_2 f_n^{+*}(y_2) f_n^-(y_2)$. The source terms $(\mathbf{S}^s, \mathbf{S}^d)$ correspond to the projection of the loading (y_1/φ and y_1) onto f_n^- and f_n^{+*}. The expressions of F_{mn} and $S_n^{s,d}$ are given below

$$F_{mn} = \begin{cases} \sqrt{\varphi}\,\mathrm{sinc}\left(b_m \dfrac{\varphi}{2}\right), & n = 0, \\ \quad = \sqrt{\dfrac{\varphi}{2}} \left[\mathrm{sinc}\left((a_n - b_m)\dfrac{\varphi}{2}\right) e^{\mathrm{i}n\pi/2} \right. & n \neq 0, \\ \qquad \left. + \mathrm{sinc}\left((a_n + b_m)\dfrac{\varphi}{2}\right) e^{-\mathrm{i}n\pi/2}\right], & \end{cases} \tag{14.49}$$

and

$$\begin{cases} S_n^s = \dfrac{e}{2}\left(\dfrac{1}{\varphi} - 1\right) \delta_{n0}, & n = 0, \ldots, N^- \\ S_n^d = -\mathrm{sinc}(n\pi\varphi), & n = -N^+, \ldots, -1, 1, \ldots N^+. \end{cases} \tag{14.50}$$

The system is of the form $\mathsf{M}\mathbf{q} = \mathbf{s}$ with the matrix M being square (this is not always the case in systems written using mode matching). Next, M is invertible if one has taken care to consider only the antisymmetric modes. Thus, the system can be solved to find $\mathbf{q}$ by inverting M or it can be solved in the least squares sense (as done by the operation $\mathsf{M}\backslash\mathbf{s}$ in Matlab). Once the system has been inverted, b_1 is obtained and thus, from (14.28), $B = e + b_1$

$$B = e - \frac{2}{\varphi}\, q(1). \tag{14.51}$$

14.6.2. *Calculation of C in the elementary problem $i = 2$*

The procedure to get C in the elementary problem $i = 2$ is similar. We consider the solution $Q = Q^{(2)} + y_2$ satisfying $\Delta Q = 0$,

$\boldsymbol{\nabla} Q.\mathbf{n}_{|\partial\mathcal{D}} = 0$ and $Q \to y_2$ for $y_1 \to \infty$. Because $q^{(2)}$ is even in y_1, we can project Q onto the bases $(f_n^{\pm})$ as follow

$$Q(\mathbf{y}) = \begin{cases} Q^-(\mathbf{y}) = \displaystyle\sum_{n=1}^{N^-} q_n^- \frac{\cosh a_n y_1}{\cosh a_n e/2} f_n^-(y_2), & -e/2 \le y_1 \le 0 \\ Q^+(\mathbf{y}) = y_2 \\ \qquad\qquad + \displaystyle\sum_{n=-N^+, n\neq 0}^{N^+} q_n^+ e^{-|b_n| y_1} f_n^+(y_2), & 0 \le y_1. \end{cases} \tag{14.52}$$

As previously, we ask to $Q^{\pm}$ to match (on average) their values and their first derivatives at $y_1 = e/2$, and we get a matrix system of the form

$$\begin{pmatrix} \mathsf{I} & -{}^t\mathsf{F}^* \\ \mathsf{F}\mathsf{A}_2 & \mathsf{B} \end{pmatrix} \begin{pmatrix} \mathbf{q}^- \\ \mathbf{q}^+ \end{pmatrix} = \begin{pmatrix} \mathbf{S} \\ \mathbf{0} \end{pmatrix}, \tag{14.53}$$

with the same definitions as previously and now, $\mathsf{A}_2 = \mathrm{diag}(a_n \tanh(a_n e/2))$ and $S_n = \int_{-\varphi/2}^{\varphi/2} \mathrm{d}y_2 \; y_2 f_n^-(y_2)$, specifically

$$S_n = -2\sqrt{\frac{2}{\varphi}} \frac{1}{a_n^2}. \tag{14.54}$$

Once the system has been inverted, we can determine

$$C = -\int \mathrm{d}\mathbf{y} \; \frac{\partial Q^{(2)}}{\partial y_2} = \int_{-e}^{0} \mathrm{d}y_1 \int_{-\varphi/2}^{\varphi/2} \left[1 - \frac{\partial Q^-}{\partial y_2} \right]. \tag{14.55}$$

where we have used that $Q^{(2)}(y_1 \ge 0, y_2) = Q^+(\mathbf{y}) - y_2$ is periodic, thus of vanishing contribution. It is now sufficient to write $C = e\varphi - q_n^- \mathrm{tanh} a_n e/a_n [f_n]_{-\varphi/2}^{\varphi/2}$ to get

$$C = e\varphi + 2\sqrt{\frac{2}{\varphi}} \frac{\mathrm{tanh} a_n e}{a_n} \; q_n^-. \tag{14.56}$$

14.6.3. *Scripts in Matlab*

Simple scripts providing (B, C) are given below.

```
function B=PbElem1(phi,e,Np)

%% version de B avec B=e+.... (elargi)
Nd=2*Np;
nd=0:2:Nd;         np=[-Np:-1,1:Np];
Nd=length(nd);  Np=length(np);
an=nd*pi/phi;   bn=2*np*pi;

for mm=1:Np,
    F(mm,1) = sqrt(phi)*sinc(bn(mm)*phi/(2*pi));
    for nn=2:Nd,
        a=an(nn); b=bn(mm); n=nd(nn); EX=exp(1i*n*pi/2);
        temp=sinc((a-b)*phi/(2*pi))*EX+sinc((a+b)*phi/(2*pi))/EX;
        F(mm,nn) = sqrt(phi/2)*temp;
    end,
end

Ac=an./tanh(an*e/2); Ac(1)=0;
M=[eye(Nd) ,-F';
    F*diag(Ac), diag(abs(bn))];

s1=0*nd'; s1(1)=e/2*(sqrt(phi)-1/sqrt(phi));
s2=-sinc(np*phi).';
S=[s1; s2];
q=M\S;

B=e-2*q(1)/sqrt(phi);

function C=PbElem2(phi,e,Np)

Nd=2*Np;
nd=1:Nd;         np=[-Np:-1,1:Np];
Nd=length(nd);  Np=length(np);
an=nd*pi/phi;   bn=2*np*pi;

    for mm=1:Np,
       for nn=1:Nd,
          a=an(nn); b=bn(mm); n=nd(nn); EX=exp(1i*n*pi/2);
          temp=sinc((a-b)*phi/(2*pi))*EX+sinc((a+b)*phi/(2*pi))/EX;
          F(mm,nn) = sqrt(phi/2)*temp;
        end,
```

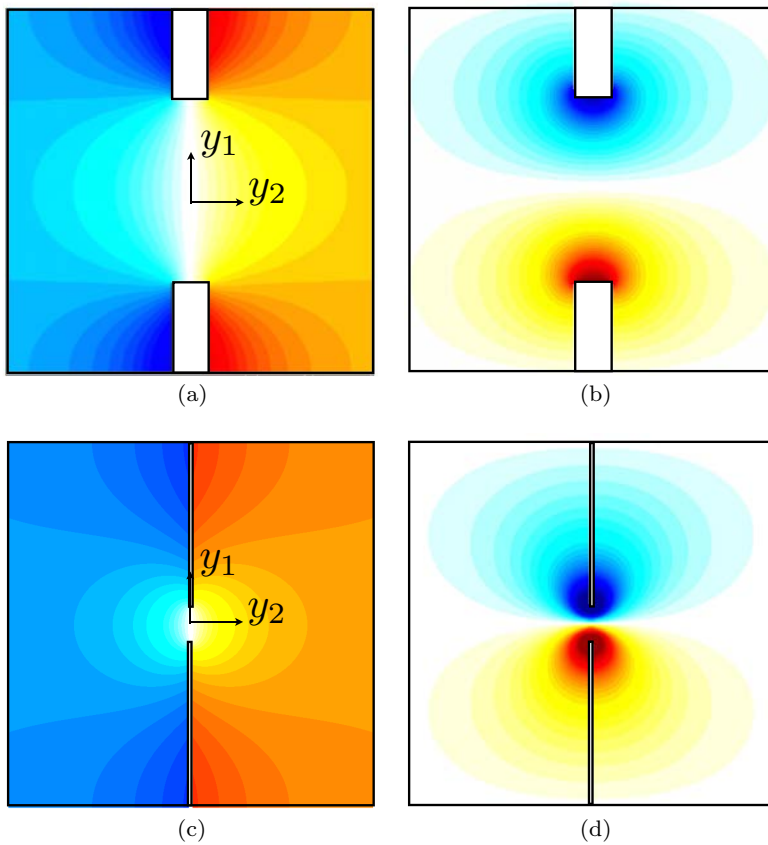


Fig. 14.14. Static fields computed for the elementary problems $i = 1$ and 2. (a) (c) $q^{(1)}(\mathbf{y})$ and (b) (d) $q^{(2)}(\mathbf{y})$. (a–b) for $e/h = 0.1$ and $\ell/h = 0.5$, (c–d) for $e/h = 0.01$ and $\ell/h = 0.9$.

```
    end

        M=[eye(Nd) ,-F';
            F*diag(an.*tanh(an*e/2)), diag(abs(bn))];

        s=sqrt(2/phi)*(((-1).^nd-1)./an.^2).'; S=[s; zeros(Np,1)];
        q=M\S;qm=q(1:Nd);

        C=phi*e+4*sqrt(2/phi)*sum( qm.'.*tanh(an*e/2)./an );
end
```

14.7. Appendix D — Reflection Coefficient Using the Classical Homogenization

In this Appendix, we consider $h = 1$ for simplicity.

14.7.1. *Scattering coefficients in the classical homogenization*

Classical homogenization relies on the same expansion of the field in power of ε. Let us consider penetrable inclusions presenting a contrast in mass density and a contrast in bulk modulus with respect to the surrounding medium. In a general way, the wave propagation is the equation of propagation in inhomogeneous media

$$\operatorname{div}\left[a(\mathbf{X})\boldsymbol{\nabla} P\right] + b(\mathbf{X})P = 0, \tag{14.57}$$

with a the inverse of the mass density and b the inverse of the bulk modulus; we denote $a(\mathbf{X}) = a$, $b(\mathbf{X}) = b$ inside the inclusion and $a(\mathbf{X}) = b(\mathbf{X}) = 1$ in the surrounding medium. At the inclusion boundaries $\partial\mathcal{D}$, P and $a(\mathbf{X})\partial_n P$ are continuous. For acoustic waves, It is easy to see that applying $a = b = 0$ in the inclusions leads to the Neumann limit, with $\Delta P + P = 0$ in the surrounding medium and $\partial_n P_{|\partial\mathcal{D}} = 0$. This is consistent with the intuitive idea that Neumann boundary condition corresponds to heavy materials, with infinite mass density and infinite bulk modulus. At leading order, $P \simeq P_0$ satisfies

$$\operatorname{div}\left[\mathsf{a}_{\text{eff}}\boldsymbol{\nabla} P\right] + b_{\text{eff}}P = 0, \tag{14.58}$$

with the tensor a_{eff} and b_{eff} given by the resolution of two elementary problems. The case of layered media, with inclusions infinitely long along X_1 and of width ℓ along X_2, we have $a(X_2)$ and $b(X_2)$ and explicit solution is possible. It ends with diagonal effective mass density tensor $\mathsf{a}_{\text{eff}} = \operatorname{diag}(a^1_{\text{eff}}, a^2_{\text{eff}})$ and effective bulk modulus given by

$$\begin{cases} a^1_{\text{eff}} = \langle a(X_2)\rangle = [\ell a + (1-\ell)] \to 1-\ell, \\ a^2_{\text{eff}} = \langle a(X_2)^{-1}\rangle^{-1} = [\ell/a + (1-\ell)]^{-1} \to 0, \\ b_{\text{eff}} = \langle b(X_2)\rangle = [\ell b + (1-\ell)] \to 1-\ell, \end{cases} \tag{14.59}$$

where the limits correspond to the Neumann case ($a = b = 0$). In this limit, the propagation inside the layered medium is described by the wave equation $a^1_{\text{eff}}\partial_{X_1^2}P + a^2_{\text{eff}}\partial_{X_2^2}P + k^2 b_{\text{eff}}P = (1-\ell)[\partial_{X_1^2}P + k^2 P] = 0$, from which we deduce that the component of the wavenumber along X_1 is k.

Imagine now that we cut the layered medium to create a slab, which occupies the space $X_1 \in [-e/2, e/2]$, and outside the slab, the medium has $a(\mathbf{X}) = b(\mathbf{X}) = 1$. For the real structure, the resulting configuration corresponds to an array of rectangular hard inclusions as studied in Section 14.2. For the equivalent homogenized medium, it corresponds to a slab of thickness e filled with the homogeneous and anisotropic material given by $(\mathsf{a}_{\text{eff}}, b_{\text{eff}})$. In this latter configuration, the problem of the scattering for an incident wave coming from $X_1 = -\infty$ with incidence θ is trivial and it has a solution of the form

$$P(X_1, X_2) = e^{\mathrm{i}k\sin\theta X_2} \begin{cases} [e^{\mathrm{i}k\cos\theta X_1} + \tilde{R}e^{-\mathrm{i}k\cos\theta(X_1+e/2)}], & \text{for } X_1 < -e/2, \\ \left[Ae^{\mathrm{i}kX_1} + Be^{-\mathrm{i}kX_1}\right], & \text{for } |X_1| \le e/2, \\ \tilde{T}e^{\mathrm{i}k\cos\theta(X_1-e/2)}, & \text{for } X_1 > e/2, \end{cases} \tag{14.60}$$

where the component of the wavenumber along X_2, $k\sin\theta$, is imposed by the continuity of $P(X_1, X_2)$ at the interfaces $X_1 = \pm e/2$. Next, applying the relations of continuity for P and $a_1(\mathbf{X})\partial_{X_1}P$ (with $a_1(\mathbf{X}) = a^1_{\text{eff}} = 1-\ell$ in the slab and $a_1(\mathbf{X}) = 1$ outside) at $X_1 = \pm e/2$, leads to usual forms of the scattering coefficients

$$\tilde{R} = \frac{2\mathrm{i}(1-c^2)\sin ke\; e^{-\mathrm{i}ke\cos\theta/2}}{(1+c)^2e^{-\mathrm{i}ke} - (1-c)^2e^{\mathrm{i}ke}}, \quad \tilde{T} = \frac{4ce^{-\mathrm{i}ke\cos\theta/2}}{(1+c)^2e^{-\mathrm{i}ke} - (1-c)^2e^{\mathrm{i}ke}}, \tag{14.61}$$

where $c \equiv \cos\theta/(1-\ell)$. The form of $\tilde{R}$ in the above equation is written in an alternative form in Eq. (14.34).

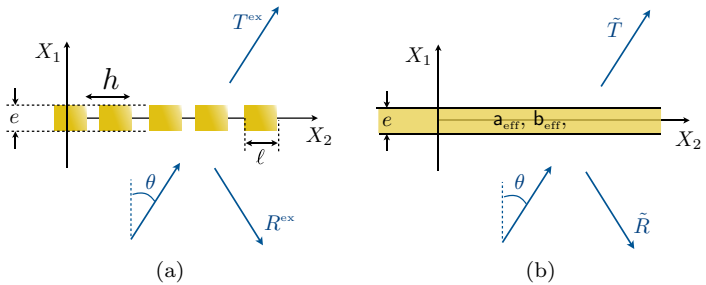


Fig. 14.15. In the classical homogenization, the actual structured film (a) is replaced by a homogeneous anisotropic film of thickness e.

The derivation of $(\tilde{R}, \tilde{T})$ is done accounting for boundary conditions being the continuity of p and of $\mathsf{a}_{\text{eff}} \boldsymbol{\nabla} p.\mathbf{n}$, which is true at this leading order (see [Marigo and Maurel (2017)]).

14.7.2. *Jump conditions in the classical homogenization*

It is tempting to inspect if classical homogenization gives jump conditions similar to those obtained from our interface model, Eq.(14.29). To that end, we consider, from Eqs. (14.60) the quantities $[\![P]\!] = \tilde{T} - \tilde{R} - 1$, $[\![\partial_{X_1} P]\!] = ik\cos\theta(\tilde{T} + \tilde{R} - 1)$, with $\overline{\partial_{X_1} P}(0, X_2) = ik\cos\theta(\tilde{T} - \tilde{R} + 1)/2$, $\overline{\partial_{X_1^2} P}(0, X_2) = (ik\cos\theta)^2(\tilde{T} - \tilde{R} + 1)/2$ and $\overline{\partial_{X_2^2} P}(0, X_2) = (ik\sin\theta)^2(\tilde{T} + \tilde{R} + 1)/2$. Expanding at leading order $(\tilde{T}, \tilde{R})$ in Eqs. (14.61) for small ke, it is easy to see that we get

$$[\![P]\!] = \frac{e\ell}{1-\ell}\overline{\frac{\partial P}{\partial X_1}}(0, X_2),$$

$$\left[\!\!\left[\frac{\partial P}{\partial X_1}\right]\!\!\right] = -\, e\ell\overline{\frac{\partial^2 P}{\partial X_1^2}}(0, X_2) + e(1-\ell)\overline{\frac{\partial^2 P}{\partial X_2^2}}(0, X_2). \tag{14.62}$$

This can be written in the form of interface parameters

$$\tilde{B} = \frac{e\ell}{1-\ell}, \quad \tilde{C} = e(1-\ell), \quad \tilde{S} = e\ell. \tag{14.63}$$

While $S = \tilde{S}$ (in the non enlarged version of the jump conditions), both $\tilde{C}$ and $\tilde{B}$ differ from C, B in Eqs. (14.30)–(14.31). The most

evident difference is that the interface parameters given by classical homogenization misses the blockage coefficients B_0 which precisely becomes dominant for $e \to 0$. Next, C and $\tilde{C}$ differ from the correction $-\pi/8(1-\ell)^2$ appearing for large enough thickness. Both corrections are due to boundary film effect, see [Marigo and Maurel (2016)].

14.8. Appendix E — Scattering Coefficients for the Modes in a Waveguide

If the film of the equivalent anisotropic medium occupies the space $X_1 \in [-e/2, e/2]$ and $X_2 \in [0, H]$ as considered in the Section 14.3.3, it is possible to calculate the reflection and transmission coefficients $(\tilde{r}_n, \tilde{t}_n)$ mode by mode (and a mode is defined in Eq. (14.36)). This is because no mode coupling occurs in this case, again owing to the orthogonality of the transverse functions. The problem first reads

$$a^1_{\text{eff}} \frac{\partial^2 P}{\partial X_1^2} + a^2_{\text{eff}} \frac{\partial^2 P}{\partial X_2^2} + k^2 b_{\text{eff}} p = 0, \tag{14.64}$$

in the film and $\Delta P + k^2 P = 0$ outside, and the relations of continuity of P and $a_1 \partial_{X_1} P$ apply at $X_1 = \pm e/2$ (with $a_1 = a^1_{\text{eff}}$ inside the layer, $a_1 = 1$ outside). Expanding the solution into the basis of transverse functions, $P(X_1, X_2) = \sum_{n \geq 0} P_n(X_1)\, \varphi_n(X_2)$, the system is decoupled, with

$$\begin{cases} a^1_{\text{eff}} P_n'' + \left[k^2 b_{\text{eff}} - a^2_{\text{eff}} \left(\dfrac{n\pi}{H} \right)^2 \right] P_n = 0, & \text{for } |X_1| \leq e/2 \\ P_n'' + \left[k^2 - \left(\dfrac{n\pi}{H} \right)^2 \right], & \text{outside} \end{cases} \tag{14.65}$$

with P_n continuous at $X_1 = \pm e/2$, and $P_n'(-e/2^-) = a^1_{\text{eff}} P_n'(-e/2^+)$, $P_n'(e/2^+) = a^1_{\text{eff}} P_n'(e/2^-)$. In the case of our layered sound hard medium, $a^1_{\text{eff}} = b_{\text{eff}} = 1 - \ell$ and $a^2_{\text{eff}} = 0$, which leads to $P_n'' + k^2 P_n = 0$ in the layer, $|X_1| \leq e/2$. Thus, the problem is identical to the previous

one (Eqs. (14.60)), owing to $k\cos\theta \to k_n$ and

$$\tilde{r}_n = \frac{2\mathrm{i}(1-c_n^2)\sin ke\; e^{-\mathrm{i}k_n e/2}}{(1+c_n)^2 e^{-\mathrm{i}ke} - (1-c_n)^2 e^{\mathrm{i}ke}},$$
$$\tilde{t}_n = \frac{4c_n e^{-\mathrm{i}k_n e/2}}{(1+c_n)^2 e^{-\mathrm{i}ke} - (1-c_n)^2 e^{\mathrm{i}ke}}, \tag{14.66}$$

with $c_n \equiv k_n/k(1-\ell)$.

14.9. Appendix F — Discussion on the Convergence of the Scattering Coefficients in the Modal Method

In the problem considered in Section 14.3.3, we have said that the high truncation $N = 2000 - 3000$ was expected and not expected. Expected because the smallest scale to be resolved is $e = H/2000$ which imposes $N \sim H/e$ (and this corresponds to the highest evanescent mode with wavenumber $k_N \sim N\pi/H$). Unexpected because we have checked that the mode coupling is negligible; the incident wave being a combination of the first 100 modes, the absence of mode coupling ensures that the modes at $n > 100$ are not excited. To be clear, we have compared $P^{\mathrm{ex}}(\mathbf{X})$ computed with 3 000 modes to the following field

$$P^{\mathrm{diag}}(\mathbf{X}) = \begin{cases} \displaystyle\sum_{n=0}^{100} P_n^{\mathrm{inc}}[e^{\mathrm{i}k_n(X_1+e/2)} + \mathsf{R}_{nn}e^{-\mathrm{i}k_n(X_1+e/2)}], & X_1 < -e/2, \\ \displaystyle\sum_{n=0}^{100} P_n^{\mathrm{inc}}\mathsf{T}_{nn}e^{\mathrm{i}k_n(X_1-e/2)}, & X_1 > e/2, \end{cases} \tag{14.67}$$

which is precisely the field composed of the 100 first modes generated directly by the incident wave. The difference $|P^{\mathrm{ex}} - P^{\mathrm{diag}}|$ is less than 0.4% in both cases, thus the evanescent field generated by higher modes is indeed negligible. The reason why high truncation

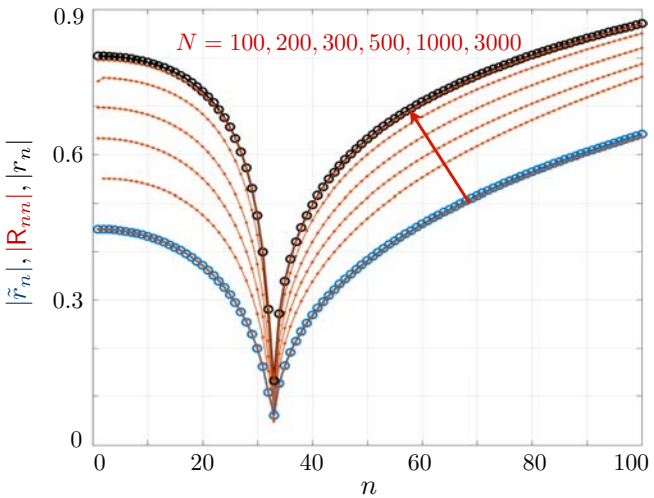


Fig. 14.16. Reflection coefficients of the modes in the guide obstructed by a perforated screen (same calculations as in Fig. 14.12). $|\mathsf{R}_{nn}|$ for $n = 0$ to 100 is shown for several truncation order N in the expansion Eq. (14.36) (orange curves). For low truncation, $N = 100$, $|\mathsf{R}_{nn}|$ are found to be close to the $\tilde{r}_n$ given by the classical homogenization in (14.66) (blue symbols); increasing N makes $|\mathsf{R}_{nn}|$ to converge toward r_n given by the interface homogenization, $|r_n|$ in Eq. (14.40) (black symbols).

has been needed is in fact related to the convergence of R_{nn} and T_{nn} for $n \leq 100$ with N, a convergence which is often disregarded. In Fig. 14.16, we have reported the variation of $|\mathsf{R}_{nn}|$, $n \leq 100$, for several truncations $N = 100$ to 3000. As it can be seen, increasing N produces a significant variation in $|\mathsf{R}_{nn}|$, 70% on average for increasing the truncation N from $N = 100$ to $N = 3000$. This low convergence has already been observed for inclusions presenting a contrast in the mass density [Maurel *et al.* (2014)]; it is a convergence law in $1/N$, due to the low regularity of the field (with discontinuous pressure gradient). Amusingly, for insufficient truncation $N = 100$, the scattering coefficients coincide with the solution $(\tilde{r}_n, \tilde{t}_n)$ given by classical homogenization, which underestimates the weight of the evanescent field (Eqs. (14.66), see Appendix 14.8).

Bibliography

[1] Abdelmoula, R., Coutris, M. and Marigo, J.-J. (1998). Comportement asymptotique d'une interphase élastique mince, *Comptes Rendus de l'Académie des Sciences-Series IIB-Mechanics-Physics-Chemistry-Astronomy* **326**, 4, pp. 237–242.

[2] Akarid, A., Ourir, A., Maurel, A., Félix, S. and Mercier, J.-F. (2014). Extraordinary transmission through subwavelength dielectric gratings in the microwave range, *Optics Letters* **39**, 13, pp. 3752–3755.

[3] Antonakakis, T., Craster, R. and Guenneau, S. (2013). Asymptotics for metamaterials and photonic crystals, in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, Vol. 469 (The Royal Society), p. 20120533.

[4] Arslanagic, S., Hansen, T. V., Mortensen, N. A., Gregersen, A. H., Sigmund, O., Ziolkowski, R. W. and Breinbjerg, O. (2013). A review of the scattering-parameter extraction method with clarification of ambiguity issues in relation to metamaterial homogenization, *Antennas and Propagation Magazine, IEEE* **55**, 2, pp. 91–106.

[5] Bonnet-Bendhia, A., Drissi, D. and Gmati, N. (2004). Simulation of muffler's transmission losses by a homogenized finite element method, *Journal of Computational Acoustics* **12**, 03, pp. 447–474.

[6] Bouchitté, G. and Felbacq, D. (2004). Homogenization near resonances and artificial magnetism from dielectrics, *Comptes Rendus Mathematique* **339**, 5, pp. 377–382.

[7] Capdeville, Y., Guillot, L. and Marigo, J.-J. (2010a). 1-d non-periodic homogenization for the seismic wave equation, *Geophysical Journal International* **181**, 2, pp. 897–910.

[8] Capdeville, Y., Guillot, L. and Marigo, J.-J. (2010b). 2-d non-periodic homogenization to upscale elastic media for p–sv waves, *Geophysical Journal International* **182**, 2, pp. 903–922.

[9] Capdeville, Y. and Marigo, J.-J. (2007). Second order homogenization of the elastic wave equation for non-periodic layered media, *Geophysical Journal International* **170**, 2, pp. 823–838.

[10] Capdeville, Y. and Marigo, J.-J. (2013). A non-periodic two scale asymptotic method to take account of rough topographies for 2-d elastic wave propagation, *Geophysical Journal International* **192**, 1, pp. 163–189.

[11] Castanié, A., Mercier, J.-F., Felix, S. and Maurel, A. (2014). Generalized method for retrieving effective parameters of anisotropic metamaterials, *Optics Express* **22**, 24, pp. 29937–29953.

[12] Craster, R. V., Kaplunov, J., Nolde, E. and Guenneau, S. (2011). High-frequency homogenization for checkerboard structures: defect modes, ultra-refraction, and all-angle negative refraction, *JOSA A* **28**, 6, pp. 1032–1040.

[13] David, M., Marigo, J.-J. and Pideri, C. (2012). Homogenized interface model describing inhomogeneities located on a surface, *Journal of Elasticity* **109**, 2, pp. 153–187.

[14] Delourme, B. (2010). *Modèles et asymptotiques des interfaces fines et périodiques en électromagnétisme*, Ph.D. thesis, Université Pierre et Marie Curie-Paris VI.

[15] Delourme, B., Haddar, H. and Joly, P. (2012). Approximate models for wave propagation across thin periodic interfaces, *Journal de Mathématiques Pures et Appliquées* **98**, 1, pp. 28–71.

[16] Dimitriadis, A., Sounas, D. L., Kantartzis, N. V., Caloz, C., Tsiboukis, T. D., *et al.* (2012). Surface susceptibility bianisotropic matrix model for periodic metasurfaces of uniaxially mono-anisotropic scatterers under oblique te-wave incidence, *Antennas and Propagation, IEEE Transactions on* **60**, 12, pp. 5753–5767.

[17] Felbacq, D. and Bouchitté, G. (2005). Left-handed media and homogenization of photonic crystals, *Optics Letters* **30**, 10, pp. 1189–1191.

[18] Flagg, C. and Newman, J. (1971). Sway added-mass coefficients for rectangular profiles in shallow water, *Journal of Ship Research* **15**, p. 4.

[19] Garcia-Vidal, F., Martin-Moreno, L. and Pendry, J. (2005). Surfaces with holes in them: New plasmonic metamaterials, *Journal of Optics A: Pure and Applied Optics* **7**, 2, p. S97.

[20] Guillot, L., Capdeville, Y. and Marigo, J.-J. (2010). 2-d non-periodic homogenization of the elastic wave equation: Sh case, *Geophysical Journal International* **182**, 3, pp. 1438–1454.

[21] Holloway, C. L., Dienstfrey, A., Kuester, E. F., O'Hara, J. F., Azad, A. K. and Taylor, A. J. (2009). A discussion on the interpretation and characterization of metafilms/metasurfaces: The two-dimensional equivalent of metamaterials, *Metamaterials* **3**, 2, pp. 100–112.

[22] Holloway, C. L., Mohamed, M., Kuester, E. F., Dienstfrey, A., *et al.* (2005). Reflection and transmission properties of a metafilm: With an application to a controllable surface composed of resonant particles, *Electromagnetic Compatibility, IEEE Transactions on* **47**, 4, pp. 853–865.

[23] Kakuno, S., Oda, K. and Liu, P. L. (1992). Scattering of water waves by vertical cylinders with a backwall, *Coastal Engineering Proceedings* **1**, p. 23.

[24] Kelders, L., Allard, J. F. and Lauriks, W. (1998). Ultrasonic surface waves above rectangular-groove gratings, *The Journal of the Acoustical Society of America* **103**, 5, pp. 2730–2733.

[25] Kuester, E. F., Mohamed, M., Piket-May, M., Holloway, C. L., *et al.* (2003). Averaged transition conditions for electromagnetic fields at a metafilm, *Antennas and Propagation, IEEE Transactions on* **51**, 10, pp. 2641–2651.

[26] Lemoult, F., Kaina, N., Fink, M. and Lerosey, G. (2013). Wave propagation control at the deep subwavelength scale in metamaterials, *Nature Physics* **9**, 1, pp. 55–60.

[27] Li, Y., Jiang, X., Li, R.-q., Liang, B., Zou, X.-y., Yin, L.-l. and Cheng, J.-c. (2014). Experimental realization of full control of reflected waves with subwavelength acoustic metasurfaces, *Physical Review Applied* **2**, 6, p. 064002.

[28] Ma, G., Yang, M., Xiao, S., Yang, Z. and Sheng, P. (2014). Acoustic metasurface with hybrid resonances, *Nature Materials*.

[29] Marigo, J. J. and Maurel, A. (2017). Second order Homogenization of subwavelength stratified media including finite size effect, *SIAM Journal on Applied Mathematics* **77**(2), pp. 721–743.

[30] Marigo, J. J. and Maurel, A. (2016). Homogenization models for thin rigid structured surfaces and films, *The Journal of the Acoustical Society of America* **140**(1), pp. 260–273.

[31] Marigo, J.-J. and Pideri, C. (2011). The effective behavior of elastic bodies containing microcracks or microholes localized on a surface, *International Journal of Damage Mechanics*, p. 1056789511406914.

[32] Martin, P. A. and Dalrymple, R. A. (1988). Scattering of long waves by cylindrical obstacles and gratings using matched asymptotic expansions, *Journal of Fluid Mechanics* **188**, pp. 465–490.

[33] Maurel, A., Félix, S. and Mercier, J.-F. (2013). Enhanced transmission through gratings: Structural and geometrical effects, *Physical Review B* **88**, 11, p. 115416.

[34] Maurel, A. and Mercier, J.-F. (2012). Propagation of guided waves through weak penetrable scatterers, *The Journal of the Acoustical Society of America* **131**, 3, pp. 1874–1889.

[35] Maurel, A., Mercier, J.-F. and Félix, S. (2014). Wave propagation through penetrable scatterers in a waveguide and through a penetrable grating, *The Journal of the Acoustical Society of America* **135**, 1, pp. 165–174.

[36] Mercier, J., Cordero, M., Félix, S., Ourir, A. and Maurel, A. (2015). Classical homogenization to analyze the dispersion relations of spoof plasmons with geometrical and compositional effects, *Applied Physics Letters*, **submitted**.

[37] Morse, P. M. and Ingard, K. U. (1968). *Theoretical Acoustics* (Princeton University Press).

[38] Oleinik, O. A., Shamaev, A. and Yosifian, G. (2009). *Mathematical Problems in Elasticity and Homogenization*, Vol. 2 (Elsevier).

[39] Pendry, J., Martin-Moreno, L. and Garcia-Vidal, F. (2004). Mimicking surface plasmons with structured surfaces, *Science* **305**, 5685, pp. 847–848.

[40] Peng, P., Xiao, B. and Wu, Y. (2014). Flat acoustic lens by acoustic grating with curled slits, *Physics Letters A* **378**, 45, pp. 3389–3392.

[41] Sanchez-Hubert, J. and Sanchez-Palencia, E. (1982). Acoustic fluid flow through holes and permeability of perforated walls, *Journal of Mathematical Analysis and Applications* **87**, 2, pp. 427–453.

[42] Simovski, C. R. (2011). On electromagnetic characterization and homogenization of nanostructured metamaterials, *Journal of Optics* **13**, 1, p. 013001.

[43] Suh, K.-D., Ji, C.-H. and Kim, B. H. (2011). Closed-form solutions for wave reflection and transmission by vertical slotted barrier, *Coastal Engineering* **58**, 12, pp. 1089–1096.

[44] Xie, Y., Wang, W., Chen, H., Konneker, A., Popa, B.-I. and Cummer, S. A. (2014). Wavefront modulation and subwavelength diffractive acoustics with an acoustic metasurface, *Nature Communications*, **5**.

[45] Zhao, J., Ye, H., Huang, K., Chen, Z. N., Li, B. and Qiu, C.-W. (2014). Manipulation of acoustic focusing with an active and configurable planar metasurface transducer, *Scientific Reports* **4**.

[46] Zhao, Y., Engheta, N. and Alù, A. (2011). Homogenization of plasmonic metasurfaces modeled as transmission-line loads, *Metamaterials* **5**, 2, pp. 90–96.

[47] Zhu, J., Chen, Y., Zhu, X., Garcia-Vidal, F. J., Yin, X., Zhang, W. and Zhang, X. (2013). Acoustic rainbow trapping, *Scientific Reports* **3**.

[48] Zhu, J., Christensen, J., Jung, J., Martin-Moreno, L., Yin, X., Fok, L., Zhang, X. and Garcia-Vidal, F. (2011). A holey-structured metamaterial for acoustic deep-subwavelength imaging, *Nature Physics* **7**, 1, pp. 52–55.

Index

World Scientific Series in Nanoscience and Nanotechnology

(Continuation of series card page)

Vol. 13 *Nanoelectronics: A Molecular View*
by Avik Ghosh (University of Virginia, USA)

Vol. 12 *Nanomaterials for Photocatalytic Chemistry*
edited by Yugang Sun (Temple University, USA)

Vol. 11 *Molecular Bioelectronics: The 19 Years of Progress Second Edition*
by Nicolini Claudio (University of Genoa, Italy)

Vol. 10 *Pore Scale Phenomena: Frontiers in Energy and Environment*
edited by John Poate (Colorado School of Mines, USA), Tissa Illangasekare (Colorado School of Mines, USA), Hossein Kazemi (Colorado School of Mines, USA) and Robert Kee (Colorado School of Mines, USA)

Vol. 9 *Handbook of Biomimetics and Bioinspiration: Biologically-Driven Engineering of Materials, Processes, Devices, and Systems (In 3 Volumes)*
edited by Esmaiel Jabbari (University of South Carolina, USA), Deok-Ho Kim (University of Washington, USA), Luke P Lee (University of California, Berkeley, USA), Amir Ghaemmaghami (University of Nottingham, UK) and Ali Khademhosseini (Harvard University, USA & Massachusetts Institute of Technology, USA)

Vol. 8 *Polyoxometalate Chemistry: Some Recent Trends*
edited by Francis Sécheresse (Université de Versailles-St Quentin, France)

Vol. 7 *Scanning Probe Microscopy for Energy Research*
edited by Dawn A Bonnell (The University of Pennsylvania, USA) and Sergei V Kalinin (Oak Ridge National Laboratory, USA)

Vol. 6 *Plasmon Resonances in Nanoparticles*
by Isaak D Mayergoyz (University of Maryland, USA)

Vol. 5 *Inorganic Nanomaterials from Nanotubes to Fullerene-Like Nanoparticles: Fundamentals and Applications*
by Reshef Tenne (Weizmann Institute of Science, Israel)

*For the complete list of volumes in this series, please visit
www.worldscientific.com/series/wssnn

World Scientific Handbook of
Metamaterials and Plasmonics

Volume 3
Active Nanoplasmonics and Metamaterials

World Scientific Series in Nanoscience and Nanotechnology*

ISSN: 2301-301X

The Series aims to cover the new and evolving fields that cover nanoscience and nanotechnology. Each volume will cover completely a subfield, which will span materials, applications, and devices.

Published

Vol. 16 *World Scientific Handbook of Metamaterials and Plasmonics (In 4 Volumes)*
Volume 1: Electromagnetic Metamaterials
Volume 2: Elastic, Acoustic, and Seismic Metamaterials
Volume 3: Active Nanoplasmonics and Metamaterials
Volume 4: Recent Progress in the Field of Nanoplasmonics
edited by Stefan A Maier (Imperial College London, UK)

Vol. 15 *Molecular Electronics: An Introduction to Theory and Experiment Second Edition*
by Juan Carlos Cuevas (Universidad Autónoma de Madrid, Spain) and Elke Scheer (Universität Konstanz, Germany)

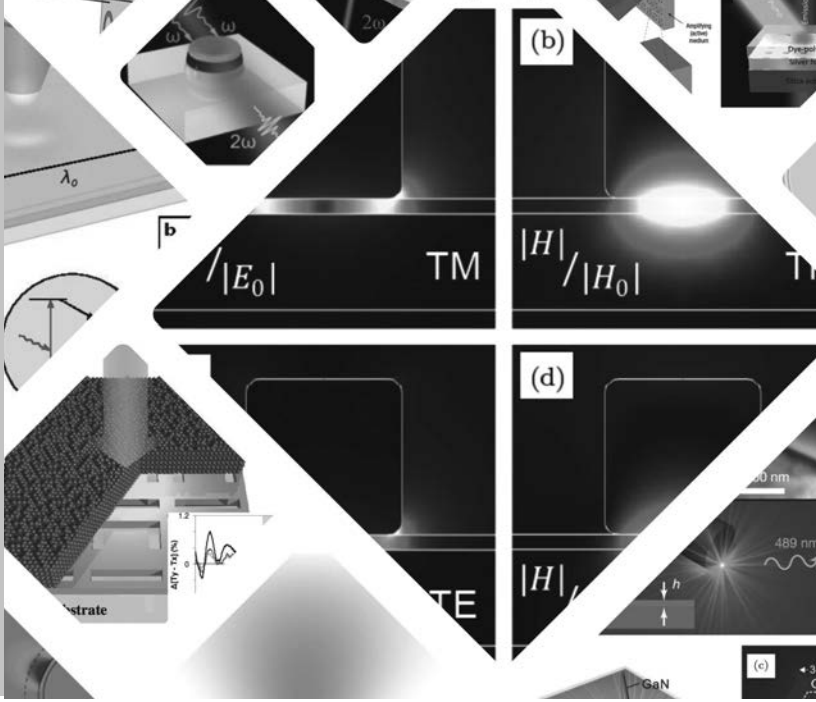
Vol. 14 *Synthesis and Applications of Optically Active Nanomaterials*
by Feng Bai (Henan University, China) and Hongyou Fan (Sandia National Laboratories, USA)

For further details, please visit: http://www.worldscientific.com/series/wssnn

(Continued at the end of the book)

Volume
16

World Scientific Series in
Nanoscience and Nanotechnology



World Scientific Handbook of
Metamaterials and Plasmonics

Volume 3
Active Nanoplasmonics and Metamaterials

Volume Editor
Ortwin Hess
Imperial College London, UK

Editor
Stefan A. Maier
Imperial College London, UK

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Cataloging-in-Publication Data

Names: Maier, Stefan A., editor.

Title: World Scientific handbook of metamaterials and plasmonics / edited by: Stefan Maier (Imperial College London, UK).

Description: Singapore ; Hackensack, NJ : World Scientific, [2017] | Series: World Scientific series in nanoscience and nanotechnology, ISSN 2301-301X ; volume 16 | Includes bibliographical references and index.

Identifiers: LCCN 2017029450 | ISBN 9789813227613 (set) | ISBN 9789813227620 (v. 1 ; hardcover ; alk. paper) | ISBN 9813227621 (v. 1 ; hardcover ; alk. paper) | ISBN 9789813227637 (v. 2 ; hardcover ; alk. paper) | ISBN 981322763X (v. 2 ; hardcover ; alk. paper) | ISBN 9789813227644 (v. 3 ; hardcover ; alk. paper) | ISBN 9813227648 (v. 3 ; hardcover ; alk. paper) | ISBN 9789813227651 (v. 4 ; hardcover ; alk. paper) | ISBN 9813227656 (v. 4 ; hardcover ; alk. paper)

Subjects: LCSH: Metamaterials. | Plasmons (Physics)

Classification: LCC TK7871.15.M48 W67 2017 | DDC 621.3028/4--dc23

LC record available at https://lccn.loc.gov/2017029450

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

Where applicable, figures in this publication will appear in color for the online version.

Desk Editor: Rhaimie Wahap

Typeset by Stallion Press
Email: enquiries@stallionpress.com

Printed in Singapore

Preface by Main Editor

It is our pleasure to present you this *Handbook of Metamaterials and Plasmonics*, charting the tremendous progress that has occurred in this exciting area of research over the last years. What continues to fascinate me about the field above all is its interdisciplinary broadness — we have arrived at a stage where metamaterials make an impact on many arrays of science where control over waves is a prominent ingredient — be they electromagnetic, acoustic, elastic, or even seismic! In these four volumes, we hence attempt to set out the richness of the field, taking metamaterials in the widest sense as artificial media with sub-wavelength structure for control over wave propagation.

Volume 1 focuses on the fundamentals of electromagnetic metamaterials in all their richness, including metasurfaces and hyperbolic metamaterials. Volume 2 widens the picture to include elastic, acoustic, and seismic systems, whereas Volume 3 presents nonlinear and active photonic metamaterials. Finally, Volume 4 includes recent progress in the field of nanoplasmonics, used extensively for the tailoring of the unit cell response of photonic metamaterials.

In its totality, we hope that this handbook will be useful for a wide spectrum of readers, from students to active researchers in industry, as well as teachers of advanced courses on wave propagation. I want to thank the volume editors Ekaterina Shamonina,

Richard Craster, Sébastien Guenneau, Ortwin Hess and Javier Aizpurua, and all the authors for their excellent and sustained work to put these four volumes together.

Stefan Maier
Imperial College London, UK
April 2017

Preface by Volume Editor

In photonics and optoelectronics, metals were for decades perceived as being rather dull, devoid of interesting or useful optical properties that one could harness for optical components and devices. Scientists likewise believed that the wave nature of light imposed seemingly fundamental constraints on the degree to which we can resolve nanoscopic details of an object, or reduce the size of photonic devices. This changed, however, with the discovery of surface-enhanced Raman scattering (SERS) in 1974, drewing appreciable attention to the optical properties of metals and, more recently, the strong need to integrate photonics with nano-electronics that propelled nanoplasmonics and nanoplasmonic metamaterials to one of the most actively researched areas of nanoscience and nanotechnology.

Indeed, when arrays of nanoplasmonic elements (each smaller than the wavelength) are assembled in two or three dimensions they may act as the 'molecules' or 'atoms' of an effective, engineered material — a metamaterial. The macroscopic optical parameters (effective permittivity, permeability, refractive index, impedance) of such a medium can be made to enter highly unusual regimes, e.g., exhibiting a permittivity close to zero, a *negative* refractive index or magnetism at infrared and optical frequencies. This leads to negative refraction and to completely surprising results, such as the 'perfect lens', the ability of a planar slab of a negative-index (e.g. $n = -1$)

metamaterial to focus lightwaves without an intrinsic limit to resolution, or the possibility to dramatically slow down and localise light pulses in metamaterial and plasmonic heterostructures. Even greater control on the propagation of electromagnetic waves is achieved when the meta-atoms are allowed to vary smoothly in all three dimensions according to 'transformation optics' design rules, forming anisotropic materials that can, e.g., cloak objects of arbitrary shape from incident radiation. Metamaterials have also been shown to allow for an exceptional degree of control on the polarisation of light, as well as for enhanced nonlinear effects that in some cases can arise entirely from the magnetic field of light — in sharp contrast to nonlinear effects in ordinary optical media, which always arise from the electric field.

However, the interaction of photons with the conduction electrons of a plasmonic metamaterial gives rise to dissipative optical losses, which for visible light can be significant (of the order of 1000 cm^{-1} or more). In essence, there exists a trade-off between localization and loss — the more tightly light fields are localized to the surface of a metal, the higher is the fraction of the modal energy inside the metal, leading to an increase of dissipation as the effective mode volume is reduced. Many applications can be negatively affected by the losses, and although this limitation has been pointed out right from the beginning of plasmonics and metamaterials research, it has only more recently become possible to identify effective strategies to mitigate losses. One of the most promising is the use of gain materials placed in close proximity to the metal surface. The enhanced electric fields associated with the plasmonic excitations strongly modify the radiative and nonradiative properties of the deployed emitters, and for suitably optimised structures loss-free operation or steady-state net amplification is possible. Higher gain densities may lead into the lasing regime if both dissipative and radiative losses are overcome.

Concurrently, (semiconductor) lasers have undergone a continuous miniaturization culminating in the recent experimental demonstration of subwavelength nanolasing. Advancing beyond traditional cavity-concepts, recent nanolasers employ plasmonic resonances for feedback, allowing them to concentrate light into mode volumes that

are no longer limited by diffraction. The use of localized surface plasmon resonances as cold-cavity modes, however, is only one route to lasing on subwavelength scales. Lasing, in fact, does not require modes predefined by geometry but merely a feedback mechanism and it has been shown that cavity-free lasing is not restricted to micrometer sizes in disordered media (using random feedback) but using nanoplasmonic waveguide structures can even be realized on subwavelength scales using *local feedback* (optical vortex states close to singularities in the local density of photonic states).

This volume on "Active Nanoplasmonics and Metamaterials" brings together the topics of nanoplasmonic lasing and spontaneous emission control and active and nonlinear metamaterials. It comprises six chapters discussing the new science and some of the surprising applications of nano-plasmonics and metamaterials when combined with quantum emitters and (quantum) gain materials. The volume shall start by providing an introduction to active plasmonics, plasmonic amplification and lasing (Chapter 1) and followed by Chapter 2, discussing plasmonic stopped-light nanolasing. Chapter 3 then reviews nonlinear metamaterials and Chapter 4 talks about controlled radiative dynamics using plasmonic microcavities while Chapters 5 and 6 discuss metamaterials and plasmonics with optical gain and active metamaterials.

Hess Ortwin
Imperial College London, UK
April 2017

This page intentionally left blank

Contents

CHAPTER 1

Active Plasmonics, Plasmonic Amplification and Lasing

ELHAM KARAMI KESHMARZI and R. NIALL TAIT

Carleton University, Canada

PIERRE BERINI*

University of Ottawa, Canada

1.1. Introduction

Plasma oscillations of conduction electrons on the surface of a metallic structure can couple to an incident electromagnetic fields at optical and near infrared frequencies to generate coupled excitations called surface plasmon polaritons (SPPs).[1] SPPs are transverse magnetic (TM) polarized excitations propagating along the surface of metal-dielectric structures. They are tightly bound to the metal's surface and are highly surface sensitive.

SPPs exhibit unique and remarkable properties, including strong field enhancement, subwavelength field localization, high sensitivity to the local dielectric environment, and an energy asymptote in their dispersion curves, which offer useful applications in integrated

*Corresponding author: pberini@uottawa.ca

photonics,[2] bio-sensing,[3] imaging,[4] spectroscopy,[5] nano-lithography[6] (see also Refs. [7] and [8] for reviews of SPP applications).

Significant improvements in micro-and nano-fabrication techniques have facilitated the study of various SPP supporting structures. Moreover, the existence of SPPs in nano-scale structures has enabled the investigation of light-matter interaction in the scale of single molecules.[9]

Unfortunately, the usefulness of SPPs is limited as these excitations have a short lifetime. The intrinsic loss in metals restricts the SPP propagation length. Inter-band transition absorption at short wavelengths and free electron collisions and scattering are the main sources of attenuation in metals. These losses could be reduced by carefully selecting the operating wavelength and via improvements in fabrication techniques. However; they cannot be fully eliminated. As the SPP mode becomes more confined, the mode's overlap with the metal is larger and so is the mode's attenuation.

Increasing the propagation length, while simultaneously maintaining confinement, has been a topic of investigation for over a decade. Loss compensation of SPPs is generally accomplished by adding optical gain to the dielectric adjacent to the metal. Optical gain has shown to reduce or even eliminate the attenuation, producing lossless SPPs propagation. It may also exceed the attenuation and produce amplification of SPPs.[10] SPP amplifiers can be integrated with passive plasmonic devices to improve their performance. SPP amplification combined with a feedback mechanism enables coherent sources of SPPs, or SPP lasers (spasers).[11] Nano-scale lasers are also feasible by coupling the surface plasmon resonance (SPR) of metallic nanoparticles to an optical gain medium.[12] A new application of amplified SPPs has emerged in optical parity-time (PT) symmetric materials where the intrinsic loss of SPPs is balanced with optical gain to produce a PT symmetric medium satisfying $\varepsilon(z) = \varepsilon^*(-z)$, where $\varepsilon(z)$ is the relative permittivity of the medium as a function of spatial coordinate z.[13] These materials exhibit extraordinary optical phenomena such as unidirectional reflectance in linear media.[14]

This chapter reviews recent studies on SPP loss compensation, amplification, and lasing. Our focus is on the structures and media

involved and developed during last ten years and so. In Section 2, SPP-supporting structures are classified based on how the SPP mode is confined in the structure. A brief description of quantum mechanical processes involved in the optical amplification of SPPs is given at the end of Section 2. In Section 3, we review developments and achievements in loss compensation, amplification, or lasing with SPPs in one-dimensional structures. In Sections 4 and 5 we consider two-dimensional and localized SPP structures respectively. In Section 6, a brief review of the applications of amplified SPPs is presented. Finally concluding remarks are given in Section 7.

1.2. Review of Surface Plasmon Polariton Structures

SPPs can be confined to the metallic surface of various structures. Metal strips, metal nano-particles, metal-cladded dielectrics are only few examples. The SPP mode size can be as small as the nano-scale, beating the diffraction limit.[15] The reason that SPPs can be concentrated to subwavelength dimensions is that their momentum can be much larger than that of photons. In fact, near the metal's plasma frequency the SPP group index grows significantly and maximum confinement can be attained, although the SPPs lifetime at this extrema is very short.

In this chapter, we classify SPP structures depending on whether they confine SPPs in one, two or three dimensions, and we highlight each structure's properties in terms of loss and confinement.

1.2.1. *One dimensional SPP structures*

In one-dimensional structures, SPPs are confined along only one direction, perpendicular to the surfaces of metal-dielectric interface(s) in the structure. One-dimensional planar structures include single or double metal-dielectric interface(s) in structures as sketched in Figs. 1.1(a) to 1.1(c).

Figure 1.1(a) shows a single metal-dielectric interface which supports a bound SPP of wavenumber $\gamma = \alpha + j\beta = \mathrm{j}2\pi\lambda_0^{-1}[\varepsilon_{\mathrm{r,m}}\varepsilon_{\mathrm{r,d}}/(\varepsilon_{\mathrm{r,m}} + \varepsilon_{\mathrm{r,d}})]^{1/2}$ where $\varepsilon_{r,m}$ and $\varepsilon_{r,d}$ are the metal and dielectric

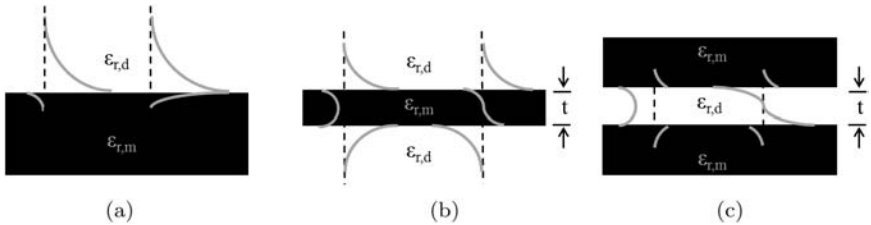


Fig. 1.1. Sketch of common one-dimensional SPP structures. (a) Single interface, (b) symmetric dielectric-cladded metal film, (c) symmetric metal-cladded dielectric. (Adapted from Ref. [10].)

relative permittivity, and λ_0 is the free space wavelength.[1] It is necessary to have $\mathrm{Re}(\varepsilon_{r,m}) < 0$ at λ_0 in order for SPPs to be excited and bound to the interface. Most metals at optical and infrared wavelengths have a negative real part of permittivity. The SPP fields decay exponentially from the interface into the metal and dielectric regions as sketched in Fig. 1.1(a).

The metal's dispersion leads to the divergence of the SPP wavenumber at the photon energy that corresponds to the energy asymptote. The dispersion characteristic of an SPP on a single Ag-SiO_2 interface is shown in Fig. 1.2.[16]

The SPP wavenumber reaches an extrema at the asymptote. However, the depth of the energy asymptote is limited due to the metallic losses and the extrema bends back to the left of the light line, where the SPP becomes radiative. The modal characteristics of the SPP on the single Ag-SiO_2 interface at three different wavelengths are summarized in Table 1.1, where $n_{eff} = \beta\lambda_0/2\pi$ is the SPP mode effective index, δ_w is the 1/e mode field width, and $L = 1/(2\alpha)$ is the 1/e propagation length.

Figure 1.1(b) shows another common one-dimensional SPP waveguide comprising a thin metal film bounded by symmetric dielectrics. This structure supports symmetric and asymmetric SPP supermodes.[16] By decreasing the metal's thickness ($t \to 0$) the attenuation of the symmetric mode decreases to a few orders of magnitude below that of the single-interface SPP of the corresponding structure. The symmetric mode on the thin metal film is thus

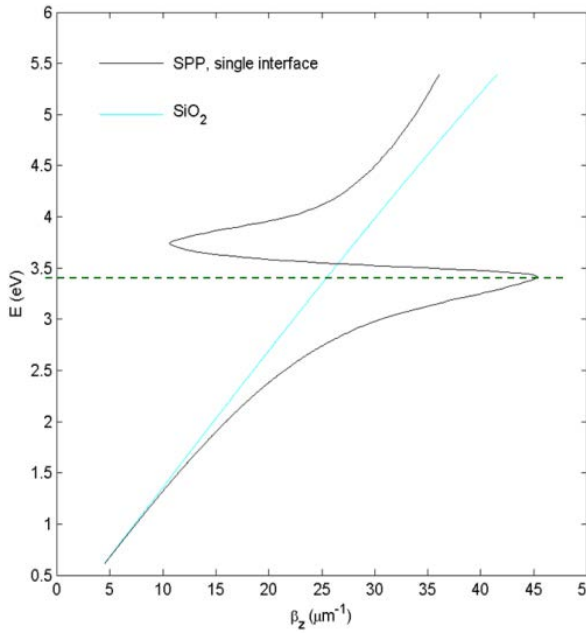


Fig. 1.2. SPP dispersion on a Ag-SiO_2 interface; the cyan curve is the light line in SiO_2 and the dashed line marks the energy asymptote. (Adapted from Ref. [16].)

Table 1.1. SPP mode's characteristics on Ag-SiO_2 interface. (Adapted from Ref. [16].)

			2α		
λ_0(nm)	n_{eff}	δ_w(nm)	(cm^{-1})	(dB/μm)	L (μm)
360	2.537	44	5×10^5	218	0.02
633	1.565	176	1.6×10^3	0.71	6
1550	1.457	1269	1×10^2	0.044	1

termed a long-range SPP (LRSPP).[17] However, the lower attenuation of the LRSPP mode is accompanied with lower confinement. On the contrary, the asymmetric mode's attenuation increases with decreasing t and is therefore termed the short-range SPP (SRSPP). When the metal film is cladded symmetrically neither mode has a

cut-off thickness t. The power attenuation (2α) of the LRSPP and SRSPP modes at $\lambda_0 = 1550$ nm in the Ag-SiO_2 system with t = 20 nm is 0.0012 and 0.45 dB/μm respectively, both increasing as λ_0 decreases.[16]

Figure 1.1(c) shows the complementary structure where a thin dielectric film is cladded by metals. If the dielectric is thick, this waveguide behaves as an ideal parallel-plate waveguide supporting many TE and TM modes. The symmetric SPP mode remains guided as the dielectric thickness shrinks ($t \to 0$), and it has no cut-off so it can be confined to an arbitrarily small width.[16] However, the mode's attenuation increases as the thickness shrinks, so it is also a SRSPP. Its attenuation in the Ag-SiO_2 system with t = 50 nm is 0.85 dB/μm at $\lambda_0 = 1550$ nm, increasing as λ_0 decreases.[16]

1.2.2. *Two dimensional SPP structures*

Two-dimensional SPP structures provide confinement along two dimensions in the plane transverse to the direction of propagation. They typically comprise thin structured slabs of metals and dielectrics on a substrate. As a natural extension of some one-dimensional structures, they enable integrated plasmonic circuits.[2]

The most common two-dimensional SPP waveguides are presented in Fig. 1.3. In Fig. 1.3(a) shows a metal stripe of limited width embedded in a symmetric dielectric medium.[18] The attenuation of the main symmetric SPP mode in this structure is about ten times lower than of that on the infinitely wide film. It is also several orders of magnitude smaller than the single-interface SPP in the limit $t \to 0$. The main symmetric SPP mode in this regime is also known as a LRSPP which offers low attenuation at the cost of less confinement. The main anti-symmetric mode has high attenuation and is also referred to as SRSPP.

Figure 1.3(b) sketches a dielectric-loaded SPP waveguide, where SPPs are confined in a high-index dielectric ridge deposited on a metal film.[19] The mode's attenuation in this structure ranges from that of the corresponding single-interface SPP to about ten times higher. Transverse SPP confinement in the structure of Fig. 1.3(c)

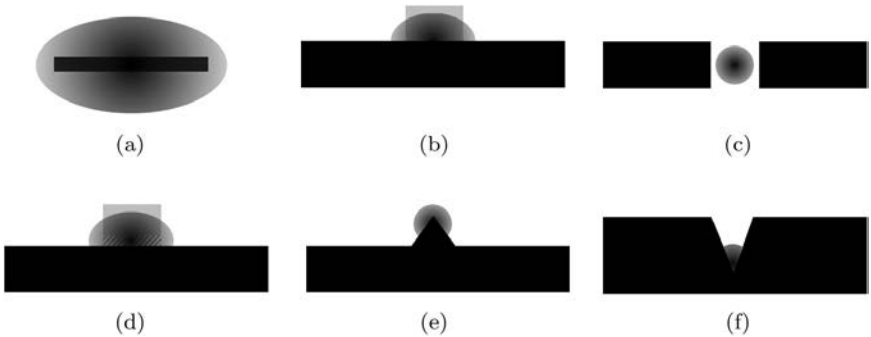


Fig. 1.3. Sketch of common two-dimensional SPP waveguides. (a) Metal stripe, (b) dielectric loaded, (c) gap, (d) hybrid, (e) wedge, and (f) channel waveguide. (Adapted from Ref. [10].)

is produced by a two-dimensional gap waveguide, which resembles rotated metal claddings.[20] SPPs are confined strongly within the gap but with a high attenuation, comparable to the corresponding metal-cladded configuration. Figure 1.3(d) shows a two-dimensional hybrid SPP waveguide[21,22] where a very thin, low-index dielectric is sandwiched between a high-index dielectric and a metal substrate. This waveguide allows strong localization to the thin low-index gap, and its attenuation is bounded by the corresponding low-and high-index SISPPs. Finally Figs. 1.3(e) and 1.3(f) sketch wedge and channel waveguides[15] consisting of a metal protrusion and a groove in a metal substrate respectively.

SPP waveguides in general trade confinement for attenuation,[16] whether by varying the wavelength (Table 1.1) or the geometry (Fig. 1.3). The trade-off is fundamental: increased SPP confinement leads to greater overlap with the metal(s), leading to greater attenuation. SPP attenuation therefore spans, in general, a very broad range (1–10^5 cm^{-1}), with the most strongly confined modes requiring the greatest gains for loss compensation and lasing.

1.2.3. *Localized SPPs*

The interaction of metallic nanoparticles with an electromagnetic field can lead to non-propagating SPP excitations which are referred

to as localized SPPs. Localized SPPs are resonant surface plasmon modes, with a resonant frequency dependent on nanoparticle shape, size, composition, and the surrounding dielectric.[23,24] As localized SPPs are resonant, a gain medium in the vicinity can give rise to lasing action through stimulated emission of radiation. This particular type of SPP laser is often referred to as a spaser which is currently of strong interest as it enables a nanoscale source of SPPs.[12,25] Metallic nanoshells (complementary to metallic nanoparticles) and metal nanoparticles on a metal substrate also confine SPPs in three dimensions and could also be used as spasers.

1.2.4. *Quantum processes involving SPP optical amplification*

Similar to light, SPPs can be quantized in energy (considering dispersion[26,27] and absorption[28]) to provide a quantitative description of microscopic interactions between SPPs and matter. The quantum nature of SPPs has been investigated and proved experimentally, for example by the excitation of single SPPs along a metallic nanowire. SPPs can be created or annihilated in a medium consisting of, *e.g.*, optically-active atoms and molecules through processes such as absorption, spontaneous emission and stimulated emission. Some quantum processes that may occur for dipolar emitters near a metal-dielectric interface supporting single-interface SPPs are sketched in Fig. 1.4.

In close vicinity to a metallic interface, dipoles may also emit radiative modes, or interact with electron-hole pairs in the metal via dipole-dipole transitions. The photonic mode density (PMD), which is a critical factor in the assessment of spontaneous emission rates, is modified due to the presence of a metallic interface.[29]

1.3. Amplification and Lasing in One-dimensional SPP Structures

As it was earlier pointed out, one-dimensional SPP structures are planar waveguides comprising a single metal-dielectric interface or

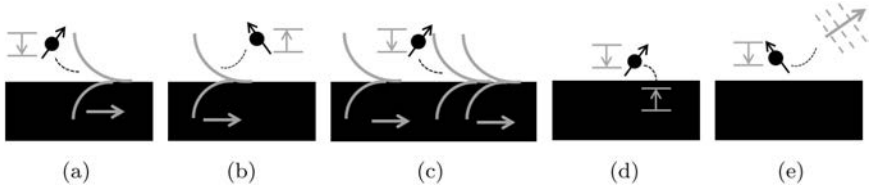


Fig. 1.4. Optical processes occurring for dipoles (filled black circles) near a single metal–dielectric interface. The magnitude of the SPP's transverse electric field is shown along with the associated dipole energy transition. The black dotted curves indicate energy transfer. (a) Spontaneous emission of SPPs, (b) absorption of SPPs, (c) stimulated emission of SPPs, (d) creation of electron-hole pairs, (e) spontaneous emission of radiation. (Adapted from Ref. [10].)

a thin metal film sandwiched in dielectric, or a thin dielectric core cladded with metals (Fig. 1.1).

SPPs are normally excited by grating or prism coupling configurations in these structures.[1] In prism coupling, such as the so-called Kretschmann-Raether configuration, a metal film is deposited onto the base of a high-index prism and the lower index dielectric covers the other side of the film. The light incident on the prism above the critical angle where total internal reflection occurs, couples to SPP excitations at the other metal-dielectric interface. In grating coupling a periodic perturbation in the metal film, such as bumps, corrugations, or slits is used to diffract the incident light and excite the SPP along the metal-dielectric interface. The role of the grating or prism in general is to increase the in-plane momentum of TM-polarized incident light to match that of the SPP, because SPPs cannot be directly excited by light (due to their larger momenta — Fig. 1.2). In the following we review the literature on SPP amplification or lasing in various one-dimensional structures.

1.3.1. *Single metal-dielectric interface*

SPP amplification in a single-interface structure comprising a thin Ag film sandwiched between a prism and a gain medium was investigated in theory by Plotz *et al.*[30] Using Fresnel reflection formulas, they computed the reflectance of light incident onto the prism versus angle of incidence and as a function of the gain of the medium.

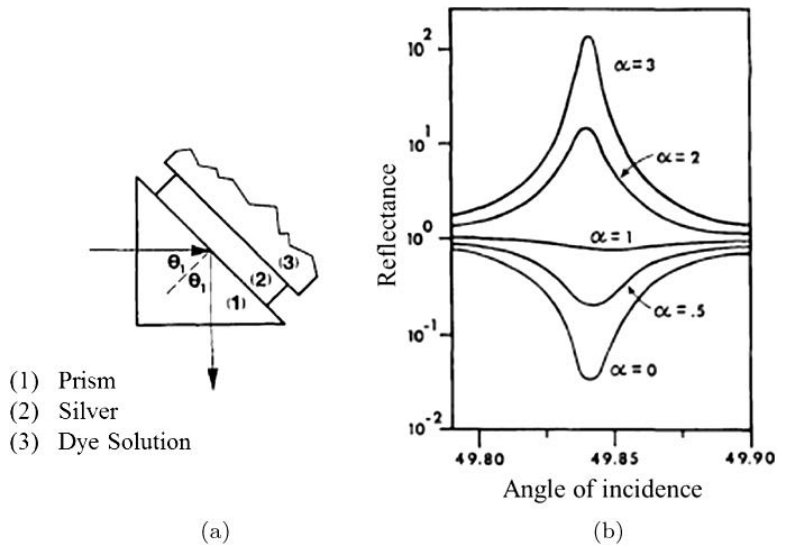


Fig. 1.5. (a) The attenuated total reflection configuration with θ_I as the angle of incidence. (b) Reflectance computed versus angle of incidence (θ_I) and as a function of gain (α). (Adapted from Ref. [30].)

In the passive case, near the angle for surface-plasmon excitation, the usual attenuated total reflection (ATR) associated with the SPP mode was observed. However, by increasing the gain, the ATR monotonically reduced such that above a certain gain threshold, enhanced reflectivity resulted. They also showed that above the gain threshold one could adjust the metal film's thickness to produce a singularity in the reflectance. Figure 1.5 shows the ATR configuration used in this modeling along with the reflectance computed as function of gain.

The same structure was also studied by Sudarkin and Demkovich[31] considering higher gain in the amplifying medium. They showed that at high gain the Fresnel formulas are not valid to calculate the reflectance of the incident beam, which is of limited width. They also predicted super-luminescence of surface waves in the high gain regime and were the first to suggest a SPP laser based on this phenomenon.

The propagation properties of SPPs along a planar single interface structure were theoretically investigated by Nezhad *et al.*[32] They analyzed the case of an infinite interface between a metal and a gain medium, rigorously working with propagation constants and Poynting vectors, and found the required gain coefficient for lossless SPP propagation. A material gain of $\sim$1260 cm^{-1} was estimated for a lossless SPP propagation for an Ag-InGaAsP interface at $\lambda_0 = 1550$ nm.

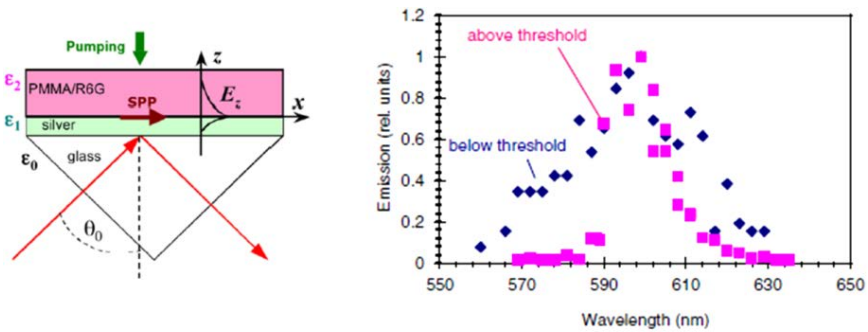
Kumar *et al.*[33] considered a metal-semiconductor diode structure comprising a p-n junction of GaAs adjacent to an Au substrate. The electron-hole recombination energy was coupled to SPPs on the Au-GaAs interface. They showed how the carrier injection level in GaAs controlled the spontaneous emission rate and provided gain for SPP amplification. Lu *et al.*[34] considered optical parametric amplification of SPPs in a nonlinear hybrid waveguide consisting of Ag-$LiNbO_3$. The seed and pump wavelengths required for efficient power conversion were found. They estimated $\sim$30 dB gain over 3 mm of coupling length for a pump intensity of $\sim$50 $\mathrm{MW/cm}^2$. Sirtori *et al.*[35] used a single interface SPP waveguide on a quantum cascade semiconductor laser at far-infrared wavelengths ($\lambda \sim 11\,\mu$m). Their goal was to increase the mode's confinement and overlap with the gain medium, while simultaneously reducing the total layer thickness in the structure. Shortly later, Tredicucci *et al.*[36] similarly proposed a single-mode SPP laser at $\lambda_0 \sim 17\,\mu$m, in a distributed feedback (DFB) configuration using two-metal (Ti/Au) grating adjacent to a quantum cascade active semiconductor.

Seidel *et al.*[37] experimentally demonstrated SPP stimulated emission at optical wavelengths. They used 39 and 67 nm thick Ag films in contact with liquid dye gain media (Cresyl violet or Rhodamine 101 in ethanol) in the Kretschmann-Raether configuration. A p-polarized probe at $\lambda_0 = 633$ nm was focused on the Ag-dye interface with a specific incidence angle near the SPP excitation angle while the dye medium was pumped at $\lambda_0 = 580$ nm. The ATR was measured as a function of the probe incidence angle, with and without pumping the gain medium. The difference between the pumped and un-pumped

cases was attributed to stimulated emission of SPPs on the Ag-dye interface. The reflectance increased over the SPP excitation angle for the 39 nm thick Ag films whereas it narrowed and deepened for the 67 nm thick Ag films, both cases in agreement with theory.[31] A similar study was conducted by Noginov *et al.*[38] but with the gain medium formed as a 10 μm-thick film doped with Rhodamine-6G. The gain medium was pumped at $\lambda_0 = 532$ nm by a Nd:YAG laser and a He-Ne laser beam at $\lambda_0 = 594$ nm was illuminating the Ag-gain interface through the prism. Increased reflectance at the SPP excitation angle was observed due to the stimulated emission of SPPs on the interface. They obtained an optical gain of 420 cm^{-1} with a dye concentration of $N = 1.2 \times 10^{22}$ cm^{-3} and a pump intensity of $I = 1.7 \times 10^7$ W/cm^2, which compensated $\sim$35% of the SPP loss. In another experiment[39] they excited SPPs through pumping a thinner ($\sim$3 μm) dye-doped polymer film at the dye's peak absorption wavelength, and decoupled spontaneously emitted SPPs to the prism in ATR configuration. They demonstrated narrowing of the SPP spectra by increasing the pump intensity, which was attributed to stimulated emission of SPPs.

Figure 1.6(a) sketches the SPP excitation and decoupling in the ATR configuration used in the experiment. The SPP emission spectrum narrowing with increasing pump intensity from 10.9 mJ/cm^2 to 81.9 mJ/cm^2 is shown in Fig. 1.6(b).

Amplified spontaneous emission of SPPs (ASE-SPP) at the interface of a 1 μm-thick PMMA layer doped with PbS quantum dots (QDs) and a 100 nm thick Au film on a Silica substrate was observed by Bolger *et al.*[40] SPPs were excited at the QDs emission peak and out-coupled through a grating on the interface. The full-width half-maximum (FWHM) SPP emission was observed to decrease with increasing pump intensity as expected for ASE-SPP. The pump intensity threshold for ASE-SPP was measured as 5 W/cm^2. It was found that ASE-SPP at high pump intensities limits the available gain such that only a 30% increase in SPP propagation length was obtained. The gain depletion due to ASE was suggested as the reason.



(a) (b)

Fig. 1.6. (a) Sketch of experimental the sample, excitation and decoupling of SPPs. (b) Spectra of SPP emission at pump intensities below and above threshold. (Adapted from Ref. [39].)

1.3.2. *Metal slab in symmetric/asymmetric dielectric*

The properties of strong optical amplification of SPPs for flat or corrugated (∼10 nm period) Ag slabs in contact with a gain medium were investigated theoretically by Avrutsky.[41] SPP resonance was obtained at $\lambda = 350$ nm, when the optical gain increased to ∼80,000 cm^{-1}. The SPP resonance was accompanied by a huge group index ($\sim 5.4 \times 10^4$), extremely low group velocity (∼1 km/s), highly localized SPPs, and an abrupt change of positive SPP loss to negative SPP loss (meaning loss overcompensation). The large material gain required in this scheme sounds challenging, but it is interesting that such extreme confinement may be possible. Although strongly enhanced ASE-SPP may compromise the available gain and noise performance.

Nezhad *et al.*[32] considered a 40 nm Ag slab embedded in InGa AsP gain media and found that ∼360 cm^{-1} of gain is required for lossless LRSPP propagation at $\lambda_0 = 1550$ nm.

Lasing in plasmonic bandgap structures was investigated[42–44] using a 2D corrugated metal film sandwiched between two symmetric dielectrics doped with dye molecules. It was shown experimentally that the dye molecules' fluorescence was highly enhanced in the vicinity of the corrugated film. SPP lasing was proposed in these structure

provided that the dye peak emission wavelength falls at the edge of the plasmonic bandgap, supporting a standing LRSPP mode due to highly enhanced SPP fields and lower absorption of this mode. The same concept was also considered by Winter *et al.*[45] who argued that fluorescent dye can also decay into asymmetric SRSPP guided modes, reducing the available gain for stimulated emission of LRSPP.

Genov *et al.*[46] developed a new method based on a quasi-metal approximation for the explicit solution of the dispersion relation of an insulator-metal-insulator (IMI) structure, where the metal slab is bounded by multiple quantum well (MQW) gain media. Using their method they could obtain the critical gain required for lossless SPP propagation in a thin Ag film cladded by InGaAsP, InGaN MQW gain media.

De Leon and Berini developed a theoretical model for treating SPP amplification in planar structures where gain is provided by a laser dye solution.[47,48] They assumed a four-level dipolar gain system, and accounted for the position dependency of dye molecules' lifetime as well as an inhomogeneous pump intensity distribution. Through their modeling, they predicted lossless LRSPP propagation on a Ag film bounded in symmetric dielectric media involving R6G in solvent using a modest pump intensity of $\sim$200 KW/cm^2 and a reasonable dye concentration of $N = 1.8 \times 10^{18}$ cm^{-3}. They also considered a single interface system using the same gain media but found that a higher dye concentration ($N = 2.4 \times 10^{19}$ cm^{-3}) and a much stronger pump intensity ($\sim$3.5 MW/cm^2) were needed for lossless SPP propagation. They later investigated the noise figure for amplified single interface SPPs and LRSPPs[49] and found that the noise figure of LRSPPs is less than that of the single interface SPPs, and that it diverges as the energy asymptote is approached in both cases due to increased spontaneous emission rate. They also showed[50] that the spontaneous decay rate into LRSPPs is lower than for decay into SRSPPs in close proximity to a metal slab in a symmetric one-dimensional structure. Figure 1.8 shows the computed decay rates into the guided modes of an infinite metal slab in symmetric dielectric.

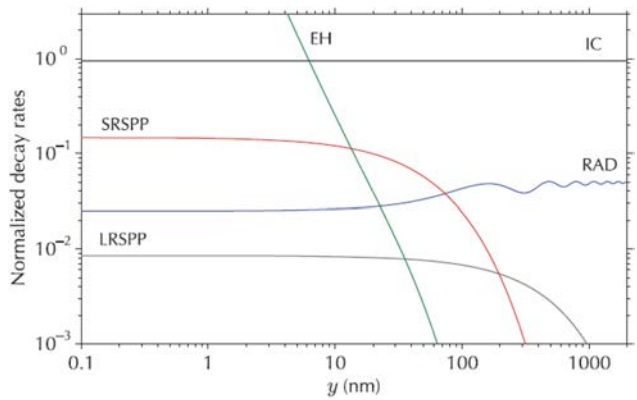
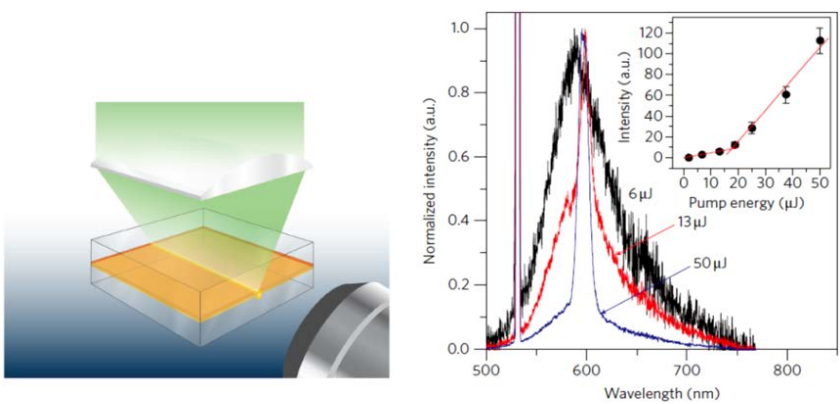


Fig. 1.7. Normalized excited state decay rates for isotropic dipole into different energy decay channels on an infinite metal slab in symmetric dielectric media. (Adapted from Ref. [50].)



(a) (b)

Fig. 1.8. (a) ASE-LRSPP measurement setup in a symmetric dielectric-metal-dielectric configuration. (b) Normalized spectrum of the TM polarized component of output emission for different pump intensities. (Adapted from Ref. [51].)

Gather *et al.*[51] reported LRSPP gain at visible wavelengths in a symmetric structure consisting of a 1 μm thick fluorescent polymeric gain layer and a 4 nm thick Au film. Figure 1.8(a) illustrates the LRSPP structure under test which was optically pumped from the top using 5 ns laser pulses at $\lambda_0 = 532$ nm. They performed

ASE-LRSPP measurements by increasing the pump stripe length and demonstrated the emission threshold and spectrum narrowing as plotted in Fig. 1.8(b). They obtained an LRSPP net gain coefficient of $\sim$8 cm^{-1} at $\lambda_0 \sim 600$ nm.

Hahn *et al.*[52] considered the interaction of the LRSPP mode with optical gain in a symmetric PMMA-Ag-PMMA structure where both claddings (top and bottom PMMA) were doped with IR-140 dye. They theoretically showed that the gain medium in the bottom cladding doubles the LRSPP gain at high pump intensities $\sim$4 MW/cm^2, even though the pump is applied from the top and is partially reflected by the Ag film. They also measured the LRSPP gain by the variable stripe length (VSL) method and found a gain coefficient of 16.7 cm^{-1} with a pump intensity of 4 MW/cm^2.

I. P. Radko *et al.*[53] implemented an asymmetric dielectric-metal-dielectric waveguide consisting of a 50 nm-thick Au film deposited on a quartz substrate and covered with a thin layer of PMMA doped with lead-sulphide (PbS) QDs. The sample was optically pumped from the top at $\lambda_0 = 532$ nm, whereas a CW probe at $\lambda_0 = 860$ nm was used to excite the SPP mode in the waveguide through a grating coupler. By collecting the leakage radiation from the quartz side of the sample at the Kretschmann angle they could estimate a SPP gain of $\sim$200 cm^{-1}.

1.3.3. *Metal-cladded dielectric*

Maier[54] computed the gain required for lossless SRSPP propagation in a symmetric metal-insulator-metal (MIM) heterostructure where the core material exhibits optical gain. Gain coefficients of 1625 cm^{-1} and 4830 cm^{-1} were required for lossless propagation at $\lambda_0 = 1550$ nm in 500 nm-thick and 50 nm-thick semiconductor cores, respectively.

Chang and Chuang[55,56] formulated quasi-orthogonality and normalization conditions suitable for computing field-enhanced spontaneous and stimulated emission rates in micro- and nano-sized metal-cladded dielectrics. They also modified the definition of the confinement factor in SPP nano-lasers, where the group velocity is very low and the SPP fields are highly confined.

SPP amplification and lasing were investigated by Li and Ning[57] in metal-semiconductor-metal (MSM) structures. They demonstrated a SPP mode power gain 1000 times greater than the semiconductor's material gain near the energy asymptote where the group velocity is very low. They also investigated the amplification and lasing threshold conditions in these structures assuming electrical injection.[58] Khurgin and Sun[59] investigated similar MSM structures near the SPP energy asymptote. They found that a significantly high current density injection of $\sim$100 kA/cm^2 is required for loss compensation due to the Purcell effect inherent to sub-wavelength confinement in these structures.

Chen *et al.*[60] used a finite difference time domain (FDTD) method to model amplified SPP propagation in MSM structures. They specifically considered an Au-InGaAs-Au structure and predicted lossless SRSPP propagation in a 75 nm-thick structure with a material gain of 2500 cm^{-1} at $\lambda_0 = 1550$ nm.

Yu *et al.*[61] investigated the reflection and transmission of a short-length, metal-cladded dielectric waveguide coupled to a small cavity filled with an InGaAsP gain medium, showing that the cavity losses can be fully compensated for a material gain of $\sim$2000 cm^{-1} at $\lambda_0 \sim$ 1550 nm. The gain assisted MDM plasmonic waveguide can work as an ON/OFF switch in the presence/absence of pumping.

Hill *et al.*[62] demonstrated electrically-pumped lasing from rectangular etched InGaAs pillars coated by 20 nm silicon nitride then by a thick Ag layer. Vertical confinement was provided by steps in refractive index in the gain region, whereas sub-wavelength confinement along the horizontal direction occurs due to the Au coating. Mirrors at the end facets define a Fabry-Perot cavity. Lasing was obtained at cryogenic temperatures (10 K) observed via leakage radiation microscopy through the substrate.

1.4. Amplification and Lasing in Two-dimensional SPP Structures

The most common two-dimensional SPP structures are displayed in Figs. 1.3(a)–(f). Thin metal stripes or films of narrow width bounded

by nearly symmetric dielectrics support LRSPPs with significantly lower loss than single-interface SPPs and metal slabs bounded by symmetric dielectrics. The reduced loss comes with the cost of less confinement so that sub-wavelength confinement cannot be achieved with thin metal stripes. Dielectric-loaded waveguides based on using a dielectric ridge of higher refractive index than the surrounding medium provide additional confinement through index contrast. There are also two-dimensional hybrid dielectric-SPP waveguides which take advantage of less loss and more confinement from dielectric and metal waveguides respectively. In the following a literature review of amplification and lasing in these common two-dimensional SPP waveguides is presented.

1.4.1. *Metal stripe in symmetric dielectric*

Alam *et al.*[63] considered a 1 μm-wide, 10 nm-thick Ag stripe on AlGaInAs multiple quantum wells (MQWs) covered by barrier material for the purpose of LRSPP amplification. The required material gain for lossless LRSPP propagation at $\lambda_0 = 1550$ nm was found $\sim$400 cm^{-1}. LRSPP whispering gallery modes in GaAs-based, microdisk cavities incorporating a thin metal film were investigated at $\lambda_0 = 1400$ nm by Chen and Guo.[64] They computed the resonator's Q and lasing gain threshold as $\sim$4000 and $\sim$200 cm^{-1} respectively.

Experimental demonstration of the stimulated emission of LRSPPs on a 8 μm-wide, 20 nm-thick Au stripe cladded by Er-doped glass was performed by Ambati *et al.*[65] at $\lambda_0 = 1532$ nm. Probe and pump signals were both coupled to the waveguide LRSPP mode. By applying a maximum pump power of 266 mW, the probe signal was enhanced to $\sim$50%.

De Leon and Berini[66] reported LRSPP amplification in a waveguide comprising a 1 μm-wide, 20 nm-thick Au stripe on a 15 μm-thick SiO_2 on Si, and covered by IR-140 dye in solvent index-matched to SiO_2. The structure was probed at $\lambda_0 = 882$ nm using butt-coupled polarization maintaining, single-mode fibers and pumped optically from the top at $\lambda_0 = 808$ nm using 20 mJ/cm^2, 8 ns laser pulses.

A small-signal mode power gain coefficient of $\gamma = 8.55$ dB/mm was obtained in this experiment.

A direct observation of plasmonic gain was reported by Kéna-Cohen *et al.*[67] in an asymmetric structure consisting of a 1 μm-wide, 21 nm-thick Au stripe of different lengths on a glass substrate and covered by a thin polymeric gain medium (Alq3:DCM). The thickness of the optically pumped gain medium was selected to support a confined LRSPP mode. Grating couplers were used to probe the waveguide using broadside incident beams and the pump was applied to the top. Figure 1.9(a) sketches the waveguides in top view with input and output grating couplers indicated. A typical loss-compensation measurement showing the pump (200 μJ/cm^2) and probe (TE and TM)

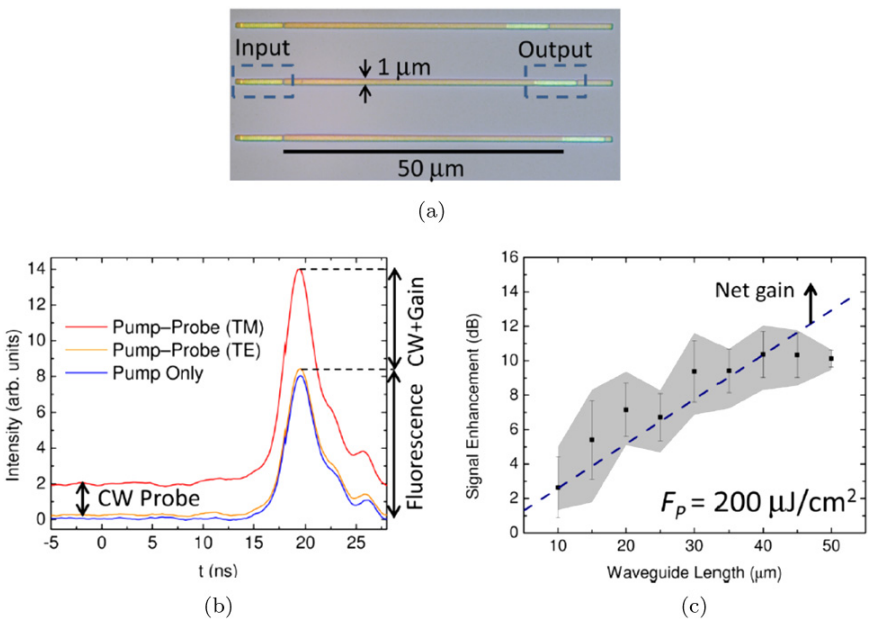


Fig. 1.9. (a) Optical micrograph of the plasmonic waveguides with locations of the input and output gratings indicated. (b) A typical loss-compensation measurement showing the pump (200 μJ/cm^2) and probe (TE and TM) for a 30 μm-long waveguide. (c) Signal enhancement as a function of length for a pump intensity of 200 μJ/cm^2. (Adapted from Ref. [67].)

for a 30 μm-long waveguide is sketched in Fig. 1.9(b). Direct fluorescence is observed for both pump−probe (TE) and pump only configurations. However, for the pump−probe (TM) arrangement, the DC contribution due to the CW probe is seen and the additional signal in the presence of the pump corresponds to both the CW contribution and gain, giving a signal enhancement factor of $F = 3.2 \pm 0.1$. Using systematic measurements of the signal enhancement as a function of waveguide length, they measured a net gain of (3.4 ± 16) dB/mm at $\lambda_0 = 633$ nm using 200 μJ/cm^2 pump intensity as shown in Fig. 1.9(c).

Flynn *et al.*[68] demonstrated LRSPP lasing at $\lambda_0 \sim 1.46$ μm in a symmetric structure based on InP and consisting of a 15 nm-thick Au film placed between MQW heterostructure layers with a minimum lateral width of 100 μm. A 1 mm long Fabry-Perot structure formed a LRSPP laser cavity which emitted a TM-polarized light under optical pumping from the top at $\lambda_0 = 1.06\,\mu$m with 140 ns pulses. The emitted spectrum narrowed with increasing pump intensity, and the emitted power was linearly dependent on the pump intensity beyond a clear threshold.

1.4.2. *Dielectric-loaded SPP waveguide*

Dielectric-loaded SPP waveguides (DLSPPW) using 600 nm thick, 400 nm-wide PMMA stripes, doped with PbS quantum dots (QDs), and deposited on a 40 nm-thick Au film were investigated by Grandidier *et al.*[69] Optical gain was provided by pumping QDs in the waveguide from top at $\lambda_0 = 532$ nm. Stimulated emission of the SPP guided mode was qualitatively confirmed by linewidth narrowing of the SPP mode. The SPP propagation length was measured versus pump irradiance exhibiting a linear increase of propagation length above the pump threshold $\sim$500 W/cm^2. A maximum increase in propagation length of $\sim$27% was obtained. The gain coefficient of the amplified waveguide versus the pump irradiance was also measured near the QD peak emission ($\lambda_0 \sim 1550$ nm). They subsequently studied PMMA stripes doped with CdSe/ZnSe QDs deposited on

a 50 nm-thick Ag film. Their work was based on leakage radiation microscopy combined with surface plasmon coupled emission. They reported a similar increase in SPP propagation length at $\lambda_0 = 632.8$ nm.[70]

C. Garcia *et al.*[71] considered a DLSPPW at $\lambda_0 \sim 860$ nm comprising a 300 nm-thick, 300 nm-wide PMMA ridge doped with PbS QDs deposited on a 70 nm-thick Au film on a glass substrate. They similarly used leakage radiation microscopy to study stimulated emission of SPPs in the waveguide. About a 33% increase in the propagation length of SPPs was reported for a probe power of 16 mW and pump intensity of $\sim$4460 W/cm^2 which produced an optical gain of $\sim$143 cm^{-1}.

Colas des Francs *et al.*[72] modeled in general the local density of guided modes for two-dimensional SPP waveguides incorporating a gain medium. They assumed a 4-level dipolar gain medium to describe spontaneous and stimulated emission rates into SPPs. They specifically applied their model to a DLSPPW reported earlier[69] and deduced the stimulated emission cross-section of PbS QDs.

Rao and Tang[73] derived conditions for lossless SPP propagation in a single metal dielectric structure assuming stimulated emission in the active dielectric. They considered a wide range of wavelengths in the visible and infrared regions, and various metals. In particular they studied lossless propagation at extremely short wavelengths near the energy asymptote. Using their model, they designed a nanoscale DLSPPW with a dielectric cross section of 22×20 nm^2 and computed a required optical gain of $\sim 5 \times 10^4$ cm^{-1} in the dielectric to produce lossless sub-wavelength propagation at $\lambda_0 = 450$ nm.

Hahn *et al.*[74] proposed a dielectric-loaded LRSPP waveguide in a symmetric PMMA-Ag-PMMA structure, where a top PMMA ridge is used to laterally confine a LRSPP mode. Both the top and bottom PMMA layers were assumed to be doped with IR-140 dye molecules and the waveguide is pumped optically from the top. They introduced single-mode DFB lasers by stepping the width of the top PMMA while optical gain is provided[75] in both PMMA regions.

1.4.3. *Hybrid SPP waveguide*

Hybrid SPPs are a coupled TM-polarized wave, guided in a nanoscale low-index dielectric gap between a high-index dielectric and a metal surface. It is a super-mode excited by coupling the SPP mode on a metal-gap interface to the photonic mode of high-index dielectric.[76–78] Loss compensation or lasing with hybrid SSPs is possible if the high-index region provides optical gain over which the hybrid SPP overlaps sufficiently.

Oulton *et al.*[79] reported an experimental demonstration of a nanoscale laser in hybrid plasmonic structures consisting of a high-gain CdS nanowire separated from a Ag substrate by a 5 nm-thick MgF_2 gap. Multi-mode lasing at $\lambda_0 \sim 489$ nm was observed from the end facets of the nanowire under a $\sim$76.25 MW/cm^2 pump irradiance. Other designs for lasing using CdS nano-squares[80] and an embedded waveguide (WEM)[81] were also reported in hybrid structures.

A hybrid SPP laser consisting of an InGaN/GaN core-shell nanorod separated from an Ag film by a thin SiO_2 layer was demonstrated by Lu *et al.*[82] The nanorod as well as the Ag film were epitaxially grown to provide the atomic-scale smoothness necessary for reducing diffraction loss and providing low threshold CW lasing. Figure 1.10(a) sketches the hybrid SPP structure showing the field concentration in the low-index region but also significant overlap with the nanorod in order to provide gain. Figure 1.10(b) shows lasing emission spectra with increasing pump irradiance at 78 K. Bimodal lasing at 510 and 522 nm was obtained with a CW pump intensity of $\sim$26 KW/cm^2 at 78 K as seen in Fig. 1.10(b).

Sidiropoulos *et al.*[83] reported ultra-fast (shorter than 800 fs) hybrid SPP lasing using ZnO nanowires separated from a Ag substrate by a thin LiF gap. ZnO excitons at room temperature lie near the SPP frequency in the hybrid SPP laser, and strong mode confinement leads to accelerated spontaneous emission recombination. The dynamics of the ZnO nanowire SPP laser were compared to those of conventional ZnO nanowire photonic lasers revealing a sub-picosecond pulse width for the SPP nanolaser

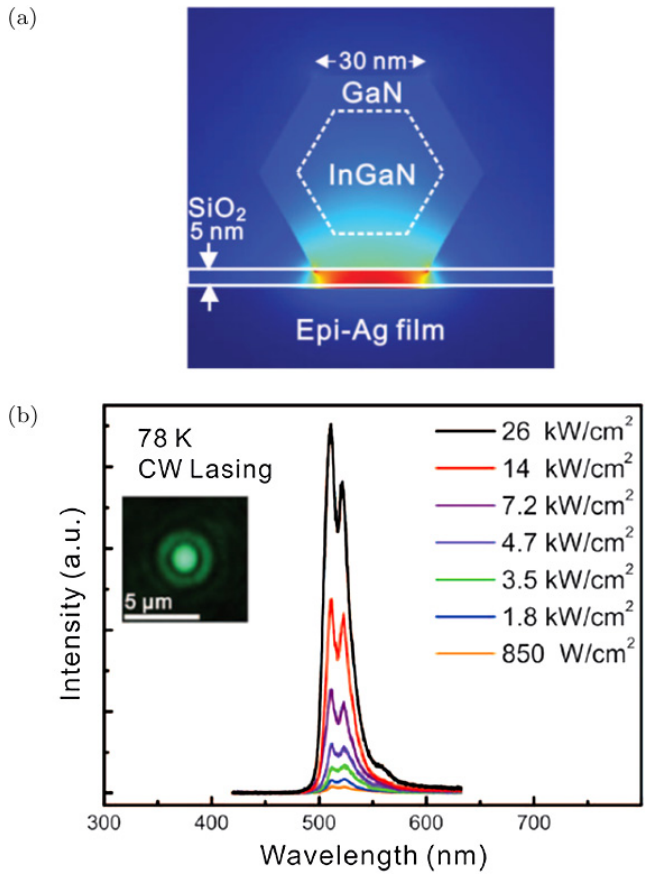


Fig. 1.10. (a) Sketch of a hybrid SPP nanolaser consisting of an InGaN/GaN core-shell gain nanorod on a thin SiO_2 layer on an epitaxial Ag film. (b) Lasing spectra at 78 K with increasing pump irradiance. Inset: Far-field laser spot with contrast fringes indicative of spatial coherence resulting from lasing. (Adapted from Ref. [82].)

but a much broader pulse width for the corresponding photonic nanolaser.

A low loss hybrid SPP laser design using MQWs in a P-N heterojunction and electric pumping was reported by Li and Huang.[84] Their design optimized the hybrid SPP mode's overlap with the P-N active region.

1.5. Amplification and Lasing in Nano-structures

Conduction electrons in metal nanoparticles irradiated by linearly polarized light form dipolar oscillations referred to as a surface plasmon resonance (SPR).[85,86] Resonant surface plasmons are supported by metal nanoparticles smaller than the skin depth (∼25 nm) so that optical fields can penetrate throughout their volume and drive electron oscillations. SPRs of high quality in proximity to a gain medium can lead to surface plasmon amplification by stimulated emission of radiation (spaser).[12,25]

Figure 1.11 sketches the spasing mechanism as explained by Stockman[25] in a spaser composed of a silver nanoshell covering a dielectric core of 10–20 nm radius, and surrounded by two dense monolayers of nanocrystal quantum dots (NQDs). Pumping the gain medium (optical or electric) excites electron–hole (e–h) pairs (solid arrow). The e–h pairs relax to excitonic levels (dotted arrow). The excitons recombine and their energy is transferred to the plasmon excitation of the metal nanoshell through resonant coupled transitions (dashed-dotted arrows).

Noginov *et al.*[87] demonstrated a spaser using spherical nanoparticles consisting of a 14 nm Au core, coated with a 15 nm SiO_2 shell and doped with Oregon Green 488 dye. The quality factor of nanoparticles was found as ∼13.2. Stimulated emission of SPPs at $\lambda \sim 531$ nm by individual nanoparticles was reported. Meng *et al.*[88] considered a monolayer of randomly oriented Au nanorods embedded in the mesopores of a silica shell dispersed on a glass substrate and covered with a 200 nm thick R6G-doped PVA film to supply gain. The spaser emission was tunable from 562 to 627 nm with a spectral line width narrowed down to 5–11 nm.

In a theoretical study by Khurgin and Sun,[89] they showed that lasing in electrically-pumped semiconductor spasers has an extremely high current threshold of ∼105 A/cm^2. They attributed the high current threshold to gain depletion caused by enhanced spontaneous emission. Dorfman *et al.*[90] proposed a low threshold and efficient spaser using quantum coherence in a three-level quantum emitter based gain medium coupled to plasmonic nanostructures. Ginzberg

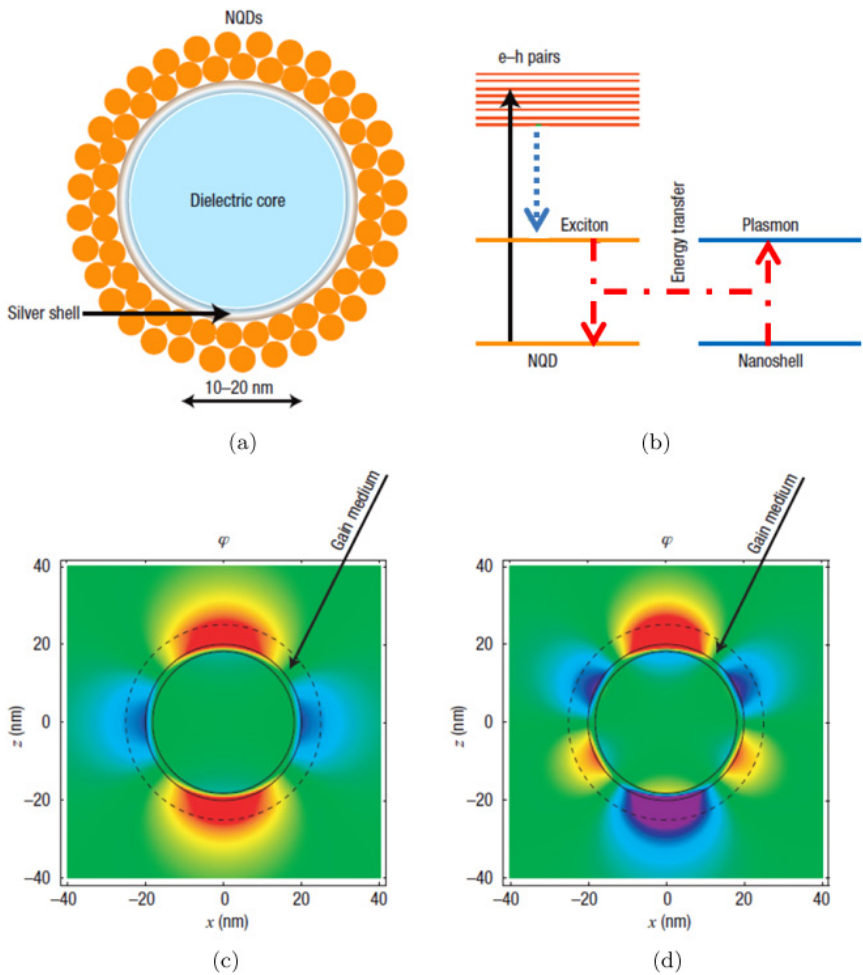


Fig. 1.11. (a) Schematic of a spaser using a silver nanoshell on a dielectric core, surrounded by two dense monolayers of NQDs. (b) Energy levels and transitions in a spaser. External radiation excites e–h pairs (vertical solid arrow). The e–h pairs relax to excitonic levels (dotted arrow). The exciton recombines and its energy is transferred to the plasmon excitation of the metal nanoshell through resonant coupled transitions (dashed-dotted arrows). (c), (d) Field amplitudes, φ, around the nanoshell in two different resonant plasmon modes. (Adapted from Ref. [25].)

and Zayats[91] developed a theory to obtain the linewidth and its enhancement in semiconductor spasers. They specifically considered bulk GaAs as the active medium of a spaser and estimated a linewidth enhancement of ∼3–6 which is about an order of magnitude larger than that predicted by the Schawlow-Townes theory. Single molecule detection based on surface enhanced Raman scattering (SERS) is possible using spacers as they can produce near-singular scattering hot spots.[92–94]

Array of nanoparticles for improved spaser quality has also been investigated. Suh *et al.*[95] reported coherent and directional emission from optically pumped arrays of Au bowtie resonators on slabs of IR-140 doped polyurethane on a Si substrate and covered with a dielectric over-layer. Zhou *et al.*[96] reports a two-dimensional plasmonic lattice consisting of a pattern of Au nanoparticles on a glass substrate, covered by a polyurethane layer doped with IR-140. Their structure features high directionality due to lattice bandedge lasing and ultrafast response due to plasmonic-enhanced localization. Loss compensation in extraordinary optical transmission (EOT) systems was theoretically studied by Marani *et al.*[97] They specifically analyzed the spectral response of an optically pumped periodic array of subwavelength slits in an opaque Au film sandwiched between two thin layers of Rh800-doped polymer. They showed that this structure can behave as an absorber, laser or amplifier due to strong spatial hole-burning effects, depending on the pump intensity. van Beijnum *et al.*[98] demonstrated laser emission in the near infrared from an array of subwavelength holes in a Au film on an optically pumped InGaAs/InP gain medium at 150 K. Linewidth narrowing and lasing threshold were observed with increasing pump power, and surface plasmon lasing was justified by experimental observations. Highly directional single-mode spasing at red wavelengths ($\lambda \sim 620$ nm) was demonstrated at room temperature by Meng *et al.*,[99] where the plasmonic nanocavity was formed by a periodic subwavelength hole array perforated in a Ag film supporting SPP Bloch waves at $\lambda \sim 620$ nm. They suggested that the optical feedback for spasing was provided

by SPP Bloch waves in the hole array nanocavity rather than the plasmonic resonance in individual holes.

1.6. Applications of SPP Amplification

1.6.1. *Amplifier*

The performance of passive SPP elements such as waveguides, couplers, splitters and Mach-Zehnder interferometers can be enhanced if their loss is compensated by SPP amplifiers. SPP amplifiers can be realized as a stand-alone component or as a gain section integrated with other SPP devices to improve their performance. Ng *et al.*[100] proposed gain assisted SPR biosensors based on the Kretschmann configuration where optical pumping leads to photon emission in a polymer which couples to surface plasmons via resonant transitions.

1.6.2. *Laser/Spaser*

SPP lasers can be constructed similar to conventional lasers, such that SPPs are directly excited at the interface of a metal with the gain medium within a cavity which allows SPPs to resonate. SPP lasers are coherent sources of confined light and can be used to produce a high-intensity, ultrafast source of light.

Tunable single-mode lasers were designed by Keshmarzi *et al.*[101] using LRSPPs on a metal stripe waveguide in the distributed feedback (DFB) and distributed Bragg reflector (DBR) configurations. Laser structures comprising step-in-width Bragg gratings in a 20 nm-thick, 1 μm-wide Ag stripe embedded in PMMA were considered. The proposed structures are thermo-optically tunable via current injected along the stripe.

1.6.3. *PT symmetric material*

PT symmetry is a concept originating from quantum mechanics which states that a non-Hermitian Hamiltonian can still have real eigenvalues provided that it respects the PT symmetry condition.[102]

It is shown that a necessary but insufficient condition for PT symmetry is that the complex potential $V(\mathbf{r})$ obeys $V(\mathbf{r}) = V^*(-\mathbf{r})$, with * denoting the complex conjugate.[103] The refractive index in photonics plays the role of the potential function in quantum mechanics[103] so that optical PT symmetric materials can be created if the refractive index n as a function of position $\mathbf{r}$ satisfies $n(\mathbf{r}) = n^*(-\mathbf{r})$, which implies a complex refractive index consisting of balanced gain and loss. These materials have attracted significant attention through the unique and extraordinary properties that they exhibit at a so-called PT spontaneous breaking point. (Reference 104 provides a review of some PT symmetry applications in optics). Indeed, some optical functionality that is not possible with natural materials can be realized by PT symmetric artificial materials. For instance, unidirectional propagation in a linear medium is only possible if a bulk magnetic device is incorporated. However, a PT symmetric medium can produce unidirectional propagation in linear media without any Faraday effect.[105] Plasmonics provides a versatile platform to realize PT symmetric materials, as loss inherently exists in SPPs and can be tuned by geometrical adjustments.[106] Gain can be incorporated and balanced with loss in SPP structures to produce a PT symmetric material. So far only a few groups have investigated PT symmetry in plasmonic systems. Benisty *et al.*[107] proposed a PT symmetric hybrid plasmonic-dielectric co-directional coupler where gain and loss were assumed in the dielectric and plasmonic waveguides respectively. They used the singular behavior of the PT structure near the spontaneous breaking point and showed that a small gain/loss modulation results in a large change in propagation properties of the supermodes. Later they proposed optical switching[108] based on this idea and emphasized on the positive role of loss provided by plasmonic waveguides. Keshmarzi *et al.*[106] proposed a PT symmetric Bragg grating based on an LRSPP waveguide using gain and loss modulations. The proposed Bragg grating was designed at the PT symmetry breaking threshold providing highly asymmetric reflectance from the end facets. Hahn *et al.*[109] designed a PT symmetric Bragg grating similarly based on an LRSPP waveguide structure. They

considered the complex modulation of effective refractive index by modulating the width of a top PMMA cladding about a certain width where the LRSPP mode couples to an antisymmetric SPP mode. The PMMA width within the grating's unit cell and the PMMA thickness were judiciously arranged to produce PT symmetry at the breaking threshold. Again, asymmetric reflectance was observed when the Bragg was probed from the left and right facets.

1.7. Concluding Remarks

SPPs are excited along the interface of metals with dielectrics at infrared or visible wavelengths. They can be concentrated to subwavelength scales beyond the light diffraction limit and offer potentially useful applications. Metallic loss in SPP structures reduce the SPP lifetime and limit their applications. Optical gain can be integrated with SPP structures to reduce the loss and maintain confinement. The gain requirement significantly varies depending on the SPP structure and mode of interest. Lossless SPP propagation using good metals (*e.g.*, Ag, Au) approximately requires a small signal material gain of 1–200 cm^{-1} for the LRSPP mode along thin metal stripes or films, 1000–2000 cm^{-1} for the single-interface SPP mode, 2000–5000 cm^{-1} for SPPs along thin metal-cladded dielectric waveguides, and 80 000 cm^{-1} for SPPs near the energy asymptote. Various gain media can be used to produce these gains (except perhaps for the latter) including good dyes, QDs and semiconductors. Dielectrics, particularly polymers, can be doped by optical dipoles such as dyes and integrated with metals. Optical pumping is usually applied to excite such materials. Epitaxial semiconductors can be pumped electrically, but their integration with metallic structures, especially buried ones, can be difficult and their gain may be polarization dependent so that pump alignment with SPP is necessary.

Research on SPP amplification and lasing to date has been mostly on the fundamentals and only primary demonstrations have been made. Many avenues are yet open to explore, especially the development of applications with usability and good performance in

mind. Future works include the development and improvement of efficient pumping techniques, room temperature operation, high quality SP resonators, low power dissipation and stability. Despite existing challenges, the prospects of SPP amplification and active plasmonics are promising. Substantial demonstrations have already been reported for various amplified SPP structures, including LRSPPs on stripes,[66–68] and slabs,[51] single interface SPPs,[36] SPPs in hybrid[80] and metal-cladded waveguides[62] through which useful applications can be developed.

References

1. Maier, S. A., *Plasmonics: Fundamentals and Applications*. Springer-Verlag, Berlin, (2007).
2. Ebbesen, T. W., Genet, C., and Bozhevolnyi, S. I., Surface plasmon circuitry. *Phys. Today* **61**, 44–50, (2008).
3. Homola, J., Surface plasmon resonance sensors for detection of chemical and biological species, *Chem. Rev.* **108**, 462–493, (2008).
4. Xu, F., Chen, G., Wang, C., Cao, B., and Lou, Y., Superlens imaging with a surface plasmon polariton cavity in imaging space, *Opt. Lett.* **138**, 3819–3822, (2013).
5. Willets, K. A. and Van Duyne, R. P., Localized surface plasmon resonance spectroscopy and sensing, *Phys. Chem.* **58**, 267–297, (2007).
6. Luo, X. and Ishihara, T., Subwavelength photolithography based on surface-plasmon polariton resonance, *Opt. Exp.* **12**, 3055–3065, (2004).
7. Luo, X. and Yan, L., Surface plasmon polaritons and its applications, *IEEE Phot. J.* **4**, 590–595, (2012).
8. Barnes, W. L., Dereux, A., and Ebbesen, T. W., Surface plasmon subwavelength optics, *Nature* **424**, 824–830, (2003).
9. Zijlstra, P., Paulo, P. M. R., and Orrit, M., Optical detection of single non-absorbing molecules using the surface plasmon resonance of a gold nanorod, *Nat. Nanotech.* **7**, 379–382, (2012).
10. Berini, P. and De Leon, I., Surface plasmon-polariton amplifiers and lasers, *Nat. Photonics* **6**, 16–24, (2012).
11. Keshmarzi, E. K., Tait, R. N., and Berini, P., Long-range surface plasmon single-mode laser concepts, *J. Appl. Phys.* **112**, 063115 (1–5), (2012).
12. Bergman, D. J. and Stockman, M. I., Surface plasmon amplification by stimulated emission of radiation: quantum generation of coherent surface plasmons in nanosystems, *Phys. Rev. Lett.* **90**, 027402 (1–4), (2003).
13. Rüter, C. E., Makris, K. G., El-Ganainy, R., Christodoulides, D. N., Segev, M., and Kip, D., Observation of parity–time symmetry in optics, *Nat. Phys.* **6**, 192–195, (2010).

14. Keshmarzi, E. K., Tait, R. N., and Berini, P., Spatially nonreciprocal Bragg gratings based on surface plasmons, *Appl. Phys. Lett.* **105**, 191110 (1–4), (2014).
15. Gramotnev, D. K., and Bozhevolnyi, S. I., Plasmonics beyond the diffraction limit. *Nat. Photonics* **4**, 83–91, (2010).
16. Berini, P., Figures of merit for surface plasmon waveguides, *Opt. Express* **14**, 13030–13042, (2006).
17. Berini, P., Long-range surface plasmon polaritons, *Adv. Opt. Photon.* **1**, 484–588, (2009).
18. Berini, P., Plasmon-polariton waves guided by thin lossy metal films of finite width: bound modes of symmetric structures, *Phys. Rev. B* **61**, 10484–10503, (2000).
19. Hohenau, A., Krenn, J. R., Stepanov, A. L., Drezet, A., Ditlbacher, H., B. Steinberger, Leitner, A., and Aussenegg, F. R., Dielectric optical elements for surface plasmons, *Opt. Lett.* **30**, 893–895, (2005).
20. Pile, D. F. P., Ogawa, T., Gramotnev, D. K., Matsuzaki, Y., Vernon, K. C., Yamaguchi, K., Okamoto, T., Haraguchi, M., and Fukui, M., Two-dimensionally localized modes of a nanoscale gap plasmon waveguide, *Appl. Phys. Lett.* **87**, 261114 (1–3), (2005).
21. Alam, M. Z., Meier, J., Aitchison, J. S., and Mojahedi, M., Super mode propagation in low index medium, *Proc. CLEO paper* JThD112, (2007).
22. Oulton, R. F., Sorger, V. J., Genov, D. A., Pile, D. F. P., and Zhang, X. A., Hybrid plasmonic waveguide for subwavelength confinement and long-range propagation, *Nature Photon.* **2**, 496–500, (2008).
23. Kelly, K. L., Coronado, E., Zhao, L. L., and Schatz, G. C., The optical properties of metal nanoparticles: the influence of size, shape, and dielectric environment, *J. Phys. Chem. B* **107**, 668–677, (2003).
24. Pelton, M., Aizpurua, J., and Bryant, G., Metal-nanoparticle plasmonics, *Laser Phot. Rev.* **2**, 136–159, (2008).
25. Stockman, M. I., Spasers explained, *Nat. Photonics* **2**, 327–329, (2008).
26. Nkoma, J., Loudon, R., and Tilley, D. R., Elementary properties of surface plasmons, *J. Phys. C* **7**, 3547–3559, (1974).
27. Archambault, A., Marquier, F., and Greffet, J., Quantum theory of spontaneous and stimulated emission of surface plasmons, *Phys. Rev. B* **82**, 035411 (1–10), (2010).
28. Matloob, R., Loudon, R., Barnett, S. M., and Jeffers, J., Electromagnetic field quantization in absorbing dielectrics, *Phys. Rev. A* **52**, 4823–4838, (1995).
29. Barnes, W. L., Fluorescence near interfaces: the role of photonic mode density, *J. Mod. Opt.* **45**, 661–669, (1998).
30. Plotz, G., Simon, H., and Tucciarone, J., Enhanced total reflection with surface plasmons, *J. Opt. Soc. Am.* **69**, 419–422, (1979).
31. Sudarkin, A. N. and Demkovich, P. A., Excitation of surface electromagnetic waves on the boundary of a metal with an amplifying medium, *Sov. Phys. Tech. Phys.* **34**, 764–766, (1988).

32. Nezhad, M. P., Tetz, K., and Fainman, Y., Gain assisted propagation of surface plasmon polaritons on planar metallic waveguides, *Opt. Express* **12**, 4072–4079, (2004).
33. Kumar, P., Tripathi, V. K., and Liu, C. S., A surface plasmon laser, *J. Appl. Phys.* **104**, 033306 (1–4), (2008).
34. Lu, F. F., Li, T., Xu, J., Xie, Z. D., Li, L., Zhu, S. N., and Zhu, Y. Y., Surface plasmon polariton enhanced by optical parametric amplification in nonlinear hybrid waveguide, *Opt. Express* **19**, 2858–2865, (2011).
35. Sirtori, C., Gmachl, C., Capasso, F., Faist, J., Sivco, D. L., Hutchinson, A. L., and A. Cho, Y., Long-wavelength ($\lambda \sim 11.5$ μm) semiconductor lasers with waveguides based on surface plasmons, *Opt. Lett.* **23**, 1366–1368, (1998).
36. Tredicucci, A., Gmachl, C., Capasso, F., Hutchinson, A. L., Sivco, D. L., and A. Y. Cho, Single–mode surface-plasmon laser, *Appl. Phys. Lett.* **76**, 2164–2166, (2000).
37. Seidel, J., Grafstrom, S., and Eng, L., Stimulated emission of surface plasmons at the interface between a silver film and an optically pumped dye solution. *Phys. Rev. Lett.* **94**, 177401 (1–4), (2005).
38. Noginov, M. A., Podolskiy, V. A., Zhu, G., Mayy, M., Bahoura, M., Adegoke, J. A., B. A. Ritzo, and Reynolds, K., Compensation of loss in propagating surface plasmon polariton by gain in adjacent dielectric medium, *Opt. Express* **16**, 1385–1392, (2008).
39. Noginov, M. A., Zhu, G., Mayy, M., Ritzo, B. A., Noginova, N., and Podolskiy, V. A., Stimulated emission of surface plasmon polaritons, *Phys. Rev. Lett.* **101**, 226806 (1–4), (2008).
40. Bolger, P. M., Dickson, W., Krasavin, A. V., Liebscher, L., Hickey, S. G., Skryabin, D. V., and Zayats, A. V., Amplified spontaneous emission of surface plasmon polaritons and limitations on the increase of their propagation length, *Opt. Lett.* **35**, 1197–1199, (2010).
41. Avrutsky, I., Surface plasmons at nanoscale relief gratings between a metal and a dielectric medium with optical gain, *Phys. Rev. B* **70**, 155416 (1–6), (2004).
42. Okamoto, T., H'Dhili, F., and Kawata, S., Towards plasmonic band gap laser, *Appl. Phys. Lett.* **85**, 3968–3970, (2004).
43. Okamoto, T., Simonen, J., and Kawata, S., Plasmonic band gaps of structured metallic thin films evaluated for a surface plasmon laser using the coupled-wave approach, *Phys. Rev. B* **77**, 115425 (1–8), (2008).
44. H'Dhili, F., Okamoto, T., Simonen, J., and Kawata, S., Improving the emission efficiency of periodic plasmonic structures for lasing applications, *Opt. Comm.* **284**, 561–566, (2011).
45. Winter, G., Wedge, S., and Barnes, W. L., Can lasing at visible wavelength be achieved using the low-loss long-range surface plasmon-polariton mode?, *New J. Phys.* **8**, 125–138, (2006).
46. Genov, D. A., Ambati, M., and Zhang, X., Surface plasmon amplification in planar metal films. *IEEE J. Quantum Electron.* **43**, 1104–1108, (2007).

47. De Leon, I. and Berini, P., Theory of surface plasmon-polariton amplification in planar structures incorporating dipolar gain media, *Phys. Rev. B* **78**, 161401(R) (1–4), (2008).
48. De Leon, I. and Berini, P., Modeling surface plasmon-polariton gain in planar metallic structures, *Opt. Express* **17**, 20191–20202, (2009).
49. De Leon, I. and Berini, P., Theory of noise in high-gain surface plasmon-polariton amplifiers incorporating dipolar gain media, *Opt. Express* **19**, 20506–20517, (2011).
50. De Leon, I. and Berini, P., Spontaneous emission in long-range surface plasmon-polariton amplifiers, *Phys. Rev. B* **83**, 081414(R) (1–4), (2011).
51. Gather, M. C., Meerholz, K., Danz, N., and Leosson, K., Net optical gain in a plasmonic waveguide embedded in a fluorescent polymer, *Nat. Photonics* **4**, 457–461, (2010).
52. Hahn, C., Song, S. H., Oh, C. H., and Berini, P., Plasmonic gain in long-range surface plasmon polariton waveguides bounded symmetrically by dye-doped polymer, *App. Phys. Lett.* **107**, 121107 (1–4), (2015).
53. Radko, I. P., Nielsen, M. G., Albrektsen, O., and Bozhevolnyi, S. I., Stimulated emission of surface plasmon polaritons by lead-sulfide quantum dots at near infra-red wavelengths, *Opt. Express* **18**, 18633–18641, (2010).
54. Maier, S. A., Gain-assisted propagation of electromagnetic energy in subwavelength surface plasmon polariton gap waveguides, *Opt. Comm.* **258**, 295–299, (2006).
55. Chang, S. W. and Chuang, S. L., Normal modes for plasmonic nanolasers with dispersive and inhomogeneous media, *Opt. Lett.* **34**, 91–93, (2009).
56. Chang, S. W. and Chuang, S. L., Fundamental formulation for plasmonic nanolasers, *IEEE J. Quantum Electron.* **45**, 1014–1023, (2009).
57. Li, D. B. and Ning, C. Z., Giant modal gain, amplified surface plasmon-polariton propagation, and slowing down of energy velocity in a metal-semiconductor-metal structure, *Phys. Rev. B* **80**, 153304 (1–4), (2009).
58. Li, D. B. and Ning, C. Z., Interplay of various loss mechanisms and ultimate size limit of a surface plasmon polariton semiconductor nanolaser, *Opt. Express* **20**, 16348–16357, (2012).
59. Khurgin, J. B. and Sun, G., Practicality of compensating the loss in the plasmonic waveguide using semiconductor gain medium, *Appl. Phys. Lett.* **100**, 011105 (1–3), (2012).
60. Chen, X., Bhola, B., Huang, Y., and Ho, S. T., Multi-level multi-thermal-electron FDTD simulation of plasmonic interaction with semiconducting gain media: applications to plasmonic amplifiers and nano-lasers. *Opt. Express* **18**, 17220–17238, (2010).
61. Yu, Z., Veronis, G., Fan, S., and Bongersma, M. L., Gain-induced switching in metal-dielectric-metal plasmonic waveguides, *Appl. Phys. Lett.* **92**, 041117 (1–3), (2008).
62. Hill, M. T., Marell, M., Leong, E. S. P., Smalbrugge, B., Zhu, Y., Sun, M., P. J. vanVeldhoven, Geluk, E. J., Karouta, F., Oei, Y. S., R. Nötzel, Ning,

C. Z., and Smit, M. K., Lasing in metal-insulator-metal sub-wavelength plasmonic waveguides. *Opt. Express* **17**, 11107–11112, (2009).

63. Alam, M. Z., Meier, J., Aitchison, J. S., and Mojahedi, M., Gain assisted surface plasmon polariton in quantum well structures. *Opt. Express* **15**, 176–182, (2007).
64. Chen, Y. H. and Guo, L. J., High Q long-range surface plasmon polariton modes in sub-wavelength metallic microdisk cavity, *Plasmonics* **6**, 183–188, (2011).
65. Ambati, M., Genov, D. A., Oulton, R. F., and Zhang, X., Active plasmonics: surface plasmon interaction with optical emitters. *IEEE J. Sel. Top. Quan. Elec.* **14**, 1395–1403, (2008).
66. De Leon, I. and Berini, P., Amplification of long-range surface plasmons by a dipolar gain medium, *Nat. Photonics* **4**, 382–387, (2010).
67. Kéna-Cohen, S., Stavrinou, P. N., Bradley, D. D. C., and Maier, S. A., Confined surface plasmon polariton amplifers, *Nano Lett.* **13**, 1323–1329, (2013).
68. Flynn, R. A., Kim, C. S., Vurgaftman, I., Kim, M., Meyer, J. R., Mäkinen, A. J., Bussmann, K., Cheng, L., Choa, F. S., and Long, J. P., A room-temperature semiconductor spaser operating near 1.5 μm, *Opt. Express* **19**, 8954–8961, (2011).
69. Grandidier, J., Colas, des Francs, G., Massenot, S., Bouhelier, A., Markey, L., J. C. Weeber, Finot, C., and Dereux, A., Gain-assisted propagation in a plasmonic waveguide at telecom wavelength, *Nano Lett.* **9**, 2935–2939, (2009).
70. Grandidier, J., Colas, des Francs, G., Massenot, S., Bouhelier, A., Markey, L., Weeber, J. C., Finot, C., and Dereux, A., Leakage radiation microscopy of surface plasmon coupled emission: investigation of gain-assisted propagation in an integrated plasmonic waveguide. *J. Microsc.* **239**, 167–172, (2010).
71. Garcia, C., Coello, V., Han, Z., Radko, I. P., and Bozhevolnyi, S. I., Partial loss compensation in dielectric-loaded plasmonic waveguides at near-infrared wavelengths. *Opt. Express* **20**, 7771–7775, (2012).
72. Colas Des Francs, G., Bramant, P., Grandidier, J., Bouhelier, A., Weeber, J. C., and Dereux, A., Optical gain, spontaneous and stimulated emission of surface plasmon polaritons in confined plasmonic waveguide, *Opt. Express* **18**, 16327–16334, (2010).
73. Rao, R. and Tang, T., Study on active surface plasmon waveguides and design of a nanoscale lossless surface plasmon waveguide, *J. Opt. Soc. Am. B* **28**, 1258–1265, (2011).
74. Hahn, C., Song, S. H., Oh, C. H., and Berini, P., Single-mode lasers and parity-time symmetry broken gratings based on active dielectric-loaded long-range surface plasmon polariton waveguides, *Opt. Express* **23**, 19922–19931, (2015).

75. Keshmarzi, E. K., Tait, R. N., and Berini, P., Near infrared amplified spontaneous emission in a dye-doped polymeric waveguide for active plasmonic applications, *Opt. Express* **22**, 12452–12460, (2014).
76. Alam, M. Z., Meier, J., Aitchison, J. S., and Mojahedi, M., Super mode propagation in low index medium, *Proc. CLEO*, JThD112, (2007).
77. Oulton, R. F., Sorger, V. J., Genov, D. A., Pile, D. F. P., and Zhang, X., A hybrid plasmonic waveguide for subwavelength confinement and long-range propagation. *Nat. Photonics* **2**, 496–500, (2008).
78. Oulton, R. F., Bartal, G., Pile, D. F. P., and Zhang, X., Confinement and propagation characteristics of subwavelength plasmonic modes, *New J. of Phys.* **10**, 105018 (1–14), (2008).
79. Oulton, R. F., Sorger, V. J., T. Zentgraf, Ma, R. M., Gladden, C., Dai, L., Bartal, G., and Zhang, X., Plasmon lasers at deep subwavelength scale, *Nature* **461**, 629–632, (2009).
80. Ma, R. M., Oulton, R. F., Sorger, V. J., Bartal, G., and Zhang, X., Room-temperature subdiffraction-limited plasmon laser by total internal reflection, *Nat. Mater.* **10**, 110–113, (2011).
81. Ma, R. M., Yin, X., Oulton, R. F., Sorger, V. J., and Zhang, X., Multiplexed and electrically modulated plasmon laser circuit. *Nano Lett.* **12**, 5396–5402, (2012).
82. Lu, Y. J., Kim, J., Chen, H. Y., Wu, C., Dabidian, N., Sanders, C. E., Wang, C. Y., M. Y. Lu, Li, B. H., Qiu, X., Chang, W. H., Chen, L. J., Shvets, G., Shih, C. K., and Gwo, S., Plasmonic nanolaser using epitaxially grown silver film. *Science* **337**, 450–453, (2012).
83. Sidiropoulos, T. P. H., Röder, R., Geburt, S., Hess, O., Maier, S. A., Ronning, C., and Oulton, R. F., Ultrafast plasmonic nanowire lasers near the surface plasmon frequency, *Nature Physics* **10**, 870–876, (2014).
84. Li, Y. and Huang, W. P., Electrically-pumped plasmonic lasers based on low-loss hybrid SPP waveguide, *Opt. Express* **23**, 24843–24849, (2015).
85. Kelly, K. L., Coronado, E., Zhao, L. L., and Schatz, G. C., The optical properties of metal nanoparticles: the influence of size, shape, and dielectric environment. *J. Phys. Chem. B* **107**, 668–677, (2003).
86. Pelton, M., Aizpurua, J., and Bryant, G., Metal-nanoparticle plasmonics. *Laser Phot. Rev.* **2**, 136–159, (2008).
87. Noginov, M. A., Zhu, G., Belgrave, A. M., Bakker, R., Shalaev, V. M., Narimanov, E. E., Stout, S., Herz, E., Suteewong, T., and Wiesner, U., Demonstration of a spaser-based nanolaser. *Nature* **460**, 1110–1113, (2009).
88. Meng, X., Kildishev, A. V., Fujita, K., Tanaka, K., and Shalaev, V. M., Wavelength-tunable spasing in the visible, *Nano Lett.* **13**, 4106–4112, (2013).
89. Khurgin, J. B. and Sun, G., How small can "Nano" be in a "Nanolaser?" *Nanophotonics* **1**, 3–8, (2012).
90. Dorfman, K. E., Jha, P. K., Voronine, D. V., Genevet, P., Capasso, F., and M. O. Scully, Quantumcoherence-enhanced surface plasmon amplification

by stimulated emission of radiation. *Phys. Rev. Lett.* **111**, 043601 (1–5), (2013).
91. Ginzburg, P. and Zayats, A. V., Linewidth enhancement in spasers and plasmonic nanolasers. *Opt. Express* **21**, 2147–2153, (2013).
92. Lawandy, N. M., Localized surface plasmon singularities in amplifying media. *Appl. Phys. Lett.* **85**, 5040–5042, (2004).
93. Gordon, J. A. and Ziolkowski, R. W., The design and simulated performance of a coated nano–particle laser, *Opt. Express* **15**, 2622–2653, (2007).
94. Li, Z. Y. and Xia, Y., Metal nanoparticles with gain toward single-molecule detection by surface-enhanced Raman scattering, *Nano Lett.* **10**, 243–249, (2010).
95. Suh, J. Y., Kim, C. H., Zhou, W., Huntington, M. D., Co, D. T., Wasielewski, M. R., and Odom, T. W., Plasmonic bowtie nanolaser arrays. *Nano Lett.* **12**, 5769–5774, (2012).
96. Zhou, W., Dridi, M., Suh, J. Y., Kim, C. H., Co, D. T., Wasielewski, M. R., Schatz, G., and Odom, T. W., Lasing action in strongly coupled plasmonic nanocavity arrays. *Nat. Nanotechnol.* **8**, 506–511, (2013).
97. Marani, R., A. D'Orazio, Petruzzelli, V., Rodrigo, S. G., L. Martin-Moreno, F. J. Garcia-Vidal, and J. Bravo-Abad, Gain-assisted extraordinary optical transmission through periodic arrays of subwavelength apertures, *New J. Phys.* **14**, 013020 (1–16), (2012).
98. van Beijnum, F., van Veldhoven, P. J., Geluk, E. J., de Dood, M. J. A., 't Hooft, G. W., and van Exter, M. P., Surface plasmon lasing observed in metal hole arrays, *Phys. Rev. Lett.* **110**, 206802 (1–5), (2013).
99. Meng, X., Liu, J., Kildishev, A. V., and Shalaev, V. M., Highly directional spaser array for the red wavelength region, *Laser Photonics Rev.* **8**, 896–903 (2014).
100. Ng, S. P., Yip, Y. Y., and Wu, C. M. L., Biosensing with gain-assisted surface plasmon-polariton amplifier: a computational investigation, *Sensors and Actuators B* **210**, 36–45, (2015).
101. Keshmarzi, E. K., Tait, R. N., and Berini, P., Long-range surface plasmon single-mode laser concepts, *J. Appl. Phys.* **112** (6), 063115, (2012).
102. Bender, C. M. and Boettcher, S., Real spectra in non-hermitian hamiltonians having PT symmetry, *Phys. Rev. Lett.* **80**, 5243–5246, (1998).
103. Rüter, C. E., Makris, K. G., R. El-Ganainy, Christodoulides, D. N., Segev, M., and Kip, D., Observation of parity–time symmetry in optics, *Nat. Phys.* **6**, 192–195, (2010).
104. Hahn, C., Keshmarzi, E. K., Song, S. H., Oh, C. H., Tait, R. N., and Berini, P., Unidirectional Bragg Gratings Using Parity-Time Symmetry Breaking in Plasmonic Systems, *IEEE J. Sel. Top. Quant. Electr., in press.*
105. Kulishov, M., Laniel, J., Bélanger, N., Azaña, J., and Plant, D., Nonreciprocal waveguide Bragg gratings, *Opt. Express* **13**, 3068–3078, (2005).
106. Keshmarzi, E. K., Tait, R. N., and Berini, P., Spatially nonreciprocal Bragg gratings based on surface plasmons, *Appl. Phys. Lett.* **105** (19), 191110, (2014).

107. Benisty, H., Degiron, A., Lupu, A., De Lustrac, A., Chénais, S., Forget, S., Besbes, M., Barbillon, G., Bruyant, A., Blaize, S., and Lérondel, G., Implementation of PT symmetric devices using plasmonics: principle and applications, *Opt. Express* **19**, 18004–18019, (2011).
108. Lupu, A., Benisty, H., and Degiron, A., Switching using PT symmetry in plasmonic systems: positive role of the losses, *Opt. Express* **21**, 21651–21668, (2013).
109. Hahn, C., Song, S. H., Oh, C. H., and Berini, P., Single-mode lasers and parity-time symmetry broken gratings based on active dielectric-loaded long-range surface plasmon polariton waveguides, *Opt. Express* **23**, 19922–19931, (2015).

CHAPTER 2

Stopped Light Nano-Lasing

A. FREDDIE PAGE and ORTWIN HESS

Imperial College London, UK
o.hess@imperial.ac.uk

2.1. Introduction

Lasers generally comprise two essential elements, gain and feedback.

In a gain material, photons (or indeed surface plasmon polaritons) are generated by spontaneous and stimulated emission. Stimulated emission can happen when a photon induces the relaxation of an electron from an energetically higher state to another state of lower energy, and, crucially, a second photon which is coherent with the first is emitted. This process can be repeated, and while the electron population is inverted (that is there are more electrons in the upper rather state, rather than in the lower state), the number of photons in a coherent state will grow exponentially. Such gain media that are available include, for example, bulk semiconductors,[11] quantum dots[19] and wells,[3] as well as organic laser dye molecules.[24]

Feedback, on the other hand, is the means by which the photons that are emitted are coupled back to interact with the same gain medium, such that they may stimulate further emission. Most lasers will use a resonant cavity for this purpose, where the cavity modes localise electromagnetic energy over a gain medium. In the field of nanolasing there have been many such examples including photonic

crystal defect modes,[1,17] microcavity resonators,[6,13,27] and even the multiple scattering of a "random laser".[2,28]

Stopped light (SL) offers an alternative mechanism for feedback than the feedback provided by a cavity. SL modes are only confined in one spatial dimension, and not in the other two. This permits a continuum of planewave solutions, i.e. SL modes are propagating waves, rather than standing waves, albeit at (or sufficiently close to) zero group velocity. Also, the lasing mode of a stopped light laser forms dynamically, based on the gain support for the continuum of modes instead of being predetermined by the geometry of a cavity. Here the feedback is provided by a balance of adjacent forward and backward power flows that form a closed-loop optical vortex on sub-wavelength scales which gives rise to the zero group velocity.

2.2. Ultraslow and Stopped Light

Storing electromagnetic energy at a single location in space over long timescales is the subject of the field of ultraslow or "stopped" light. Such a phenomenon would find applications in data storage, quantum information, and light harvesting.[8] In this regard, there are three properties of a wave-packet that must be reduced by a structure or material in order to have stopped light: The speed of the wave-packet itself, dispersion which causes the wave-packet to distort, and energy loss to the material. We should note here however, that strictly, there is not a single speed to be associated with the movement of a wave-packet because a wave-packet can be composed of plane waves which, in general, will each move at different speeds in a dispersive material. Indeed, there are thus a number of measures which seek to characterise the average speed of a wave-packet, such as the energy transport velocity or the centrovelocity.[23] Perhaps the most familiar measure is the group velocity, which is well defined when energy is transferred through a medium which is uniform in the propagation direction, and is the speed that the mean of a Gaussian wave-packet would travel through such a medium.

In this chapter, we shall consider light propagating through planar hybrid metal-dielectric waveguide structures, such as in Fig. 2.1.

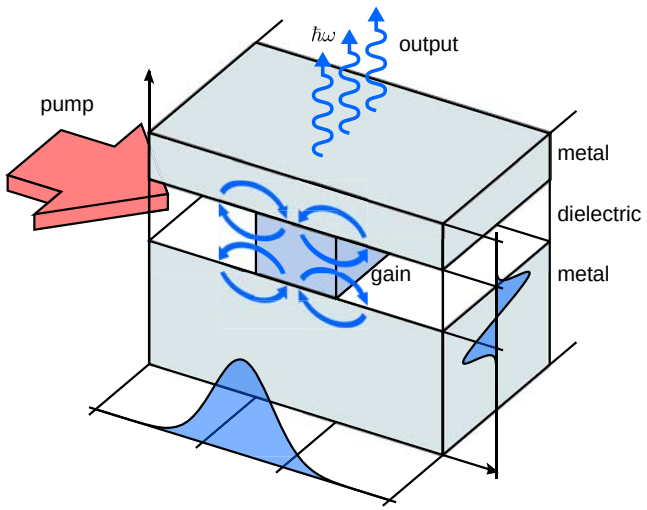


Fig. 2.1. Nanoplasmonic stopped-light laser. Light, guided in a hybrid metallo-dielectric waveguide, is localised in the vicinity of stopped-light singularities formed by the vertical metal-dielectric-metal heterostructure. In the overlap region with a gain medium (blue box) a lasing vortex (blue arrows) is formed, providing (in spite of an absence of a cavity) lateral feedback (i.e. in horizontal direction) on sub-wavelength scales. In this setup, the gain is pumped from the side and light is emitted as leaky wave towards the top. A field strength profile is also shown superimposed on the structure.

Particularly we shall look at the waveguide modes, or bound modes, as therein lie the greatest potential to alter the velocity, dispersion, and loss by changing the structural parameters. These waveguide modes can be solved for with a transfer matrix method.[31] Bound modes are such that there can be energy within the structure without it being incident from the outside.

There are two ways to theoretically describe the propagation of light within a waveguide. The first one is based on defining a driving field at a single point at all times and propagate the field through the waveguide. The second way implies that one knows the field configuration throughout the system at a single instant and calculates the time evolution. Both descriptions are in principle equivalent and describe the same field evolution. However it is the second approach that is the more natural picture when discussing ultraslow and stopped

light as one analyses light that effectively does not or only very slowly propagates. In the second picture there is a well defined wavepacket which evolves in time, and this picture is more suited to understanding translation and distortion than the first where indeed it is more difficult to set up the wavepacket initially by exciting from a single point in space. This choice has implications for the description of the system in the Fourier domain: A wavepacket that is well defined in space at an instant in time will have a well defined spatial Fourier transform in that instant, i.e. parametrised by a real valued in-plane wavevector β. This means however, that in dispersive media where the permittivity ε is complex, the frequency ω is also rendered complex. A complex frequency has two parts; the real part which determines the phase change of waves in time evolution, and an imaginary part which is the loss — the rate of exponential decay of energy in a material.

The dispersion relation $\omega(\beta)$ can be expanded about a central wavevector β_0 as,

$$\omega(\beta) = v_\mathrm{p}\beta_0 + v_\mathrm{g}(\beta - \beta_0) + v_\mathrm{d}\frac{\sigma_0}{2}(\beta - \beta_0)^2 + \ldots \tag{2.1}$$

Here three velocities v_p, v_g, and v_d have been introduced, which are the *phase*, *group*, and *dispersion* velocities. respectively. Their effects are most clearly explained when describing the time evolution of a Gaussian wavepacket with a spatial bandwidth σ_0 and carrier wavevector β_0. At time $t = 0$,

$$\varphi(x, t = 0) = \exp\left(-\frac{x^2}{2\sigma_0^2} + i\beta_0 x\right). \tag{2.2}$$

Then the evolution of such a wavepacket under the dispersion relation $\omega(\beta)$ of Eq. 2.1 becomes,

$$\varphi(x,t) \propto \exp\left(-\frac{(x - v_\mathrm{g}t)^2}{2\sigma_0^2 + 2(v_\mathrm{d}t)^2} + i\beta_0(x - v_\mathrm{p}t)\right)$$
$$\times \exp\left(i\frac{(x - v_\mathrm{g}t)^2}{2\sigma_0^2 + 2(v_\mathrm{d}t)^2}\frac{v_\mathrm{d}t}{\sigma_0}\right), \tag{2.3}$$

where the first term indicates a Gaussian, the mean of which translates at a speed v_g, with a standard deviation that increases in width

with a speed that limits to v_{d}, all whilst the phase propagates with the velocity v_{p}. The second term is a chirp term associated with the dispersion velocity.

These velocities in general are complex quantities, with further reshaping effects, though it is the imaginary component of the central frequency i.e. the *loss* $\gamma = \mathrm{Im}\,\omega(\beta_0)$, that is of most interest. This will have the effect of multiplying the wavepacket by the damping factor $\exp(-\gamma t)$.

It is clear for a stopped light application, a dispersion relation is required that has vanishing group velocity, and minimised dispersion velocity and losses at a particular central wavevector. Therefore it becomes a necessity to be able to design waveguides which have such properties.

2.3. Plasmonic Stopped Light Structures

2.3.1. *Planar waveguides*

The structures that will be considered here are planar slab waveguide structures with metallic layers. These will support surface plasmon modes that, as shall be shown, can be tuned to optimise the dispersion properties of the structure. For the purposes of lasing, light of a narrow frequency range needs to be confined to narrow mode volumes with a gain medium. This can be achieved using stopped light, by designing a structure with *two* stopped light points, i.e. points of zero group velocity $\mathrm{Re}\frac{\mathrm{d}\omega}{\mathrm{d}\beta} = 0$.

The dispersion relation between any two *adjacent* stopped light points is necessarily monatonic since by definition, there are no additional turning points within the interval. If the two adjacent points are optimised to have large wavevector separation and small frequency separation, then light can be confined within to widths inversely proportional to the wavevectors within the interval. This can be quantified as the band velocity, which is the average group velocity between the ZGV endpoints,

$$v_{\mathrm{b}} = \frac{\omega_2 - \omega_1}{\beta_2 - \beta_1} = \frac{\Delta\omega}{\Delta\beta}. \tag{2.4}$$

The wavevector bandwidth determines the minimum width that a light pulse can be confined to, i.e. $\sim 2\pi/\Delta\beta$, whereas the frequency bandwidth will set the overall flatness of the band, and additionally ensures the operation of the stopped light device to remain quasi-monochromatic, which becomes important as the stopped light mode is coupled to inverted emitters in a narrow frequency band. Hence for the purposes of a stopped light laser, a heterostructure must be found that minimises the band velocity.

2.3.2. *Optimisation*

The dispersion of plasmonic structures is ultimately determined by how its layers are composed, i.e. their thickness, material, and relative ordering. It is possible to predict, and even control, how the dispersion relation will look before explicitly calculating it. Having such a model becomes useful in reducing the search space of parameters in optimisation methods, since calculating the dispersion relation consists of solving a transcendental equation, which requires numerical methods. The predictive power stems from the spatial profile of the plasmonic mode having peaks on metal/dielectric interfaces with exponential tails,

$$\phi(z) \propto \begin{cases} \exp\left(-\mathrm{Im}\,\kappa_+(\beta)(z-z_0)\right) & z > z_0 \\ \exp\left(\mathrm{Im}\,\kappa_-(\beta)(z-z_0)\right) & z < z_0 \end{cases}, \tag{2.5}$$

where $\mathrm{Im}\,\kappa(\beta) > 0$, and broadly increases with β. For small values of β, the tails are broad and the mode overlaps with the rest of the structure, whereas for large β the mode only is only sensitive to the interface. Each metal/dielectric interface hosts a surface plasmon, and at low wavevectors they will overlap and hybridise, while for high β they will decouple.

Early studies to seek dispersion relations with a stopped light point were described in Ref. [10], using a metal substrate underneath two dielectric layers, a high index beneath a low index dielectric. The dispersion relation initially followed the steeper plasmon dispersion of the metal/low-index overshooting the lower frequency of the asymptotic metal/high-index surface plasmon frequency which

it would tend to for high wavevectors. This produced a ZGV point between the two regimes, that was tunable by the thickness of the middle layer.

A similar approach can be used for introducing *two* stopped light points. The two modes to hybridise are that of a *metal-insulator-metal* (MIM) system,[5] and a *metal-air* (MA) plasmon. The structures and dispersion relations are plotted in Fig. 2.2. The MIM

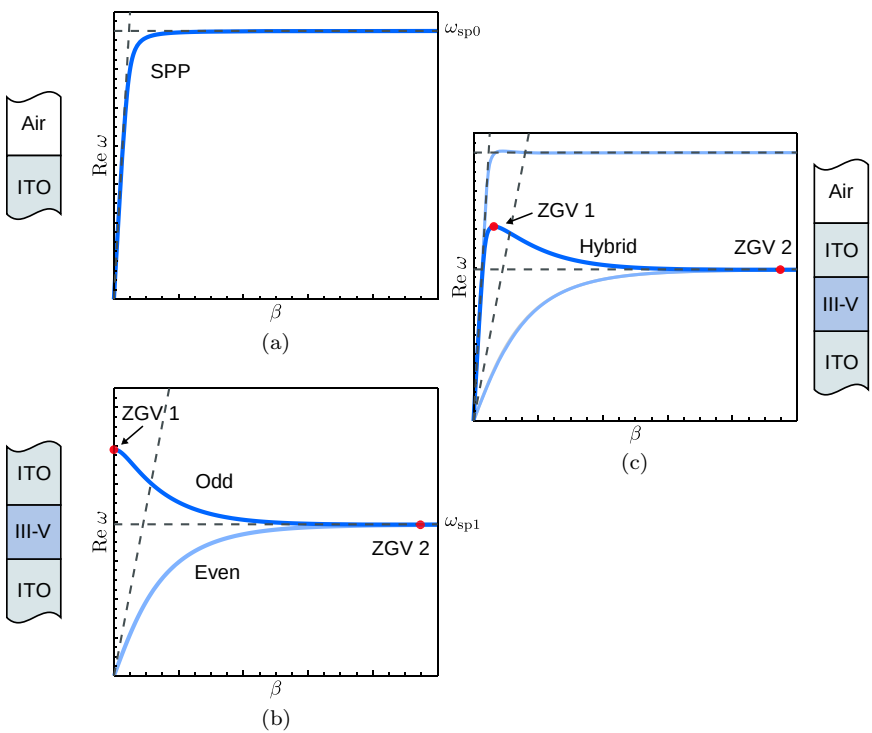


Fig. 2.2. Hybridisation of modes. Dispersion relations, $\mathrm{Re}\,\omega(\beta)$, are plotted, with the asymptotic light lines in vacuum $\omega = c\beta$, and in dielectric $\omega = c\beta/\varepsilon_{\mathrm{bg}}^2$, and surface plasmon frequencies, $\omega_{\mathrm{sp}_0} = \omega_{\mathrm{p}}/\sqrt{\varepsilon_\infty + 1}$ and $\omega_{\mathrm{sp}_1} = \omega_{\mathrm{p}}/\sqrt{\varepsilon_\infty + \varepsilon_{\mathrm{bg}}}$. Modes of interest are in black, and ZGV points are marked as red circles. a. The SPP mode of a metal-air interface, is hybridised with b. the odd mode of a metal-insulator-metal structure, c. forming a metal-insulator-metal-air structure in order to produce a mode with two zero group velocity points at finite wavevector.

system has two modes, one even and one odd; the odd mode (in the z component of the **E**-field) contains two ZGV points itself, the first one is at zero wavevector and is not useful being inside the light cone, the second one is at a high wavevector near where the even and odd modes become degenerate, and is preserved.

In the combination structure, that is a *metal-insulator-metal-air* (MIMA) system, the odd mode of the MIM system hybridises with the MA plasmon, removing the zero-wavevector ZGV point but introducing a new point in the overlap. The second ZGV point can also be perturbed and is being pulled in to a lower wavevector value, depending of the precise structural configuration. Indeed, the positions of the ZGV points can be tuned by adjusting the thicknesses of each layer, while further fine control can be attained by introducing additional dielectric layers.[9]

Reducing the band index of a planar plasmonic heterostucture is a problem suited for engaging an Evolutionary Algorithm (EA). EAs are used to optimise systems that are specified by a set of discreet or numeric parameters, which determine a figure of merit. In this case of a plasmonic stopped light laser, the material composition of layers, thicknesses, and order will determine the plasmon modes that are supported and hence the band velocity, which is optimised to reduce. The EA will compare multiple variants of a structure, each with parameters that are randomly perturbed, hence having slightly different dispersion relations. The structures are ranked by their band velocity, and a subset of the structures with the lowest band velocity are kept to the next round for further mutation. The best structure in each round is continually improved upon until a structure is found within required tolerances.

2.4. Properties of a Stopped Light Laser

In this chapter, we consider structures with materials characteristic of a III-V semiconductor, i.e. InGaAsP for dielectric layers, and a transparent conducting oxide,[14] such as Indium Tin Oxide (ITO) for metallic layers. The dielectric layer has a constant permittivity

$\varepsilon = 11.68$. ITO, which has a plasma resonance in the visible which can be tuned by doping, allowing for operation in the near infrared (i.e. at telecoms wavelength, $\lambda \approx 1550$ nm). It is modelled by a Drude model with parameters as provided by experimental data,[15] with a reduced loss which can be achieved through high fabrication quality and at low temperatures.[12]

A stopped-light heterostructure, optimised by an evolutionary algorithm (EA) for a low band velocity, is presented in Fig. 2.3. It is composed of an ITO substrate on the bottom, and a semiconductor layer sandwiched between an ITO strip on top with layers optimised to the nearest 10 nm. The modes of this structure are shown alongside in Fig. 2.3(b), and indeed the bound TM1 mode hosts two ZGV points. As a result it exhibits a very flat band with a band velocity, in this case, of $v_{\mathrm{b}} = -c/262$.

The mode profile of the energy density of a planewave in the mode is also shown in dependence on the wavevector β. Being plasmonic in nature, the field is peaked on the interface between the metallic and dielectric layers, and for low β values, where the dispersion follows the light cone, it is primarily located in the air layer, entering the structure for higher β values.

For completeness, the complex wavevector (β) modes are also presented in Fig. 2.3. The first four complex β modes are listed here. The modes map to the complex frequency set in particular places on the dispersion curve, but pick up large loss where they diverge.

Having, to a large degree, removed the effects of drift and dispersion from our system, by choosing one with a wide flat band, we are left with loss as the mechanism that will contribute to hindering energy confinement. For all wavevectors where the dispersion relation departs from the light line, the system experiences a loss of around $-2\mathrm{Im}\,\omega \approx \gamma_{\mathrm{p}} \approx 11\mathrm{ps}^{-1}$. This will dissipate any energy within the system on such a timescale. Naturally, the question arises, whether the loss can be compensated for, such as by replacing the dielectric layer with a semiconductor medium, or by explicitly adding emitters such as quantum dots or gain molecules.

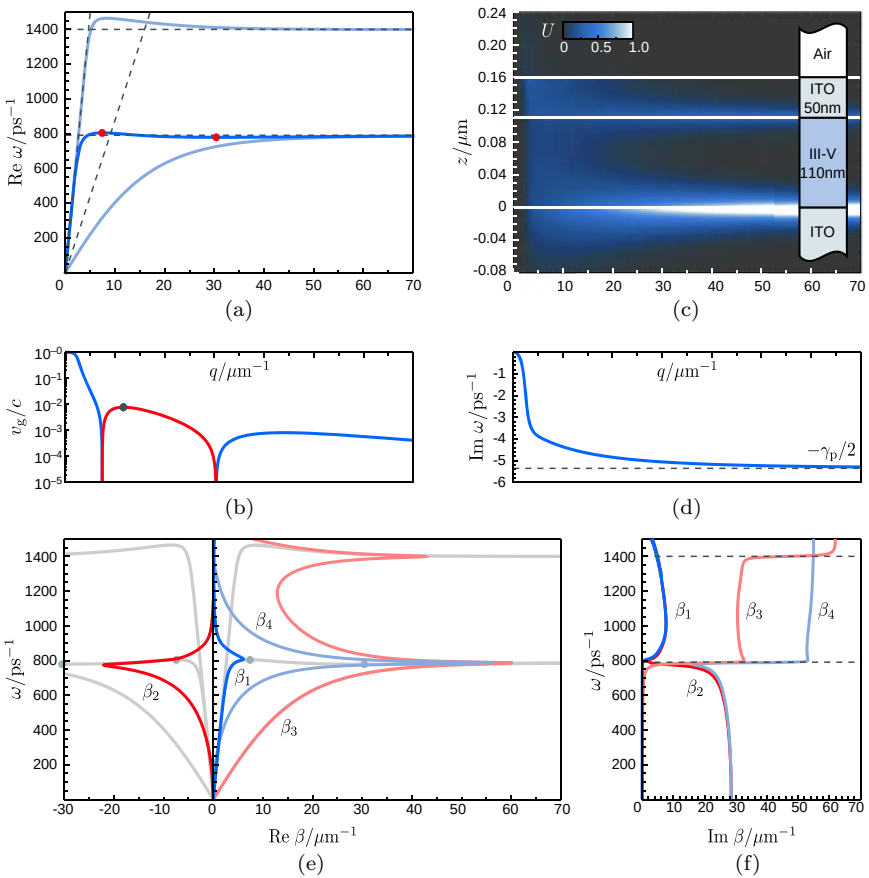


Fig. 2.3. Dispersion relation of a stopped light structure. (a) Complex frequency mode dispersion relations — The mode of interest is highlighted in blue and its ZGV points marked as red dots. Also marked are the light line and surface plasmon frequency, both for the vacuum and dielectric. (b) Group velocity log plot of the mode of interest - Positive group velocities in blue, negative in red. The point of zero dispersion is marked. (c) Description of the structure and energy density throughout the structure of a planewave of wavevector β in the bound mode. (d) Modal loss, $\mathrm{Im}\,\omega(\beta)$, up-to the limiting value of $-\gamma_p/2$. (e) Complex wavevector modes $\mathrm{Re}\,\beta(\omega)$. Of the first four modes which are shown the second mode (red) is of negative phase velocity. The complex frequency modes are plotted in the background (in light grey) to illustrate the correspondence. (f) Modal loss, $\mathrm{Im}\,\beta(\omega)$, of corresponding complex wavevector modes.

2.4.1. *Excitation of stopped light modes*

Given a structure, such as the example presented in the previous section, it is not immediately obvious how to excite light into the stopped light mode. Indeed, the usual methods for coupling to a plasmonic waveguide become unsuitable in the stopped light case. Bound modes are by definition not coupled to external radiation modes, which is to say that light energy cannot be transferred to a bound mode by a beam of light incident on the surface. This is also linked to the secondary reason, that there is a mismatch between the wavevector of incident light and the bound modes of a plasmonic structure, along the direction of propagation. Incident light is located exactly on the light cone in energy-momentum space and has a projection in the propagation direction within it, whereas plasmonic modes are on a line that sits outside the light cone, and indeed need not be bounded at all in momentum.

In plasmonic structures, one method of coupling is achieved by adding a local spatial inhomogeneity, such as a prism, or a grating. In the case of the prism, light is sent down the prism, within the prism's shallower light cone ($\beta \leq n\omega/c$), this allows for points where the energy and momentum of both the incoming beam, and the plasmon mode match.[16] In addition to this condition, the prism must be finite in extent, as the incoming light is part of the radiation spectrum of the "waveguide with prism" system that will ultimately by transmitted, reflected, or absorbed. This mode will mix with both the radiation modes and bound modes of the "waveguide without prism" system, and will decouple from the prism further down the waveguide.

A grating, which is a periodic patterning of the waveguide along the propagation direction, acts in a similar way.[21] The regular patterning allows for scattering between wavevector modes that are integer multiples of the grating wavevector, $2/d$, where d is the period. This gives a *momentum kick* to the incident light field, allowing it to match with the dispersion relation of the bound mode. Again, the radiation mode of the "waveguide with grating" system then mixes with the bound and radiation modes of the "waveguide without grating".

However, these methods are ineffective with stopped light structures because the bound mode has a low group velocity such that the incident light is unable to get sufficiently far away from the grating or prism and instead is out-coupled back through the radiation modes of the combined structures, instead of travelling far enough for the system to be described by modes without a grating or prism.

Another possible scheme for coupling light into a waveguide is end-fire coupling.[25] Here the waveguide is assumed to terminate in a plane perpendicular to the direction of propagation. If a light pulse is incident on this terminal plane, its profile can be decomposed into a mixture of bound modes and radiation modes of the waveguide; the radiation modes propagate away, leaving the bound modes in the system. This too is unsuitable for exciting stopped light modes, as zero group velocity modes will not propagate down the structure, they will not enter and instead be reflected back. This method of excitation is equivalent to using the complex wavevector picture, where the temporal profile of excitation at one point along the axis (the terminal) is known, however ZGV points are not well described in the complex wavevector picture.

Finally, a way to excite the bound modes of a plasmonic waveguide, that is compatible with stopped light, is to have them emitted directly from within the structure. Here, an emitter would be placed within the SL structure and these would be pumped to an excited state, such that when they relax, by spontaneous or stimulated emission, they are able to emit directly into the bound stopped light mode.

2.5. Small Signal Gain Properties

Adding gain to a nanoplasmonic heterostructure will alter the bound modes that are supported by the passive structure. A material that can spontaneously emit photons will also be available to emit light via stimulated processes, thereby adding gain to the system. This gain will be dispersive in two ways.

Firstly, different frequencies experience different amounts of gain. Secondly, any change to the imaginary part of the refractive index

(i.e. the material gain) will in turn induce a change in the real part of the refractive index due to causality and the Kramers-Kronig relations. Thus, designing an active SL structure requires particular care that the introduction of a gain medium does not damage the stopped light character of the waveguide. The induced change in mode structure can be analysed using the same transfer matrix methods of the previous section with the inclusion of a Lorentz resonance at a defined emission frequency, ω_e, to one of the dielectric layers. The Lorentz resonance represents the transitions between the levels of a two-level system with an energy difference $\hbar\omega_e$ In this picture, there is an occupation density of upper and lower levels, averaged over space, of single two-level emitters.

The strength of the resonance is proportional to the *inversion density* of the two levels, i.e. how much more the higher level is occupied than the lower one, $\Delta N = N_2 - N_1$, such that a layer with emitters embedded may be represented by:

$$\varepsilon = \varepsilon_{bg} + \frac{\Delta N}{N} \frac{\omega_{pe}^2}{\omega_e^2 - \omega(\omega + i\gamma_e)} \tag{2.6}$$

$$\omega_{pe}^2 = -N\sqrt{\varepsilon_{bg}}\gamma_e \sigma_e c, \tag{2.7}$$

where ε_{bg} is the permittivity of the layer hosting the resonance, σ_e is the emission cross-section, and γ_e is the width of the resonance.[30] Note that for negative inversion, the emitter becomes an absorber as there is a higher density of emitters in the lower state. The SL structure as presented in the previous section is modified by replacing the dielectric layer with a Drude-Lorentz emitter (with a 10 nm buffer on both ends of zero inversion density to simulate quenching by the metal layer). The emission frequency is set to match either one of the ZGV points, and the other parameters are set as in Ref. [18] representing the inclusion of realistic laser dye molecules.[24] The resulting dispersion and loss relations are shown in Fig. 2.4 for inversion densities ΔN varying between 0 and N, with a fixed emitter density, and excitation about ZGV1 and ZGV2. The first point of note is that, even on full inversion, the addition of gain does not significantly

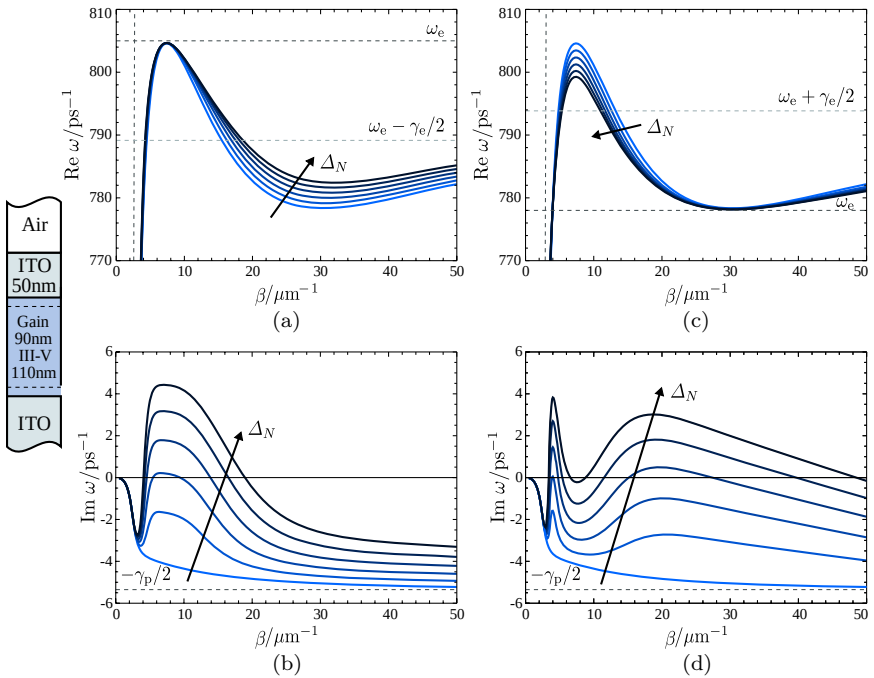


Fig. 2.4. Perturbation of dispersion relation and loss with Lorentzian gain. The change of mode shape is plotted as Lorentzian gain is introduced to the system. For subfigures on the left, emission is about ZGV1, and on the right about ZGV2. (a) and (b) show a zoom in of how the dispersion relation changes with increasing inversion density, while (c) and (d) show the corresponding loss/gain of the mode.

change the dispersion, with the maximum shift in frequency being around 0.6%. The presence of ZGV points is preserved, though they may drift slightly, i.e. ZGV2 moves slightly right with emission about ZGV1. There is no change at the frequency that is being excited, this is because the permittivity change of a Lorentzian is zero at the resonant frequency. Adding gain has the side effect of making the structure a slightly better SL structure as the band velocity is marginally reduced. The key change, however, is to be found in the loss. As the inversion density increases, the loss decreases for wavevectors where the dispersion is within the gain width. Initially the plasmons' loss is reduced as the inversion increases, then for an inversion of

around $\Delta N/N \approx 0.4$ some wavevectors become undamped, and even eventually experience gain, Plasmons sitting in these modes will grow exponentially in amplitude whilst small enough to remain in the small signal gain regime. As ZGV2 is flatter than ZGV1, when the emission is about this point, a wider range of frequencies fall within the gain width, leading to a larger range of β values that can become undamped.

2.6. Dynamics

A frequency domain analysis can only bring valuable insight on aspects such as the small signal gain, but it is unable to describe (dynamic) nonlinearities. When plasmons are emitted, electrons are demoted from higher energetic states to lower ones. This depletes the population inversion, reducing the available gain over time, leading to a nonlinear field-dependence. In addition, there are spatial effects to consider, such as spatial hole burning; the TMM has assumed uniformity in the direction of propagation, whereas the level of inversion can vary both in this direction and perpendicular to the stacking. Depending on the mode formed, some regions may host higher field densities which can deplete the local gain. Thus a dynamic, spatially resolved, time domain picture and simulation is required to capture all the aspects of emission into a stopped light mode. In this section we will thus present and discuss results from FDTD simulations that fully take on board the spatio-temporal dynamics of the light field, the gain system as well as the nanoplasmonic environment.

2.6.1. *Four-level gain model system*

The two level system, which we used in our discussion in the previous section, is appropriate for modelling a single electronic transition. However, its inversion density had been set "by hand". Moreover, it is well known that there is no way in which a two level system can reach a state of inversion by relying solely on the processes of spontaneous emission and absorption (the best that can be achieved is equal occupation of the levels). A *four-level system* on the other hand

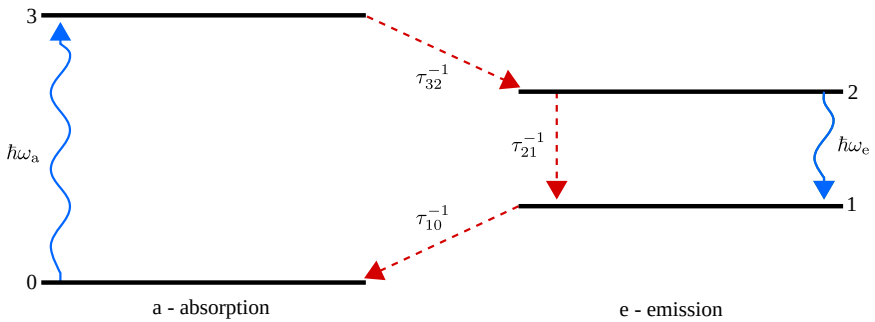


Fig. 2.5. Schematic of the four-level system. A four-level system is composed of two two-level systems, a and e, with active transitions ($0 \leftrightarrow 3$) and ($1 \leftrightarrow 2$) respectively, which are coupled by nonradiative transition rates τ_{32}^{-1} and τ_{10}^{-1}. In this scheme, the emission subsystem permits radiative transitions at an energy of $\hbar\omega$ and has a slow nonradiative recombination rate τ_{21}^{-1}, while the absorption subsystem is electrically pumped at a constant rate r_p.

can be constructed such that there can be a dynamically maintained population inversion between two of its levels, from where stimulated emission can occur. Let us consider a four-level system that contains two two-level systems, labelled e and a for emission and absorption; e is a system with energy levels 1 and 2 and energy gap $\hbar\omega_e$, in-between a, a system with energy levels 0 and 3 and greater energy gap $\hbar\omega_a$, as depicted in Fig. 2.5. The two upper levels, 2 and 3 are coupled by a fast nonradiative relaxation channel, with rate τ_{32}^{-1}, as are the two lower levels, 0 and 1 (τ_{10}^{-1}). This has the effect of rapidly depleting the 1st and 3rd level shortly after they become occupied. Between the levels of the emission two-level subsystem, there is additionally a slow nonradiative channel (τ_{21}^{-1}).

The key point of a four-level system which is useful to achieve lasing is that electrons that are pumped from levels 0 and 3, will quickly decay to level 2, leading to an inversion density of level 2 over level 1, which is available for stimulated emission. Fortunately there are may gain material systems such as laser dyes that meet this condition. In general this will then allow optical pumping between levels 0 and 3 to generate inversion between levels 1 and 2. Alternatively, a constant electrical pump rate r_p can be used, placing the emphasis

on the emitted fields, to which the stopped light singularities are tuned.

Four-level systems are incorporated into the FDTD framework using time-domain differential equation for the polarisation,

$$\frac{\partial^2 \mathbf{P}_\mathrm{e}}{\partial t^2} + \gamma_\mathrm{e} \frac{\partial \mathbf{P}_\mathrm{e}}{\partial t} + \omega_\mathrm{e}^2 \mathbf{P}_\mathrm{e} = \frac{\Delta N}{N} \omega_\mathrm{pe}^2 \varepsilon_0 \mathbf{E} \,, \tag{2.8}$$

Here the polarisation has been split off into a part $\mathbf{P}_\mathrm{e}$ that is connected with the radiative resonances, it is added to the total polarisation when entering into the electric field update equations. The corresponding level occupation densities update with the auxiliary differential equations,[30]

$$\frac{\partial N_3}{\partial t} = r_\mathrm{p} N_0 - \frac{N_3}{\tau_{32}} \tag{2.9}$$

$$\frac{\partial N_2}{\partial t} = \frac{N_3}{\tau_{32}} + \frac{1}{\hbar \omega_\mathrm{e}} \left(\frac{\partial \mathbf{P}_\mathrm{e}}{\partial t} + \frac{\gamma_\mathrm{e}}{2} \mathbf{P}_\mathrm{e} \right) \cdot \mathbf{E} - \frac{N_2}{\tau_{21}} \tag{2.10}$$

$$\frac{\partial N_1}{\partial t} = \frac{N_2}{\tau_{21}} - \frac{1}{\hbar \omega_\mathrm{e}} \left(\frac{\partial \mathbf{P}_\mathrm{e}}{\partial t} + \frac{\gamma_\mathrm{e}}{2} \mathbf{P}_\mathrm{e} \right) \cdot \mathbf{E} - \frac{N_1}{\tau_{10}} \tag{2.11}$$

$$\frac{\partial N_0}{\partial t} = \frac{N_1}{\tau_{10}} - r_\mathrm{p} N_0. \tag{2.12}$$

Quantum noise and resulting *amplified spontaneous emission* (ASE) can modelled on the basis of this semiclassical description by including spatially resolved dynamic Langevin noise terms to the system, which accounts for the dissipative reservoirs feeding back stochastically on the system. The noise couples to the four-level system and induces incoherent transitions, which then become amplified, allowing for the triggering of the lasing regime.[20]

2.6.2. *Lasing dynamics*

In order to use in FDTD, the layered geometry of Fig. 2.3 is recreated to fit in a simulation box with the top and bottom layers (that were unbounded in the analytic study) having a finite height of 500 nm,

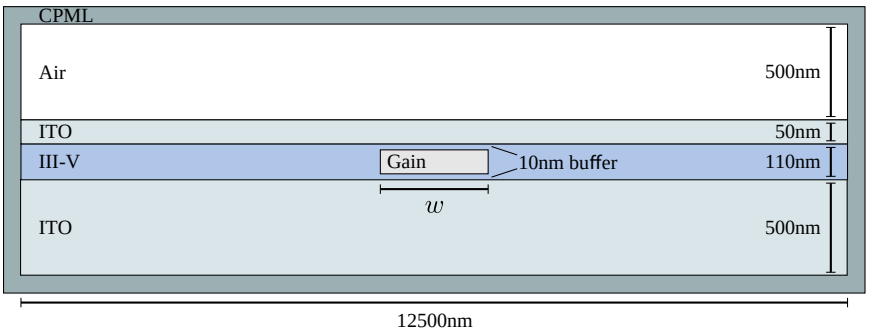


Fig. 2.6. Schematic of the FDTD setup. The layered structure described in the frequency domain analysis is truncated to fit in a simulation domain surrounded by CPML layers. The pumped gain region is placed in the horizontal centre of the simulation, and extends over a width w; vertically it takes up the height of the III-V layer which is embedded in up-to a 10 nm buffer on either side.

and the entire structure having a width of 12500 nm. The simulation domain is surrounded with *convolutional perfectly matched layers* (CPML) layers, which quickly attenuate incident fields without introducing reflection such that they do not interact with the simulation boundary.[22] This is depicted in Fig. 2.6.

In order to demonstrate the effect of a stopped light band on lasing, we shall consider two cases: the first with the stopped light structure as described, and the second with a control structure with the effect of removing the stopped light points whilst keeping a TM1 mode that is in the same frequency range.

Figures 2.7(a–c) shows the mean inversion and energy density of a point in the centre of the emitter region.For the cases in the SL structure that did enter a lasing regime, characteristic relaxation oscillations can be seen where the inversion initially builds up and spontaneous emission events are induced. The fields then are amplified as they stimulate further emission, increasing the field in the mode coherently. From here the emitted fields start to grow exponentially, and when they are of sufficient strength will deplete the inversion density. This reduction of inversion feeds back by decreasing the available gain, leading to a decrease in field energy as the energy in the field is lost to dissipative processes in the metal layers.

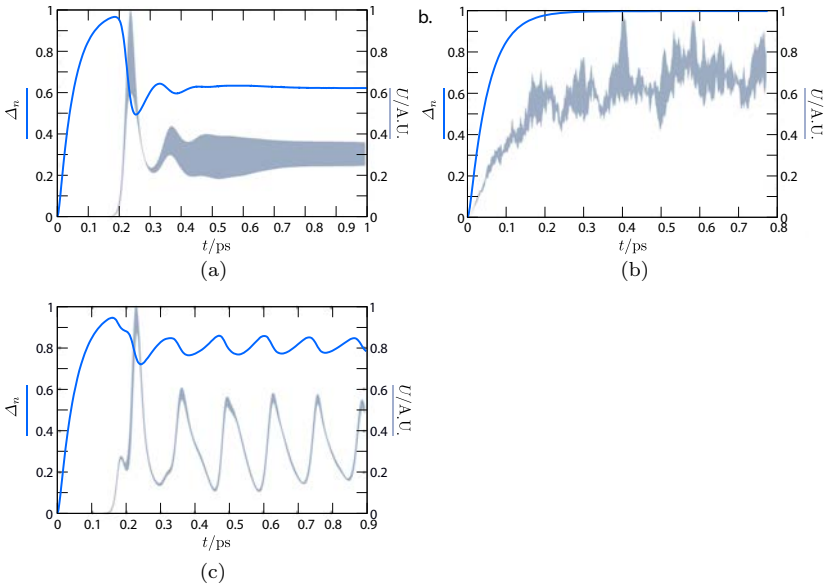


Fig. 2.7. Lasing onset with increasing gain density. (a), (b), (c). Energy density and mean inversion in the stopped-light structure for (a) a gain density $N = 0.002\,\mathrm{nm}^{-3}$ and (b) ta gain density $N = 0.001\ \mathrm{nm}^{-3}$. (c) shows the dynamics of the control structure with a gain density $N = 0.004\ \mathrm{nm}^{-3}$.

The decrease in energy density in the field allows for the inversion to rebuild. This continues in an oscillatory manner with the energy density lagging behind the inversion by 90°. The amplitude of the inversion and energy oscillations decrease with each cycle until a stable steady value for both is reached.

If the inversion density is less than a threshold value then the system cannot enter a lasing regime and is only able to support occasional bursts of amplified sponaneous emission (ASE). This case is shown in the inversion and energy density of Fig 2.7(b). Once the lasing threshold has been passed, the energy density in the mode increases linearly with the gain density.

In contrast, the control structure, albeit entering a regime of relaxation oscillations, is more erratic and the oscillations do not settle to a steady state. Instead the oscillations continue with a factor of 4 between the peak energy density and the trough.

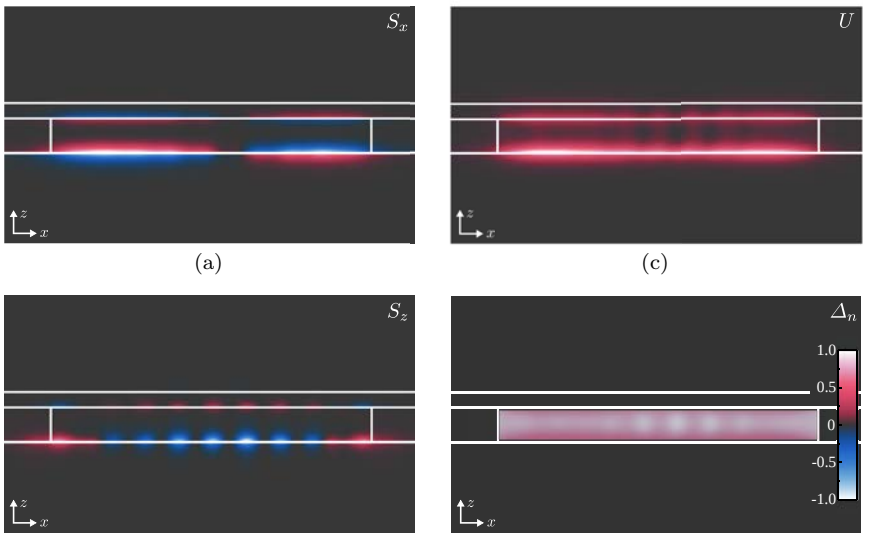


Fig. 2.8. Energy flux and density of lasing mode. The lasing mode of a SL structure with a width 1000nm. a. and b. show the x and z component of the cycle-averaged Poynting vector respectively, with positive values in red, and negative in blue. Focusing on the bottom right lobe of the flux in (a), energy flows into the gain region in the dielectric layers and out of it in the metal layer. For the same lobe in (b), energy moves downwards out of the gain region towards the centre, and upwards into the region at the edge. The counter-clockwise energy vortex in this corner is mirrored in the other corners. The cycle-averaged energy density is plotted in (c). and inversion in (d). The inversion is the complement of the energy density, being highest in areas with low energy density, and depleted where the mode sits.

2.6.3. *Lasing mode*

In this section, the lasing modes are investigated for SL structures with varying width of the gain region in the range $w \in [200, 1500]$ nm, using a fixed gain density of $N = 0.002$ nm^{-3}. The spatially resolved, cycle averaged Poynting vector, energy density, and inversion are shown in Fig. 2.8 for a structure with width 1000 nm. The energy is concentrated on metal-dielectric interfaces, strongest on the lower interface, and is localised around the gain medium despite the absence of a cavity along the horizontal direction. It is the Poynting

vector which shows how energy circulates in the structure to generate a feedback which is required for lasing. There are four energy flux vortices, one in each corner of the structure. These have energy moving out of the gain region into the metal layers and into it in dielectric layers. Considering the x component of the Poynting flux, (Fig. 2.8(a)), the forward and backward flows are in exact balance. The balanced counter-propagation of energy in the negative-permittivity (metal) layers against the dielectric layer. This is the basis of feedback in the system.

The inversion is shown in Fig. 2.8(d), there are areas of spatial hole burning where the energy density is highest. The spatial modulation seen is explained when considering the field profile and its formation. It is important to note that the modes that are formed when the structure enters the lasing regime are a dynamic synthesis of the spectrum of planewave modes available rather than that of a predefined cavity mode. Indeed, these dynamic modes that form are propagating waves, with an advancing phase, rather than purely standing waves, as shown in Fig. 2.9. It will be shown that this results from modes being centred independently on finite positive and negative β points, the relative excitation of each competing with each other, rather than having symmetric excitation at $\beta = 0$. This is in contrast even to the stopped light lasing structure considered in Ref. 18, which is photonic in nature rather than plasmonic, and emitted symmetrically about $\beta = 0$ as a standing wave.

In all cases, the modes are inwardly propagating, that is the leftmost half propagates right and vice versa. These two halves form a standing wave where they meet, and this node need not be in the centre of the structure, instead being randomly chosen by the spontaneous symmetry breaking at the transition from ASE to lasing. The asymmetry of the mode profile will be discussed later in the chapter.

Lasing is shown to be possible for structures with a gain region width down to 200 nm, at which point the steady state inversion rises to around 0.9. For this case one observes a standing wave in the interior of the gain stripe, i.e. a symmetric excitation of positive

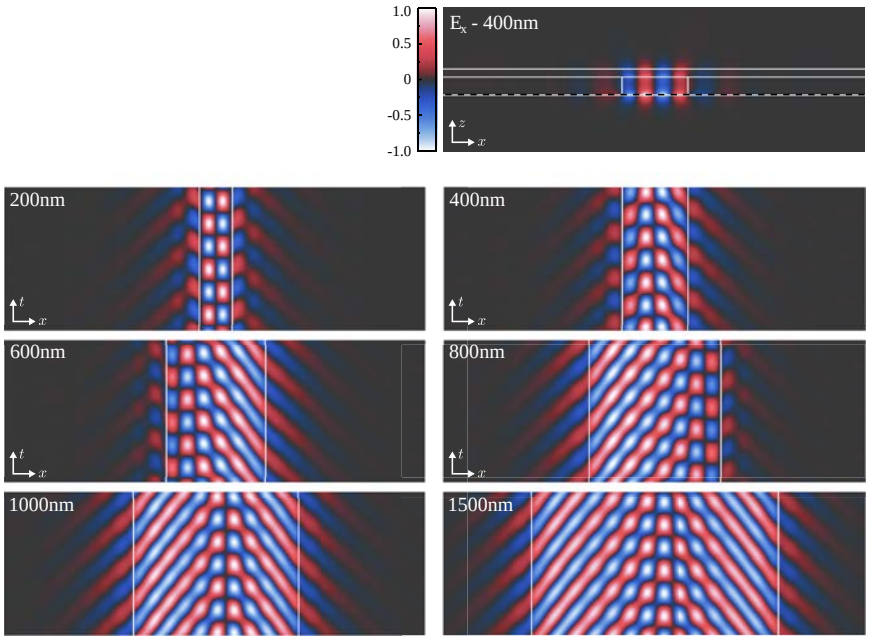


Fig. 2.9. Field evolution in steady state. The dynamics of the E_x field along a line slice at the bottom metal interface is displayed . For the 200 nm structure, the field inside the gain region is a standing wave. The other structures have two inward propagating waves which meet at a standing node. This node can appear towards the edges, i.e. for 600 and 800 nm, or positioned towards the centre for 1000 and 1500 nm. A stream of plasmons is emitted at either side of the gain region, with a predominantly negative phase velocity.

and negative β modes. The confinement of such a structure is at its limit here, with significant parts of the field profile outside the gain region. For widths smaller than this, there will not be enough gain to pass the lasing threshold. This thinnest confinement at 200 nm is 12× the free-space wavelength, and indeed 3.5× the bulk wavelength in the semiconductor layer.

Outside of the gain region, a stream of plasmons is emitted to either side with a negative phase velocity. To describe these emitted plasmons, it is convenient to use a complex-wavevector picture, as we effectively have a steady state oscillating source for which the field is known at a fixed point in space (the edges of the gain region) and such

that we can consider the spatial evolution away from this point. Once the system has entered into a steady state one can then analyse the spectral content of the fields. Using *discrete Fourier transform* (FFT) methods, the peak frequency component, ω_{peak}, of the electric field is identified and isolated. This returns a spatially resolved, complex valued function $\tilde{\mathbf{E}}(\mathbf{x}, \omega_{\text{peak}})$. Taking a spatial Fourier transform along the waveguide (x) axis, allows the extraction of the spatial power spectrum, which is averaged over z positions within the gain layer, $I(\beta) = |\tilde{\mathbf{E}}(\beta, \omega_{\text{peak}})|^2$.

The wavevector of the emitted plasmon can be extracted by taking the FFT over positions outside of the emitting gain region. Figure 2.10 shows the spectrum of fields emitted from the right

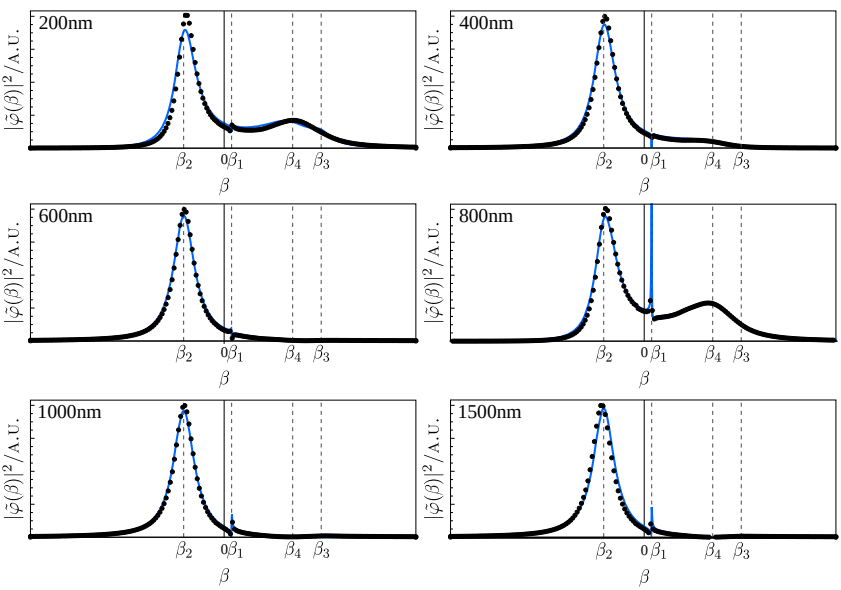


Fig. 2.10. Complex wavevector plasmons. Power spectra of fields emitted from the right terminal of the gain region (black dots). This is fit to the lineshape of the first four complex β plasmons (black line), i.e. a Lorentzian function, $|\tilde{\phi}(\beta)|^2 = |\sum_i^4 \phi_i/(\beta - \beta_i)|^2$. The amplitudes are varied to fit, but the wavevectors are fixed from the complex β plasmon dispersion at the lasing frequency. The imaginary part becoming the width in the spectral density. In each, the negative phase velocity β_2 plasmon is strongest, with β_4 having a large amplitude for 200 nm and 800 nm which can be also seen in the field profiles of Fig. 2.9.

terminal, i.e. integrating from the edge of the gain region up to but excluding the CPML layer. These data are fit to the first four analytically determined complex-wavevector modes (see Fig. 2.3) at the lasing frequency. The wavefunction of such modes is the Fourier transform of a complex decaying exponential,

$$\phi(x) = \theta(x) \exp(i\mathrm{Re}\,\beta_i\, x) \exp(-\mathrm{Im}\,\beta_i\, x) \tag{2.13}$$

$$\tilde{\phi}(\beta) \propto \frac{1}{\beta - \beta_i}, \tag{2.14}$$

where β_i is the complex wavevector of the bound mode SPP. This is summed over each of the first four plasmons, each with a complex amplitude, then the absolute square of this is compared with the data, i.e.

$$|\phi(\beta)|^2 = \left| \sum_{i=1}^{4} \frac{\phi_i}{\beta - \beta_i} \right|^2 . \tag{2.15}$$

The fit allows the amplitude of each resonance to change, but keeps the wavevectors constant. Excellent agreement is found, confirming the presence and applicability of the description of complex-wavevector plasmons.

The negative group velocity plasmon β_2, has the strongest amplitude in all cases, though the narrow width β_1 plasmon is present too. In cases where the standing wave node is close to the terminal, i.e. for $w = 200, 400, 800$ nm there is significant excitation of the short propagation length β_4 plasmon. Thus, the stopped light lasing principle can be utilised as a source of coherent plasmons at a single frequency and discrete wavevector. Alternatively by adding a grating to the structure away from the gain region, the plasmons may be coupled out, converting this into a photonic SL laser.

To capture the profile of the lasing mode, the spatial power spectrum, $I(\beta)$, can also be taken over the entire domain, rather than in the interval to the right of the gain region. We note that this function is not symmetric about $\beta = 0$ since it was transformed from a

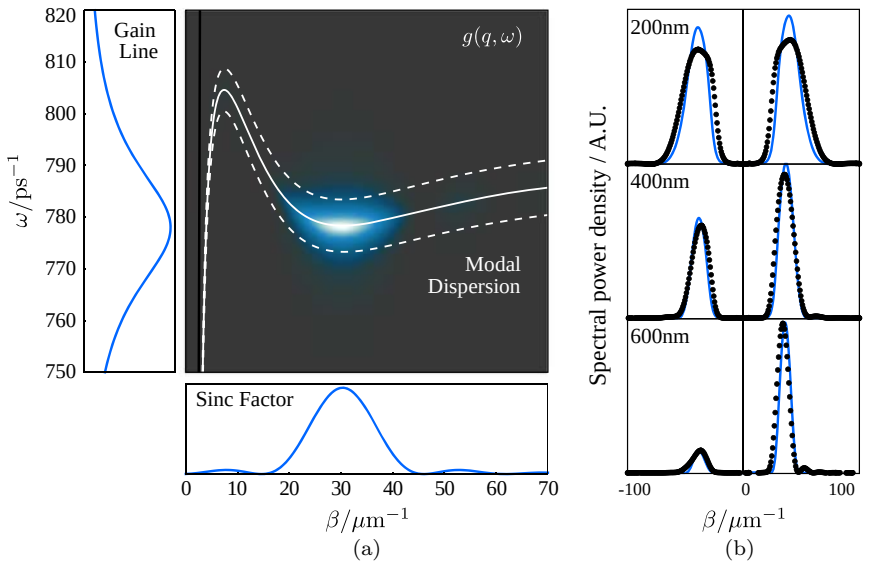


Fig. 2.11. Power spectrum of lasing modes. (a) Coupling strength $g(\beta, \omega)$, which is composed of the emission lineshape (left), the *sinc* function of the gain region (bottom), and the mode support (overlayed). The frequency which maximises the the integral of g over wavevector is *picked* by the system to lase at. (b) temporal and spatial power spectrum of the lasing mode for model (black lines) and FDTD data (black points). The spatial mode profile is modelled by a modified form of the coupling strength at the lasing frequency, as given by Eq. 2.16. Results are in excellent agreement with the FDTD results.

complex valued function. Hence it does not follow the usual evenness properties of the Fourier transform of a real function. This allows for inspection of how the positive and negative wavevector (and hence phase velocity) modes are independently excited. Figure 2.11(c) shows the power spectra of lasing modes in the range of gain widths considered. It can be seen that in each case, the power spectrum is a bimodal distribution with peaks about $\beta \approx \pm 30.7\ \mu\mathrm{m}^{-1}$, which is exactly the wavevector of the second stopped light point. As one would expect, the width of each peak is inversely proportional to the gain width chosen, e.g. the smallest gain section at 200 nm, the width is around 5 $\mu\mathrm{m}^{-1}$. There is no preferential excitation of the peaks, and seemingly no correlation between their relative amplitude. This

is because they are two competing modes, one of which will initially take the lead in growth due to the spontaneous symmetry breaking of amplified spontaneous emission. The first mode will dominate the gain, being able to exponentially grow in field strength. The second mode may still be allowed to grow but spatially separated from the first, harvesting areas where the field and gain are less strongly coupled. This picture is corroborated when viewing the time evolution of the steady state fields, which show two lobes, inward propagating, with a finite overlap width where the wave becomes standing, this is depicted in Fig. 2.9. This is developed further in Ref. 29, where it is shown how the location of the nodes move on longer time scales.

It is possible to predict the spectral content of the lasing mode, $I(\beta)$, up to a factor of the weights of the mode in the $+\beta$ and $-\beta$ excitation (I^+, I^-), by a slight modification of $g(\beta, \omega)$, to include both excitations,

$$I(\beta) \propto \frac{\gamma_{\rm pl}(\beta)}{(\omega - \omega_{\rm pl}(\beta))^2 + \gamma_{\rm pl}(\beta)^2} \left(I^+ \mathrm{sinc}\left(\frac{w(\beta - \beta_2)}{2}\right) + I^- \mathrm{sinc}\left(\frac{w(\beta + \beta_2)}{2}\right)\right)^2, \tag{2.16}$$

which is in excellent agreement with the FDTD spectra.

2.7. Conclusions

In this chapter we introduced *plasmonic stopped light lasing*, whereby surface plasmon polaritons are localised by reducing the group velocity of a wavepacket to zero whilst within a gain medium. This is in contrast to traditional nanolasing schemes, which localise energy in a resonant cavity.

In order to explain the concept, topics of dispersion, and the nuances of the complex-frequency/complex-wavevector pictures have been discussed. Frequency domain methods have been employed alongside an evolutionary optimisation algorithm in order

to characterise the properties and quality of a structure and select for optimal structures within constraints.

An exemplary structure was introduced, composed of realistic materials, whose properties were studied throughout the rest of the chapter. Analysis in frequency domain was continued to investigate how, in the small signal gain regime, a two-level emitting resonator could compensate for the inherent material losses in metal layers, determining idealised threshold values of inversion required to free a mode of damping.

Equipped with this analysis, finite difference time domain simulations capture the dynamics of spatio-temporal and nonlinear effects. The previous frequency domain analysis of the threshold inversion was corroborated by varying the gain density in each time domain simulation, and it was found that lasing is indeed possible in this scheme.

The lasing mode was investigated, and a model for predicting its modal content proposed. It was found that the feedback provided from stopped light lasing ultimately derives from a dynamically formed vortex of power flow, with propagation and counter-propagation balancing between the dielectric and metal layers. Despite being stopped, the lasing mode carries a finite phase velocity, with an inward propagating phase modulation that can be detected from the top of the structure.

As an output of the lasing process, coherent plasmon polaritons (sitting on the cusp of the complex-wavevector dispersion curve) are emitted from the sides of the gain region.

This is a new type of sub-wavelength laser, where the active component is smaller than a few hundred nanometres, and coherently emits plasmons directly into a waveguide without relying on external coupling mechanisms. The dynamic formation of the cavity-free lasing mode is a new physical feature, the implications of which are open. It could conceivably become the basis for single frequency coherent SPP generation in quantum plasmonic applications,[7,26] or function as the basis of a quantum fluid such as a photonic Bose-Einstein Condensate.[4]

Bibliography

[1] Altug, H., Englund, D. and Vučković, J. (2006). Ultrafast photonic crystal nanocavity laser. *Nature Phys.*, **2**(7):484–488.

[2] Cao, H. (July 2003). Lasing in random media. *Waves in Random Media*, **13**(3):R1–R39.

[3] Carrère, H., Marie, X., Lombez, L. and Amand, T. (2006). Optical gain of InGaAsNInP quantum wells for laser applications. *Applied Physics Letters*, **89**(18).

[4] Carusotto, I. and Ciuti, C. (February 2013). Quantum fluids of light. *Reviews of Modern Physics*, **85**(1):299–366.

[5] Economou, E. N. (1969). Surface plasmons in thin films. *Phys. Rev.*, **182**:539–554.

[6] Iga, K., Koyama, F. and Kinoshita, S. (1988). Surface emitting semiconductor lasers. *IEEE J. Quantum Electron.*, **24**:1845–1855.

[7] Jacob, Z. and Shalaev, V. M. (October 2011). Plasmonics goes quantum. *Science*, **334**(6055):463–464.

[8] Jang, M. S. and Atwater, H. A. (November 2011). Plasmonic rainbow trapping structures for light localization and spectrum splitting. *Physical Review Letters*, **107**(20):207401.

[9] Karalis, A., Joannopoulos, J. and Soljačić, M. (2009). Plasmonic-dielectric systems for high-order dispersionless slow or stopped subwavelength light. *Phys. Rev. Lett.*, **103**:43906.

[10] Karalis, A., Lidorikis, E., Ibanescu, M., Joannopoulos, J. D. and Soljačić, M. (August 2005). Surface-plasmon-assisted guiding of broadband slow and subwavelength light in air. *Physical Review Letters*, **95**(6):63901.

[11] Kenyon, A. J., Chryssou, C. E., Pitt, C. W., Shimizu-Iwayama, T., Hole, D. E., Sharma, N. and Humphreys, C. J. (2002). Luminescence from erbium-doped silicon nanocrystals in silica: Excitation mechanisms. *Journal of Applied Physics*, **91**(1).

[12] Khajavikhan, M., Simic, A., Katz, M., Lee, J. H., Slutsky, B., Mizrahi, A., Lomakin, V. and Fainman, Y. (February 2012). Thresholdless nanoscale coaxial lasers. *Nature*, **482**(7384):204–207.

[13] McCall, S. L., Levi, A. F. J., Slusher, R. E., Pearton, S. J. and Logan, R. A. (1992). Whispering-gallery mode microdisk lasers. *Appl. Phys. Lett.*, **60**:289–291.

[14] Naik, G. V., Shalaev, V. M. and Boltasseva, A. (2013). Alternative plasmonic materials: beyond gold and silver. *Adv. Mater.*, **25**:3264–3294.

[15] Noginov, M. A., Gu, L., Livenere, J., Zhu, G., Pradhan, A. K., Mundle, R., Bahoura, M., Barnakov, Yu. A. and Podolskiy, V. A. (2011). Transparent conductive oxides: Plasmonic materials for telecom wavelengths. *Applied Physics Letters*, **99**(2):021101.

[16] Otto, A. (1968). Excitation of nonradiative surface plasma waves in silver by method of frustrated total reflection. *Zeitschrift Fur Physik*, **216**(4):398–410.

[17] Painter, O. (1999). Two-dimensional photonic band-gap defect mode laser. *Science*, **284**:1819–1821.
[18] Pickering, T. W., Hamm, J. M., Page, A. F., Wuestner, S., and Hess, O. (September 2014). Cavity-free plasmonic nanolasing enabled by dispersionless stopped light. *Nature Communications*, **5**(May):4972.
[19] Plum, E., Fedotov, V. A., Kuo, P., Tsai, D. P. and Zheludev, N. I. (2009). Towards the lasing spaser: controlling metamaterial optical response with semiconductor quantum dots. *Optics Express*, **17**(10):8548–8551.
[20] Pusch, A., Wuestner, S., Hamm, J. M., Tsakmakidis, K. L. and Hess, O. (March 2012). Coherent amplification and noise in gain-enhanced nanoplasmonic metamaterials: A Maxwell-Bloch Langevin approach. *ACS Nano*, **6**(3):2420–2431.
[21] Ritchie, R. H., Arakawa, E. T., Cowan, J. J. and Hamm, R. N. (November 1968). Surface-plasmon resonance effect in grating diffraction. *Physical Review Letters*, **21**(22):1530–1533.
[22] Roden, J. A. and Gedney, S. D. (December 2000). Convolution PML (CPML): An efficient FDTD implementation of the CFSPML for arbitrary media. *Microwave and Optical Technology Letters*, **27**(5):334–339.
[23] Smith, R. L. (1970). The velocities of light. *American Journal of Physics*, **38**:978.
[24] Sperber, P., Spangler, W., Meier, B. and Penzkofer, A. (1988). Experimental and theoretical investigation of tunable picosecond pulse generation in longitudinally pumped dye laser generators and amplifiers. *Optical and Quantum Electronics*, **20**(5):395–431.
[25] Stegeman, G. I., Wallis, R. F. and Maradudin, A. A. (1983). Excitation of surface polaritons by end-fire coupling. *Optics Letters*, **8**(7):386–388.
[26] Tame, M. S., McEnery, K. R., Ozdemir, S. K., Lee, J., Maier, S. A. and Kim, M. S. (June 2013). Quantum plasmonics. *Nat Phys*, **9**(6):329–340.
[27] Vahala, K. J. (2003). Optical microcavities. *Nature*, **424**:839–846.
[28] Wiersma, D. S. (May 2008). The physics and applications of random lasers. *Nat. Phys.*, **4**(5):359–367.
[29] Wuestner, S., Pickering, T. W., Hamm, J. M., Page, A. F., Pusch, A. and Hess, O. (2015). Ultrafast dynamics of nanoplasmonic stopped-light lasing. *Faraday Discussions*, **178**:307–324.
[30] Wuestner, S., Pusch, A., Tsakmakidis, K. L., Hamm, J. M. and Hess, O. (September 2010). Overcoming losses with gain in a negative refractive index metamaterial. *Physical Review Letters*, **105**(12):127401.
[31] Yeh, P., *Optical Waves in Layered Media*. Wiley Series in Pure and Applied Optics. Wiley (1988).

CHAPTER 3

Nonlinear Metamaterials

MARTTI KAURANEN

Tampere University of Technology, Finland
martti.kauranen@tut.fi

STEFAN LINDEN

Rheinische Friedrich-Wilhelms–Universität Bonn, Germany
linden@physik.uni-bonn.de

MARTIN WEGENER

Karlsruhe Institute of Technology, Germany
martin.wegener@kit.edu

Nonlinear metamaterials have attracted considerable interest in recent years. Here, we review the current status of the research in this field and discuss the nonlinear optical properties of metal-based metamaterials, hybrid metamaterials, and purely dielectric metamaterials. In doing so we concentrate on the near-infrared and visible parts of the spectrum. We critically discuss the prospects and limitations of nonlinear metamaterials for future applications.

3.1. Introduction

Metamaterials are rationally designed artificial solids composed of sub-wavelength building blocks ("meta-atoms") that are densely

packed into a man-made crystal.[1] In this spirit, metamaterials obtain their properties from structure rather than chemistry. This concept has enriched optics and photonics with new and fascinating aspects and has led to the development of artificial materials with unprecedented linear optical properties such as magnetism at optical frequencies, negative index materials, and strong chirality.

Regarding optical properties, not only linear but also nonlinear properties are of great technological interest. For example, nonlinear frequency conversion, optical switching, and electro-optic modulation are readily used in today's telecommunication systems. Also, a compact and fairly efficient frequency-doubling crystal is used in every green "laser pointer" emitting at 532 nm wavelength.

Here, we review the current status of the field of nonlinear metamaterials. At this, we will concentrate on the near-infrared and visible spectral ranges. Nonlinear metamaterials operating at microwave frequencies will not be covered. Readers interested in this spectral range are referred to a recent colloquium.[2] Moreover, we will not consider the nonlinear optical properties of propagating surface plasmons and single particles. For these structures we point the reader to a recent review article.[3] Finally, we will not cover resonant gratings or photonic crystals except for cases where they have played a role in the development of metamaterial concepts.

This review is organized as follows. After a brief introduction to nonlinear optics, we discuss prospects as well as principal limitations of nonlinear metamaterials and address the origin of the nonlinear response of metamaterials. In Section 3.3, we review metal-based nonlinear metamaterials. Here, we concentrate on the second- and third-order nonlinear response of such metamaterials. Higher-order nonlinearities will not be considered. Section 3.4 is devoted to hybrid metamaterials in which metallic meta-atoms are combined with an efficient nonlinear dielectric medium. Finally, the emerging field of nonlinear dielectric metamaterials is covered in Section 3.5. Finally, in Section 3.6, we present an outlook of the field.

3.2. Basic Concepts

3.2.1. *Nonlinear optics*

The optical response of materials is described by their polarization. In everyday life, the polarization depends linearly on the applied optical field. When the optical field is sufficiently strong, however, the scaling is not linear anymore and we enter the regime of nonlinear optics.[4] The nonlinear optical effects can often be described by expanding the material polarization $P(t)$ as power series in the optical electric field $E(t)$ as

$$P(t) = \varepsilon_0[\chi^{(1)}E(t) + \chi^{(2)}E^2(t) + \chi^{(3)}E^3(t) + \cdots], \tag{3.1}$$

where ε_0 is the vacuum permittivity and $\chi^{(n)}$ are susceptibilities of various orders. For a time-harmonic field at frequency ω, it is evident that the nonlinear terms in Eq. (3.1) give rise to polarization components that oscillate at the harmonics of the incident field. The polarization acts as a source of radiation, providing output at the harmonic frequencies.

The second-order term in Eq. (3.1) is by far the most important for applications in frequency conversion. For example, second-harmonic generation (SHG) forms the basis for the aforementioned green "laser pointers". The simple SHG can be extended to sum- and difference-frequency generation by applying two different fields to the medium. Another application is to use a low-frequency field to modify the refractive index for an optical field, allowing electro-optic modulation of light. However, all these second-order (and other even-order) effects require additional considerations. Such effects can occur only in non-centrosymmetric materials within the electric-dipole approximation of the light-matter interaction. This symmetry rule is a major challenge in the search for new second-order materials.

In addition to third-harmonic generation, the third-order response gives rise to a response also at the original frequency. This can be interpreted as a nonlinear contribution to the refractive index of materials. Again, such responses can also be considered for fields at different frequencies or fields propagating in different directions, giving rise to

various four-wave-mixing schemes. An important application of this is to use a control beam to modify the refractive index of a signal beam, which provides the basis for all-optical switching.

The optical field and material polarization are vectorial quantities. In consequence, the susceptibilities are tensorial quantities. In addition, the optical responses depend on frequency. By expressing the optical field as a sum of several frequency components ω_n

$$\mathbf{E}(t) = \sum_n E(\omega_n) e^{-i\omega_n t}, \tag{3.2}$$

where the summation is over both positive and negative frequencies, we can then express the second-order response as

$$P_i(\omega_n) = \varepsilon_0 \sum_{j,k} \chi^{(2)}_{ijk}(\omega_n; \omega_m, \omega_l) E_j(\omega_m) E_k(\omega_l). \tag{3.3}$$

Here, the indices ijk refer to the field (polarization) components of the field and material polarization and the output frequency is $\omega_n = \omega_m + \omega_l$. The structure of the susceptibility tensor is determined by the symmetry group of the material.

3.2.2. *Conventional nonlinear optical materials*

It is instructive to start by reminding ourselves about the design strategies for conventional nonlinear materials, which will make it easier to understand the additional issues that need to be considered for metamaterials. For this purpose, organic nonlinear materials provide an ideal model case, because significant effort on molecular engineering has been taken to optimize the nonlinear responses of such materials.[5] This is also in the spirit of metamaterials where the individual nanoparticles can be considered as metamolecules or meta-atoms.

The nonlinear responses of individual molecules are defined by their hyperpolarizabilities of various orders. The hyperpolarizabilities are tensorial quantities defined in the molecular frame such that the structure of the tensors can be determined from the symmetry properties of the molecular structure.

By considering either classical or quantum theories of the nonlinear responses, we find that the hyperpolarizabilities can have resonances associated with transitions between the various energy levels of the molecules. These resonances can enhance the nonlinear responses whenever any of the interacting frequencies is close to any of the molecular transition frequencies. The resonance behavior is also related to the so-called Miller's rule, which states that the nonlinear response of conventional materials can be understood to a good accuracy from their linear optical responses at all the interacting frequencies.[6]

As a collection of molecules are brought together to form a macroscopic sample, the individual molecules will never have exactly the same orientation. The macroscopic nonlinear response, described by the molecular susceptibility tensors, is therefore obtained by projecting the response of individual molecules from the molecular to sample frame and then averaging this over the orientational distribution of the molecules.

It is also important to note that the electromagnetic field acting on a given molecule is not the same as the external field applied on the macroscopic sample. Instead, the local field acting on the molecule is obtained from the external field through multiplying by a local-field correction factor, often treated using the Lorentz local-field factor L.[4] In principle, this factor can be different for all the interacting frequencies $L(\omega)$. However, for the most common cases, where the nonlinear molecules are supported by a dielectric matrix (e.g, a polymer) and the molecular number density remains sufficiently low, the correction factors are essentially the same for all frequencies. For a second-order process, for example, which describes the interaction of three fields, the total local-field correction is thus L^3.

Finally, we also have to consider the growth of the nonlinear signal in the nonlinear material. Such phase-matching considerations are particularly important for frequency conversion processes, where the frequencies of the interacting fields can be widely different. In consequence, the fields experience different refractive indices and their phases accumulate differently in propagation. The incident and

generated fields therefore maintain a favorable phase relation only over a limited distance known as coherence length.[4]

Second-order (and other even-order) effects require additional considerations. Such effects can occur only in non-centrosymmetric materials within the electric-dipole approximation of the light-matter interaction. This symmetry rule, which has to be taken into account on both the molecular and sample levels, has been a major challenge in the search for new second-order materials. Nevertheless, the symmetry is necessarily broken at the surface of any material, giving rise to a non-vanishing second-order surface nonlinearity.[7] On the other hand, second-order effects become possible even in the bulk of centrosymmetric materials once higher-multipole (magnetic-dipole and electric-quadrupole) interactions are taken into account.[8]

3.2.3. *Prospects and limitations of nonlinear metamaterials*

What aspects of nonlinear optics can be improved by investigating metamaterials? After all, near 100% frequency conversion efficiency from second-order nonlinear effects or 180° nonlinear phase shifts (necessary for destructive interference, e.g., in a Mach-Zehnder interferometer) from third-order nonlinear effects in ordinary nonlinear optical materials are routine already.

First, consider an optical-waveguide architecture in silicon photonics. Silicon has centrosymmetry and thus exhibits zero second-order nonlinear susceptibility. To get a linear electro-optic effect or to generate second-harmonic for measuring pulse durations on chip or for second-order frequency mixing, one does need a material without centrosymmetry. Growing such ordinary crystals on silicon appears hopeless because of the lack of lattice matching. One can, e.g., add organic molecules which can be poled. By locally straining the silicon, the symmetry can be broken. Another option is to lithographically fabricate a non-centrosymmetric metamaterial at locations where it is needed. This appears like a reasonable idea because, after all, the entire chip is fabricated lithographically anyway. Absorption losses have to be reasonably low though.

Second, we can hope for larger effective nonlinear optical coefficients from metamaterials, which would allow for making nonlinear optical devices smaller or for reducing the necessary input powers. This aspect is relevant especially when aiming towards compact integrated optical-chip architectures. However, caution has to be exerted that these larger nonlinear coefficients do not come at the prize of largely increased losses. A meaningful quantity to address this aspect is the so-called figure of merit, which can be defined as the frequency conversion efficiency per absorption length or as the nonlinear phase shift per absorption length. We will come back to this aspect below.

The first aspect should be clear by itself. Regarding the second aspect, one might ask: On the basis of which physical effects can we hope for enhanced effective nonlinear optical coefficients as compared to advanced ordinary nonlinear optical crystals? Suppose we do not change the microscopic electronic properties of the constituent materials in their bulk when assembling them into a metamaterial. In this case, an effective enhancement of the optical nonlinearities can be due to (i) local-field enhancements (by metals), (ii) due to (dielectric) cavity effects, or (iii) due to surface effects. To get some intuition into

(i) local field-enhancement effects, let us consider a metamaterial based on a cubic lattice with lattice constant a. Light with intensity I impinges along one of the cubic axes. Suppose that every unit cell acts like a funnel for light in the sense that it concentrates photons by some means from the cross section of the unit cell, a^2, to a smaller area given by a^2/f^2, where $f^2 \geq 1$ is a dimensionless factor. Without absorption, the photon number is conserved. This means that the photon current density (photons per unit area and time), and hence the local intensity of light I in steady state increases according to $I \to f^2 I$. Correspondingly, the electric field amplitude of light increases like $E \to |f|E$. Obviously, we can identify the factor $|f|$ as the so-called field-enhancement factor. It is a critical parameter in any discussion on nonlinear optical metamaterials. Here we neglect for a while that f is space-dependent and might be unequal for the different components of the electric near field.

We now position a nonlinear material with macroscopic optical polarization, P, at the output of the funnel. Inspection of Eq. (1) shows that the n-th order contribution $\propto E^n$ in this expression is enhanced by a factor of $|f|^n$ and the corresponding intensity by a factor of $|f|^{2n}$. However, as this nonlinear contribution is now emerging from an area that is reduced with respect to that of the unit cell a^2 by a factor of f^2, the overall emerging power of the n-th order contribution increases only by a factor of $|f|^{2(n-1)}$, and the corresponding nonlinear electric field by the square root $|f|^{(n-1)}$. As this field is proportional to the associated n-th order nonlinear optical susceptibility, we can finally say that the effective nonlinear susceptibility of the metamaterial gets enhanced like $\chi^{(n)} \rightarrow f^{(n-1)}\chi^{(n)}$. For example, for a field-enhancement factor of $f = 10$, the second-order susceptibility for SHG or for optical rectification would get enhanced ten-fold, that for third-order processes like self-phase modulation or third-harmonic generation hundred-fold. The expected increase for a high-harmonic generation process of order $n = 11$ would be ten orders of magnitude — a truly gigantic factor.

For ideal phase matching, at a given incident power, one could equivalently make the nonlinear medium thinner. Alternatively, keeping the medium thickness fixed, one could go to much lower incident powers. Both would be extremely useful for practical applications — as pointed out above.

However, there is a catch. We have tacitly assumed that absorption is negligible. To get large field enhancements, metal nanostructures are attractive. Due to causality (the fact that we cannot change the past), which leads to the Kramers-Kronig relations in optics, we fundamentally cannot get field enhancements in passive media via resonances without absorption ("no real parts without imaginary parts"). One might still hope that absorption only slightly reduces the expected advantage. However, detailed quantitative theoretical model calculations for metal-based optical metamaterials have shown[9] that one cannot even get anywhere close to unity-efficiency frequency conversion for four-wave mixing — regardless of metamaterial thickness and metal filling fraction within. Likewise,[9] one

cannot get even close to a 180° phase shift from self-phase modulation. Both examples avoid further problems from possibly insufficient phase matching between the fundamental wave and the nonlinear wave. These findings mean that it is completely hopeless to think of nonlinear optical metamaterials in terms of potential replacements of existing ordinary nonlinear optical materials. Unfortunately, most of the examples that we shall cover below fall into this category. The situation is different when thinking in terms of integrating metamaterials into optical chips, as already discussed above.

(ii) Effective field-enhancements can also be obtained in dielectric cavities, for which the loss arguments outlined above do not apply. One can imagine artificial materials composed of many internal cavities with high quality factors. The Bragg resonance in dielectric photonic crystals behaves much like that. Near a photonic band edge, the group velocity of light can be slowed down, leading to enhanced light-matter interactions and enhanced effective nonlinear optical coefficients. We shall not cover nonlinear photonic crystals in this Chapter and rather refer the reader to Ref. 10. A related situation occurs when using one- or two-dimensional grating couplers on top of constituent materials. Just like in a distributed-feedback laser, the grating effectively acts like a cavity, thereby increasing the effective optical nonlinearity. Examples like that shall be presented below.

(iii) For surface effects to become comparable to bulk effects, the metamaterial must essentially be mainly composed of surfaces between different constituent materials. This means that the metamaterial lattice constant must be on the scale of one nanometer or just a few nanometers. We shall cover examples like that in Section 3.5.1.

3.2.4. *Nonlinear response of metamaterials*

Metamaterials usually consist of two- or three-dimensional arrays of nanoparticles. Early works on metamaterials were mainly based on metal nanoparticles,[1] whereas high-index dielectric (in particular, silicon) materials are receiving increasing present attention.[11,12] In

both cases, the particles support morphological (Mie-type) electromagnetic resonances, which depend on the size and shape of the particles as well as their environment. Such resonances are associated with strong local fields, which are expected to be beneficial for enhancing optical interactions, as explained in the previous section. For metals, the resonances are also referred to as localized surface-plasmon resonances.

Due to the character of the morphological resonances of nanoparticles, the design strategies for nonlinear metamaterials are very different from those for conventional materials. First, the intrinsic nonlinearities of the most common materials are spectrally flat in the usual visible and near-infrared spectral regions. We can therefore not take advantage of the resonances in the material response in order to optimize the nonlinear responses of metamaterials. On the other hand, we can treat the material nonlinearities by the respective susceptibilities, because metamaterials are already on the level beyond the atomic or molecular responses.

We therefore need to focus on optimizing the strong local fields supported by the morphological resonances. However, these local fields cannot be treated in terms of the simple Lorentz local-field factors. Instead, one needs to take into account that the local-field factors are space-dependent, i.e., they are different at different locations $\mathbf{r}$ around the nanoparticles. In addition, the factors are tensorial, i.e., they take into account the possibility that the local field contains polarization components not present in the applied field. The local field and the applied field are thus related by

$$E_{loc,i}(\omega, \mathbf{r}) = \sum_j L_{ij}(\omega, \mathbf{r}) E_j(\omega), \tag{3.4}$$

where i and j refer to polarization components of the fields.

Once the local fields at all the interacting frequencies are taken into account, we find that the effective susceptibility of an individual nanoparticle for a second-order process, for example, is[13]

$$\boldsymbol{\chi}(\omega_n; \omega_m, \omega_l) = \int \mathbf{L}(\omega_n, \mathbf{r}) \mathbf{L}(\omega_m, \mathbf{r}) \mathbf{L}(\omega_l, \mathbf{r}) \boldsymbol{\chi}^{\text{material}} d\mathbf{r}, \tag{3.5}$$

where tensorial notation has been used.

The most common materials (gold, silver, silicon) used to fabricate metamaterials are centrosymmetric with no dipolar second-order bulk nonlinearity. In addition, for the metals, the strongest hot spots occur at the metal-dielectric interfaces. We can therefore expect that the nonlinearity of an individual nanoparticle originates from the local surface response that needs to be integrated over the shape of the nanoparticle. The symmetry of the surface nonlinearity is such that equal local responses from opposite sides of a nanoparticle tend to be out of phase, i.e., they cannot radiate into far field. One therefore needs to tailor the local-field distributions in such a way that they are in some sense asymmetric.

If the second-order nonlinearity originates from the surface response, how can we ever expect it to give rise to an appreciable effective response? The effective bulk susceptibility obtained from the surface susceptibility is $\chi^{\text{bulk}} = \chi^{\text{surface}}/d$, where d is the thickness of the interfacial layer between the two materials. Here, the value of $d = 1$ nm is certainly a conservative estimate. The experimental values for the dominant components of the surface susceptibility of the glass-air and gold-air interfaces are on the order of 10^{-20} m^2/V[14] and 10^{-16} m^2/V,[15] respectively. We then find that the respective bulk susceptibilities would be on the order of 10 pm/V and 10^5 pm/V. The former value is comparable to the best conventional crystals, whereas the latter value is much higher. It is therefore not completely unreasonable to expect that well-designed metamaterial structures could be built up to give rise to appreciable nonlinear responses.

Higher-multipole (magnetic and quadrupole) effects are another concept that can play an important role in the nonlinear responses of metamaterials. However, these terms can mean very different things depending on the context. For example, second-order effects become allowed in the bulk of centrosymmetric materials due to magnetic and quadrupole effects in the atomic-level optical response.[8] Although such effects are expected to play a minor role compared to the dipolar surface response, their role is not yet fully understood in the effective response of metamaterials. On the other hand, the Mie theory

describes the responses of nanoparticles in terms of (effective) electric and magnetic multipoles of various orders even when the atomic response is strictly dipolar. A particular example here is the fact that the fundamental morphological resonance of silicon spheres has magnetic-dipole character.[11] Furthermore, the magnetic resonances of split-ring resonators are prominent in giving rise to negative permeability and refractive index of metamaterials.[1]

Metamaterials with effective properties consist of spatial arrangements of individual nanoparticles. Their properties therefore also depend on electromagnetic coupling between the individual particles. The near-field coupling between the two particles in a dimer, for example, can give rise to very strong hot spots in the gap between the two particles,[16] which could further boost the nonlinear responses. On the other hand, the coupling can also have a detrimental effect on the quality of the resonances of the individual particles.[17] In the most common cases, the spatial arrangement of the individual particles (or equivalently dimers or oligomers) is periodic. The typical periods of a few 100 nm are smaller than the optical wavelengths used in experiments, but not really so much smaller. This opens up the possibility that the resonances of the individual particles are modified by the electromagnetic resonances of the periodic lattice.[18,19] Such surface-lattice resonances can also have either a beneficial or detrimental effect on the quality of the resonances.

Finally, metamaterials also allow the effective optical parameters of the material to be tailored. Of particular interest here is the possibility of tailoring the refractive index at the interacting wavelengths. In particular, zero index or negative index could be used to advantage in novel phase-matching schemes.

3.3. Metal-Based Nonlinear Metamaterials

3.3.1. *Second-order nonlinearities*

A large majority of the research on nonlinear metamaterials has relied on the second-order response of metal-based metamaterials. There are two main reasons for this. First, the whole research field

of metamaterials was initiated because of the unique optical properties of metal nanostructures. Second, SHG is perhaps the simplest nonlinear process of all, which is very easy to implement experimentally. In addition, the designer aspects of metamaterials can be nicely connected to the rather stringent symmetry requirements for second-order effects. Such effects have therefore provided a convenient platform to test how the fundamental considerations of nonlinear materials should be modified for metamaterials.

As already mentioned, even centrosymmetric materials, including gold and silver, have a dipolar second-order surface nonlinearity. Coupling to this part of the nonlinearity requires that at least one of the interacting optical beams has a field component along the surface normal. In the plane-wave approximation, this requires that the beam be applied at oblique angle of incidence on the sample. In order to avoid such conventional surface effects, most experiments on second-order metamaterials have relied on SHG at normal (or near normal) incidence. Here, the beam at the fundamental frequency, typically at a near-infrared wavelength, is applied on the sample and the transmitted SHG signal is detected (see Fig. 3.1).

In this geometry, the symmetry rule for second-order response is fulfilled as long as the sample appears non-centrosymmetric at normal incidence. Such geometry therefore allows the role of the designer features of the metamaterial sample to the second-order response to

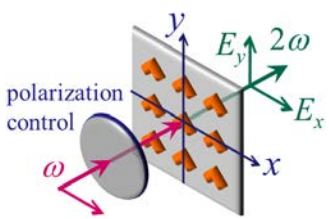


Fig. 3.1. Schematic representation of the experimental setup for second-harmonic generation from metamaterials. The field at the fundamental frequency ω is derived from a femtosecond laser at a near-infrared wavelength. The second-harmonic signal at 2ω is usually detected in the transmitted direction. Various polarization components of the fields are used for excitation and detection. A very similar setup could be used for third-harmonic generation.

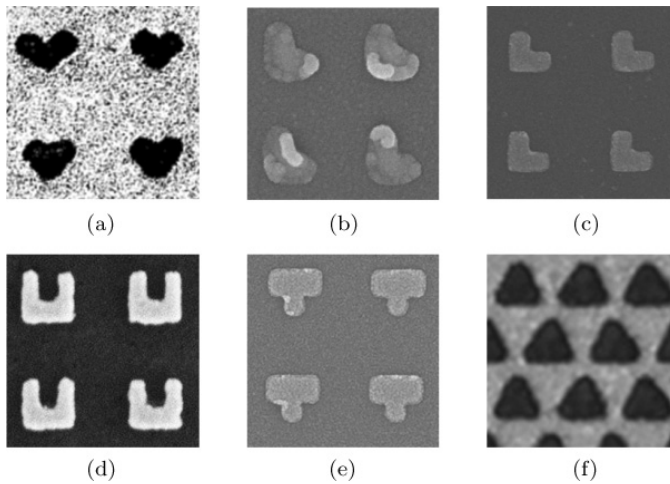


Fig. 3.2. Basic shapes of nanoparticles for SHG-active metamaterials. L-shaped (or equivalently V-shaped) particles of different generations (a–c), showing the evolution in sample quality. (a) Year 1997 (adapted with permission from Ref. [22], copyright 1999 Springer-Verlag), (b) Year 2004 (adapted with permission from Ref. [28], copyright 2008 Optical Society of America), and (c) Year 2011 (adapted with permission from Ref. [30], copyright 2011 Optical Society of America). Other shapes used for SHG (d–f). (d) U-shaped split-ring resonators (adapted with permission from Ref. [17], copyright 2011 American Physical Society), (e) T-shaped particles (adapted with permission from Ref. [37], copyright 2015 American Chemical Society), and (f) triangular nanoholes in metal film (adapted with permission from Ref. [45], copyright 2014 American Physical Society).

be investigated. Typical samples have consisted of gold nanoparticles with lateral dimensions of a 100–200 nm, lattice period of a few 100 nm, and gold thickness of 20–50 nm (see Fig. 3.2).

After the early work on SHG from rough metal surfaces,[20,21] the first example of a metal nanostructure designed for second-order nonlinear response consisted of an array of L-shaped silver nanoparticles (see Figure 3.2a), which was used for SHG autocorrelation measurements to determine the plasmon dephasing time in the particles.[22] Similar structures made of gold (see Figs. 3.2(b,c)) were then used for more detailed investigations of their SHG response.[23] As expected, the SHG response was found to be strongest when the fundamental wavelength was close to the plasmonic resonance of the particles. In

addition, the response was also found to depend on the ordering of the particles in the array.

The connection to metamaterials was made by studies of the SHG response of U-shaped split-ring resonators (SRR, see Fig. 3.2(d)), which have the same overall symmetry as the L-shaped particles.[24] The resonances of SRRs can be classified as having either electric or magnetic character. The first results suggested that the magnetic resonances are favorable for SHG. However, the SHG response from U-shaped nanoholes in a gold film was found to be comparable.[25] For such complementary structures, the roles of electric and magnetic resonances are interchanged, and the question about the importance of the magnetic resonances for the SHG response was left open. SHG from other types of nanohole arrays has also been investigated.[26,27]

The early investigations showed that the expected symmetry rules of SHG were fulfilled only approximately but not completely. These results were explained by the breaking of the sample symmetry from ideal by the deviation of the particle shape from the design and by the presence of unavoidable fabrication-related defects in the particles (see Fig. 3.2(b)).[28] Interestingly, such effects can be interpreted in terms of effective higher-multipole (magnetic and quadrupole) effects to the effective second-order nonlinearity of the structure.[29] Interference between the various multipoles can be used to control the directional properties of radiation,[12] and this was thus observed in the context of nonlinear optics relatively early. A few years later, samples of much higher quality (see Fig. 3.2(c)) were found to fulfill the expected symmetry rules of SHG at much higher precision, thereby reaching the dipole limit in the effective response.[30]

An important question, of course, is the role of resonance enhancement in the nonlinear response. This becomes evident in the comparison of the early, low-quality (see Fig. 3.2(b)) samples to newer, high-quality (see Fig. 3.2(c)) samples.[30] For the latter, the plasmon resonances exhibit much less inhomogeneus broadening, thereby improving the quality of the plasmon resonance and the SHG yield by one order of magnitude.

The role of plasmon resonances has been addressed more carefully by studying the strength of SHG as the fundamental or second-harmonic wavelength is tuned over the plasmon resonance of the structure. The initial experiments in this direction suggested that a resonance at the fundamental wavelength is clearly beneficial for the SHG response but that a resonance at the second-harmonic wavelength acts as a loss mechanism that just absorbs the generated SHG light.[31] The nonlinear response could also be increased be increasing the density of the metamolecules in the array. However, an attempt towards this was ultimately limited by the broadening of the plasmon resonance for sufficiently high metamolecular densities due to interparticle coupling effects.[17]

The role of the resonance at the SHG wavelength has recently been revisited, first in the context of the response of individual metal dimers, where both theory[32] and experiments[33] suggested that even such resonances can be beneficial for SHG. These studies were then extended to metasurfaces. First, the response from an array of aluminum nanobars was shown to be enhanced by the resonance at the second-harmonic wavelength.[34] However, here the origin of the SHG response is not completely clear, because the nanobars are ideally centrosymmetric. At about the same time, an array of V-shaped gold nanoparticles demonstrated exactly the same effect.[35]

There are a few additional items that need to be considered when discussing the resonance enhancement of the second-order response. As already mentioned, the early results on U-shaped SRRs suggested that their magnetic resonance is favorable for SHG. This result was emphasized by comparing the responses of SRRs with T-shaped gold particles, where the latter do not support the circulating currents required for the magnetic resonance.[36] On the other hand, the results from U-shaped nanoholes contradicted this result.[25] More recently, the SHG response of L-shaped particles was compared to that of different T-shaped particles (see Figure 3.2e).[37] The overall properties of the L-shape are equivalent to those of the SRRs, i.e., they also support magnetic resonances. Nevertheless, the strongest responses from the T-shaped particles were comparable or even higher than

those from the L-shaped particles. These results also showed that the plasmonic resonances alone are not sufficient for a strong response. At the same time, one needs to consider carefully the geometry of the particles and the detailed local-field distributions supported by the geometry, as also discussed in the early work on T-shaped nanodimers.[38]

The nonlinear response of conventional materials can be predicted quite well by the Miller's rule, which links the linear optical properties of a material at the interacting frequencies to the nonlinear properties.[6] Miller's rule has also been discussed in the context of nonlinear metamaterials. At this time, it seems that the linear and nonlinear properties are linked through the plasmonic resonances[35] but not as universally as for conventional materials.[39] In particular, there is evidence that the constant that links the linear and nonlinear properties depends on the type of a metamaterial, whereas the constant is essentially universal for conventional materials.

The resonance properties of metamaterials can be tailored through electromagnetic interactions between the nanoparticles, which can give rise to, e.g., Fano-type lineshapes of the resonances.[40] Another possibility is to rely on lattice interactions between the particles in the array. Such interactions can give rise to very sharp spectral features whenever a diffraction order opens in the ambient material or the substrate of the sample, as demonstrated in the linear response of plasmonic arrays.[18,19] For SHG, such surface lattice resonances can come into play at the fundamental or the SHG wavelength. Considering the facts that the period of typical arrays is on the order of 500 nm and the fundamental wavelength is in the range of 800–1100 nm, diffraction orders at the SHG wavelength can easily be opened in the substrate even when no orders propagate in free space.

The lattice interactions have influenced the SHG response of metamaterials since the very early work,[23] where the response was found to depend on the ordering of the particles in the array. However, the lattice interactions are becoming more prominent only more recently. The role of the orientational distribution

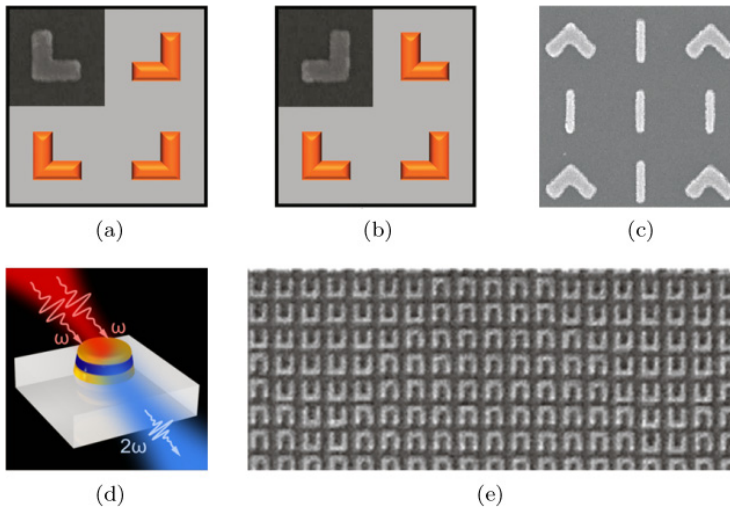


Fig. 3.3. Advanced structures for SHG-active metamaterials. Variations in the ordering of L-shaped particles (a,b adapted with permission from Ref. [41], copyright 2012 American Chemical Society), combination of SHG-active L-shaped particles with SHG-passive nanobars (c, adapted with permission from Ref. [42], copyright 2013 American Physical Society), gold disks separated by a dielectric layer (d, adapted with permission from Ref. [47], copyright 2015 American Chemical Society), and a grating where the phase of the SHG wavelets is reversed by the orientation of split-ring resonantors (e, adapted with permission from Ref. [49], copyright 2015 Nature Publishing Group).

of the metamolecules in the array was investigated for L-shaped metamolecules.[41] The result was that, for a similar orientational distribution, the response can be either enhanced or suppressed compared to a reference sample, depending on detailed ordering of the particles in the array (see Figs. 3.3(a,b)). Here, the variations in the ordering also changed the size of the unit cell of the array, thus opening diffraction orders either in one or two directions. Surprisingly, the lattice interactions were detrimental for the quality of the plasmon resonances in one case and beneficial in the other. In a different context, the response of L-shaped particles was enhanced by the presence of centrosymmetric nanobars with no SHG response as such (see Fig. 3.3(c)).[42] The enhancement was well explained by the lattice interactions between the two types of particles.

Metamaterials also provide a convenient way to investigate the fundamental symmetry rules of nonlinear interactions. For example, nanobars are centrosymmetric and should not give rise to SHG at normal incidence and direct transmission. The reason for this is that the local responses from the symmetrically opposite locations cancel in the far field. However, the local responses can add up to other directions, as first demonstrated early in Ref. [43] and discussed theoretically recently in Ref. [44].

A more subtle issue is that when the nonlinear responses are analyzed in the basis of circular polarizations for the fundamental and SHG frequencies, the in-plane orientation of the sample gives rise to a topological phase in the SHG response (see Fig. 3.2(f)),[45,46] which could be used to control the phase of the emitted SHG field (see also Section 3.3.2). In addition, the circular basis gives rise to new types of selection rules for the allowed SHG signals, which would not be easily revealed by an analysis in the Cartesian basis. In this context, metasurfaces have made it easy to address fundamental symmetry principles, because the analysis is not complicated by propagation effects within the sample.

The role of electric and magnetic resonances to the SHG response has also been investigated for dimers consisting of two gold nanodisks separated by a dielectric layer.[47] Such metamolecules are centrosymmetric. However, when a metasurface of such dimers is fabricated, the symmetry is broken in the direction of the surface normal (as in conventional surface nonlinear optics), and the SHG response can be accessed at oblique angles of incidence. The results suggest that, for these particular samples, the SHG response is the strongest when the fundamental wavelength is tuned close to the magnetic resonance of the structure.

Very recent developments with regard to nonlinear studies of metasurfaces have been closely connected to the more general developments in the field of metasurfaces. A particular advance was the development of semi-continuos phase control of the optical response of metasurfaces, which gives rise to the generalized Snell's law.[48] From a very general viewpoint, the phase control, of course, is

associated with holographic principles and diffractive surfaces. This was first demonstrated by flipping the phase of the SHG signal by two opposite orientations of U-shaped SRRs and constructing basic diffractive elements for SHG light (see Fig. 3.3(e)).[49] The same principle was further extended for more general shaping of the SHG beam.[50]

Most of the concepts discussed above are based on metasurfaces that are carefully designed to appear non-centrosymmetric at normal incidence. However, there are recent examples of more random structures, consisting of densely-packed metal nanorods. Such structures exhibit hyperbolic dispersion in their optical properties and have also been shown to possess interesting second-order nonlinear properties.[51]

3.3.2. *Third-order nonlinearities*

The promise of metamaterials for nonlinear optics is related to the enhanced local fields supported by the material structures. In this development, third-order effects have been particularly important, because they are not constrained by the non-centrosymmetry requirement and can therefore occur in any material. The most important third-order effects are third-harmonic generation (THG), intensity-dependent refractive index and absorption, as well as various four-wave-mixing (FWM) processes.

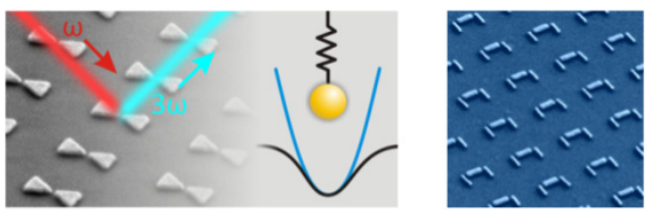
The historical development also here goes back to the observations of surface enhancement of these processes. Surface-plasmon-enhanced THG from silver films was observed in 1996.[52] This work was later extended to enhanced THG from silver nano-island films, where the key differences between SHG and THG, arising from symmetry considerations, were also highlighted.[53] Gold-silica nanocomposite films near the percolation threshold have been used to obtain large third-order susceptibility for FWM.[54]

Bulk-type nanocomposites have played an important role in the development of materials with tailored nonlinear refractive index or nonlinear absorption. Here, the interest has been in three main classes. The Maxwell-Garnett model is applicable to cases where one

can make a clear distinction between a host material and inclusions with low fill fraction.[55] The Bruggeman model describes composites where the constituents are treated equally, and is thus applicable to arbitrary fill fractions.[56] The multilayer structures typically have layer thicknesses well below the wavelength.[57] Note that these structures fulfill the definition of metamaterials quite well, perhaps even better than some of the arrays used in more contemporary studies.

Experimentally, these effects have been mainly studied using metal nanoparticles embedded in a dielectric host. For example, the magnitude of the nonlinear absorption coefficient was controlled by varying the fill fraction of gold nanoparticles in a nonlinear host material.[58] Later on, even sign reversal of nonlinear absorption was observed in gold-silica composites with high fill fractions,[59] and the local-field enhancement was extended to other types of nonlinear processes.[60,61]

With the development of metamaterial concepts, third-order effects have received new attention. THG was first observed from arrays of split-ring resonators.[36] The role of resonance enhancement in THG was investigated for bowtie dimer nanoantennas with a nanogap between the two parts (see Fig. 3.4(a)).[62] It was shown that the THG response can be predicted well on the basis of the linear response of the structure, with deviations occurring only for gap sizes below 20 nm. A similar study was then extended for nanorod antennas.[63] In both cases, the results were interpreted in terms of



(a) (b)

Fig. 3.4. Metamaterials used for THG experiments. (a) Bowtie nanodimers (adapted with permission from Ref. [62], copyright 2012 American Chemical Society). (b) Dolmen-shaped trimers with Fano resonances (adapted with permission from Ref. [64], copyright 2014 American Chemical Society).

a classical anharmonic oscillator model. THG has also been investigated in more complex structures where the resonances exhibit Fano-type characteristics (see Fig. 3.4(b)).[64]

Although third-order effects are allowed in all materials, their properties do also depend on the symmetry of the sample. The topological phase arising from the interaction between circular polarizations and anisotropic structures was used to achieve a continuous control of the phase of the THG signal.[46] By varying the orientation of the metamolecules in different parts of the sample, a semi-continuous change of the phase was achieved, allowing the direction of the THG signal to be controlled.

FWM was first investigated in a designed surface structure by comparing the response from a continuous gold film and a film with nanostructured grating[65] (see Fig. 3.5(a)). The grating was found to enhance the FWM process by more than one order of magnitude. Even higher enhancement of 2000 was achieved by designing the grating grooves to support cavity modes, which were further coupled by surface waves (see Fig. 3.5(b)).[66] Enhancements reaching six orders of magnitude have been predicted theoretically from nanogratings[67] and a systematic design procedure to optimize surfaces consisting of nanocavities has been presented.[68] FWM from a nanostructured metal film has also been used to achieve negative refraction, which arises quite naturally from the phase matching considerations of the process.[69] The more recent concepts of phase gradients along the surface have also been demonstrated for FWM, giving rise to anomalous phase matching and nonlinear lensing (see Fig. 3.5(c)).[70]

Third-order effects also form the basis for all-optical switching where a pump pulse is used to modify the optical properties experienced by a signal pulse. For instance, a modulation of 60% was observed by using a pump pulse to vary the coupling of the probe to propagating surface plasmons in a gold grating.[71] This required however relatively large pump fluences in the range of 60 mJ/cm^2. The switching efficiency could be improved by using plasmonic nanorod metamaterials (see Fig. 3.6(a)) with non-local response,[72] resulting in up to 80% modulation and fluence of 7 mJ/cm^2. A further

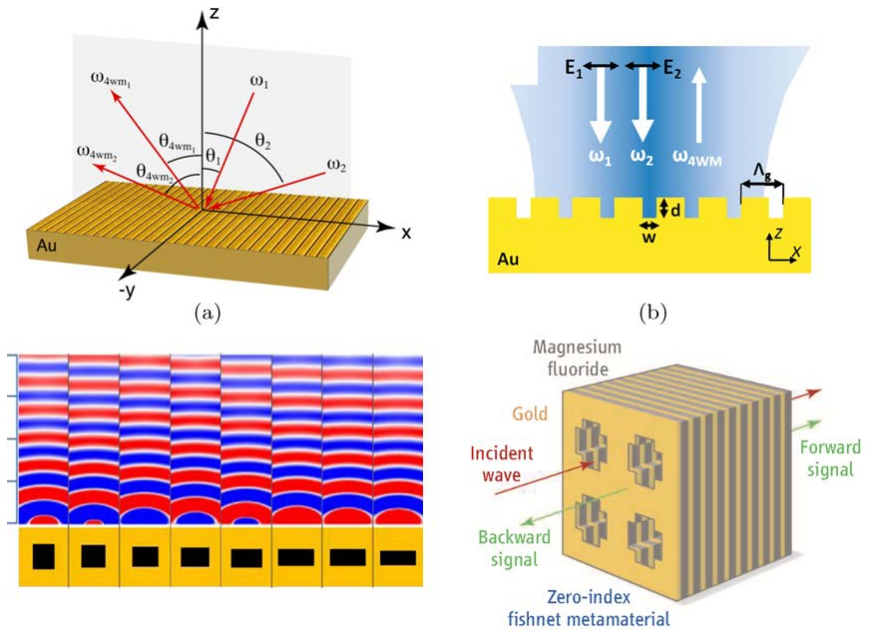
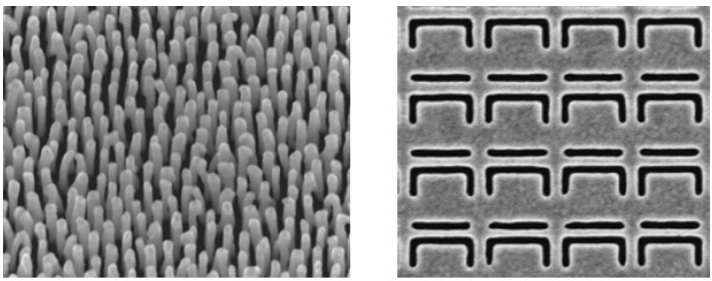


Fig. 3.5. Metal nanostructures used for various FWM processes. (a) A metal grating for surface-enhanced FWM (reprinted with permission from Ref. [65], copyright 2010 American Physical Society). (b) A metal grating whose grooves support cavity resonance, giving rise to even higher enhancement (adapted with permission from Ref. [66], copyright 2010 American Chemical Society). (c) Metal nanocavities whose shape is used to control the phase of the emitted FWM wavelets (adapted with permission from Ref. [70], copyright 2016 Nature Publishing Group). (d) A fishnet metamaterial structure[75] with zero refractive index, allowing simultaneous phase matching of the FWM signal in the forward and backward directions (adapted with permission from Ref. [76], copyright 2013 American Association for the Advancement of Science).

advance was achieved by using a gold film with specially designed holes (see Fig. 3.6(b)),[73] which allowed sub-100 fs switching speeds and 40% modulation at the estimated pump fluence of 270 μJ/cm^2 and relied on interband transitions of gold. Very recently, it has been shown that the nonlinear response of nanorod metamaterials can be tailored over a broad range only through its geometrical design parameters,[74] although the response time was not yet determined.



(a) (b)

Fig. 3.6. Metamaterials for all-optical switching. (a) A hyperbolic metamaterials consisting of plasmonic nanorods (adapted with permission from Ref. [72], copyright 2011 Nature Publishing Group). (b) A metal film with specially designed hole pattern (adapted with permission from Ref. [73], copyright 2011 John Wiley and Sons).

Most of the effects described above have been based on metasurfaces, i.e., structures that are optically thin. For thicker structures, one also needs to be concerned about phase matching, i.e., that the nonlinear signals from different parts of the sample add up in phase. This is an important issue for frequency conversion where the waves at the interacting frequencies can experience very different indices of refraction. However, metamaterials also allow their linear properties to be tailored. In particular, if the indices are zero, the phase does not accumulate in propagation, allowing efficient nonlinear interactions. This has been demonstrated in a fishnet metamaterial where near-zero index was achieved in a narrow wavelength range (see Fig. 3.5(d)).[75] The experiment therefore relied on nearly-degenerate FWM in order to ensure that all wavelengths were in the zero-index band.

3.4. Hybrid Nonlinear Metamaterials

We have seen in the previous section that one can tailor the intrinsic nonlinear optical properties of metallic metamaterials by controlling the geometry as well as the arrangement of the meta-atoms. A promising strategy to further enhance the nonlinear response is to combine the metallic meta-atoms with an efficient nonlinear dielectric

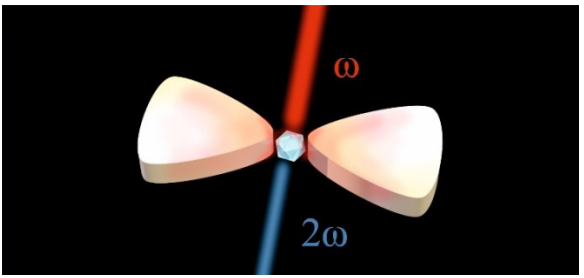


Fig. 3.7. Scheme of a nonlinear hybrid meta-atom consisting of a metallic antenna and a nonlinear dielectric particle.

medium into a hybrid nonlinear metamaterial (see Figure 3.7). Interestingly, this idea was already proposed in John Pendry's seminal paper on the split-ring resonator.[77]

Hybrid nonlinear metamaterials potentially offer several benefits: (i) The resonant excitation of plasmonic modes by the pump beam can give rise to a strong local-field enhancement in the vicinity of the metallic meta-atoms (see above). By either embedding the metallic meta-atoms in the nonlinear dielectric medium or by depositing nonlinear dielectric nanoparticles in the "hot spots", one can hope to significantly enhance the interaction of the nonlinear dielectric material with the light field. (ii) Another interesting aspect of hybrid nonlinear metamaterials is that the electromagnetic near-field in the vicinity of the meta-atoms can exhibit vector components that are not present in the far-field, e.g., a component of the electric field in the direction of the wave vector of the incident pump beam. Hence, hybrid nonlinear metamaterials have the potential to access components of the nonlinear susceptibility tensor of the dielectric nonlinear material which cannot be excited in conventional far-field experiments. (iii) The metallic meta-atoms can not only be used to concentrate the incident pump light in small volumes at the position of the nonlinear dielectric medium but can also help to control the coupling of the generated nonlinear signal to the far field. To combine both effects in a single metamaterial, one has to design metallic meta-atoms such that they exhibit plasmonic modes resonant to both the incoming pump light and the generated nonlinear signal.

In the following, we will discuss selected examples of rationally designed hybrid nonlinear metamaterials. We will not cover rough metal surfaces or random ensembles of metal nanoparticles embedded in a nonlinear dielectric medium or mixed with nonlinear dielectric nanoparticles.

3.4.1. *Nonlinear frequency conversion*

3.4.1.1. *Hybrid metamaterials containing bulk dielectric nonlinear media*

Gallium arsenide (GaAs) is a non-centrosymmetric semiconductor that is known to exhibit a large second-order nonlinear optical susceptibility. This property makes GaAs an attractive choice as the nonlinear dielectric medium in hybrid nonlinear metamaterials designed for near- and mid-IR pump frequencies. For instance, W. Fan *et al.* used optical interference lithography in combination with several etching steps to fabricate arrays of coaxial holes in a 70-nm thick gold film filled with 140-nm thick GaAs annuli.[78] Excitation of the structures with femtosecond mid-IR-pulses tuned close to the cut-off frequency of the coaxial waveguide mode resulted in a strong SHG signal. The strength of the SHG signal was attributed to a combination of the local field enhancement in the coaxial holes and the large nonlinear response of GaAs. In contrast, no SHG signal was observed when GaAs was replaced by silicon, which exhibits a vanishing bulk second-order nonlinear optical susceptibility. This control experiment indicated that the intrinsic nonlinearity of the gold did not significantly contribute to the SHG signal of the GaAs sample. Subsequent experiments performed by the same group on nanoaperture arrays in gold films filled with GaAs posts led to even larger SHG signals.[79]

As stated above, the electromagnetic near-field of metallic nanostructures can be employed to access components of the nonlinear susceptibility tensor of a dielectric nonlinear material which cannot be excited in conventional far-field experiments. F.B.P. Niesler *et al.* demonstrated this effect in a series of SHG experiments on split ring resonator arrays deposited on a (110) GaAs wafer.[80] By varying the

orientation of the split-ring resonators relative to the crystallographic axes of the GaAs wafer and by analysing the polarization of the generated second-harmonic signal, they could identify a second-harmonic signal that originated specifically from the interplay of the electromagnetic near-field component of the split ring resonators normal to the wafer surface and an element of the bulk GaAs second-order nonlinear susceptibility tensor, which can be only accessed by this field component.

A promising strategy for nonlinear optics in the mid-IR spectral range is to combine metallic metasurfaces with multi-quantum well (MQW) semiconductor heterostructures. By properly designing the MQW structure, one can engineer the intersubband transitions and achieve an extremely large second-order nonlinear response.[81] However, there is a catch. In order to benefit from this large nonlinearity, all light fields must be polarized perpendicular to the semiconductor layers. Hence, one cannot simply access the strong nonlinear response of the MQW semiconductor heterostructures from the far-field. In order to overcome this problem, Lee *et al.* sandwiched a 400-nm thick MQW structure between a gold mirror and an array of asymmetric gold nanocrosses.[82] Excitation of the plasmonic mode of the long arm of the nanocrosses with a quantum cascade laser operating at a wavelength of 8 μm under normal incidence induced a strong local electric field in the MWQ structure with the desired field orientation, i.e., perpendicular to the quantum well layers. The generated second-harmonic light, which was also polarized perpendicular to the quantum-well layers, coupled to the plasmonic mode of the short arm of the nanocrosses and was efficiently radiated in backwards direction to the far-field. A control experiment on the MWQ structure without nanocrosses resulted in no significant nonlinear response for normal incidence of the pump beam. Later it was shown that the coupling of metamaterial resonators to the nonlinearity of MWQ structures can be used to create second-harmonic phased-array sources.[83]

Electric-field-induced second-harmonic generation (EFISH) is a nonlinear process that results from mixing of an optical pump beam with a dc-electric field in a third-order nonlinear medium. This effect

is potentially very useful as it allows to electrically control the nonlinear signal strength. However, the corresponding third-order nonlinear process is usually very weak in bulk materials and requires large voltages (typically kilovolts). To overcome this problem, Cai *et al.* combined a plasmonic metasurface with a cavity filled with a nonlinear polymer.[84] By pumping this device with a near-infrared laser, they observed a voltage dependent second-harmonic signal with a relative modulation of the nonlinear conversion efficiency as high as 7% per volt.

3.4.1.2. *Hybrid metamaterials containing nonlinear nanoparticles*

As a consequence of an increased demand for subwavelength coherent light sources, an interest for subwavelength frequency conversion schemes has also emerged. In principle, one could simply use nanometer sized nonlinear dielectric crystals for these applications. Practically, however, nonlinear frequency conversion based on individual nonlinear dielectric nanocrystals is usually not very efficient.

A strategy to overcome the poor conversion efficiency of individual nonlinear dielectric nanocrystals could be to place the nanocrystals in the hot spots of plasmonic nanostructures. By doing so, one could aim at taking advantage of both the plasmonic field-enhancement effect and the large nonlinear optical susceptibility of the dielectric nanoparticles. Following this idea, Pu *et al.* fabricated core-shell nanocavities consisting of 100-nm barium titanate ($BaTiO_3$) nanoparticles enclosed by a thin gold layer.[85] Compared to bare $BaTiO_3$ nanoparticles, the core-shell nanocavities showed an enhancement in the second-harmonic radiation power of more than 500 times. Control experiments on 200-nm thick solid gold nanoparticles only gave rise to a weak, spectrally broad two-photon photoluminescence signal. This observation led the authors to the conclusion that the SHG enhancement of the core-shell nanocavities was not dominated by the intrinsic nonlinearity of the gold. However, one could argue on this occasion that the proper reference should be a core-shell nanocavity with a linear dielectric nanoparticle of the same size.

In a recent experimental study on THG from gap nanoantennas containing indium-tin-oxide (ITO) nanoparticles, the authors reported on an upconversion enhancement of the hybrid nanostructure of up to 10^6-fold compared with an isolated ITO nanoparticle.[86] However, this enhancement factor has to be interpreted with great caution as already a bare gap nanoantenna gave rise to a THG signal that was approximately five orders of magnitude larger than that of the isolated ITO nanoparticle (Supplement of Ref. [86]). A second study on THG from gap nanoantennas containing ITO nanoparticles casted serious doubts on the interpretation that the nonlinear upconversion enhancement was caused by the third-order susceptibility of the ITO nanoparticle.[87] By performing THG spectroscopy on different sized gap nanoantennas, Metzger *et al.* found strong indications that the nonlinear enhancement of the investigated hybrid plasmonic/dielectric nanoantennas was mainly related to changes in the linear optical properties of the gap nanoantenna resonances due to the presence of the ITO nanoparticles. Hence, the authors of the second study identified the gold nanoantennas as the dominant source of the THG emission. This interpretation is also supported by SHG experiments on different hybrid plasmonic-dielectric nanoantennas.[88] By using both linear and nonlinear dielectric nanoparticles, the authors could show that an increase of the SHG efficiency of plasmonic nanoantennas obtained by filling their feed gaps with a dielectric nanoparticle was independent of the nonlinear properties of the nanoparticles.

3.4.2. *Nonlinear switching and modulation*

The strong electromagnetic near-field of metallic nanostructures in combination with a third-order nonlinear medium can be also employed to increase the nonlinear refractive index. For instance, Fu *et al.* studied the optical nonlinearity of hybrid structures composed of cadmium telluride (CdTe) quantum dots and gold nanoparticle arrays using the Z-scan method.[89] They observed an 8-fold increase of the nonlinear refractive index of the hybrid structure compared to the bare CdTe quantum dots when the plasmonic resonance of the

Au surface was tuned to be in resonance with the exciton transition in the CdTe quantum dots.

The resonance frequency and the electromagnetic near-field distribution of the plasmonic modes of a given metallic nanostructure often react sensitively on minute changes in its dielectric environment. This feature is for instance employed in plasmonic sensor applications. In the context of nonlinear optics, one can use this effect in combination with the nonlinear refractive index of appropriate dielectric media to efficiently modulate and switch the optical properties of hybrid metamaterials. This has been demonstrated in a fishnet metamaterial based on silver and amorphous silicon.[89] Photo excitation of the amorphous Si layer at visible wavelength resulted in the modulation of negative refractive index in the near-IR. The switching speed was found to be limited by carrier relaxation in silicon to 58 ps.[90] A much faster, sub-ps, relaxation times were subsequently observed in similar structures.[91,92] Wurtz *et al.* investigated metallic nanoaperture array covered with a nonlinear polymer.[93] They observed a bistable behavior of the transmission properties which depended both on the wavelength and the intensity of the control light field. In the control experiments performed under identical illumination conditions on bare metallic nanoaperture arrays and on bare polymer films no optical nonlinearity was observed.

3.5. Dielectric-Based Metamaterials

3.5.1. *Second-order nonlinearities*

As discussed above, to obtain a non-zero bulk effective second-order nonlinear optical susceptibility, the metamaterial crystal must globally break inversion symmetry. At first sight, we can use the same approaches as for metals, i.e., we could, for example, consider a lattice of dielectric split-ring resonators or other motifs breaking centrosymmetry.

Before proceeding, we should like to specify better though what we mean by "dielectrics" within this section. Optical phonons and other resonances in undoped semiconductors or isolators can lead to

large positive and even to negative values of the electric permittivity. Negative values occur for frequencies of light above the corresponding eigenfrequency. In this regime, the physics of the dielectric is closely similar to what we have said about metals. One might hope that the unwanted effects of damping and dissipation are less pronounced for dielectrics in this frequency regime compared to metals. However, experiments have rather shown similar behavior to metals. For a fair comparison, one must, of course, compare the imaginary parts of the electric permittivity in the same frequency regime. For example, pronounced phonon resonances in silicon carbide (SiC) occur at around 10 μm wavelength. There, the imaginary parts of the permittivity are smaller than for metals in the visible regime, but they are comparable to those of good metals like gold at around 10 μm wavelength. The same holds true for doped semiconductors or isolators, where the plasma frequency of the free-electron gas is much lower than in metals. Because of the overall similarity to metals, we refer the reader to the above sections for this regime of dielectrics with negative electric permittivity.

At frequencies of light below the corresponding eigenfrequency of undoped semiconductors or isolators, the electric permittivity is positive. Upon nanostructuring of the material, large positive values allow for appreciable local-field enhancement effects and for localized Mie resonances. We will come back this possibility in the next section in the context of third-order optical nonlinearities. One should be aware though that these large positive values in the vicinity of a resonance are unavoidably connected to unwanted imaginary parts via the Kramers-Kronig relations. Thus, the above general arguments regarding possible nonlinear optical figures of merit in the bulk apply here, too.

Notably, all ordinary high-end commercially available nonlinear optical crystals pretty much avoid resonances and operate in the off-resonant regime where dispersion and hence absorption are very small. What good can metamaterials do for us within this regime?

We have seen that surfaces or interfaces between two different materials locally break inversion symmetry, allowing for, e.g., surface

second-harmonic generation. However, these effects involve only a small fraction of the atoms, and, hence, the effects are much weaker than those from the bulk.

ABCABC... nanolaminates composed of three different off-resonant dielectrics A, B, and C are an interesting concept. Here, each of the three dielectrics by itself shall exhibit centrosymmetry. Laminates are stacks of layers and have already been discussed by James Clerk Maxwell. Obviously, the ABC stacking globally breaks inversion symmetry, whereas an ABAB... stacking does not. It is intuitively clear right away that the number of atomic layers per metamaterial layer plays a crucial role: If the ABC layers are very thick, one has only few atomic interfaces for a given thickness and the effective nonlinearity is expected to be low. On the other end, the individual ABC layers just cannot be thinner than one atomic lattice constant of the constituents A, B, and C. Furthermore, if the layers are not atomically smooth, roughness will play an important role if the layers become very thin. On this basis, one intuitively expects an optimum in regard to the effective bulk second-order nonlinear optical susceptibility at some intermediate period or metamaterial lattice constant of the ABC stack.

Broadly speaking, ABC stacks of atomically thin layers can be realized in many different ways. They have even been realized long before the notion of metamaterials was born, e.g., by Langmuir-Blodgett techniques — even in the context of nonlinear optics. Recently, ABC nanolaminates have been realized using atomic-layer deposition (ALD). ALD is generally accepted as a CMOS compatible technology. This aspect has important implications for the use of such ABC laminate metamaterials in platforms like silicon photonics. Furthermore, ALD, unlike physical high-vacuum evaporation or other forms of epitaxy like molecular-beam epitaxy, leads to conformal deposition. This means that all (inner) surfaces are coated the same way regardless of their orientation. This aspect might allow for fabricating completely different metamaterial architectures in the future.

Recently, two groups published results on second-harmonic generation on ABC nanolaminate metamaterials made by ALD. One

group[94] used the three dielectrics A=aluminum oxide (Al_2O_3), B=oxide (TiO_2), and C=hafnium oxide (HfO_2) as ingredients (see Fig. 3.8), the other group[95] used A=TiO_2, B=Al_2O_3, and C=indium oxide (In_2O_3). Both groups excited the stack under oblique incidence of light and at a fundamental wavelength in the near infrared (800 nm[94] and 980 nm,[95] respectively) and detected the emerging second harmonic in the deep blue or green spectral region. Reference [95] reported a *zzz*-component of the second-order susceptibility tensor as large as 5 pm/V, more than ten times larger than what was reported in Ref. [94]. However, recent sample exchange among the two groups led to values that are significantly less apart.[96] Reference [94] carefully studied the dependence on the thickness of one ABC period and found an optimum at around 3 nm thickness, which

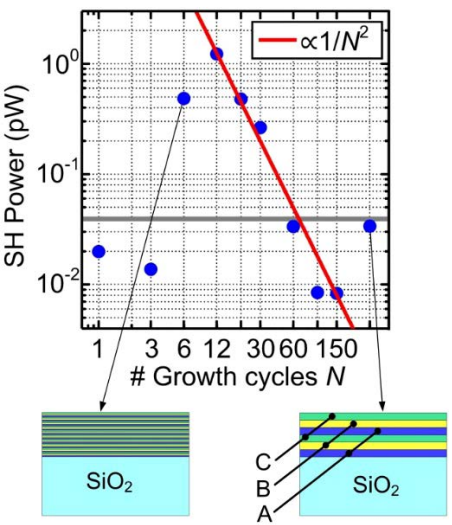


Fig. 3.8. Second-harmonic generation from ABC nanolaminate metamaterials. The ABC stacking breaks centro-symmetry. A=Al_2O_3, B=TiO_2, and C=HfO_2. The upper part shows the measured second-harmonic power versus the number of ALD growth cycles for each individual layer. The total thickness of the stack is kept constant. The maximum at around 12 layers corresponds to a thickness of one ABC laminate period of 2.7 nm. The maximum is due to a trade-off between the number of interfaces per thickness on the one hand and the quality of the interfaces on the other hand. The insets below illustrate two cases. (Reprinted with permission from Ref. [94], copyright 2015 AIP Publishing LLC.)

is more than hundred times smaller than the wavelength of both the fundamental and that of the second harmonic. In this regime, the effective-medium approximation is well justified. Correspondingly, the measured second-harmonic signal scales quadratically with the overall stack thickness, as expected for phase-matched second-harmonic generation from a bulk metamaterial.

Intuitively, the ABC stacks can be seen as a bulk metamaterial composed of densely packed interfaces, breaking centrosymmetry both locally and globally. Unfortunately, no theory has been worked out to date to explain these findings on a microscopic footing. Such theory would be highly desirable to be able to rationally optimize these structures. The situation is remotely similar to that investigated theoretically years ago for asymmetric semiconductor quantum-well systems.[97] There, the envelope approximation was used for the electronic wave function. This approximation does not appear appropriate though for the discussed ABC nanolaminate metamaterials.

3.5.2. *Third-order nonlinearities*

Gustav Mie showed more than a century ago that dielectric nanoparticles with positive electric permittivity exhibit Mie resonances just like metallic nanoparticles. It is well known that the Mie resonances exhibit alternating magnetic and electric character, with the fundamental (i.e., lowest frequency) resonance having a magnetic-dipole character. Clearly, one can pack these nanoparticles into arrays. However, for low refractive-index contrast, such arrays or crystals do not qualify as an effective medium because the lattice constant or period is comparable to the wavelength of light. For example,[1] the free-space wavelength of the first magnetic resonance of a sphere is given by the product of its diameter and its refractive index. Thus, for a refractive index of 3 and a period of twice the diameter, the free-space magnetic-resonance wavelength is only 50% larger than the metamaterial period. The situation is yet worse for the next electric resonance. The situation is slightly better for silicon within the visible spectral regime. For example, at 450 nm free-space wavelength, the

real part of the electric permittivity of silicon is around 20, leading to a real part of the refractive index between 4 and 5.

Reference [98] studied a square array of amorphous-silicon nanocylinders on a thick silica layer on silicon. The disks had a diameter of 500 nm, a height of 220 nm, and the lattice constant was 800 nm. Under these conditions, the fundamental magnetic-dipole resonance occurs at about 1240 nm free-space wavelength. The corresponding wavelength of the third-harmonic is 413 nm. For normal incidence of light with respect to the array, the authors found that the field localization at the magnetic resonance results in two orders of magnitude enhancement of the harmonic intensity with respect to an unstructured bulk silicon substrate.

By the same physics, the magnetic-dipole resonance of such metamaterials composed of silicon nanodisks also shows enhanced two-photon absorption, which is proportional to the imaginary part of the third-order nonlinear optical susceptibility. In degenerate pump-probe experiments using 65 fs pulses, a recent paper[98] shows an 80-fold enhancement with respect to bulk silicon, albeit on a comparably large thermal background. Herein, the center wavelength of the pump and the probe pulses derived from a Ti:sapphire laser oscillator has been tuned such that the corresponding photon energy is slightly below the band edge of amorphous silicon. Correspondingly, self-phase modulation, which is proportional to the real part of the third-order nonlinear optical susceptibility, is also enhanced by the magnetic resonance of the dielectric metamaterial. Standard z-scan experiments also yield consistent results.[99]

3.6. Outlook

Based on the broad body of work discussed in this review one can doubtlessly assert that the metamaterial concept is not restricted to linear optics but can be also applied to nonlinear optics. As shown above, quite different metamaterial designs can be used for nonlinear frequency conversion and modulation schemes. Despite the rapid progress in recent years there are however, a number of elementary aspects of nonlinear metamaterials which are not fully understood

yet. For instance, the role of surface and bulk terms in the local second-order nonlinear response of metallic metamaterials has not been completely clarified. Another open question is the origin of the nonlinear response of hybrid metamaterials containing metallic nanostructures and nonlinear dielectric nanoparticles.

Losses in metal based nonlinear metamaterials constitute a serious limitation for their nonlinear performance. This might be tolerable for proof-of-principle experiments and for building the fundamental understanding as discussed in this review. However, if one aims at replacing conventional nonlinear crystals by metal based nonlinear metamaterials in real-world applications, losses turn out to be a show stopper. In contrast, the situation appears more promising when thinking in terms of using nonlinear metamaterials as localized light sources, e.g., in optical chips. Another promising approach is to use nonlinear metamaterials based on low-loss dielectrics. Here, we expect to see new experiments and further progress in the future.

References

1. Soukoulis, C. M. and Wegener M. (2011). Past achievements and future challenges in the development of three-dimensional photonic metamaterials, *Nature Photon.* **5**, 523–530 (2011).
2. Lapine, M., Shadrivov, I. V., and Kivshar, Y. S. (2014). Colloquium: Nonlinear metamaterials, *Rev. Mod. Phys.* **86**, 1093–1123.
3. Kauranen, M. and Zayats, A. V. (2012). Nonlinear plasmonics, *Nature Photon.* **6**, 737–748.
4. Boyd, R. W., *Nonlinear Optics* (*3rd edition*). (Academic Press, London, 2008).
5. Prasad, P. N. and Williams, D. J., *Introduction to Nonlinear Optical Effects in Molecules and Polymers*. (John Wiley & Sons, New York, 1991).
6. Miller, R. C. (1964). Optical second harmonic generation in piezoelectric crystals, *Appl. Phys. Lett.* **5**, 17–18.
7. Shen, Y. R. (2000). Surface nonlinear optics: A historical perspective, *IEEE J. Sel. Topics Quantum Electron.* **6**, 1375–1379.
8. Heinz, T. F. Second-order nonlinear optical effects at surfaces and interfaces, in *Nonlinear surface electromagnetic phenomena*, H.-E. Ponath and G. I. Stegeman, eds. (Elsevier, Amsterdam, 1991), Chapter 5, pp. 353–415.
9. Khurgin, J. B. and Sun G. (2013). Plasmonic enhancement of the third order nonlinear optical phenomena: Figures of merit, *Opt. Express* **21**, 27460.

10. Slusher, R. E. and Eggleton, B. J. (Eds.), *Nonlinear Photonic Crystals.* (Springer, Heidelberg, 2003).
11. Kuznetsov, A. I., Miroshnichenko, A. E., Fu, Y. H., Zhang, J. B., and Luk'yanchuk B. (2012). Magnetic light, *Sci. Rep.* **2**, 492.
12. Staude, I., Miroshnichenko, A. E., Decker, M., Fofang, N. T., Liu, S., Gonzales, E., Dominguez, J., Luk, T. S., Neshev, D. N., Brener, I., and Kivshar, Y. (2013). Tailoring directional scattering through magnetic and electric resonances in subwavelength silicon nanodisks, *ACS Nano* **7**, 7824–7832.
13. Kim, E., Wang, F., Wu, W., Yu, Z., and Shen, Y. R. (2008). Nonlinear optical spectroscopy of photonic metamaterials, *Phys. Rev. B* **78**, 113102.
14. Rodriguez, F. J., Wang, F. X., and Kauranen, M. (2008). Calibration of the second-order nonlinear optical susceptibility of surface and bulk of glass, *Opt. Express* **16**, 8704–8710.
15. Wang, F. X., Rodríguez, F. J., Albers, W. M., Ahorinta, R., Sipe, J. E., and Kauranen, M. (2009). Surface and bulk contributions to the second-order nonlinear optical response of a gold film, *Phys. Rev. B* **80**, 233402.
16. Muhlschlegel, P., Eisler, H., Martin, O. J. F., Hecht, B., and Pohl, D. W. (2005). Resonant optical antennas, *Science* **308**, 1607–1609.
17. Linden, S., Niesler, F. B. P., Förstner, J., Grynko, Y., Meier, T., and Wegener, M. (2012). Collective effects in second-harmonic generation from split-ring-resonator arrays, *Phys. Rev. Lett.* **109**, 015502.
18. Auguié, B. and Barnes, W. L. (2008). Collective resonances in gold nanoparticle arrays, *Phys. Rev. Lett.* **101**, 143902.
19. Chu, Y., Schonbrun, E., Yang, T., and Crozier, K. B. (2008). Experimental observation of narrow surface plasmon resonances in gold nanoparticle arrays, *Appl. Phys. Lett.* **93**, 181108.
20. Chen, C. K., de Castro, A. R. B., and Shen, Y. R. (1981). Surface-enhanced second harmonic generation, *Phys. Rev. Lett.* **46**, 145–148.
21. Wokaun, A., Bergman, J. G., Heritage, J. P., Glass, A. M., Liao, P. F., and Olson, D. H. (1981). Surface second-harmonic generation from metal island films and microlithographic structures, *Phys. Rev. B* **24**, 849–856.
22. Lamprecht, B., Leitner, A., and Aussenegg, F. R. (1999). Femtosecond decay-time measurement of electron-plasma oscillation in nanolithographically designed silver particles, *Appl. Phys. B* **68**, 419–423.
23. Tuovinen, H., Kauranen, M., Jefimovs, K., Vahimaa, P., Vallius, T., Turunen, J., Tkachenko, N. V., and Lemmetyinen, H. (2002). Linear and second-order nonlinear optical properties of arrays of noncentrosymmetric gold nanoparticles, *J. Nonlinear Opt. Phys. Mater.* **11**, 421–432.
24. Klein, M. W., Enkrich, C., Wegener, M., and Linden, S. (2006). Second-harmonic generation from magnetic meta-materials, *Science* **313**, 502–504.
25. Feth, N., Linden, S., Klein, M. W., Decker, M., Niesler, F. B. P., Zeng, Y., Hoyer, W., Liu, J., Koch, S. W., Moloney, J. V., and Wegener, M. (2008). Second-harmonic generation from complementary split-ring resonators, *Opt. Lett.* **33**, 1975–1977.

26. Xu, T., Jiao, X., Zhang, G. P., and Blair, S. (2006). Second-harmonic emission from subwavelength apertures: Effects of aperture symmetry and lattice arrangement, *Opt. Express* **15**, 13894–13906.
27. Lesuffleur, A., Swaroop, Kumar, L. K. and Gordon, R. (2006). Enhanced second harmonic generation from nanoscale double-hole arrays in a gold film, *Appl. Phys. Lett.* **88**, 261104.
28. Canfield, B. K., Kujala, S., Jefimovs, K., Vallius, T., Turunen, J., and Kauranen, M. (2004). Linear and nonlinear optical responses influenced by broken symmetry in an array of gold nanoparticles, *Opt. Express* **12**, 5418–5423.
29. Kujala, S., Canfield, B. K., Kauranen, M., Svirko, Y., and Turunen, J. (2007). Multipole interference in the second-harmonic optical radiation from gold nanoparticles, *Phys. Rev. Lett.* **98**, 167403.
30. Czaplicki, R., Zdanowicz, M., Koskinen, K., Laukkanen, J., Kuittinen, M., and Kauranen, M. (2011). Dipole limit in second-harmonic generation from arrays of gold nanoparticles, *Opt. Express* **19**, 26866–26871.
31. Niesler, F. B. P., Feth, N., Linden, S., and Wegener, M. (2011). Second-harmonic optical spectroscopy on split-ring-resonator arrays, *Opt. Lett.* **36**, 1533–1535.
32. Thyagarajan, K., Rivier, S., Lovera, A., and Martin, O. J. F. (2012). Enhanced second-harmonic generation from double resonant plasmonic antennae, *Opt. Express* **20**, 12860–12865.
33. Celebrano, M., Wu, X., Baselli, M., Großmann, S., Biagioni, P., Locatelli, A., De Angelis, C., Cerullo, G., Osellame, R., Hecht, B., Duò, L., Ciccacci, F., and Finazzi, M. (2015). Mode matching in multiresonant plasmonic nanoantennas for enhanced second harmonic generation, *Nature Nano.* **10**, 412–417.
34. Metzger, B., Gui, L., Fuchs, J., Floess, D., Hentschel, M., and Giessen, H. (2015). Strong enhancement of second harmonic emission by plasmonic resonances at the second harmonic wavelength, *Nano Lett.* **15**, 3917–3922.
35. Linnenbank, H. and Linden, S. (2015). Second harmonic generation spectroscopy on second harmonic resonant plasmonic metamaterials, *Optica* **2**, 698–701.
36. Klein, M. W., Wegener, M., Feth, N., and Linden, S. (2007). Experiments on second- and third-harmonic generation from magnetic metamaterials, *Opt. Express* **15**, 5238–5247.
37. Czaplicki, R., Mäkitalo, J., Siikanen, R., Husu, H., Lehtolahti, J., Kuittinen, M., and Kauranen, M. (2015). Second-harmonic generation from metal nanoparticles — resonance enhancement versus particle geometry, *Nano Lett.* **15**, 530–534.
38. Canfield, B. K., Husu, H., Laukkanen, J., Bai, B. F., Kuittinen, M., Turunen, J., and Kauranen, M. (2007). Local field asymmetry drives second-harmonic generation in noncentrosymmetric nanodimers, *Nano Lett.* **7**, 1251–1255.
39. O'Brien, K., Suchowski, H., Rho, J., Salandrino, A., Kante, B., Yin, X., and Zhang, X. (2015). Predicting nonlinear properties of metamaterials from the linear response, *Nature Mater.* **14**, 379–383.

40. Luk'yanchuk, B., Zheludev, N. I., Maier, S. A., Halas, N. J., Nordlander, P., Giessen, H., and Chong, C. T. (2010). The fano resonance in plasmonic nanostructures and metamaterials, *Nature Mater.* **9**, 707–715.
41. Husu, H., Siikanen, R., Lehtolahti, J., Laukkanen, J., Kuittinen, M., and Kauranen, M. (2012). Metamaterials with tailored nonlinear optical response, *Nano Lett.* **12**, 673–677.
42. Czaplicki, R., Husu, H., Siikanen, R., Mäkitalo, J., Kauranen, M., Laukkanen, J., Lehtolahti, J., and Kuittinen, M. (2013). Enhancement of second-harmonic generation from metal nanoparticles by passive elements, *Phys. Rev. Lett.* **110**, 093902.
43. McMahon, M. D., Lopez, R., and Haglund, R. F., Ray, Jr. E. A. and Bunton, P. H. (2006). Second-harmonic generation from arrays of symmetric gold nanoparticles, *Phys. Rev. B* **73**, 041401.
44. de Ceglia, D., Vincenti, M. A., De Angelis, C., Locatelli, A., Haus, J. W., and Scalora, M. (2015). Role of antenna modes and field enhancement in second harmonic generation from dipole nanoantennas, *Opt. Express* **23**, 1715–1729.
45. Konishi, K., Higuchi, T., Li, J., Larsson, J., Ishii, S., and Kuwata-Gonokami, M. (2014). Polarization-controlled circular second-harmonic generation from metal hole arrays with threefold rotational symmetry, *Phys. Rev. Lett.* **112**, 135502.
46. Li, G., Chen, S., Pholchai, N., Reineke, B., Wong, P. W. H., Pun, E. Y. B., Cheah, K. W., Zentgraf, T., and Zhang, S. (2015). Continuous control of the nonlinearity phase for harmonic generations, *Nature Mater.* **14**, 607–612.
47. Kruk, S., Weismann, M., Bykov, A. Y., Mamonov, E. A., Kolmychek, I. A., Murzina, T., Panoiu, N. C., Neshev, D. N., and Kivshar, Y. S. (2015). Enhanced magnetic second-harmonic generation from resonant metasurfaces, *ACS Photonics* **2**, 1007–1012.
48. Yu, N., Genevet, P., Kats, M. A., Aieta, F., Tetienne, J.-P., Capasso, F., Gaburro, Z. (2001). Light propagation with phase discontinuities: Generalized laws of reflection and refraction, *Science* **334**, 333–337.
49. Segal, N., Keren-Zur, S., Hendler, N. and Ellenbogen, T. (2015). Controlling light with metamaterial-based nonlinear photonic crystals, *Nature Photon.* **9**, 180–184.
50. Keren-Zur, S., Avayu, O., Michaeli, L., and Ellenbogen, T. (2016). Nonlinear beam shaping with plasmonic metasurfaces, *ACS Photonics* **3**, 117–123.
51. Segovia, P., Marino, G., Krasavin, A. V., Olivier, N., Wurtz, G. A., Belov, P. A., Ginzburg, P., and Zayats, A. V. (2015). Hyperbolic metamaterial antenna for second-harmonic generation tomography, *Opt. Express* **23**, 30730–30738.
52. Tsang, T. Y. F. (1996). Surface-plasmon-enhanced third-harmonic generation in thin silver films, *Opt. Lett.* **21**, 245.
53. Kim, E. M., Elovikov, S. S., Murzina, T. V., Nikulin, A. A., Aktsipetrov, O. A., Bader, M. A., and Marowsky, G. (2005). Surface-enhanced optical third-harmonic generation in Ag island films, *Phys. Rev. Lett.* **95**, 227402.

54. Liao, H. B., Xiao, R. F., Fu, J. S., Yu, P., Wong, G. K. L., and Sheng, P. (1997). Large third-order optical nonlinearity in Au:SiO2O2 composite films near the percolation threshold, *Appl. Phys. Lett.* **70**, 1.
55. Sipe, J. E. and Boyd, R. W. (1992). Nonlinear susceptibility of composite optical materials in the Maxwell Garnett model, *Phys. Rev. A* **46**, 1614–1629.
56. Zeng, X. C., Bergman, D. J., Hui, P. M., and Stroud, D. (1988). Effective-medium theory for weakly nonlinear composites, *Phys. Rev. B* **38**, 10970–10973.
57. Boyd, R. W. and Sipe, J. E. (1994). Nonlinear optical susceptibilities of layered composite materials, *J. Opt. Soc. Am. B* **11**, 297–303.
58. Smith, D. D., Fischer, G., Boyd, R. W., and Gregory, D. A. (1997). Cancellation of photoinduced absorption in metal nanoparticle composites through a counterintuitive consequence of local field effects, *J. Opt. Soc. Am. B* **14**, 1625–1631.
59. Piredda, G., Smith, D. D., Wendling, B., and Boyd, R. W. (2008). Nonlinear optical properties of a gold–silica composite with high gold fill fraction and the sign change of its nonlinear absorption coefficient, *J. Opt. Soc. Am. B* **25**, 945–950.
60. Gomez, L. A., de Araujo, C. B., Brito-Silva, A. M., and Galembeck, A. (2008). Solvent effects on the linear and nonlinear optical response of silver nanoparticles, *Appl. Phys. B* **92**, 61–66.
61. Souza, R. F., Alencar, M. A. R. C., da Silva, E. C., Meneghetti, M. R., and Hickmann, J. M. (2008). Nonlinear optical properties of Au nanoparticles colloidal system: Local and nonlocal responses, *Appl. Phys. Lett.* **92**, 201902.
62. Hentschel, M., Utikal, T., Giessen, H., and Lippitz, M. (2012). Quantitative modeling of the third harmonic emission spectrum of plasmonic nanoantennas, *Nano Lett.* **12**, 3778–3782.
63. Metzger, B., Hentschel, M., Lippitz, M., and Giessen, H. (2012). Third-harmonic spectroscopy and modeling of the nonlinear response of plasmonic nanoantennas, *Opt. Lett.* **37**, 4741.
64. Metzger, B., Schumacher, T., Hentschel, M., Lippitz, M. S., and Giessen, H. (2014). Third harmonic mechanism in complex plasmonic fano structures, *ACS Photonics* **1**, 471–476.
65. Renger, J., Quidant, R., Van Hulst, N., and Novotny, L. (2010). Surface-enhanced nonlinear fourwave mixing, *Phys. Rev. Lett.* **104**, 046803.
66. Genevet, P., Tetienne, J. P., Gatzogiannis, E., Blanchard, R., Kats, M. A., Scully, M. O., and Capasso, F. (2010). Large enhancement of nonlinear optical phenomena by plasmonic nanocavity gratings, *Nano Lett.* **10**, 4880–4883.
67. Poutrina, E., Ciraci, C., Gauthier, D. J., and Smith, D. R. (2012). Enhancing four-wave-mixing processes by nanowire arrays coupled to a gold film, *Opt. Express* **20**, 11005.
68. Almeida, E. and Prior, Y. (2015). Rational design of metallic nanocavities for resonantly enhanced four-wave mixing, *Sci. Reports* **5**, 10033.

69. Palomba, S., Zhang, S., Park, Y., Bartal, G., Yin, X., and Zhang, X. (2012). Optical negative refraction by four-wave mixing in thin metallic nanostructures, *Nature Mater.* **11**, 34.
70. Almeida, E., Shalem, G., and Prior, Y. (2016). Subwavelength nonlinear phase control and anomalous phase matching in plasmonic metasurfaces, *Nature Comm.* **7**, 10367.
71. Rotenberg, N., Caspers, J. N. and van Driel, H. M. (2009). Tunable ultrafast control of plasmonic coupling to gold films, *Phys. Rev. B* **80**, 245420.
72. Wurtz, G. A., Pollard, R., Hendren, W., Wiederrecht, G. P., Gosztola, D. J., Podolskiy, V. A., and Zayats, A. V. (2011). Designed ultrafast optical nonlinearity in a plasmonic nanorod metamaterial enhanced by nonlocality, *Nature Nanotech.* **6**, 107.
73. Ren, M., Jia, B., Ou, J.-Yu, Plum, E., Zhang, J., MacDonald, K. F., Nikolaenko, A. E., Xu, J., Gu, M., and Zheludev, N. I. (2011). Nanostructured plasmonic medium for terahertz bandwidth all-optical switching, *Adv. Mater.* **23**, 5540–5544.
74. Neira, A. D., Olivier, N., Nasir, M. E., Dickson, W., Wurtz, G. A., and Zayats, A. V. (2015). Eliminating material constraints for nonlinearity with plasmonic metamaterials, *Nature Commun.* **6**, 7757.
75. Suchowski, H., O'Brien, K., Wong, Z. J., Salandrino, A., Yin, X., Zhang, X. (2013). Phase mismatch–free nonlinear propagation in optical zero-index materials, *Science* **342**, 1223.
76. Kauranen, M. (2013). Freeing nonlinear optics from phase-matching, *Science* **342**, 1182–1183.
77. Pendry, J. B., Holden, A. J., Robbins, D. J., and Stewart, W. J. (1999). Magnetism from conductors and enhanced nonlinear phenomena, *IEEE Trans. Microwave Theory Tech.* **47**, 2075–2084.
78. Fan, W., Zhang, S., Malloy, K. J., Brueck, S. R. J., Panoiu, N. C., and Osgood, R. M. (2006). Second harmonic generation from patterned GaAs inside a subwavelength metallic hole array, *Opt. Express* **14**(21), 9570–9575.
79. Fan, W., Zhang, S., Panoiu, N.-C., Abdenour, A., Krishna, S., Osgood, R. M., Malloy, K. J., and Brueck, S. R. J. (2006). Second harmonic generation from a nanopatterned isotropic nonlinear material, *Nano Lett.* **6**(5), 1027–1030.
80. Niesler, F. B. P., Feth, N., Linden, S., Niegemann, J., Gieseler, J., Busch, K., and Wegener, M. (2009). Second-harmonic generation from split-ring resonators on a GaAs substrate, *Opt. Lett.* **34**(13), 1997–1999.
81. Rosencher, E., Fiore, A., Vinter, B., Berger, V., Bois, Ph., Nagle, J. (1996). Quantum engineering of optical nonlinearities, *Science* **271**, 168–173.
82. Lee, J., Tymchenko, M., Argyropoulos, C., Chen, P. Y., Lu, F., Demmerle, F., Boehm, G., Amann, M. C., Alù, A., and Belkin, M. A. (2014). Giant nonlinear response from plasmonic metasurfaces coupled to intersubband transitions, *Nature* **511**, 65–69.

83. Wolf, O., Campione, S., Benz, A., Ravikumar, A. P., Liu, S., Luk, T. S., Kadlec, E. A., Shaner, E. A., Klem, J. F., Sinclair, M. B., and Brener, I. (2015). Phased-array sources based on nonlinearmetamaterial nanocavities, *Nature Comm.*, doi:10.1038/ncomms8667.
84. Cai, W., Vasudev, A. P., and Brongersma, M. L. (2011). Electrically controlled nonlinear generation of light with plasmonics, *Science* **333**, 1720–1723.
85. Pu, Y., Grange, R., Hsieh, C. L., Psaltis, D. (2010). Nonlinear optical properties of core-shell nanocavities for enhanced second-harmonic generation, *Phys. Rev. Lett.* **104**, 207402 .
86. Aouani, H., Rahmani, M., Navarro-Cía, M., Maier, S. A. (2014). Third-harmonic-upconversion enhancement from a single semiconductor nanoparticle coupled to a plasmonic antenna, *Nat. Nanotechnol.* **9**, 290–294.
87. Metzger, B., Hentschel, M., Schumacher, T., Lippitz, M., Ye, X., Murray, C. B., Knabe, B., Buse, K., and Giessen, H. (2014). Doubling the efficiency of third harmonic generation by positioning ITO nanocrystals into the hot-spot of plasmonic gap-antennas, *Nano Lett.* **14**, 2867–2872.
88. Linnenbank, H., Grynko, Y., Förstner, J., and Linden, S. (2016). Second harmonic generation spectroscopy on hybrid plasmonic/dielectric nanoantennas, *Light: Science & Applications* **5**, e16013.
89. Fu, M., Wang, K., Long, H., Yang, G., Lu, P., Hetsch, F., Susha, A. S., and Rogach, A. L. (2012). Resonantly enhanced optical nonlinearity in hybrid semiconductor quantum dot–metal nanoparticle structures, *Appl. Phys. Lett.* **100**, 063117.
90. Kim, E., Shen, Y. R., Wu, W., Ponizovskaya, E., Yu, Z., Bratkovsky, A. M., Wang, S.-Y., and Williams, R. S. (2007). Modulation of negative index metamaterials in the near-IR range, *Appl. Phys. Lett.* **91**, 173105.
91. Cho, D. J., Wu, W., Ponizovskaya, E., Chaturvedi, P., Bratkovsky, A. M., Wang, S.-Y., Zhang, X., Wang, F., and Shen, Y. R. (2009). Ultrafast modulation of optical metamaterials, *Opt. Express* **17**, 17652.
92. Dani, K. M., Ku, Z., Upadhya, P. C., Prasankumar, R. P., Brueck, S. R. J., and Taylor, A. J. (2009). Subpicosecond optical switching with a negative index metamaterial, *Nano Lett.* **9**, 3565–3569.
93. Wurtz, G. A., Pollard, R., and Zayats, A. V. (2006). Optical bistability in nonlinear surface-plasmon polaritonic crystals, *Phys. Rev. Lett.* **97**, 057402.
94. Alloatti, L., Kieninger, C. M., Frölich, A., Lauermann, M., Frenzel, T., Köhnle, K., Freude, W., Leuthold, J., Wegener, M., and Koos, C. (2015). Second-order nonlinear optical metamaterials: ABC-type nanolaminates, *Appl. Phys. Lett.* **107**, 121903.
95. Clemmen, S., Hermans, A., Solano, E., Dendooven, J., Koskinen, K., Kauranen, M., Brainis, E., Detavernier, C., and Baets, R. (2015). Atomic layer deposited second-order nonlinear optical metamaterial for back-end integration with CMOS-compatible nanophotonic circuitry, *Opt. Lett.* **40**, 5371.
96. Hermans, A., Kieninger, C., Koskinen, K., Wickberg, A., Solano, E., Dendooven, J. Kauranen, M. Clemmen, S., Wegener, M., Koos, C., and Baets,

R. (2017). On the determination of $\chi(2)$ in thin films: A comparison of one-beam second-harmonic generation measurement methodologies, *Sci. Rep.* **7**, 44581.

97. Khurgin, J. B. (1988). Second-order nonlinear effects in asymmetric quantum-well structures, *Phys. Rev. B* **38**, 4056.
98. Shcherbakov, M. R., Neshev, D. N., Hopkins, B., Shorokhov, A. S., Staude, I., Melik-Gaykazyan, E. V., Decker, M., Ezhov, A. A., Miroshnichenko, A. E., Brener, I., Fedyanin, A. A., and Kivshar, Y. S. (2014). Enhanced third-harmonic generation in silicon nanoparticles driven by magnetic response, *Nano Lett.* **14**, 6488.
99. Shcherbakov, M. R., Vabishchevich, P. P., Shorokhov, A. S., Chong, K. E., Choi, D.-Y., Staude, I., Miroshnichenko, A. E., Neshev, D. N., Fedyanin, A. A., and Kivshar, Y. S. (2015). Ultrafast all-optical switching with magnetic resonances in nonlinear dielectric nanostructures, *Nano Lett.* **15**, 6985.

CHAPTER 4

Controlled Radiative Dynamics Using Plasmonic Nanocavities

GLEB M. AKSELROD* **and**
MAIKEN H. MIKKELSEN†

Duke University, USA
*gleb.akselrod@duke.edu
†m.mikkelsen@duke.edu

4.1. Introduction to Purcell Enhancement Using Plasmonic Nanocavities

Spontaneous emission from quantum emitters such as organic molecules, quantum dots and bulk semiconductors is a fundamental process central to many modern technologies including light emitting diodes, lasers, single photon sources, and biological imaging. The radiative properties of these emitters are determined by their transition dipole moment and the electromagnetic environment of the emitter, as determined by Fermi's Golden Rule. The dipole moment is intrinsic to the electronic structure of the emitters and hence cannot be readily changed. On the other hand, the electromagnetic environment of emitters can be controlled by coupling them to optical cavities that have a high density of optical states. This effect of modifying the radiative dynamics by changing the optical density of states is known as the Purcell effect[1] and has been recognized and

studied for decades. In addition to the fundamental importance of the Purcell effect, it is also being explored in the context of developing high-speed devices based on spontaneous emission, in which the slow intrinsic emission lifetime of typical emitters is a serious limitation. For example, ultrafast light-emitting diodes could be useful for on-chip optical communication where stimulated emission light sources (lasers) are not suitable. Likewise, the repetition rate of single photon emitters is currently limited by the spontaneous emission lifetime and could thus benefit from large Purcell enhancements.[2–4]

The Purcell enhancement, defined as the enhancement in the spontaneous emission rate relative to emitters in free space, is related to the geometrical parameters of the cavity, namely the ratio of the cavity quality factor, Q, to the cavity mode volume, V. Thus the quest for large Purcell enhancements generally focuses on two directions: increasing the quality factor of the cavity and reducing the mode volume. The optical structures typically take the form of cavities based on either dielectric or metallic materials. Work on dielectric cavities[5–10] typically focuses on increasing the cavity quality factor as the mode volume is limited to a substantial fraction of a cubic wavelength. Despite advances in dielectric cavity fabrication resulting in extremely high quality factors,[7,8] the maximum Purcell enhancements have been limited to 10–100 due to limitations on the mode volume. Plasmonic nanocavities based on noble metals are particularly powerful for Purcell enhancement because the mode volumes are no longer diffraction-limited. Another advantage of plasmonic cavities is their relative large linewidth ($Q \sim 10$) which enables efficient coupling to broadband room-temperature emitters, something not possible with narrow-linewidth dielectric cavities. Plasmonic cavities have been widely studied for absorption enhancement[11,12] as well as spontaneous emission rate enhancement,[3,11,13,14] and for modification of the emission directionality.[15–17] These demonstrations have led to a range of proof-of-principle devices based on enhancement of absorption and spontaneous emission, including single photon sources,[2–4] nanolasers,[18,19] and photodetectors.[20–22]

One downside to plasmonic cavities is that coupled emitters are prone to large non-radiative rates due to Ohmic losses in the metal, which must be managed through nanocavity design, as will be discussed below. Of particular note is the fact that the Purcell factor becomes poorly defined in such lossy cavities. While in dielectric cavities the enhancement in the total emission rate is equal to the enhancement in the radiative rate, this is no longer true in metallic cavities. Thus it becomes necessary to specify not only the Purcell factor but also the quantum yield for free space photon emission of the cavity, from which the radiative rate can be calculated given the total emission rate. For the sake of consistency with literature, in this chapter we will refer to the enhancement in the total emission rate as the Purcell enhancement, while simultaneously being careful to specify the quantum efficiency for emission.

One of the most well-studied metal nanocavities is the bowtie nanoantenna,[11] which combines a small metal-insulator-metal gap along with a geometry that acts as an antenna, allowing coupling of the cavity mode to free space. However, such structures are often made using top-down nanofabrication techniques including electron beam lithography[11,23] or ion beam milling,[24] making it difficult to create sub-10 nm lateral gaps in which the largest electric field enhancements occur. In addition to small gaps, the nanocavities must also act as optical antennas by efficiently coupling the confined optical mode to free space or a waveguide. The patch antenna is a promising geometry that can overcome the lateral gap fabrication challenge and at the same time act as an efficient optical antenna.[4,17,25,26] The patch antenna structure consists of a metal disk or faceted metal nanoparticle which is situated over a metal ground plane, separated vertically by a dielectric spacer. This vertical geometry enables the use of planar fabrication techniques, allowing the metal-insulator-metal gap to be controlled with nanometer[27] and even sub-nanometer[28] precision. Patch antennas with micrometer diameters have been used to enhance the total spontaneous emission rate by a factor of $\sim$80 and showed directional emission, although the

radiative quantum efficiency was only a few percent.[17] The optical patch antenna has also been theoretically proposed as an efficient single photon source with a high repetition rate.[4]

In this chapter we will review recent work on optical patch antennas, a system that has proved to be a very versatile and well-controlled system for tailoring the spontaneous emission rate of fluorescent emitters such as molecules,[29] quantum dots,[30,31] and two-dimensional materials.[32] The structure consists of a metal nanocube situated over a metal film, separated by a $\sim$ 10 nm spacer consisting of a passive dielectric or an emissive material. This structure forms a plasmonic nanocavity with an ultra-small effective mode volume of $V_{eff} \approx 0.001(\lambda/n)^3$ and a typical quality factor of $Q \approx 15$ (Figure 4.1(a)). Fluorescent emitters embedded into this nanocavity have shown fluorescence intensity enhancement of up to a factor of

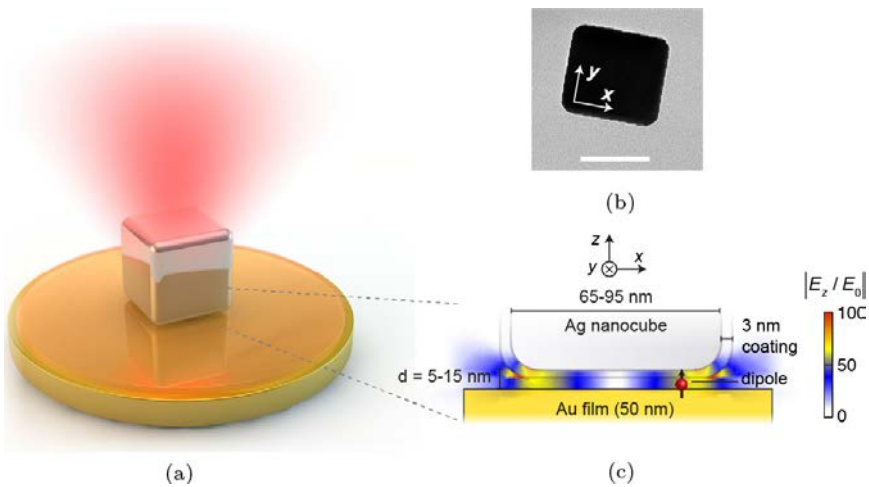


Fig. 4.1. (a) Schematic of the nanocavity formed by a colloidal silver nanocube situated over a metal ground plane, separated by a dielectric spacer. Upward beam illustrates the directional emission from the cavity due to its action as a nanopatch antenna. (b) Transmission electron micrograph of a single nanocube, showing atomically flat facets. Scale bar, 50 μm. (c) Cross-sectional schematic of the nanocavity, showing the electric field enhancement in the z direction at the fundamental resonance of the structure. An optimally oriented and positioned dipole is illustrated.

30,000[33] as well as ∼1,000-fold enhancements in the spontaneous emission rate.[29,34]

The fundamental mode of the nanocavity is confined laterally by the edges of the nanocubes and vertically by the nanoscale gap formed between the nanocube's bottom facet and the metal film (Figures 4.1(a–b)). This unique plasmonic mode has a number of key advantages. First, the nanoscale gap creates electric field enhancements of up to ∼100-fold, and hence a ∼10,000-fold Purcell enhancement over a large fraction of the nanocavity area[27] (Figure 4.1(c)). These large field enhancements can subsequently be used for the enhancement of the spontaneous emission rate of emitters embedded in the cavity.[29,35] Second, the resonance wavelength can be tuned over a wide wavelength range by changing the thickness of the spacer layer in the gap or by changing the nanocube size.[27,36] Third, despite the large field enhancements and the proximity of emitters to metal surfaces, the radiative quantum yield of the structure is high, up to ∼50%, due to the antenna action of the nanocavity.[29,33,37] This is in contrast to many plasmonic structures in which significant non-radiative quenching occurs for sub-10 nm gap sizes. Finally, the patch antenna emission is directional in the vertical direction, with a calculated collection efficiency of 84% using a standard microscope objective.[29]

4.2. Optical Properties of the Nanocavity

We first review the basic optical and spectroscopic properties of the nanoantenna, also referred to as the nanocavity. The nanocavity formed by the film-coupled nanocubes have several unique features that make it excellent for Purcell enhancement. The gap between the nanocube and metal ground plane can be thought of as a Fabry-Pérot resonator, with the longitudinal length of the cavity determined by the side length of the nanocube, hence determining the fundamental resonance wavelength. The vertical (transverse) dimension of the cavity is determined by the thickness of the polymer spacer layer, or other material filling the resonator. The optical properties can be predicted theoretically using a recently developed coupled-mode

theory.[38] Alternatively, the structure can be interpreted using transmission line theory, in analogy with microwave transmission lines.[39] In this chapter, we will only consider the optical properties of the structure as obtained from full-wave simulations.[37]

4.2.1. *Nanocavity resonances*

The nanocavity resonances can be analyzed using full-wave numerical simulations.[27] A typical scattering spectrum from simulations is shown in Figure 4.2(a), for a cavity with an $L = 81$ nm nanocube side length and a $d = 8$ nm gap. Several narrowband resonances are observed, which have the same spectral positions under TM and TE polarized excitation. The fundamental mode of the cavity has a Lorentzian line shape with a resonance at $\lambda = 675$ nm and a linewidth of ∼35 nm. These modes are analogous to waveguide cavity modes that have been observed in other plasmonic structures having a metal-insulator-metal structure.[36,40–42] The modes can be excited from free space and simultaneously observed in the scattering spectrum due to the leaky nature of the waveguide. Modes 1–3, as

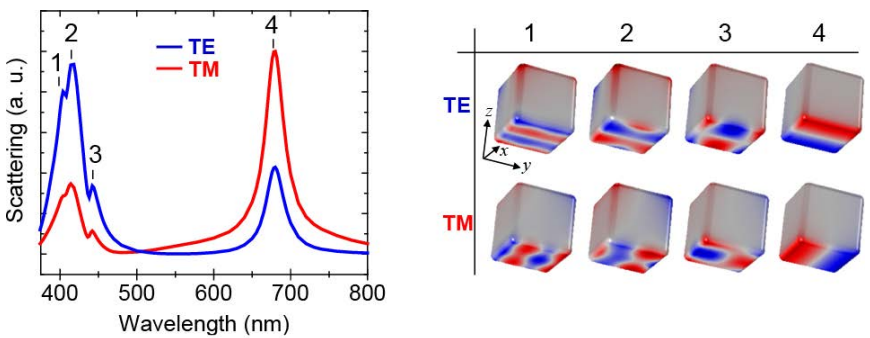


Fig. 4.2. (a) Simulated scattering spectrum of a typical nanocavity under TM and TE excitation, showing multiple resonances. Peak 4 is the fundamental resonance of the nanocavity, while peaks 1,2, and 3 are the second order modes. (b) Electric displacement fields on the surface of the nanocube at the resonance wavelength indicated in (a).

labeled in Figure 4.2(a) occur at shorter wavelengths in the scattering spectrum and correspond to higher order modes that have multiple field nodes. These modes have strong spectral overlap with each other but are distinct from the fundamental resonance (mode 4), which has been the focus of experimental work. To visualize these modes, in Figure 4.2(b), the normal component of the electric displacement field on the surface of the nanocube is shown. The complex pattern of mode 3 is attributed to both x and y propagating components. The modes have nearly identical profiles for both TE and TM polarized excitation, indicating that these modes are independent of polarization.

The electric field distribution of the fundamental mode has dipolar character, with the maximum electrical field near the edges of the cavity (Figures 4.3(a,c)). The dominant electric field is in the

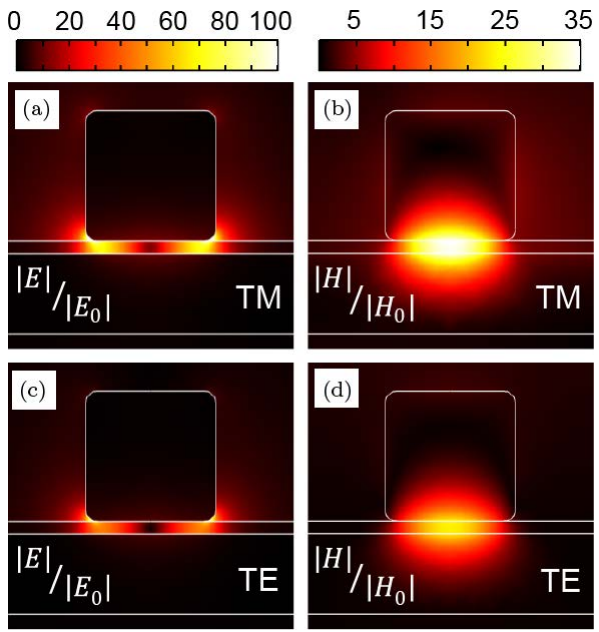


Fig. 4.3. (a,c) Electric field enhancement and (b,d) magnetic field enhancement relative to free space at the fundamental resonance of the nanocavity. The field enhancement is higher under (a,b) TM excitation, as compared to (c,d) TE excitation.

vertical direction, normal to the metal surfaces. The maximum field enhancement relative to free space is ∼100×, corresponding to an optical intensity enhancement of 10,000×. While the field distribution is found to be similar for TE and TM excitation, TM excitation results in larger field enhancements. Unlike the electric field, the magnetic field distribution has a maximum in the center of the cavity with the dominant field oriented in the plane, parallel to the metal surfaces (Figures 4.3(b,d)).

4.2.2. *Directional scattering and emission*

While large field enhancement has been observed in other metal-insulator-metal structures,[11,43] the film-coupled nanocube system is unique in its ability to efficiently couple the confined plasmons to free space by acting as an optical antenna, thus resulting in high radiative efficiency. Figure 4.4 shows schematics of two well-studied plasmonic antenna geometries — the bowtie antenna and the film-coupled

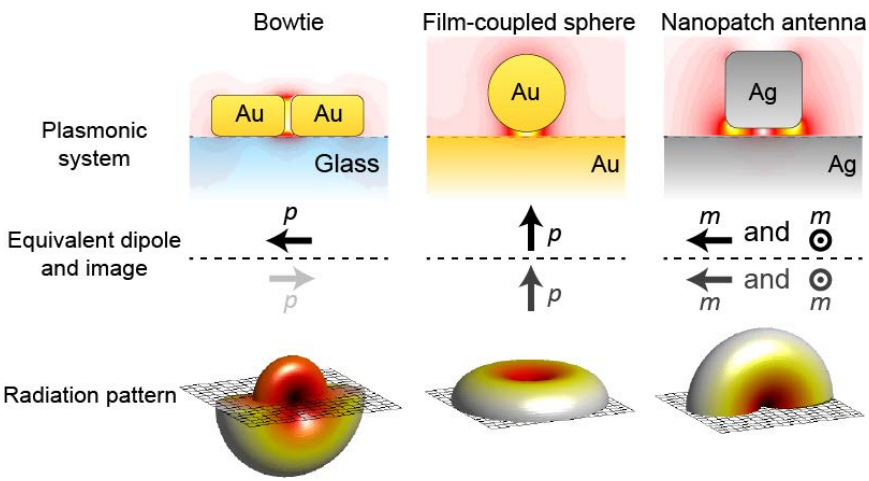
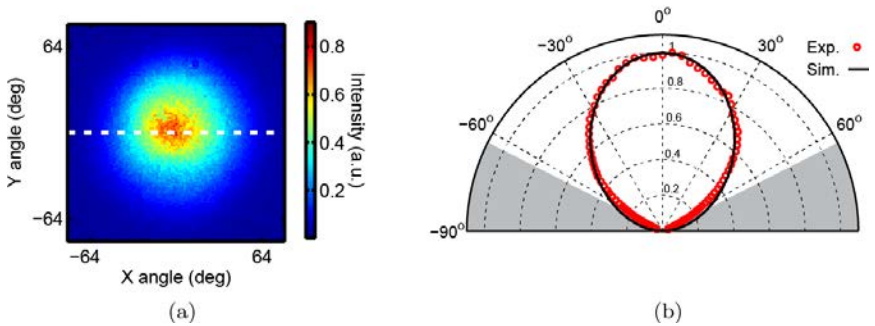


Fig. 4.4. Geometries and the radiation patterns of two plasmonic antenna geometries — bowtie antenna and film-coupled sphere — and their radiation patterns as compared to the film-coupled nanocube (nanopatch antenna). The effective dipole for each structure and its image dipole in the substrate are shown. The nanopatch antenna, which has an effective magnetic dipole, shows the highest coupling into free space.

sphere—and their free space radiation patterns, with a comparison to the film-coupled nanocube. The plasmonic modes in these structures result from the hybridization of plasmonic resonances that occur in each of the constituent elements. This hybridization creates either symmetric or antisymmetric modes. The highly confined antisymmetric mode couples to free space, having the radiation pattern of an electric dipole. In the case of the bowtie antenna on glass, the image dipole is weak, and most of the radiation couples to the glass substrate, although the main emission lobe is normal to the antenna. The film-coupled sphere system has a vertical effective electric dipole with a strong image dipole. These electric dipoles radiate primarily along the substrate, making collection with free space optics difficult.

In contrast to these structures, the film-coupled nanocube has an effective magnetic dipole due to the opposing currents in the top and bottom surface of the cavity. Two degenerate magnetic dipoles are present, one along each in-plane axis. These dipoles radiate out of plane, where the radiation lobe is perpendicular to the surface, emitting into free space. This geometry is analogous to the microwave patch antenna and has similar radiation characteristics. Consequently, the film-coupled nanocube is also referred to as a nanopatch antenna. In addition to acting like an antenna, the gap region can also be considered a small mode volume cavity formed by the nanoscale gap between the nanocube and the metal film.

The radiation pattern of a single nanopatch antenna was measured using momentum (Fourier) space imaging of the scattered light resonant with the fundamental mode.[29] The measurement involves imaging the back aperture of the objective lens (the Fourier plane) onto a CCD camera, and subsequently converting the intensity at each pixel to the scattering angle.[44] The measured radiation pattern shows good agreement with the simulated far-field radiation pattern obtained from finite-difference time-domain simulations (Figure 4.5). Using an objective lens with an NA = 0.9, it is estimated that 84% of plasmons that decay as free space photons can be collected by the objective.



(a) (b)

Fig. 4.5. (a) Momentum space (angle) image of scattering from a single nanocavity with $L = 75$ nm and $g = 8$ nm. (b) Measured and simulated radiation pattern from a single nanopatch antenna, where the measurement is taken from the cross-section of the image in (a). Gray regions indicate angular regions outside the collection cone of the objective lens. Using an objective lens with NA = 0.9, it is estimated that 84% of light is collected, based on the simulations.

4.3. Theory of Emitters Coupled to Nanocavity

The coupling of emitters to the nanocavity can be understood by considering the fluorescent emitters as four-level systems (Figure 4.6), which is typical for organic dyes and colloidal semiconductor quantum dots. In the absence of the nanocavity, excitation in free space occurs at a rate γ_{ex}^0 from the ground state (4.1) to the excited state (4.2), with transition energy $\hbar\omega_{ex}$. In the case of molecular or quantum dot emitters, the absorption is followed by fast (∼ps) vibrational relaxation to the relaxed excited state (4.3). The relaxed excited state subsequently decays back to the ground state radiatively by spontaneously emitting a photon with energy $\hbar\omega_{sp}$ at a rate γ_r^0. Likewise, the decay to the ground state can occur nonradiatively with an internal decay rate γ_{int}^0 through emission of phonons. This spontaneous decay transition occurs at an energy lower than the excitation energy, allowing for control of excitation and emission independently by tuning the plasmonic resonance. The internal nonradiative decay is an intrinsic property of the emitter that is determined by factors such as the solid-state environment or coupling of the excited electron to the vibrational states of the molecule. The intrinsic radiative quantum yield (QY_0), also referred as the quantum efficiency, is defined as the

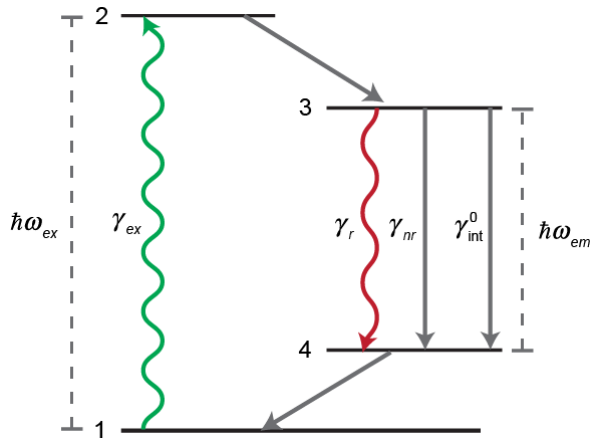


Fig. 4.6. The energy diagram of a typical emitter with a four-level energy structure. Wavy lines indicate radiative transitions (absorption or emission), while straight lines indicate non-radiative transitions. The relevant rates in the system are labeled, including the excitation rate γ_{ex}, the radiative decay rate due to the plasmonic structure γ_r, the non-radiative decay rate due to the plasmonic structure γ_{nr}, and the internal decay rate of the emitter γ^0_{int}. The excitation transition energy is $\hbar\omega_{ex}$, while the emission transition energy is $\hbar\omega_{em}$.

fraction of excited molecules that emit a photon, given by

$$QY_0 = \frac{\gamma_r^0}{\gamma_r^0 + \gamma_{\text{int}}^0} = \frac{\gamma_r^0}{\gamma_{sp}^0}, \tag{4.1}$$

where γ_{sp}^0 is the intrinsic spontaneous decay rate, also referred to as the spontaneous emission rate.

4.3.1. *Simulation of nanocavity*

When the emitter is coupled to the nanocavity, the emission rate can be enhanced by controlling the spectral overlap between the nanocavity resonance and the radiative transition energy of the emitter. Likewise, the absorption rate can be enhanced by overlapping the absorption resonance with the nanocavity resonance. To quantify both absorption and emission, a full understanding of the emitters' electromagnetic environment is obtained from finite-element simulations of the structure in COMSOL Multiphysics. In this section, we

describe the general approach for the simulations used to support the experimental results described in Refs. 29, 32, and 33, although details varied slightly in each of these works.

The nanocavity structure was first analyzed in the scattering formalism in order to obtain the scattering spectrum and the spatial maps of the field enhancement under the nanocube. In the scattering formalism, the analytical solution for the incident fields in the absence of the nanocube is used as the background field. In the presence of the nanocube, modifications to this background field are then attributed to the nanocavity structure. With this method, it is possible to obtain the enhancement of the electric and magnetic fields relative to the free space values, which is a convenient feature for comparison with experimental results.

To model the structure, a large spherical domain was created around a single nanocavity, and perfectly matched layer boundary conditions were utilized to simulate an open boundary. The corners of the nanocube were given a radius of curvature of $\sim$8 nm, in accordance with the corner radius obtained from transmission electron microscopy images of the silver nanocubes. A 3 nm shell with a refractive index of $n = 1.45$ was placed around the nanocube to simulate the polymer (poly(vinyl pyrrolidone)) layer which remains after chemical synthesis. The spacer layer was simulated as a lossless dielectric with a refractive index of $n = 1.45$. The spacer layer thickness and nanocube size was varied to match the experimental conditions as determined by ellipsometry and SEM analysis. The thickness of the metal ground plane is not found to be critical as long as the thickness is greater than $\sim$50 nm. The substrate under the metal film is semi-infinite glass with a refractive index of $n = 1.47$. The excitation was accomplished with a monochromatic plane wave either at normal incidence, or at $\sim 62^\circ$ relative to normal with TM polarization. The excitation angle was not found to affect the resonance wavelength or the electric field spatial distribution, but it does affect the maximum field enhancement.[27] This simulation domain was also used to simulate the spontaneous emission rate enhancement based on the Green's function formalism.

The results from the scattering simulations produce the spatial field distribution in the nanocavity (Figure 4.3) as well as its spectral response, indicating the energy of the fundamental mode (Figure 4.2). The excitation rate enhancement for a molecule placed in the nanocavity at a position $\mathbf{r}$ is related to the field enhancement by

$$\frac{\gamma_{ex}(\mathbf{r})}{\gamma_{ex}^{0}(\mathbf{r})} = \frac{|\mathbf{E}_{ex}(\mathbf{r}) \cdot \hat{\mathbf{n}}_p|^2}{|\mathbf{E}_{ex}^{0}(\mathbf{r}) \cdot \hat{\mathbf{n}}_p|^2} \tag{4.2}$$

where γ_{ex} is the excitation rate in the cavity, $\mathbf{E}_{ex}^{0}$ is the electric field experienced for the molecule in free space or on a reference substrate, $\mathbf{E}_{ex}$ is the excitation rate in the nanocavity, and $\hat{\mathbf{n}}_p$ is the orientation vector of the molecules absorption transition dipole moment. The absorber/emitter molecules are assumed to be small perturbations to the cavity system and hence are not included in the simulations. The dominant field of the fundamental mode is in the vertical direction, normal to the metal surfaces, and hence emitters with this vertical orientation experience the highest enhancements.

4.3.2. *Enhancement in the spontaneous emission rate*

After vibrational relaxation following absorption, the emission transition can also be enhanced due to interaction with the nanocavity via the Purcell effect. Emission and absorption in four-level systems are incoherent processes and hence can be treated independently. To simulate the spontaneous emission dynamics of emitters coupled to the nanocavity, the Green's function formalism was employed.[14,37] In this analysis, collective effects of multiple emitters was neglected and the system was assumed to be in the weak coupling regime. The weak coupling assumption is valid, as discussed in Ref. 34. Following Fermi's golden rule, the spontaneous emission rate for a point dipole at position $\mathbf{r}$ is

$$\gamma_{sp} = \frac{\pi\omega}{3\hbar\varepsilon_0}|\mathbf{p}|^2\rho(\mathbf{r},\omega_{em}) + \gamma_{\text{int}}^{0} \tag{4.3}$$

where ω_{em} is the transition frequency, $\mathbf{p}$ is the transition dipole moment, and $\rho(\mathbf{r},\omega)$ is the local density of optical states (LDOS).

The first term in Eq. (4.3) constitutes the coupling of the transition dipole moment to free space or a photonic structure, determined by the LDOS. The transition dipole moment of an emitter is assumed to be unchanged in different electromagnetic environments, which is the assumption of weak coupling. Meanwhile γ_{int}^0 is the internal decay rate of the dipole, which is also not affected by the electromagnetic environment. The LDOS, on the other hand, can be controlled using photonic structures and is expressed in terms of the dyadic Green's function, $\mathbf{G}$,

$$\rho(\mathbf{r},\omega) = \frac{6\omega}{\pi c^2}[\hat{\mathbf{n}}_p \cdot \text{Im}\{\mathbf{G}(\mathbf{r},\mathbf{r})\} \cdot \hat{\mathbf{n}}_p] \tag{4.4}$$

where $\hat{\mathbf{n}}_p$ is the unit vector for the transition dipole orientation. The Green's function is obtained from

$$\nabla \times \nabla \times \mathbf{G}(\mathbf{r},\mathbf{r}_0) - \varepsilon(\mathbf{r})k_0^2\mathbf{G}(\mathbf{r},\mathbf{r}_0) = \mathbf{I}\delta(\mathbf{r}_0) \tag{4.5}$$

where ε is the spatially-dependent permittivity in the structure, and k_0 is the free-space wavevector.

The LDOS for an emitter in free space is $\rho_0(\omega) = \omega^2/(\pi^2 c^3)$, which gives the free-space spontaneous emission rate of

$$\gamma_{sp}^0 = \gamma_r^0 + \gamma_{\text{int}}^0 = \frac{\omega^3|\mathbf{p}|^2}{3\hbar\pi\varepsilon_0 c^3} + \gamma_{\text{int}}^0. \tag{4.6}$$

The presence of the plasmonic structure modifies the decay of the emitter. When coupled to the nanocavity, the decaying emitter generates a plasmon excitation in the gap region. This plasmon subsequently decays either into free space by emitting a photon at a rate γ_r, or decay by Ohmic losses in the metal with a rate γ_{nr}. The total spontaneous decay rate of the emitter coupled to the nanocavity is then given by

$$\gamma_{sp} = \gamma_r + \gamma_{nr} + \gamma_{\text{int}}^0. \tag{4.7}$$

The Green's function formalism above does not make it possible to distinguish between the radiative and non-radiative decay. Instead

only the total coupling rate to the plasmonic mode can be calculated, given by

$$\gamma_r + \gamma_{nr} = \frac{2\omega^2}{\hbar\varepsilon_0 c^2}[\hat{\mathbf{n}}_p \cdot \mathrm{Im}\{\mathbf{G}(\mathbf{r},\mathbf{r})\} \cdot \hat{\mathbf{n}}_p]|\mathbf{p}|^2. \tag{4.8}$$

To determine the radiative and nonradiative rates in the nanocavity explicitly, the Ohmic losses in the plasmonic structure are first calculated directly by integrating the absorbed power over the entire simulation domain Ω

$$\gamma_{nr} = \frac{1}{2}\frac{\gamma_r^0}{W_r^0}\int_\Omega \mathrm{Re}\{\mathbf{J}^* \cdot E\}dV. \tag{4.9}$$

Here $\mathbf{J}$ is the current density in the metal induced by the oscillating molecular dipole, and $W_r^0 = \omega^4|\mathbf{p}|^2/(12\pi\varepsilon_0 c^3)$ is the total power radiated by a dipole in free space. From knowledge of the non-radiative losses, it is then possible to calculate the radiative decay rate using $\gamma_r = \gamma_{sp} - \gamma_{nr} - \gamma_{\mathrm{int}}^0$. Alternatively, the radiative rate could be calculated by integrating the emission from the dipole over the upper half sphere of the simulation domain, but this approach makes it difficult to separate the contribution from surface plasmons that are launched by the nanocavity. In the present simulations the contribution from surface plasmon generation is not treated separately, but is rather included in the radiative decay channel.

The enhancement in the spontaneous emission rate for an emitter coupled to the nanocavity relative to free space or a control substrate is $\gamma_{sp}/\gamma_{sp}^0$. This value is often referred to as the Purcell factor, although this terminology has more precise meaning in the context of low-loss dielectric cavities in which the enhancement in the total spontaneous emission rate is equal to the enhancement in the radiative decay rate, γ_r/γ_r^0. In the context of emitters coupled to plasmonic structures, the Purcell factor in most works refers to the enhancement in the total rate. We follow this convention in the following sections. However, plasmonic structures often have large non-radiative losses, making a short decay lifetime an insufficient metric of the radiative performance of the nanocavity. Instead, it is also important to calculate the complementary metrics of radiative

rate enhancement and the radiative quantum yield, QY, as measures of how well the structure enhances and radiates the fluorescence into free space.

To simulate the quantum yield and Purcell enhancement, the emitters were modeled as monochromatic point dipoles in COMSOL Multiphysics. These dipole sources were swept in a grid pattern laterally across the nanocavity, calculating the Green's function and decay rates at each position. The same geometry was used to calculate both the scattering and absorption properties, and the emission rates. The four-fold symmetry of the structure was used to reduce the number of simulation points.

4.4. Summary of Fabrication

Here we give a general summary of the techniques used in the experiments described in this chapter, although the details of the fabrication varied slightly in each work. We refer readers to the individual papers for details on methods and also to a detailed visual version of the fabrication procedure.[45] The key features that enables fabrication of reproducible plasmonic nanocavities with small mode volumes is the use of colloidally-synthesized silver nanocubes and a nanometer-precision polymer spacer layer. The nanocubes, originally developed by Sun *et al.*,[46] are crystalline with atomically-flat facets of tunable size that form one side of the nanocavity. The nanocubes are chemically synthesized using CF_3COOAg as a precursor using the method described by Zhang *et al.*[47] As a byproduct of the synthesis, the nanocubes are coated with a $\sim$3 nm thick layer of poly(vinyl pyrrolidone) (PVP), which contributes to the volume of nanocavities formed with the nanocubes.

The nanocavities are fabricated by first depositing a 50–75 nm thin film of Au on a glass substrate by electron beam evaporation using a 5 nm Cr adhesion layer, or by creating a flat gold surface using template-stripping. The spacer layer between the nanocube and the metal, which constitutes the volume of the nanocavity, is typically composed of a polymer film assembled using a layer-by-layer dip coating technique. To form this polymer film, the metal films are

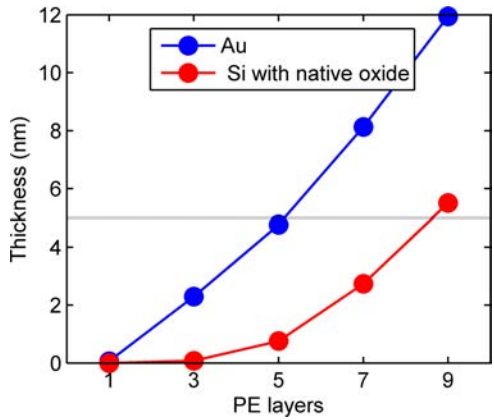


Fig. 4.7. Thickness of PAH/PSS polymer film as a function of number of PE layers on a glass surface and on a gold surface measured by ellipsometry. The PE layer number is defined as the total number of polymer solution dips, including PAH and PSS.

immersed in alternating positively and negatively charged polymer solutions of cationic poly(allylamine) hydrochloride (PAH) (3 mM) and anionic polystyrenesulfonate (PSS) (3 mM) for 5 minutes each. The samples are rinsed with water between each polymer deposition. The alternating charge of each polyelectrolyte (PE) layer insures that only a single monolayer of polymer is deposited at each step. The thickness of the spacer layer is determined by the number of PE layers used and the thickness is measured using variable angle spectroscopic ellipsometry. Each PAH-PSS bilayer contributes an approximate thickness of 2 nm, although initial layers are thinner due to lower adhesion to the gold surface (Figure 4.7). This vertical assembly technique for forming the cavity spacer gives nanometer control of the nanocavity dimensions, which is not possible with lithographic techniques. The final layer of the polymer films is always the positively charged PAH layer to promote adhesion of the negatively charged nanocubes. In order to integrate molecular emitters into the cavity, the polymer films are immersed in an aqueous solution of Ru dye, which intercalate into the top layers of the polymer film. Alternatively, the dielectric spacer can be formed by atomic layer deposition

of an oxide such as aluminum oxide, or the emissive material itself, in particular semiconductor quantum dots, can act as the spacer. To complete the nanocavities, a solution of silver nanocubes is deposited on the polymer film allowing them to attach to the PAH layer, and the residual nanocubes are rinsed away with water. The concentration of nanocubes in solution and the deposition time is controlled such that the final surface density gives a mean separation of $>3\,\mu$m, allowing individual nanocavities to be addressed optically with a focused laser beam. The nanocubes can also be deposited at a high density to form perfectly absorbing metasurfaces.[48]

4.5. Probing the Mechanisms of Purcell Enhancement

In this section, we describe experiments aimed at understanding the ability of the nanocavity to control the radiative dynamics of quantum emitters. Measuring large Purcell factors with conventional emitters such as fluorescent molecules and semiconductor quantum dots is challenging due to their short intrinsic fluorescence lifetimes of 1–10 ns. According to simulations, which will be discussed below, the nanocavity is expected to produce Purcell enhancements of 4500-fold for optimally coupled emitters. This Purcell enhancement would result in a spontaneous emission lifetime of several picoseconds, which is beyond the resolution limit of single photon detectors with a temporal response of $\sim$30 ps. This temporal resolution barrier was encountered in the time-resolved measurements in the work by Rose *et al.*[33]

To overcome this challenge, we integrated into the nanocavity a phosphorescent ruthenium metal complex which has an intrinsically long lifetime of $\tau_0 = 600$ ns.[29] The nanocavities were excited through an objective lens non-resonantly with the nanocavity at $\lambda_{ex} = 535$ nm with a pulsed laser (150 fs pulse duration) at a repetition rate of 20 MHz (Figure 4.8(a)). The emission was collected by the same objective lens, passed through a long-pass filter and detected using time-resolved single photon counting. Only nanocavities with a resonance of $\lambda_{cav} \approx 620-650$ nm were selected, regardless of the spacer layer thickness. The transition dipole moment of the molecules was

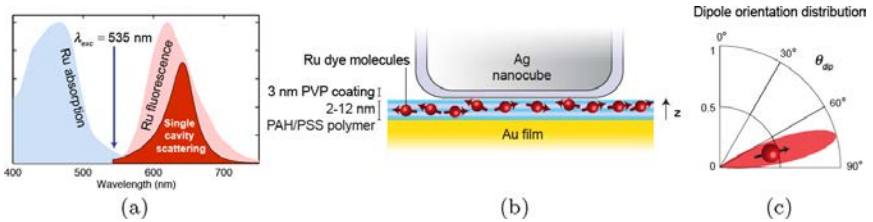


Fig. 4.8. (a) Absorption and emission spectra of the Ru dye, showing overlap between the emission spectrum and the nanocavity scattering spectrum. (b) Schematic of the nanocavity structure with embedded Ru dye molecules. The arrows represent the orientation of their transition dipole moments. (c) Measured distribution of orientations of the Ru transition dipole moments.

found to be oriented at $\sim 75°$ relative to normal, as is illustrated in Figures 4.8(b–c). The orientation was determined using angle- and polarization-resolved fluorescence measurement of the dye on a dielectric substrate.[29] This orientation distribution was found to be critical to accurately model the emission dynamics, as will be discussed below.

Figure 4.9(a) shows a representative emission decay curve from a nanocavity with an 8 nm total gap thickness (including the PVP layer surrounding the nanocube). In contrast to the long intrinsic decay with a lifetime of $1/\gamma_{sp}^{0} = \tau_0 = 600$ ns on glass, the nanocavity displays a much faster decay with a highly non-exponential time dependence. The fast initial decay is attributed to molecules that are optimally positioned within the cavity and have the most favorable angular orientation. The non-exponential decay occurs because each molecule coupled to the nanocavity experiences a different emission rate depending on its position and orientation. To obtain the underlying distribution of rates, the decay curve can be fitted to a stretched exponential function, a standard technique for analyzing time-resolved fluorescence in complex environments.[49] The decay constant for the stretched exponential fit is found to be $1/\gamma_{sp}^{\max} = \tau_{\max} = 0.7$ ns, which characterizes the fastest decay rate present in the distribution of rates. This maximum decay corresponds to a Purcell factor of $\gamma_{sp}^{\max}/\gamma_{sp}^{0} = 860$. A Laplace transform is then applied to the stretched exponential to obtain the entire distribution

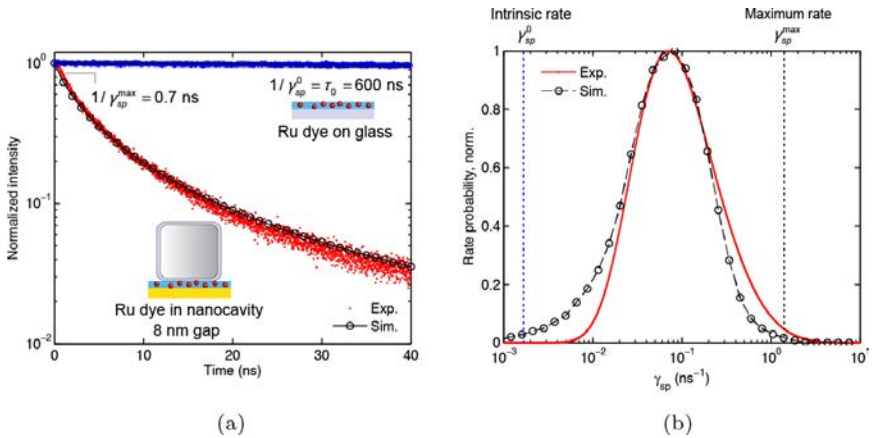
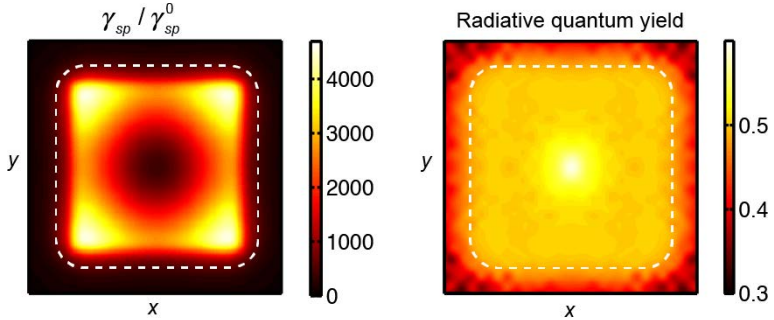


Fig. 4.9. (a) Measured and simulated decay dynamics from Ru dye molecules coupled to a nanocavity with an 8 nm spacer thickness. Also shown is the intrinsic decay of Ru dye on glass, having a lifetime of 600 ns. (b) Distribution of decay rates obtained from time-resolved emission data in (a). Experimental distribution is obtained from a stretched exponential fit, while the simulated distribution is obtained from the spatial map of the spontaneous emission decay rate.

of decay rates (Figure 4.9(b)). This distribution shows that while a small fraction of molecules are optimally coupled to the cavity, the most likely emission occurs with a Purcell factor of ∼60.

An important parameter in correctly simulating the emission dynamics is the orientation of the emitter transition dipole moment in the cavity. The orientation is critical because the emitter-cavity coupling scales with $\cos^2\theta$, where θ is the angle between the dipole direction and the electric field direction. In many studies of plasmon-enhanced fluorescence, the distribution of orientations is assumed to be isotropic[12,17] or to have an optimal orientation.[50] In our work, an experimental technique was used to directly measure the distribution of Ru dipole orientations. To measure the dipole orientation, a sample was prepared consisting of a 5 PE layer polymer film embedded with Ru dye deposited on a Si wafer with a ∼1000 nm thermal oxide layer. Fluorescence from this sample was measured as a function of incidence angle, observation angle, and polarization using a goniometer setup as described in Ref. 51. Based on these measurements, it



(a) (b)

Fig. 4.10. Simulated map of the (a) spontaneous emission rate enhancement and the (b) radiative quantum yield, as a function of position in a nanocavity with an 8 nm spacer and a resonance of $\lambda_{\text{cav}} = 650$ nm. The dipole orientation is assumed to be vertical, along the direction of the dominant field. White dashed outlines represent lateral extent of the nanocube.

was determined that most of the molecular emission dipoles are oriented at 75° relative to the surface normal, with a distribution width of ±15°.

In order to predict the emission dynamics, the structure was simulated following the procedure outlined in Sections 4.3. Figure 4.10 shows the spontaneous emission rate enhancement, $\gamma_{sp}/\gamma^0_{sp}$, and the radiative quantum yield, QY, as a function of position in a nanocavity with an 8 nm total gap thickness and a $\lambda_{\text{cav}} = 650$ nm resonance. The largest Purcell enhancement from these simulations is ∼4,500 for a vertically oriented dipole located near the corners of the nanocavity. The exponential decay for a molecule located at a position $\mathbf{r}$ in the nanocavity and oriented at an angle θ relative to normal is given by

$$I(\mathbf{r}, \theta, t) \propto \gamma_r(\mathbf{r}) \exp[-\gamma_{sp}(\mathbf{r}) \cos^2 \theta \, t]. \tag{4.10}$$

By performing averaging over all dipole positions and over the measured orientation distribution, it is possible to predict the emission dynamics from the nanocavity. Averaging was also performed over the vertical positions of the emitters, which were assumed to be in the top 2 nm of the polymer film. The calculated dynamics for a cavity with an 8 nm gap shows excellent agreement with the

experimental data (Figure 4.9(a–b)). To obtain these decay curves, no fit parameters were used other than normalization. The simulated emission rates can be plotted as a distribution of rates, which also shows good agreement with the distribution extracted from the stretched exponential fit. As seen in Figure 4.9(b), a fraction of molecules are expected to have Purcell factors of ∼1000.

To further demonstrate control over spontaneous emission, nanocavities were fabricated with gap thicknesses ranging from 5 to 15 nm. Figure 4.11(a) shows the measured time-resolved emission from a representative nanocavity for each gap thickness. With decreasing gap thickness a strong decrease in the initial emission lifetime is observed. The maximum Purcell enhancement reaches a value of ∼2000 for the $d = 5$ nm gap thickness (Figure 4.11(b)). The most likely rate enhancement, $\gamma^*_{sp}/\gamma^0_{sp}$ shows a similar trend, but with lower values due to the non-optimal position of most emitters.

A critical question to answer in metal-enhanced fluorescence experiments is whether short emission lifetimes, as those observed

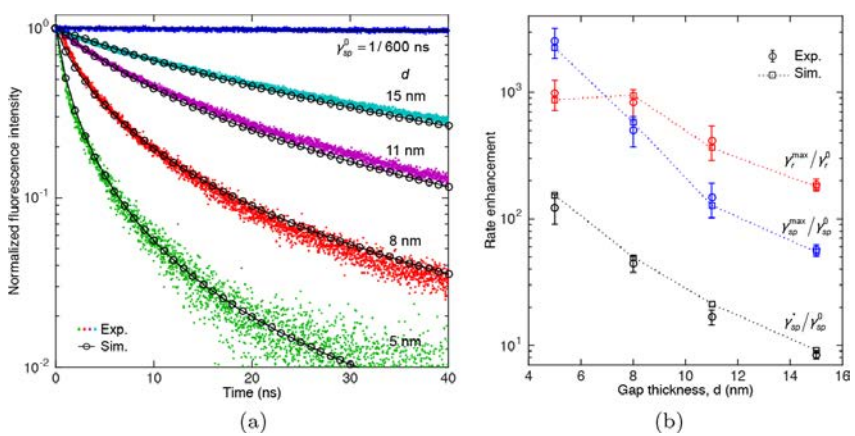


Fig. 4.11. (a) Time-resolved emission from four representative nanocavities with varying gap sizes from 5 to 15 nm. Simulated emission dynamics (black lines with open circles) show good agreement with the experiment. (b) Enhancement in the maximum spontaneous emission rate $\gamma^{\max}_{sp}/\gamma^0_{sp}$, enhancement in the mostly likely rate $\gamma^*_{sp}/\gamma^0_{sp}$, and enhancement in the maximum radiative rate $\gamma_r^{\max}/\gamma_r^0$, as a function of gap thickness.

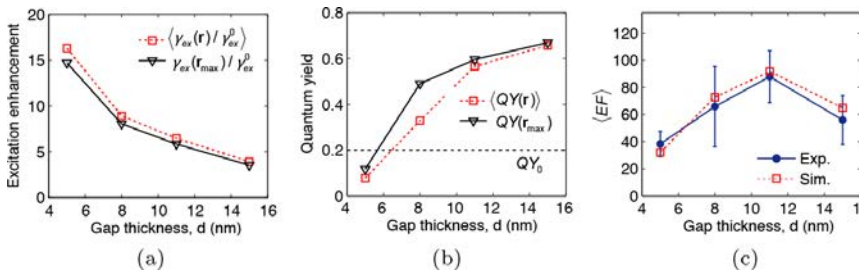


Fig. 4.12. Simulated spatially averaged enhancement in the (a) excitation rate at $\lambda_{ex} = 535$ nm, (b) the quantum yield at the emission wavelength of $\lambda = 650$ nm (red squares). Also shown is the excitation enhancement and quantum yield at the position of maximum Purcell enhancement (black triangles). (c) The total fluorescence enhancement factor obtained from simulations and from experiments, as a function of gap thickness.

in this system, are the result of non-radiative quenching or radiative rate enhancement. While radiative quantum yield cannot be measured directly, it can be inferred from a combination of simulations and a measurement of the fluorescence enhancement factor $\langle EF \rangle$. This factor is defined as the enhancement in fluorescence intensity from a single molecule in the nanocavity relative to the same molecule on glass. Figures 4.12(a–b) shows the calculated excitation rate enhancement and the quantum yield as a function of gap thickness, averaged over the nanocavity. The simulated quantum yield is highest for larger gap sizes, with quenching occurring for gaps $d < 8$ nm. The intrinsic quantum yield of the dye in the polymer film is assumed to be 20%. Using the calculated collection efficiency from the nanocavity of $\eta = 84\%$ and the collection efficiency from dipoles on glass of $\eta_0 = 15\%$, the spatially averaged fluorescence enhancement factor is calculated (Figure 4.12(c)). To obtain the experimental values for $\langle EF \rangle$, the same polymer films embedded with Ru dye were deposited on a glass substrate and the fluorescence was measured under the same experimental conditions as the nanocavities. The experimentally measured enhancement factor as a function of gap thickness is shown in Figure 4.12(c), normalized for measurement area. The excellent agreement between the measured and predicted enhancement factor implies the accuracy of the

constituent simulations — namely the excitation rate enhancement and radiative quantum yield. Consequently, knowing the quantum yield of the nanocavities, which is >50% for gap sizes d >8 nm, it is now possible to calculate the radiative rate enhancement based on the measured total emission rate. Using the quantum yield values, Figure 4.11(b) shows that the maximum radiative rate enhancement of $\gamma_r^{\max}/\gamma_r^0 \approx 1000$ occurs for gap sizes of $d = 5$ and 8 nm. These large radiative rate enhancements are achieved due to the tightly controlled geometry of the nanocavity, which simultaneously provides large field enhancement and high quantum efficiency due to efficient outcoupling. The experimental realization is enabled by the bottom up fabrication method which gives nanometer scale control over the spacer thickness and nanocube size. The inclusion of the slow emitting Ru dye into the structure allowed for full characterization and understanding of the radiative and non-radiative rates which would not have been possible with fast intrinsic emitters due to experimental limitations.

4.6. Ultrafast Spontaneous Emission

Integrating Ru dye into the nanocavities demonstrated the capability of this plasmonic structure for achieving large Purcell enhancements while maintaining a high quantum yield due to the antenna action of the cavity. For future device applications such as modulated LEDs and single photon sources, it is important to develop ultrafast spontaneous emission sources based on stable and technologically relevant materials. One such set of materials is colloidal semiconductor quantum dots (QDs). These emitters have a wavelength-tunable room temperature emission, excellent photostability, and can be easily integrated with other materials due to their solution processability.[52] For example, QDs have been used as stable, room-temperature single photon sources,[53] but their intrinsic radiative lifetime of ∼20 ns limits the obtainable repetition rate. Likewise, light emitting diodes are not suitable for high data rate telecommunications due to the slow switching time of LEDs, limited by the spontaneous emission lifetime.

Various plasmonic structures have been used over the past several years to control the spontaneous emission from QDs, but so far Purcell factors of at most 145 have been demonstrated.[17,26,54–56] However, these Purcell enhancements are often accompanied by large non-radiative losses. For example, structures assembled by atomic force microscopy consisting of a QD coupled to a Au nanoparticle shows a Purcell factor of up to 145, but the radiative decay enhancement is only $\sim 8\times$.

In our recent work,[34] we demonstrated coupling of QDs to the nanocube nanocavities, resulting in ultrafast spontaneous emission with a lifetime of $<$11 ps, corresponding to Purcell factors of 880. While a similar Purcell enhancement was achieved in this work as compared to the work on Ru dye enhancement, the enhancement of QDs results in an absolute lifetime in the picosecond time regime. This produces a source of spontaneous emission that operates at ultrafast frequencies and with high radiative efficiency. The structures are produced by first depositing a single PAH layer on a 50 nm gold film which is evaporated on a glass substrate. CdSe/ZnS QDs core-shell QDs, with an emission spectrum centered at $\lambda_{em} = 620\,\text{nm}$ are spin coated on a gold film coated with PAH at a density such that $\sim$10 QDs are located in each nanocavity. The QD coated sample is then immersed in a solution of nanocubes which adhere to the top of the QDs, completing the nanocavities (Figure 4.13(a)). The nanocavities with a resonance of $\lambda_{\text{cav}} = 625{-}635$ nm were identified using darkfield microscopy.

To demonstrate the overall fluorescence enhancement of the QDs, the nanocavities were excited with a continuous-wave laser non-resonantly at $\lambda_{\text{ex}} = 514$ nm. Figure 4.13(b) shows the fluorescence intensity as a function of excitation power from three representative nanocavities as well as a control sample consisting of the QDs on glass. The fluorescence intensity from QDs coupled to the nanocavity is substantially higher than from QDs on the control sample. A linear dependence between excitation power and fluorescence intensity is observed over the power regime probed. At higher excitation powers, permanent photobleaching of the QDs occurs, before the saturation

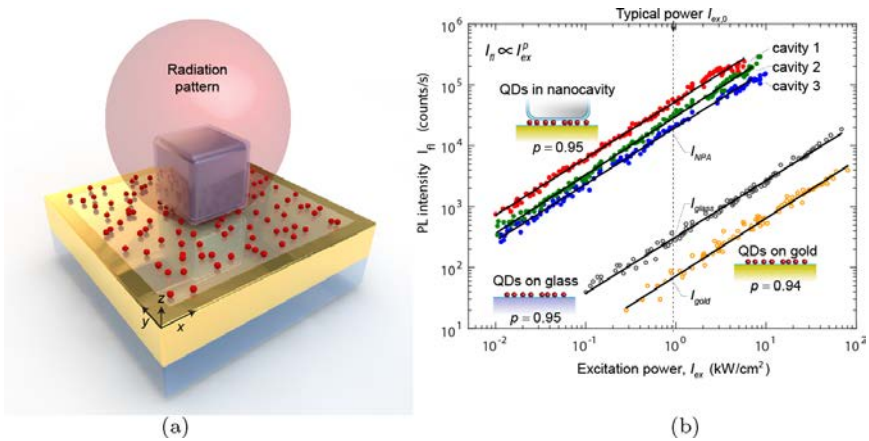


Fig. 4.13. (a) Schematic structure of colloidal QDs integrated into the nanocavity, showing the directional radiation pattern of the structure. (b) Fluorescence intensity as a function of excitation power with a continuous-wave laser for QDs coupled to three representative nanocavities, as well as for QDs on glass, and QDs on a gold film. Fits are to power laws, where p is the exponent. The vertical dashed line indicates the power at which the time-resolved measurements were performed.

regime can be reached. QDs that are on the gold film but not coupled to a nanocavity show fluorescence quenching of 70% as compared to the control sample. The spatially averaged fluorescence enhancement factor is found to vary from 177 to 2,300 with an average value of $\langle EF \rangle_{avg} = 831$. This variation in $\langle EF \rangle$ is attributed to: (1) the random spatial distribution of the small number of QDs ($\sim$10) within the nanocavity, with QDs near the cavity edges experiencing higher field enhancements (Figure 4.3); and (2) the random orientation of the QD transition dipoles, with optimally oriented QDs experiencing larger enhancement. As with the Ru molecular dyes, this fluorescence enhancement is attributed to an increased excitation rate, modified quantum yield, and a higher collection efficiency. To compare with the predicted enhancement factor, the structure was simulated using the method described in Section 3. Assuming an intrinsic quantum yield for QDs on glass of 10%, the spatially averaged enhancement factor was found to be $\langle EF \rangle = 660$, which is in good agreement with

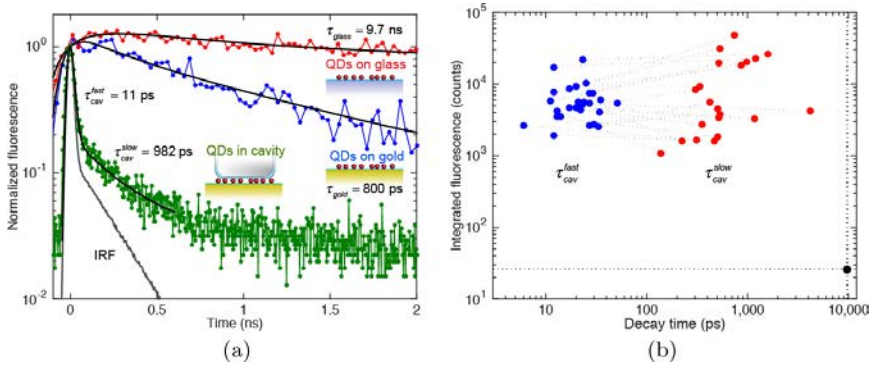


Fig. 4.14. (a) Normalized time-resolved emission from QDs coupled to a nanocavity (green), QDs on glass (red), and QDs on gold without nanocubes (blue). The instrument resosponse function (IRF) is shown in gray. A bi-exponential function convoled with the IRF is used to fit the nanocavity emission, while a single exponential convolved with the IRF is used to fit emission from the other two samples. (b) Fast and slow components of the fluorescence decay times for ∼30 nanocavities. Some decays showed more robust fits to a single exponential function, in which case the slow component is absent.

the experimentally obtained value for ∼10 different nanocavities of $\langle EF \rangle_{avg} = 831$. This agreement indicates that the high quantum yield of ∼50% predicted by simulations is close to the value observed in the experimental system.

Having established the emission efficiency, we now turn to the emission dynamics. Figure 4.14(a) shows the time-resolved emission from a nanocavity with embedded QDs under pulsed excitation at $\lambda_{ex} = 535$ nm, as well as emission from QDs on glass. The control sample shows emission with a single-exponential lifetime of $\tau_{glass} = 9.7$ ns. For QDs on the gold film without the presence of the nanocubes, the emission has a lifetime of $\tau_{gold} = 0.8$ ns, but with a substantially reduced intensity, as seen in Figure 4.13(b). When the QDs are coupled to the nanocavity, the emission is dramatically shorter than from the control sample, which occurs concurrently with a large intensity increase (Figure 4.13(b)). The spatial and orientational distribution of the QDs is expected to produce a non-exponential decay curve, as observed with Ru dye coupled to the nanocavity.[29] However, the instrument response function

of the avalanche photodiode used to detect the fluorescence has a ∼30 ps resolution, which prevents the extraction of the entire decay distribution. Instead, a bi-exponential function convolved with the instrument response function is used to extract the dominant decay components. This fitting results in a fast decay lifetime of $\tau_{NPA}^{fast} = 11$ ps and slow decay time of $\tau_{NPA}^{slow} = 981$ ps. The fast component, which contributes approximately 50% of the emission, corresponds to a Purcell factor of $F_p = \tau_{glass}/\tau_{NPA}^{fast} = 880$. This Purcell factor is only a lower bound due to the limits of the detector and the amount of deconvolution that can be performed. The lifetimes and the relative fluorescence contributions of the fast and slow components were measured for ∼30 other nanocavities (Figures 4.14(b–c)). All nanocavities showed a fast decay lifetime between 11 and 51 ps, corresponding to Purcell factors of 190 to 880. Some nanocavities were found not to have a slower component, perhaps due to the absence of QDs that are poorly coupled to the cavity. On average the fast and slow component contribute equal amounts of fluorescence, when integrated over time.

The simulated Purcell factor was obtained by calculating the Purcell factor at each position for a randomly orientated QD dipole, and then by performing averaging over all positions. This results in an average Purcell enhancement of $F_p = 990$, corresponding to a lifetime of ∼10 ps, which agrees well with the measured fast decay lifetime. However, the simulations predict Purcell enhancements of to 4,000 × for optimally positioned and oriented QDs. This suggests that such large enhancements are likely occurring in the experimental system but are not resolved by the current fluorescence detection system. Overall, this work demonstrates that these nanocavities can be used to generate spontaneous emission on the pico-second time scale with high efficiency. This emission lifetime points towards the possibility of ultrafast (>100 GHz) electrical modulation of such a spontaneous emission source, which is on the order of or beyond the capabilities of semiconductor lasers but without the threshold requirements associated with stimulated emission sources.

4.7. Ultrafast Single Photon Generation

The Purcell enhancement of QDs can also be extended to the regime of single photon generation by coupling individual QDs to a nanocavity.[31] Using the same fabrication procedure as in the previous section, single QD coupling can be achieved by simply reducing the QD concentration during deposition. By chance, when nanocubes are deposited, certain nanocavities will have exactly one QD coupled to the gap, and a fraction of these will have an optimally positioned QD near one of the corners. Individual nanocubes to be used for detailed experiments are first identified using wide field fluorescence imaging. Those nanocavities exhibiting bright photoluminescence (PL) are selected for single particle measurements. In particular, to identify cavities with single QDs, second order photon correlation measurements are performed using two single photon detectors using continuous wave (cw) excitation at a wavelength of 488 nm. Figure 4.15(a) shows the photon correlation function $g^2(t)$ from a typical nanocavity exhibiting anti-bunching behavior at $t = 0$, indicating that a single

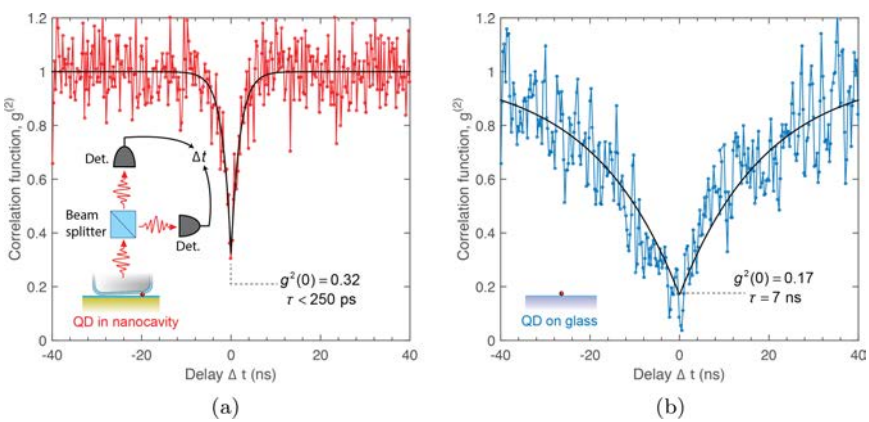


Fig. 4.15. (a) The second order photon correlation function $g^2(t)$ from a single QD coupled to a nanocavity. Inset shows the measurement configuration using a beamsplitter and two time-resolved single photon detectors. The correlation function has a minimum value of $g^2(0) = 0.32$ indicating single photon emission, with a binning-limited lifetime of $\tau < 250$ ps. (b) The same correlation measurement for a single QD on glass, also showing anti-bunching behavior ($g^2(0) = 0.17$) but with a longer lifetime of $\tau = 7$ ns.

QD is coupled to the cavity and is behaving like a quantum emitter. The lifetime of the anti-bunching decay is <250 ps, limited by the time bin size, and sets an upper bound on the QD excited state lifetime. Similar correlation measurements were performed for ∼12 other nanocavities coupled to single QDs. In comparison, a single QD on glass shows similar anti-bunching behavior, but with a much longer lifetime of 7 ns (Figure 4.15(b)). These measurements hence give a lower bound for the Purcell enhancement for single QDs of $F_p > 28$, limited by the binning time used in this experiment.

To fully resolve the emission dynamics and determine the true Purcell factor, time-resolved fluorescence measurements were performed on the same nanocavities. The same nanocavity as shown in Figure 4.15(a) is excited with a pulsed laser at a wavelength of 535 nm and shows a biexponential PL decay with a fast lifetime of $\tau_{fast} = 13$ ps and a slow component of $\tau_{fast} = 680$ ps (Figure 4.16(a)). These lifetimes are obtained from a fit of the data to a biexponential function convolved with the measured instrument response function, which has a width of 30 ps (Figure 4.16(a)). From

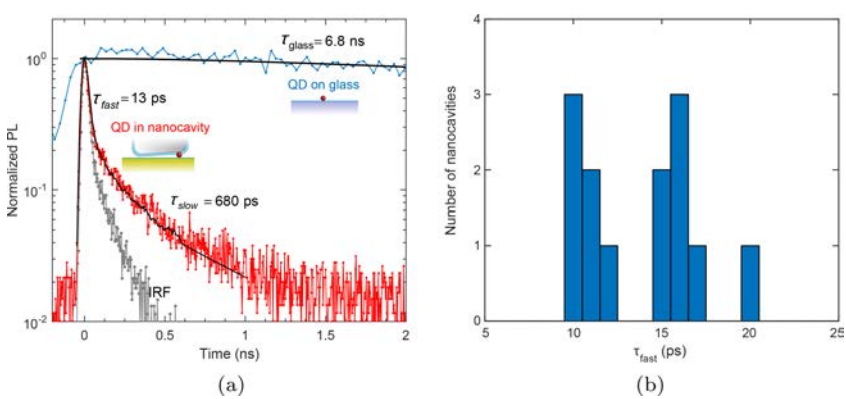


Fig. 4.16. (a) Time-resolved PL from the same single QD as in Figure 4.16, showing a biexponential decay, with a fast lifetime of 13 ps (extracted through deconvolution) and a slow lifetime of 680 ps. The fast lifetime is limited by the instrument response function (IRF) of the detector, show in gray. The lifetime of a single QD on glass is 6.8 ns. (b) Histogram of the fast lifetime components for 13 nanocavities coupled to single QDs, as verified by anti-bunching measurements.

the fit it was determined that a majority of the emission (>97%) originates from the fast component. In contrast, a single QD on glass shows single exponential decay with a lifetime of 6.8 ns, as determined by averaging over several quantum dots. The fast component of the coupled QD decay, which is still limited by the resolution of the detector, gives a lower bound to the maximum Purcell factor of $F_P > 540$. The increase in the radiative rate can then be calculated by using the estimated quantum yield for a QD on glass of $QY_0 = 20\%$ (based on integrating sphere measurements) and the simulated quantum yield of the nanocavity of $QY = 50\%$. This results in a radiative rate enhancement of $\gamma_r/\gamma_r^0 = (\tau_{glass}/\tau_{fast})(QY/QY_0) = 1350$. This number is slightly below the predicted value from simulations of $F_P = 2000$ for an optimally positioned, randomly oriented dipole. Similar decay dynamics were observed for ∼12 other nanocavities coupled to single QDs (Figure 4.16(b)), as confirmed by photon correlation measurements.

As with ensemble emitter measurements discussed earlier, a key consideration is whether the short emission lifetimes are due to radiative or non-radiative enhancement. To address this, the time-integrated fluorescence from a coupled QD and a QD on glass was measured under cw excitation at a wavelength of 488 nm as shown in Figure 4.17. The coupled QD shows a dramatic enhancement of 1,900-fold enhancement in the PL intensity at an excitation power of 1000 nW. Both the coupled and uncoupled QD show a linear dependence between excitation power and PL intensity, for powers below ∼2000 nW. The large PL enhancement is attributed to three factors: (1) enhancement in the collection efficiency due to the antenna, (2) enhancement in the excitation rate, and (3) enhancement in the quantum yield. Simulating all three effects, including spatial and orientation averaging of the emitters, results in a predicted enhancement factor per QD of $EF = 2500$, which is in qualitative agreement with the measured value. This agreement indicates that the quantum yield of the experimental structure is indeed high at ∼50%, as predicted by simulations.

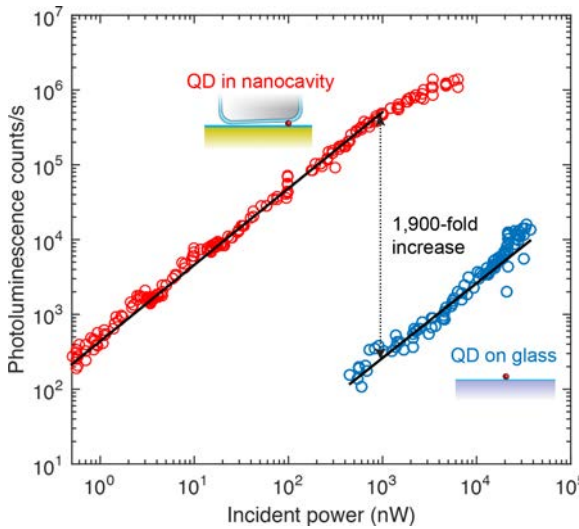


Fig. 4.17. Time-integrated PL intensity as a function of incident power for two cases: a single QD coupled to a nanocavity (shown in red) and a single QD on glass (shown in blue), performed using cw excitation. The nanocavity produces an enhancement of 1,900-fold relative to the QD on glass.

The maximum detected count rate obtained from a single enhanced QD was ∼1 MHz. Given that the collection, transmission and detection efficiency of the optical setup is ∼1%, the actual emitted count rate is estimated to be ∼100 MHz. Saturation of the QD was not possible because irreversible photodamage occurred at intensities below the saturation level. However, given improvements in material stability and possible encapsulation, the maximum count rate expected from a single QD could be as high as 80 GHz, additionally assuming that multiphoton emission effects can be avoided. Such high count rates could be achieved by integrating more stable quantum emitters into the nanocavities, such as fluorescence crystal color centers in diamond[57] or silicon carbide.[58]

4.8. From Nanocavities to Metasurfaces

Thus far we have shown how individual nanocavities can be used to control the radiative properties of luminescent emitters. For a

number of applications, such as wavelength conversion and solid state lighting, emission must be controlled over microscopic and macroscopic areas. Due to the reproducible nature of our fabrication method and the good size uniformity of the nanocubes, the film-coupled nanocubes can be assembled into large scale metasurfaces consisting of many nanocavities. As has been shown recently,[36,48] these nanocube-based metasurfaces can act as nearly-perfect spectrally selective absorbers. This absorption phenomenon arises from destructive interference between reflection from the metal ground plane and reflection (scattering) from the nanocavities. In addition to acting as absorbers, these surfaces can act as large-area luminescent emitters for future applications in wavelength conversion and light sources. In this context, we review our recent results on large-area metasurface absorbers.

Figure 4.18(a) shows an SEM image of a typical metasurface composed of 75 nm nanocubes, with a mean spacing of ∼200 nm. The metasurfaces are fabricated by the same colloidal, solution-based

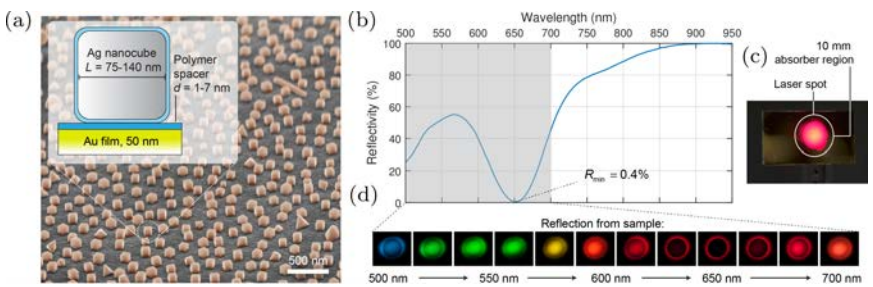


Fig. 4.18. (a) SEM image of perfect absorber surface consisting of a gold film covered with silver nanocubes. Inset shows a cross-sectional schematic of the sub-wavelength resonators. (b) Reflection spectrum from an absorber with a resonance at 650 nm, utilizing 75 nm nanocubes and a 7 nm spacer, showing a minimum reflectivity of 0.4%. (c) Image of sample with a 10 mm region coated with nanocubes on top of gold, while surrounding region is coated only with gold. A defocused laser with a diameter of ∼13 mm illuminates the absorber region. (d) Images of the laser beam on a screen after reflection from the absorber sample shown in (c), as the laser wavelength is swept from 500 to 700 nm. When the laser is resonant with the absorption at 650 nm, no reflection is seen from the nanocube coated region.

method as described for assembly of individual nanocavities. The only exception is that the concentration of nanocubes in the deposition solution is increased such that the final nanocavity surface coverage is ∼20%. Despite the non-periodic arrangement, the metasurface exhibits narrowband and nearly-perfect absorption at a center wavelength of 650 nm (Figure 4.18(b–d)), with a minimum reflection of only 0.4%. To visually demonstrate the strong absorption, a defocused laser spot was used to illuminate a sample having a 10 mm round perfect absorber area. As the laser wavelength is tuned on-resonance with the perfect absorber metasurface, the reflection from the sample exhibits a black region indicating full absorption, while the areas outside the metasurface region show full reflection.

As with individual nanocavities, the resonance of the entire metasurface can be controlled by changing the size of the nanocube and the thickness of the polymer spacer layer. Figure 4.19(a–b) shows

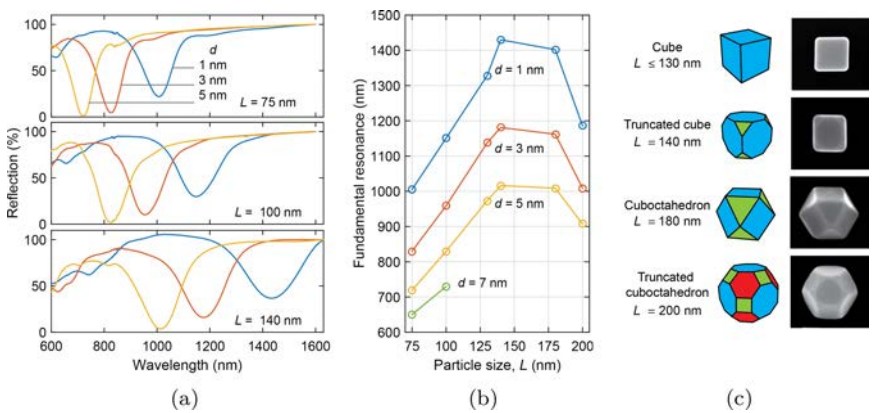


Fig. 4.19. (a) Reflection spectra of metasurfaces with resonances ranging from the visible to the near infrared using a range of nanocube sizes (L) and spacer layer thicknesses (d). (b) Resonance wavelength as a function of particle size and spacer thickness for the metasurfaces in (a). (c) SEM images of typical particles from each synthesis having average particle size L. For L $\geq$ 140 nm the particles become non-cubic, transitioning to a truncated cube, to a cuboctahedron, to a truncated cuboctahedron. Schematics of each shape are also shown. As the particle size increases, the facets become smaller, which blue-shifts the absorber resonance, seen in (b).

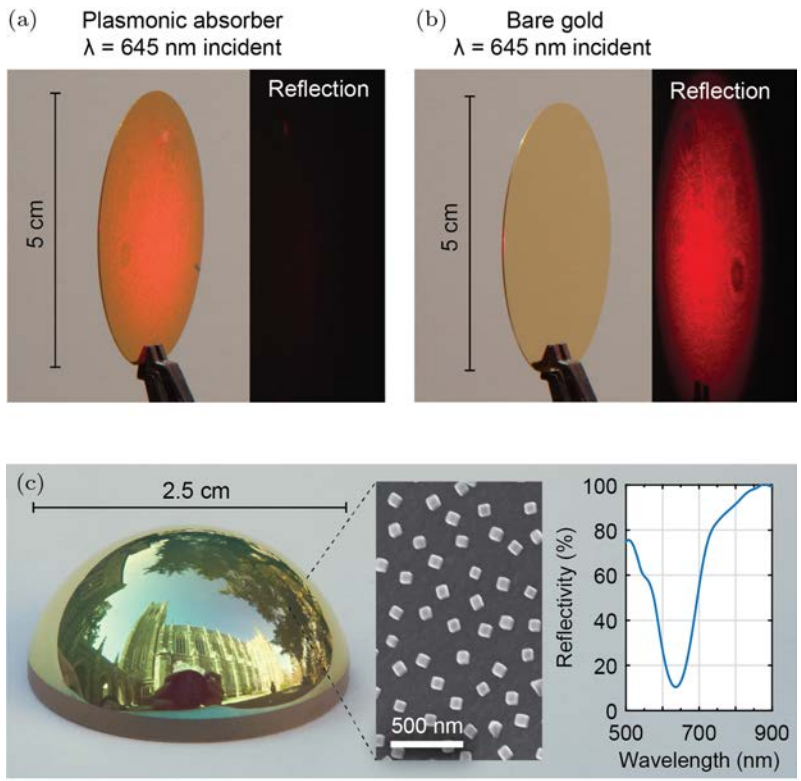


Fig. 4.20. (a) A 2-inch diameter wafer coated over the entire surface with a gold film, a 7 nm spacer, and 75 nm nanocubes, with a resonance of 645 nm. A defocused 645 nm laser beam is incident on the sample and the reflection is imaged on a screen, showing almost no visible reflection. (b) For comparison with (a), a 5 cm wafer coated only with a gold film shows strong reflection of the laser beam, imaged using the same camera exposure settings. (c) A glass half-sphere coated with 100 nm of silver and a 7 nm PAH-PSS spacer layer, followed by conformal deposition of 75 nm nanocubes. Insets show SEM image taken on the side slope of the sample, and the near-normal incidence reflectance spectrum showing ∼90% absorption.

that the resonance can be tuned from the visible to the near-IR spectrum using nanocubes ranging from 75 to 140 nm and spacer thicknesses from 1–5 nm. The best absorption performance is achieved from thicker spacer layers. The resonances cannot be moved beyond ∼1400 nm because larger nanocubes cannot be synthesized using

our current procedure. As the nanocube synthesis reaction time is increased, the particles develop new facets, decreasing the effective area of the cavity under the nanocube even though the total diameter of the particle is larger (Figure 4.19(c)).

The fabrication method for forming these metasurfaces is entirely solution-based, including the spacer and nanocube deposition. As a result, the metasurfaces can be coated onto objects of arbitrary size and shape. Figures 4.20(a–b) show a two-inch wafer coated with the metasurface and for comparison a gold-coated wafer. A defocused laser beam at a wavelength of 645 nm is reflected from the bare gold substrate while the beam is almost completely absorbed by the metasurface. The metasurface can also be coated on non-planar objects, such as the half-sphere shown in Figure 4.20(c). The deposition of the spacer layer and nanocubes was done using dip coating, as with the planar wafer in Figures 4.20(a–b), showing the generality and scalability of the approach.

4.9. Conclusions and Outlook

In conclusion, the nanocavity based on film-coupled nanocubes is a powerful platform for controlling the radiative properties of emitters such as molecules, quantum dots, and two-dimensional semiconductors. The large Purcell enhancements are enabled by the small mode volume cavity and the efficient antenna action of the structure. The very good agreement of the experimental results with simulations is the result of excellent control of the nanocavity dimensions using bottom-up colloidal fabrication, in particular the nanoscale control over the gap thickness. Future work still remains to improve on the current theoretical understanding of the structure by accounting for possible losses from the nanocavity into surface plasmons.[59]

Even larger Purcell enhancements could be enabled by controlling the orientation and position of the emitter dipoles in the nanocavity using chemical or nanofabrication methods. For optimally positioned and oriented dipoles, Purcell enhancements of $\sim$4000 are possible, corresponding to terahertz-frequency modulation of fluorescence

using an emitter with a 10 ns intrinsic lifetime. This platform is also compatible with electric excitation, potentially allowing these structures to function as light-emitting diodes or as a high repetition-rate single photon source, if only one dipole is coupled to the cavity.

References

1. Purcell, E. (1946). Spontaneous emission probabilities at radio frequencies, *Phys. Rev.* **69**, 681.
2. Schietinger, S., Barth, M., Aichele, T., and Benson, O. (2009). Plasmon-enhanced single photon emission from a nanoassembled metal-diamond hybrid structure at room temperature, *Nano Lett.* **9**(4), 1694–1698.
3. Choy, J., Hausmann, B., Babinec, T., *et al.* (2011). Enhanced single-photon emission from a diamond-silver aperture, *Nat. Photonics* **5**, 738–743.
4. Esteban, R., Teperik, T. V., and Greffet, J. J. (2010). Optical patch antennas for single photon emission using surface plasmon resonances, *Phys. Rev. Lett.* **104**(2), 026802.
5. Yablonovitch, E. (1987). Inhibited spontaneous emission in solid-state physics and electronics, *Phys. Rev. Lett.* **58**(20), 2059–2062.
6. Claudon, J., Bleuse, J., Malik, N., *et al.* (2010). A highly efficient single-photon source based on a quantum dot in a photonic nanowire, *Nat. Photonics*, **4**, 174–177.
7. Hennessy, K., Badolato, A., Winger, M., *et al.* (2007). Quantum nature of a strongly coupled single quantum dot-cavity system, *Nature*, **445**(7130), 896–899.
8. Englund, D., Fattal, D., Waks, E., *et al.* (2005). Controlling the Spontaneous Emission Rate of Single Quantum Dots in a Two-Dimensional Photonic Crystal, *Phys. Rev. Lett.* **95**(1), 013904.
9. Gérard, J., Sermage, B., Gayral, B., *et al.* (1998). Enhanced spontaneous emission by quantum boxes in a monolithic optical microcavity, *Phys. Rev. Lett.* **81**(5), 1110–1113.
10. Bleuse, J., Claudon, J., Creasey, M., *et al.* (2011). Inhibition, enhancement, and control of spontaneous emission in photonic nanowires, *Phys. Rev. Lett.* **106**(10), 103601.
11. Kinkhabwala, A., Yu, Z., Fan, S., *et al.* (2009). Large single-molecule fluorescence enhancements produced by a bowtie nanoantenna, *Nat. Photonics*, **3** (11), 654–657.
12. Chen, Y., Munechika, K., and Ginger, D. S. (2007). Dependence of fluorescence intensity on the spectral overlap between fluorophores and plasmon resonant single silver nanoparticles, *Nano Lett.* **7**(3), 690–696.
13. Russell, K., Liu, T., Cui, S., and Hu, E. (2012). Large spontaneous emission enhancement in plasmonic nanocavities, *Nat. Photonics*, **6**, 459–462.

14. Anger, P., Bharadwaj, P., and Novotny, L. (2006). Enhancement and quenching of single-molecule fluorescence, *Phys. Rev. Lett.* **96**(11), 3–6.
15. Curto, A. G., Volpe, G., Taminiau, T. H., *et al.* (2010). Unidirectional emission of a quantum dot coupled to a nanoantenna, *Science*, **329**(5994), 930–933.
16. Kosako, T., Kadoya, Y., and Hofmann, H. (2010). Directional control of light by a nano-optical Yagi–Uda antenna, *Nat. Photonics*, **4**, 312–215.
17. Belacel, C., Habert, B., Bigourdan, F., *et al.* (2013). Controlling spontaneous emission with plasmonic optical patch antennas, *Nano Lett.* **13**(4), 1516–1521.
18. Noginov, M. A., Zhu, G., Belgrave, A. M., *et al.* (2009). Demonstration of a spaser-based nanolaser, *Nature*, **460**(7259), 1110–1112.
19. Oulton, R. F., Sorger, V. J., Zentgraf, T., *et al.* (2009). Plasmon lasers at deep subwavelength scale, *Nature*, **461**(7264), 629–632.
20. Brongersma, M. L., Halas, N. J., and Nordlander, P. (2015). Plasmon-induced hot carrier science and technology, *Nat. Nanotechnol.* **10**(1), 25–34.
21. Sobhani, A., Knight, M. W., Wang, Y., *et al.* (2013). Narrowband photodetection in the near-infrared with a plasmon-induced hot electron device, *Nat. Commun.* **4**, 1643.
22. Li, W., and Valentine, J. (2014). Metamaterial perfect absorber based hot electron photodetection, *Nano Lett.* **14**(6), 3510–3514.
23. Aćimović, S. S., Kreuzer, M. P., González, M. U., and Quidant, R. (2009). Plasmon near-field coupling in metal dimers as a step toward single-molecule sensing, *ACS Nano*, **3**(5), 1231–1237.
24. Kollmann, H., Piao, X., Esmann, M., *et al.* (2014). Toward plasmonics with nanometer precision: Nonlinear optics of helium-ion milled gold nanoantennas, *Nano Lett.* **14**(8), 4778–4784.
25. Yi, M., Zhang, D., Wang, P., *et al.* (2011). Plasmonic interaction between silver nano-cubes and a silver ground plane studied by surface-enhanced raman scattering, *Plasmonics*, **6**(3), 515–519.
26. Yuan, C., Wang, Y., and Cheng, H. (2013). Modification of fluorescence properties in single colloidal quantum dots by coupling to plasmonic gap modes, *J. Phys. Chem. C*, **117**(24), 12762–12768.
27. Lassiter, J. B., McGuire, F., Mock, J. J., *et al.* (2013). Plasmonic waveguide modes of film-coupled metallic nanocubes, *Nano Lett.* **13**(12), 5866–5872.
28. Ciracì, C., Hill, R. T., Mock, J. J., *et al.* (2012). Probing the ultimate limits of plasmonic enhancement, *Science*, **337**(6098), 1072–1074.
29. Akselrod, G., Argyropoulos, C., Hoang, T., *et al.* (2014). Probing the mechanisms of large Purcell enhancement in plasmonic nanoantennas, *Nat. Photonics*, **8**, 835–840.
30. Hoang, T. B., Akselrod, G. M., Argyropoulos, C., *et al.* (2015). Ultrafast spontaneous emission source using plasmonic nanoantennas, *Nat. Commun.* **6**, 7788.

31. Hoang, T. B., Akselrod, G. M., and Mikkelsen, M. H. (2015). Ultrafast room-temperature single photon emission from quantum dots coupled to plasmonic nanocavities, *Nano Lett.* **16**(1), 270–275.
32. Akselrod, G. M., Ming, T., Argyropoulos, C., *et al.* (2015). Leveraging nanocavity harmonics for control of optical processes in 2D semiconductors, *Nano Lett.* **15**(5), 3578–3584.
33. Rose, A., Hoang, T. T. B., McGuire, F., *et al.* (2014). Control of radiative processes using tunable plasmonic nanopatch antennas, *Nano Lett.* **14**(8), 4797–4802.
34. Hoang, T. B., and Akselrod, G. M. (2015). Ultrafast spontaneous emission source based on plasmonic nanopatch antenna, *Nat. Commun.* **6**, 7788.
35. Rose, A., Huang, D., and Smith, D. R. (2011). Controlling the second harmonic in a phase-matched negative-index metamaterial, *Phys. Rev. Lett.* **107**(6), 063902.
36. Moreau, A., Ciracì, C., Mock, J. J., *et al.* (2012). Controlled-reflectance surfaces with film-coupled colloidal nanoantennas, *Nature*, **492**(7427), 86–89.
37. Ciraci, C., Rose, A., Argyropoulos, C., and Smith, D. R. D. (2014). Numerical studies of the modification of photodynamic processes by film-coupled plasmonic nanoparticles, *J. Opt. Soc. Am. B*, **31**, 2601–2607.
38. Bowen, P. T., and Smith, D. R. (2014). Coupled-mode theory for film-coupled plasmonic nanocubes, *Phys. Rev. B*, **90**, 195402.
39. Ciraci, C., Britt Lassiter, J., Moreau, A., and Smith, D. R. (2013). Quasi-analytic study of scattering from optical plasmonic patch antennas, *J. Appl. Phys.* **114**(16), 163108.
40. Dionne, J. A., Lezec, H. J., and Atwater, H. A. (2006). Highly confined photon transport in subwavelength metallic slot waveguides, *Nano Lett.* **6**, 1928–1932.
41. Hill, M. T., Marell, M., Leong, E. S. P., *et al.* (2009). Lasing in metal-insulator-metal sub-wavelength plasmonic waveguides, *Opt. Express*, **17**(13), 11107–11112.
42. Gordon, R. (2006). Light in a subwavelength slit in a metal: Propagation and reflection, *Phys. Rev. B — Condens. Matter Mater. Phys.* **73**, 8–10.
43. Zhang, J., Fu, Y., Chowdhury, M. H., and Lakowicz, J. R. (2007). Metal-enhanced single-molecule fluorescence on silver particle monomer and dimer: Coupling effect between metal particles, *Nano Lett.* **7**, 2101–2107.
44. Richard, M., Kasprzak, J., Romestain, R., *et al.* (2005). Spontaneous coherent phase transition of polaritons in CdTe microcavities, *Phys. Rev. Lett.* **94**(18), 187401.
45. Hoang, T. B., Huang, J., and Mikkelsen, M. H. (2016). Colloidal synthesis of nanopatch antennas for applications in plasmonics and nanophotonics, *J. Vis. Exp.*, e53876, doi:10.3791/53876.
46. Sun, Y., and Xia, Y. (2002). Shape-controlled synthesis of gold and silver nanoparticles, *Science*, **298**(5601), 2176–2179.

47. Zhang, Q., Li, W., Wen, L.-P., *et al.* (2010). Facile synthesis of Ag nanocubes of 30 to 70 nm in edge length with CF(3)COOAg as a precursor, *Chem. Eur. J.* **16**(33), 10234–10239.
48. Akselrod, G. M., Huang, J., Hoang, T. B., *et al.* (2015). Large-area metasurface perfect absorbers from visible to near-infrared, *Adv. Mater.*,**27**, 8028-8034.
49. Berberan-Santos, M. N., Bodunov, E. N., and Valeur, B. (2005). Mathematical functions for the analysis of luminescence decays with underlying distributions 1. Kohlrausch decay function (stretched exponential), *Chem. Phys.* **315**(1–2), 171–182.
50. Cang, H., Liu, Y., Wang, Y., *et al.* (2013). Giant suppression of photobleaching for single molecule detection via purcell effect, *Nano Lett.* **13**(12), 5949–5953.
51. Barritault, P., Gétin, S., Chaton, P., *et al.* (2002). Determination of surface-bound-fluorophore orientation by goniometric fluorescence polarization: Application to quantification of DNA-chip readouts, *Appl. Opt.* **41**(22), 4732–4738.
52. Kim, J. Y., Voznyy, O., Zhitomirsky, D., and Sargent, E. H. (2013). 25th anniversary article: Colloidal quantum dot materials and devices: A quarter-century of advances, *Adv. Mater.* **25**, 4986–5010.
53. Vittorio, M. De, Pisanello, F., Martiradonna, A., *et al.* (2010). Recent advances on single photon sources based on single colloidal nanocrystals, *Opto-Electronics Rev.* **18**(1), 1–9.
54. Ureña, E. B., Kreuzer, M. P., Itzhakov, S., *et al.* (2012). Excitation enhancement of a quantum dot coupled to a plasmonic antenna, *Adv. Mater.* **24**, 314–320.
55. Ratchford, D., Sha, F., Kim, S., *et al.* (2011). Manipulating coupling between a single semiconductor quantum dot and single gold nanoparticle, *Nano Lett.* **11**, 1049–1054.
56. Okamoto, K., Vyawahare, S., and Scherer, A. (2006). Surface-plasmon enhanced bright emission from CdSe quantum-dot nanocrystals, *J. Opt. Soc. Am. B*, **23**(8), 1674.
57. Babinec, T. M., Hausmann, B. J. M., Khan, M., *et al.* (2010). A diamond nanowire single-photon source, *Nat. Nanotechnol.* **5**(3), 195–199.
58. Castelletto, S., Johnson, B., Ivády, V., *et al.* (2014). A silicon carbide room-temperature single-photon source, *Nat. Mater.* **13**(11), 151–156.
59. Pors, A., and Bozhevolnyi, S. I. (2015). Quantum emitters near layered plasmonic nanostructures: Decay rate contributions, *ACS Photonics* **2**, 228–236.

CHAPTER 5

Metamaterials and Plasmonics with Optical Gain

MIKHAIL A. NOGINOV

Norfolk State University, USA
mnoginov@nsu.edu

5.1. Introduction

As it has been shown in last fifteen years, metamaterials — engineered composite materials with rationally designed composition, shapes, sizes and mutual orientation of subwavelength inclusions — can manipulate electromagnetic waves[1–4] (as well as acoustic waves[5,6] and heat flows[7]) with unprecedented efficiency, enabling a variety of nearly fantastic phenomena and applications ranging from negative index of refraction[8,9] and subdiffraction imaging[8,10–12] to transformation optics[13–17] and invisibility cloaking.[16,17] Another technological effort of paramount importance is aimed at the development of the metamaterials-based electronics operating at optical frequencies,[18–23] whose operation speed will exceed that of conventional semiconductor electronics by several orders of magnitude. In first several years of the metamaterial development, the main focus was at the design of passive nanocircuit components analogous to conductors, inductances and capacitors in traditional electronics.[18,19] However,

electronic circuits also require active elements, such as switches and transistors.[20–23] Their nanophotonic counterparts include metamaterials and plasmonic structures with optical gain, nonlinearity and tunability. The former phenomenon (gain) and its effect on optical loss in metallic inclusions and nanostructures, stimulated emission, and active control of metamaterials and plasmonic systems is the major focus of this chapter. The chapter is not designed to provide the comprehensive review of the research field, but it rather outlines the major concepts, ideas, and experimental demonstrations.

5.2. Problem — Loss, Solution — Gain

Optical loss in metallic (and alternative non-metallic[24]) plasmonic components of metamaterials presents a big challenge to the widespread use of metamaterials and plasmonics and hinders many existing and prospective applications. (Here and below, we will refer to any plasmonic material as to "metal".) The research field of plasmonics covers two major classes of phenomena. The first of them is the localized surface plasmon (SP), which is an oscillation of free electrons in a metallic particle or nanostructure, whose resonance frequency is the plasma frequency adjusted by the size and shape of the particle.[25,26] The second, relevant to it phenomenon is the surface plasmon polariton (SPP) — a surface electromagnetic wave propagating along the interface between metal and dielectric, coupled with collective oscillations of the free electron gas density. In both localized surface plasmons and propagating surface plasmon polaritons, electric field partly resides in a metal, where it experiences strong loss, and partly — in an adjacent nominally loss-less dielectric medium. In 1989, Sudarkin and Demkovich have predicted that optical gain added to the dielectric can partly compensate SPP's propagation (and radiative) loss and, if gain overcompensates loss, lead to a stimulated emission of SPPs.[27] (The details of SPP propagation in the presence of gain have been studied theoretically in Refs. [28 and 29].) A similar concept of the SP loss compensation in metallic nanospheres surrounded by a dielectric medium with gain has been proposed by Lawandy in 2004.[30])

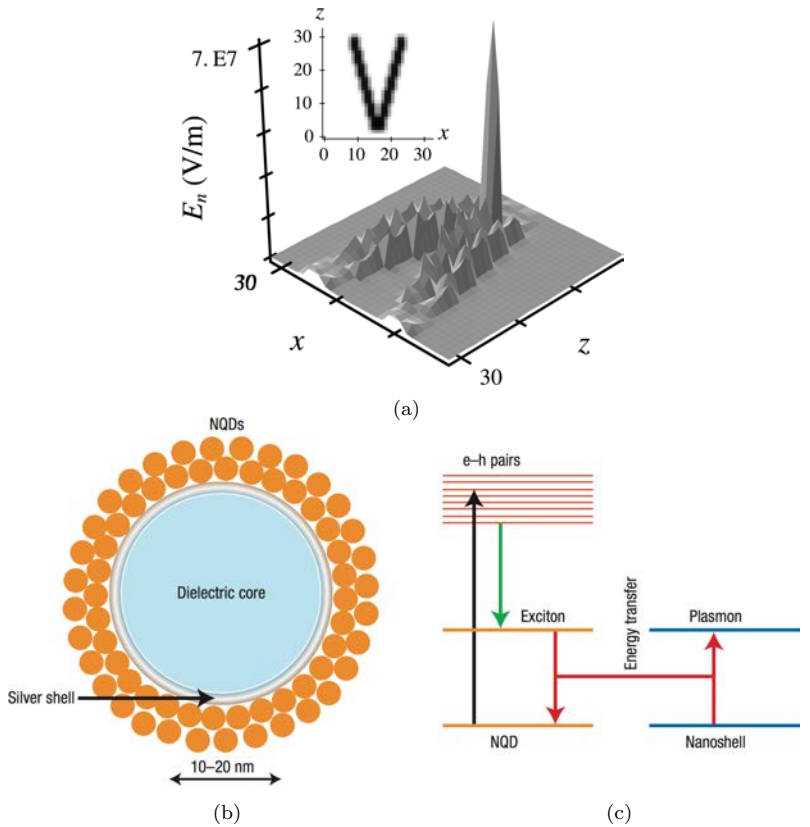


Fig. 5.1. (a) Local field amplitude in the plane of the V-shaped metallic nanostructure (inset) for one of the spaser eigenmodes. Adopted from Ref. [31]. (b) Alternative schematic of a spaser consisting from a silver nanoshell deposited onto a dielectric core (with a radius of 10–20 nm), and surrounded by two dense monolayers of quantum dots. (c) Schematic of energy levels and transitions in a spaser. The external radiation produces electron–hole pairs (vertical arrow). The e–h pairs relax to excitonic levels. The exciton recombines and its energy is transferred (without radiation) to the plasmon excitation of the metal nanoparticle (nanoshell) through resonant coupled transitions. Adopted from Ref. [32].

In 2003, Bergman and Stockman have theoretically predicted stimulated emission of localized surface plasmons supported by a V-shaped metallic nanostructure interfacing a (quantum dot based) gain medium,[31] Fig. 5.1(a). The proposed source of generated

coherent SPs was termed SPASER, in analogy to LASER, in which the letter L, standing for light, was replaced by the letters SP, standing for surface plasmon. In a spaser, like in a regular laser, active medium with population inversion provides for optical gain and amplification. However, the feedback is provided not by the two mirrors, whose separation cannot be smaller than half wavelength ($\sim$0.25 μm for visible light), but by the SP oscillation, which can be supported by plasmonic nanostructures ten-to-hundred times smaller than the wavelength. Correspondingly, spasers can be used in the future nanocircuitry operating at optical frequency as ultra-compact generators and amplifiers of coherent SPs.[23] Surface plasmon oscillations outcoupled to photonic modes make spaser a nanoscopic laser, whose size can be substantially smaller than the diffraction limit ($\sim$half wavelength). An alternative design of the spaser nanoparticle, consisting on a dielectric core, silver shell and two monolayers of quantum dots (which is easier to synthesize using wet chemistry methods) has been proposed in Ref. [32], Fig. 5.1(b). The excitation scheme of quantum dots and surface plasmons in a spaser is depicted in Fig. 5.1(c). The scattering and the stimulated emission of nanoparticles composed of an active SiO_2 core (with gain) and gold shell have been analyzed in Ref. [33].

The pioneering theoretical works[27–33] laid the foundation to the comprehensive studies of surface plasmons and gain undertaken over the last decade.

5.3. Evaluation of the Critical Threshold Gain for Simple Plasmonic Systems

5.3.1. *Surface plasmon polaritons in the Kretchmann geometry*

In the Kretschmann configuration, thin metallic film (usually silver or gold) is deposited on the high-index prism, interfacing with air or low-index dielectric, Fig. 5.2(a). The surface plasmon polariton propagating at the metal/dielectric interface is excited by p polarized light incident onto the metallic film from the side of the prism at the

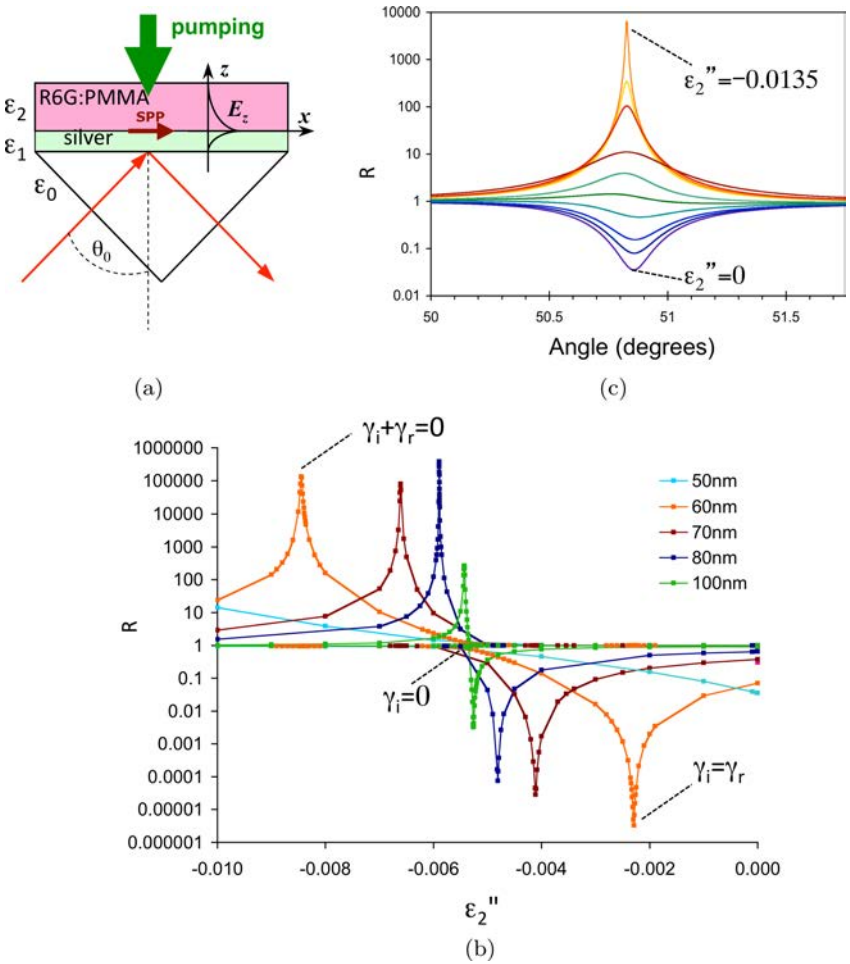


Fig. 5.2. (a) Schematics of the experimental sample (Kretschmann geometry) depicting high index glass prism, metallic film, and dye-doped polymer that is optically pumped to produce gain. (b) Reflectance R^0 calculated in the singularity point ($k_x - k_x^0 - \Delta k_x^0 = 0$) as the function of the imaginary part of the dielectric permittivity of the dye-doped polymer ε_2'' (gain) at several different thicknesses of the silver film. (c) Angular reflectance profiles $R(\theta)$ calculated for 55 nm Ag film at gradually changing values of ε_2''. This film, as well as the 50 nm film in Fig. 5.2(b) are substantially thin. Correspondingly, no reduction of reflection has been predicted with increase of gain in the dielectric medium.

critical angle θ, at which the in-plane component of the photon wave-vector $k_x(\theta)$ matches the wave-vector of SPP k_x^0,

$$k_x(\theta) = k_x^0, \tag{5.1}$$

where

$$k_x(\theta) = (\omega/c)n_0 \sin(\theta), \tag{5.2}$$

$$k_x^0 = \frac{\omega}{c}\sqrt{\frac{\varepsilon_1(\omega)\varepsilon_2(\omega)}{\varepsilon_1(\omega)+\varepsilon_2(\omega)}}, \tag{5.3}$$

n_0 is the index of refraction of the prism, ε_1 is the (negative) real part of the dielectric permittivity of metal, ε_2 is the dielectric permittivity of the dielectric medium above the metallic film, ω is the angular frequency, and c is the speed of light. (The directions of the Cartesian axes are shown in Fig. 5.2(a).) In a typical SPP excitation experiment, the incidence angle or the wavelength is scanned while the reflectance R is measured. At the critical angle, when $k_x(\theta) = k_x^0$, SPP is excited, the energy is transferred from the incident light wave to the SPP, and the reflectance drops significantly. This is the manifestation of the excitation of SPP, which is routinely observed experimentally. In the expressions above, all dielectric permittivities and wave-vectors, in the first approximation, were assumed to be real. This assumption is not universal and does not hold for the analysis presented below.

The reflection (in p polarization), measured in the setup of Fig. 5.2(a) as the function of the incidence angle θ or the angular frequency ω is given by[34]

$$R(\theta,\omega) = \left|\frac{r_{01} + r_{12}\exp(2ik_{z1}d)}{1 + r_{01}r_{12}\exp(2ik_{z1}d)}\right|, \tag{5.4}$$

where

$$r_{ik} = \frac{(k_{zi}\varepsilon_k(\omega) - k_{zk}\varepsilon_i(\omega))}{(k_{zi}\varepsilon_k(\omega) + k_{zk}\varepsilon_i(\omega))}, \tag{5.5}$$

and

$$k_{zi} = \pm\sqrt{\kappa} = \pm\sqrt{\varepsilon_i(\omega)\left(\frac{\omega}{c}\right)^2 - k_x^2}, \quad (i = 0, 1, 2). \tag{5.6}$$

The sign selection in Eq. 5.6 is based on cutting the k_z^2 complex plane along the negative imaginary axis (meaning that the phase of complex numbers is defined between $-1/2\pi$ and $3/2\pi$).[35]

In the limit of small plasmonic loss/gain, when the characteristic length of the SPP propagation (decay length) L is much greater than $2\pi/k_x^0$, and in the vicinity of the resonance angle θ_0, Eq. 5.4 can be simplified, revealing the physics behind the gain-assisted plasmonic loss compensation[35]:

$$R(\theta) \approx |r_{01}^0| \left[1 - \frac{4\gamma_i\gamma_r}{(k_x - k_x^0 - \Delta k_x^0)^2 + (\gamma_i + \gamma_r)^2}\right], \tag{5.7}$$

where

$$r_{01}^0 = r_{01}(\theta_0), \tag{5.8}$$

$$\delta(\theta) = 4(k_x - k_x^0 - \Delta k_x^0)\mathrm{Im}(r_0)\mathrm{Im}(e^{i2k_z^0 d})/\xi, \tag{5.9}$$

and

$$\xi = \frac{c(\varepsilon_2' - \varepsilon_1')}{2\omega}\left(\frac{\varepsilon_2' + \varepsilon_1'}{\varepsilon_2'\varepsilon_1'}\right)^{3/2}. \tag{5.10}$$

(Here and below, ε_i' and ε_i'' stand for real and imaginary parts of the dielectric permittivity. The shape of $R(\theta)$ is dominated by the Lorentzian term in Eq. (5.7). Its width is determined by the propagation length of SPP,

$$L = [2(\gamma_i + \gamma_r)]^{-1}, \tag{5.11}$$

which, in turn, is defined by the sum of the *internal* (or *propagation*) loss

$$\gamma_i = \frac{\omega}{2c}\left(\frac{\varepsilon_1'\varepsilon_2'}{\varepsilon_1' + \varepsilon_2'}\right)^{3/2}\left(\frac{\varepsilon_1''}{\varepsilon_1'^2} + \frac{\varepsilon_2''}{\varepsilon_2'^2}\right), \tag{5.12}$$

and the *radiation* loss caused by SPP leakage into the prism,

$$\gamma_r = \mathrm{Im}(r_{01}e^{i2k_z^0 d})/\xi. \tag{5.13}$$

The radiation loss also leads to the shift of the extremum of the Lorentzian profile from its resonant position k_x^0,

$$\Delta k_x^0 = \mathrm{Re}(r_{01} e^{i2k_z^0 d})/\xi. \tag{5.14}$$

As gain in the dielectric medium is getting larger (ε_2'' becomes more negative), the propagation loss γ_i, evaluated in the singularity point $k_x - k_x^0 - \Delta k_x^0 = 0$, monotonously changes from positive values (loss) to negative values (gain). At the same time, the radiative loss γ_r remains practically unchanged, as long as $|\varepsilon_2''| \ll |\varepsilon_2'|$. Correspondingly, in relatively thick metallic films, when the initial internal loss γ_i (without gain) is larger than the radiative loss γ_r, by gradual increase of the gain in the dielectric, the system transitions through the following regimes: (i) At very small gain $|\varepsilon_2''| \ll |\varepsilon_1''\varepsilon_2''^2/\varepsilon_1''^2|$, its increase makes γ_i and the reflectance R^0 in the singularity point ($k_x - k_x^0 - \Delta k_x^0 = 0$) smaller, see Eq. 5.7 and Fig. 5.2(b). (ii) At $\gamma_i = \gamma_r$, the reflectance R^0 becomes zero, $R^0 = 0$. (Note that in very thin metallic films, $\gamma_i < \gamma_r$ and zero reflectance is achieved when a dielectric medium has the loss rather than the gain. In this case, the regimes (i) and (ii) are not present and the transformation of the reflectance spectra starts from the step (iii) below, see Figs. 5.2(b) and 5.2(c)) (iii) With further increase of the gain, γ_i reduces even further and the minimal reflectance R^0 becomes larger. (iv) At $\gamma_i = 0$, the reflectance R^0 becomes equal to unity, $R^0 = 1$. (v) As the gain becomes even larger, γ_i is getting negative (propagation *gain* instead of propagation *loss*) and a dip in the reflectance profile changes to a peak. (vi) At $\gamma_i + \gamma_r = 0$ (gain in the dielectric compensates both propagation and radiative losses), the denominator in Eq. 5.7 becomes equal to zero and the reflectance is predicted to be infinitely high, $R^0 \to \infty$. It is apparent that the regime vi (and possibly regime v) cannot be adequately described in the steady state approximation, calling for a more advanced time dependent model taking indo account amplified spontaneous emission and gain saturation.

Depending the dielectric permittivity of metal (Ag) and dielectric used in the calculations, full compensation of the *internal*

(propagation) loss γ_i requires optical gain in the dielectric ranging from 420 cm^{-1} to 1310 cm^{-1}.[35] These values are achievable in liquid solutions and polymeric films doped by laser dyes with emission cross sections of $\sim 10^{-16}$ cm^2 in concentration $\sim 2 \times 10^{-2}$ M, pumped with $\sim$10 ns Q-switched laser pulses at pumping density $\sim$10 MW/cm^2.[35]

5.3.2. *Localized surface plasmons in metallic nanoparticles*

Similarly to the case of surface plasmon polaritons, localized surface plasmon resonance in spheroidal metallic nanoparticles is predicted to get enhanced when the surrounding dielectric medium has critical optical gain.[30,36] Its value can be estimated as follows. The polarizability (per unit volume) for isolated metallic nanoparticles is given by

$$\beta = (4\pi)^{-1} \frac{[\varepsilon_m - \varepsilon_d]}{[\varepsilon_d + p(\varepsilon_m - \varepsilon_d)]}, \tag{5.15}$$

where ε_m and ε_d are the dielctric permittivities of metal and dielectric, respectively, and p is the depolarization factor, equal to 1/3 for a sphere and smaller values for oblong spheroids.[37] (Note that a fractal aggregate of Ag nanoparticles discussed in Section 4.2 can be roughly thought of as a collection of spheroids, with different aspect ratios, formed by various chains of nanoparticles in the aggregate.[37]) If at the wavelength λ_0,

$$\varepsilon_d' = -p\varepsilon_m'/(1-p), \tag{5.16}$$

then the real part of the denominator in Eq. 5.15 is nullified, leading to a singularity. The latter, however, is limited by imaginary parts of dielectric permittivities ε_m'' and ε_d'', which are always positive in passive media. Using the Drude formula

$$\varepsilon_m = \varepsilon_b - \frac{\omega_p^2}{\omega(\omega + i\Gamma)}, \tag{5.17}$$

where ε_b, ω_p, and Γ are the interband dielectric constant, plasma frequency, and relaxation rate, respectively, one gets

$$\lambda_0 = \lambda_p \sqrt{\varepsilon_b + \varepsilon'_d(1-p)} \tag{5.18}$$

(here λ_p is the plasma frequency).

The situation changes dramatically if the dielectric medium has optical gain, $\varepsilon''_d < 0$. In this case, at

$$\varepsilon''_d = -p\varepsilon''_m/(1-p), \tag{5.19}$$

the condition, which can be reached at appropriate population inversion and gain, both real and imaginary terms in the denominator of Eq. 5.15 are canceled, leading to extremely large local fields, restricted only by saturation effects.[38] According to Ref. [30], this singularity can be evidenced by an increase of the Rayleigh scattering.

The gain coefficient can be written as

$$g = 4\pi n''/\lambda_0 = -(2\pi/\lambda_0)\varepsilon''_d/\sqrt{\varepsilon'_d}, \tag{5.20}$$

where the basic relation $n = n' + in'' = \sqrt{\varepsilon' + i\varepsilon'}$ and the approximation $n \approx \sqrt{\varepsilon'_d}$ have been used.

Thus the gain needed to compensate for the loss of the localized SP is given by

$$\begin{aligned} g &= (2\pi/\lambda_0)[p/(1-p)]\varepsilon''_m/n \\ &= (2\pi/\lambda_0)(\Gamma/\omega_p)[p/(1-p)][\varepsilon_b + n^2(1-p)/p]^{3/2}. \end{aligned} \tag{5.21}$$

For the resonant wavelength $\lambda_0 = 0.56$ μm, $\varepsilon''_m = 0.405$, and $p = 0.114$ (oblong spheroids), the required $\varepsilon''_d = -0.052$, and the gain coefficient is $g = 4 \times 10^3$ cm^{-1}. (Here $n = 1.33$ and known optical constants from Ref. [39] were used). For spherical nanoparticles ($p = 1/3$), the critical gain is several times larger, $g \sim 10^4$ cm^{-1}. Note that the critical gain of $\sim 10^4$ cm^{-1} has also been predicted in silver-coated SiO_2 nanospheres doped with rare earth ions.[33]

The gain $g \sim 10^3$ cm^{-1} is within the limits of highly concentrated laser dyes and dye-doped polymers pumped with $\sim$10 ns pulses of a

Q-switched laser.[36] The gain of the order of $\sim 10^4$ cm^{-1} is still not impossible but more challenging.

5.4. Loss Compensation: Experiment

5.4.1. *Propagating surface plasmon polaritons (SPPs)*

The first experimental demonstration of partial compensation of the SPP loss by optical gain in the dielectric medium has been reported in Ref. [40]. The Kretschmann schematics of the experiment, featuring flow cell with the laser dye solution serving as the gain medium, is depicted in Fig. 5.3. The laser dye above the silver film was optically pumped (with the cw 10 mW dye laser at $\lambda = 580$ nm, at the angle corresponding to SPP excitation at the pumping wavelength) to produce the population inversion and optical amplification. The reflectance of the optically pumped sample was measured as the function of the incidence angle $R(\theta)$ at $\lambda = 632$ nm. The slight change

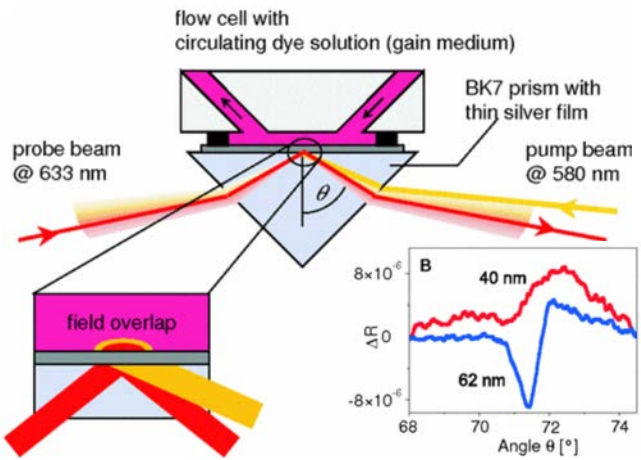


Fig. 5.3. Schematics of the first SPP loss compensation experiment. The signature of the partial loss compensation is found in the slight change of the angular reflectance profile $R(\theta)$, right inset. The depicted differential reflectance curves have been measured in the samples with 40 nm and 62 nm thick Ag films and rhodamine 101 laser dye. Qualitatively similar results have been obtained using cresyl violet dye. Adopted from Ref. [40].

of the angular reflectance profile $R(\theta)$ (right inset of Fig. 5.3) manifested partial compensation of the SPP loss by optical gain. Note that in this first experimental demonstration, the pumping and the gain were very low, resulting in a very small change of reflectance (of the order of $\Delta R \sim 10^{-5}$).

The substantially larger SPP loss compensation has been demonstrated in Ref. [35]. Experimentally, the poly(methyl methacrylate) polymer (PMMA) doped with rhodamine 6G laser dye (R6G) was deposited on top of the Ag film (Fig. 5.2(a)). It was pumped from the top with the Q-switched frequency-doubled Nd:YAG laser at $\lambda = 532\,\text{nm}$, $t_{pulse} \sim 10$ ns, and the time-dependent reflectance kinetics was measured before, during, and after the pumping pulse (Fig. 5.4). Note that the peak pumping power in this particular experiment, $\geq$1 MW, exceeded that in Ref. [40] by eight orders of magnitude.) The reflectance values measured before the laser pulse and during the laser pulse resulted in two distinctly different angular reflectance profiles $R(\theta)$ depicted in Fig. 5.4(b). Their ordering corresponded to that of the theoretical curves $R(\theta)$ calculated in the presence of gain and in absence of gain, Fig. 5.2(c). By fitting the experiment with the model predictions, it has been shown that the achieved optical gain in the dielectric medium ($\approx$420 cm^{-1}) compensated nearly one third of the SPP loss. (Note that the corresponding change of reflectance $\Delta R \approx 280\%$ was much larger than that in Ref. [40], $\Delta R \approx 0.001\%$.)

Partial compensation of the SP propagation loss by gain at the telecom wavelength $\lambda = 1532$ nm has been demonstrated in Ref. [41]. In this study, thin (20 nm) gold nanostrip plasmonic waveguide was sandwiched between two optically pumped slabs of Er^{3+}-doped phosphate glass providing for gain (in the long-range surface plasmon polariton configuration), Fig. 5.5. It has been shown that in this regime of the relatively low loss, even moderate gain available in the Er^{3+}-doped glass (not exceeding several tens cm^{-1}) is sufficient to produce a reasonably strong amplification of the propagating signal, 1.73 dB at the waveguide length of 8 mm. This leads to the (rather obvious) conclusion of practical importance: the compensation and

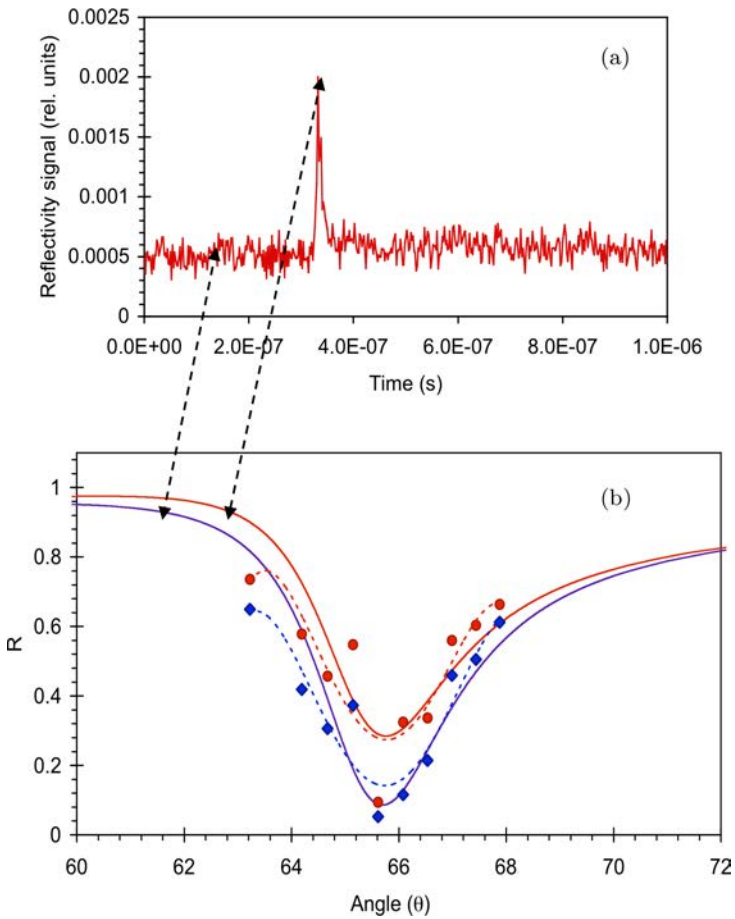


Fig. 5.4. (a) SPP reflectance kinetics showing substantial enhancement of reflectance during the pumping pulse. (b) Angular dependences of reflectance measured (at $\lambda = 594$ nm) before the pumping pulse (diamonds) and during the pumping pulse (circles). Solid lines — model fitting. Dashed lines — guide for eyes. Adopted from Ref. [35].

overcompensation of the SP loss by gain is easier to implement (taking into account required pumping energy and laser damage) in the systems, in which the loss is sufficiently low. This makes the development of low-loss plasmonic waveguides (with advanced material properties and low-loss geometries) of paramount importance.

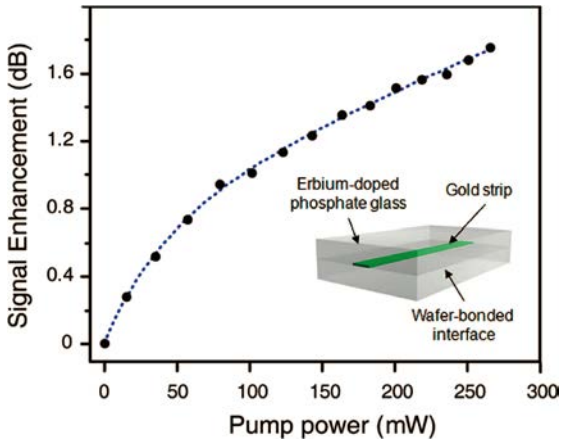


Fig. 5.5. Signal enhancement in the surface plasmon polariton waveguide as a function of pumping power in continuous mode. At the maximum pumping power of 266 mW, the signal is ∼50% enhanced. Inset: Schematics of the experimental sample. Adopted from Ref. [41].

5.4.2. *Localized surface plasmons*

As it has been predicted in Refs. [30 and 33], compensation of SP loss by optical gain in the surrounding dielectric medium can lead to enhancement of the Rayleigh scattering. Experimentally, the Rayleigh scattering in suspensions of aggregated Ag nanoparticles in methanol solution of R6G laser dye has been studied in the pump/probe experiments in Refs. [36 and 42], Fig. 5.6(a). The Q-switched laser pumping ($\lambda = 532$ nm, $t_{pulse} \sim 10$ ns) approximately corresponded to the maximum of the dye's absorption, while the wavelength of the probing laser beam ($\lambda = 562$ nm) was close to the maximum of the dye's gain band. The pumping (and the corresponding optical gain in the dye molecules) were varied, while the intensity of the probing light was kept constant. It has been shown that with increase of the pumping intensity, the intensity of the scattered probing light could increase up to six times, Fig. 5.6(b). In agreement with Refs. [30 and 33], this effect was explained by (partial) compensation of the plasmon loss in Ag aggregates by optical gain, leading to increase of the quality factor of the SP oscillations,

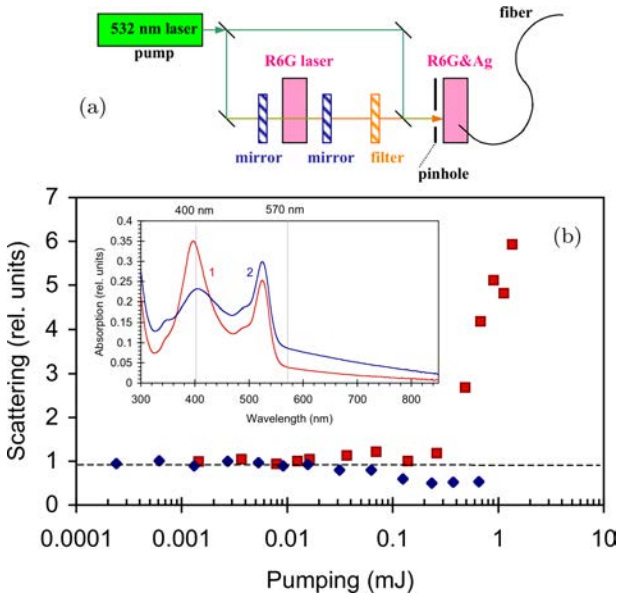


Fig. 5.6. (a) Pump-probe setup for the Rayleigh scattering experiment. The home-build R6G laser was pumped by the same frequency doubled Nd:YAG laser, which excited dye molecules in the dye/Ag aggregate suspension. (b) Intensity of the Rayleigh scattering as the function of the pumping energy in two different dye-Ag aggregate mixtures. Suspension # 1 (squares and absorption spectrum #1 in the inset) demonstrated strong enhancement Rayleigh scattering, while suspension #2 (diamonds and absorption spectrum #2 in the inset) did not. Adopted from Refs. [36, 42].

increase of the effective scattering cross section, and corresponding increase of the Rayleigh scattering.

The regimes of mixing and steering of the dye/Ag aggregate suspension, time, temperature and other ambient conditions could make the absorption spectra and the Rayleigh scattering of two nominally identical mixtures different from each other. Correspondingly, gain-induced increase of the Rayleigh scattering could be large (Fig. 5.6(b), squares), small, or even negative (Fig. 5.6(b), diamonds). The strength of the effect was linked to the shape of the Ag aggregate's absorption spectrum, which depended on the nanoscopic details of the aggregation. The latter appeared to be very important

to the compensation of loss by gain. Thus, no loss compensation has been observed in suspensions of isolated Ag nanoparticles in the dye solution. Most probably, the major contribution to the loss compensation came from the dye molecules adsorbed to Ag aggregate in some critical hot spots of the electric field. (See the relevant discussion in Section 5.4.3 and Ref. [43].)

5.4.3. *Metamaterials*

The concept of compensation of optical loss with gain, is, of course, not limited to simplest plasmonic structures but can be extended to more complex composite systems and metamaterials.

As an example, the effect of optically pumped PbS quantum dots, providing for optical gain at ∼1050 nm, on the transmission of the metamaterial depicted in Fig. 5.7 has been studied in Ref. [44]. At the available pumping density, 50 W/cm^2, the optical transmission

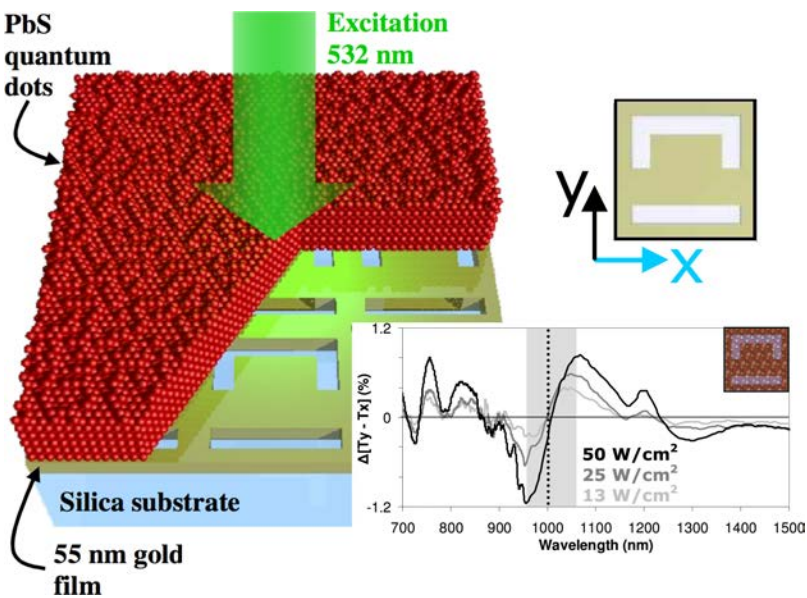


Fig. 5.7. Schematics of the metamaterial with gain (PbS quantum dots). Top right: metamaterial's meta-atom. Bottom right: Difference between pump-induced changes of the transmission signal for x and y-polarizations for different levels of pumping. Adopted from Ref. [44].

of the metamaterial changed by ∼1%. Although this demonstration served as a proof of principle of compensation of loss with gain in a metamaterial, much larger levels of pumping and gain are required for full compensation of loss with optical gain and achieving the threshold for coherent stimulated emission of plasmonic meta-atoms.

As it has been argued in Ref. [43], the effect of the optical gain in a dielectric medium on the metallic absorption loss in a metamaterial or composite plasmonic material is maximized when the gain is not distributed uniformly but rather placed in judiciously chosen hot spots of the electric field. This idea has been experimentally realized in Ref. [45], in which the gain medium (epoxy doped with the rhodamine 800 dye) has been placed inside the fishnet metamaterial with negative index of refraction, Fig. 5.8.

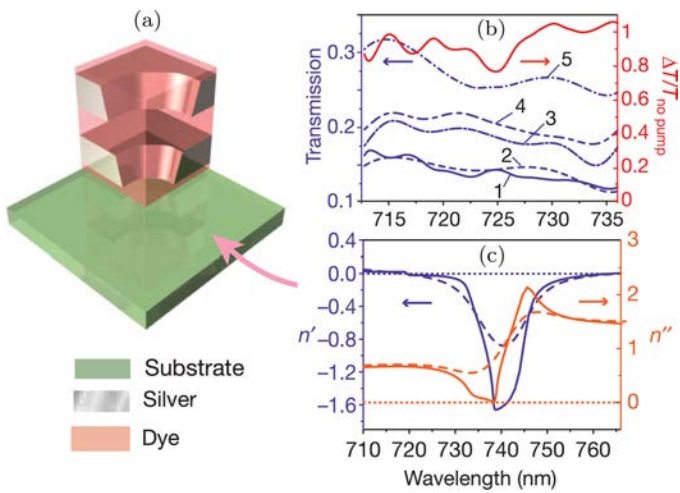


Fig. 5.8. (a) Unit cell of the fishnet metamaterial, which has the dye-doped epoxy in the spacer region and above the fishnet structure. (b) The transmission spectra without pumping (line 1), with the optimized delay between pump and probe (54 ps) and 1-mW pumping power (line 5), with the optimized delay and 0.12-mW pumping power (line 3), with the optimized delay and 0.16-mW pumping power (line 4), and with the pump preceding the probe and 1-mW pumping power (line 2). The wavelength-dependent relative transmission change from the pump–probe experiment is shown by the top-most solid line. (c) The effective refractive index, $n = n' + in''$, evaluated with (solid) and without (dashed) gain. Adopted from Ref. [45].

It has been experimentally demonstrated that optical pumping, providing for population inversion and amplification in the gain medium, doubles the sample's transmittance, from ∼0.15 to ∼0.3.[45] The comparison of the experiment with the theoretical model has shown that the pumping makes the negative real part of the index of refraction n' more negative and fully compensate the absorption loss n'' (at $\lambda = 737$ nm, n' changes from -0.66 to -1.017) wheares n'' decreases from 0.66 to 0.039). This study has demonstrated the possibility of creating an optical negative-index metamaterial (and probably a much broader class of photonic metamaterials), which are not limited by inherent losses in their metallic constituents.

5.5. Stimulated Emission

5.5.1. *Surface plasmon polariton lasers (1D mode confinement)*

An even stronger pumping, resulting in full compensation of loss by gain, can lead to stimulated emission of SPPs — the effect experimentally demonstrated in Ref. [46] in the setup similar to that in Fig. 5.2(a). First, when the pumping is weak, excited R6G dye molecules spontaneously emit to the available modes, including SPP modes. The generated SPPs propagate at the metal/polymer interface and partly out-couple to the photonic modes in the glass prism. The broad spectrum of the out-coupled SPP emission is depicted in Fig. 5.9(a). With increase of the pumping intensity ($\lambda = 532$ nm, $t_{pulse} \sim 10$ ns), the spectrum of emission (measured approximately along the direction of the reflected beam in Fig. 5.2(a)) is getting much narrower (Fig. 5.9(a)) and the input-output emission curve features the threshold characteristic of lasers, Fig. 5.9(b).

Furthermore, the threshold pumping density depends on the detection angle, reaching its minimum at the same angle as the SPP reflection profile $R(\theta)$, Fig. 5.9(c). The first two phenomena — narrowing of the emission line and threshold in the input/output curve — suggest onset of the stimulated emission. At the same time, the similarity of the angular reflectance profile and the angular

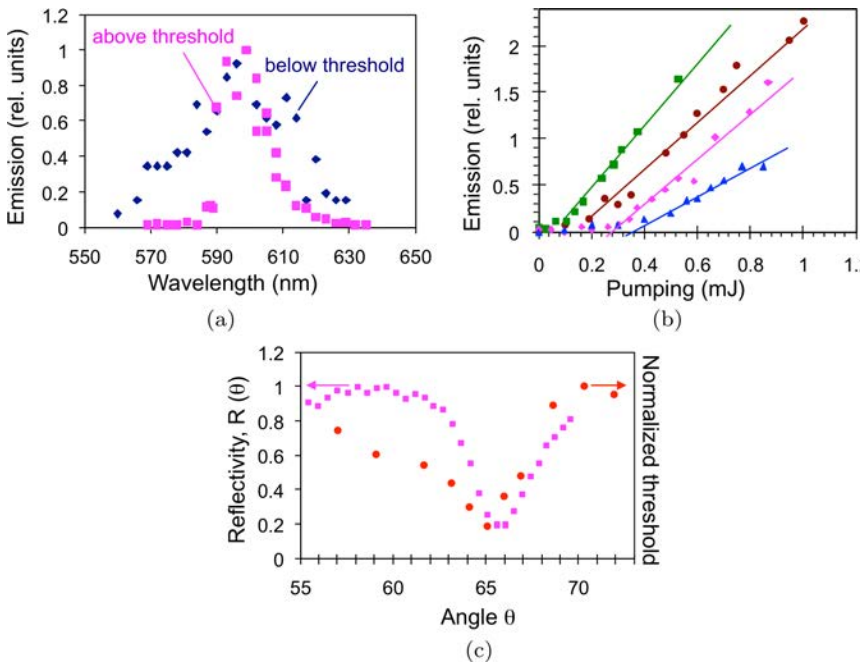


Fig. 5.9. (a) Emission spectra of SPPs excited by optically pumped R6G dye molecules and outcoupled to the glass prism at $\theta = 68.7°$ (in the setup of Fig. 5.2(a)). Blue diamonds — spontaneous emission below the threshold; pink squares — stimulated emission above the threshold. (b) Input/output curves of the SPP stimulated emission recorded at different out-coupling angles. (c) Normalized threshold of the SPP stimulated emission measured at different out-coupling angles (circles) plotted together with the angular reflectance profile $R(\theta)$ (squares). Adopted from Ref. [46].

dependence of the threshold indicates that the observed stimulated emission is supported by surface plasmon polaritons.

The stimulated emission of SPPs discussed above[46] did not have any intentional feedback and was analogous to amplified spontaneous emission (ASE) known in dielectric gain media. To the contrary, the stimulated emission of surface plasmon polaritons in Ref. [47], did have the feedback provided by the microcyllindrical cavity, as described below. Experimentally, thin (d = 10 μm diameter) gold wire was dipped into the dichloromethane solution of R6G dye and PMMA polymer and then dried in air. This simple procedure allowed

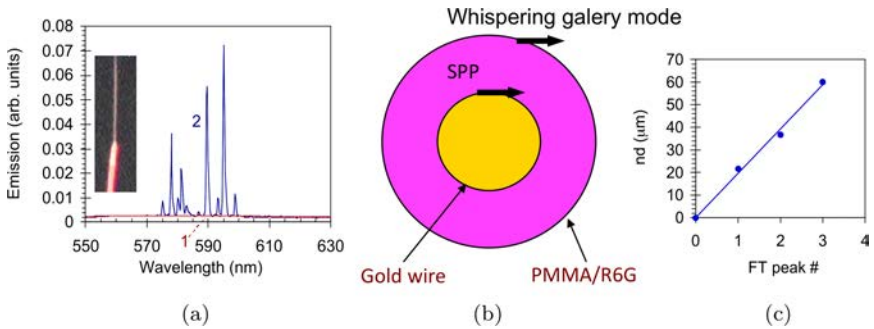


Fig. 5.10. (a) Spectra of dye emission below (trace 1) and above (trace 2) the stimulated emission threshold. Inset: microphotograph of the Au wire (10 μm diameter) with deposited R6G:PMMA cylindrical cavity ($\sim$30 μm outer diameter). (b) Schematic cross section of the cylindrical microcavity. (c) Values nd corresponding to the Fourier Transform peaks. Adopted from Ref. [47].

one to obtain a microcyllindrical SPP cavity as shown in inset of Fig. 5.10(a) and Fig. 5.10(b). At weak pumping of the dye with the frequency doubled Q-switched Nd:YAG laser, only very dim spontaneous emission has been observed, trace 1 in Fig. 5.10(a). With increase of the pumping energy, several sharp and intense emission lines (characteristic of multimode stimulated emission) appeared in the spectrum, trace 2 in Fig. 5.10(a). The Fourier transform of the multi-line emission spectrum resulted in a series of equidistant peaks corresponding to the effective optical length of the circumference traveled by the SPP equal to $(nd)_{\text{exp}} = 19.6$ μm (Fig. 5.10(c)), where d was the wire diameter and n was the effective refractive index for the supported wave. This figure was close to the one calculated based on the known values of d and n for SPP, $(nd)_{calc}^{SPP}$ =17.4, and very far from that expected of the whispering gallery mode wave propagating at the interface of R6G:PMMA and air $(nd)_{calc}^{whisper} = 45$. Therefore, it was concluded that the experimentally observed stimulated emission was supported by SPPs propagating at the interface of Au and dye-doped polymer. The stimulated emission originating from the whispering gallery mode was never detected, probably because of high roughness of the air/polymer interface.

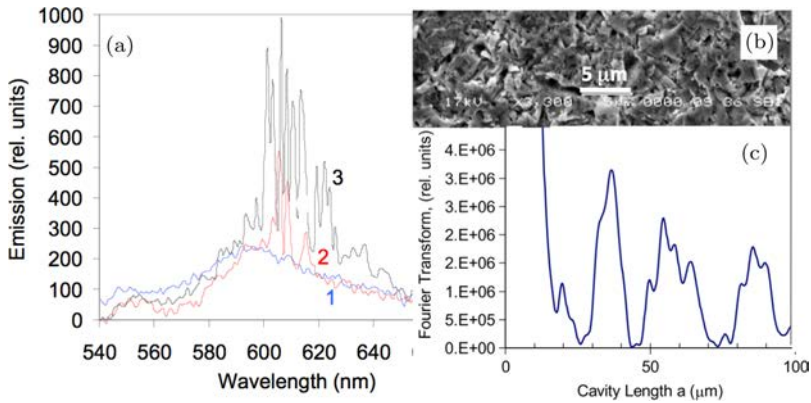


Fig. 5.11. (a) Emission spectra in the RB:PMMA film deposited on roughened bulk silver (with surface roughness equal to 234 nm), pumped with 7 mJ (1), 13 mJ (2) and 20 mJ (3) laser pulses. (b) Scanning electron microscope (SEM) image of the roughened silver substrate. (c) Equidistant peaks in the Fourier Transform of the stimulated emission spectrum, corresponding the cavity length ≈21 μm. Adopted from Ref. [48].

The easiest to realize SPP laser feedback has been demonstrated in Ref. [48], in which the PMMA film doped with rhodamine B dye (RB:PMMA) was deposited on top of intentionally roughened bulk silver, Fig. 5.11. In this series of samples, the evolution of the emission spectra with increase of the pumping energy (Fig. 5.11) qualitatively resembled that observed in the microcyllinder cavity in Ref. [47] (Fig. 5.10) — in both cases, a series of sharp laser lines emerged above the well-defined threshold. In analogy with photonic random lasers,[49] in which the lasing is supported by random cavities formed by multiple scatterers, the observed stimulated emission served as a manifestation of the Surface Plasmon Random Laser, the phenomenon first reported in Ref. [50]. Contrary to the expectations, the Fourier Transforms of the stimulated emission spectra have shown that the lasing effect on top of randomly roughened (meta)surfaces is dominated by one particular cavity size, whose length (ranging between ∼10 μm and ∼30 μm) increased with increase of the pumping density. This intriguing observation, similar to that in photonic random lasers,[51,52] calls for further theory development.

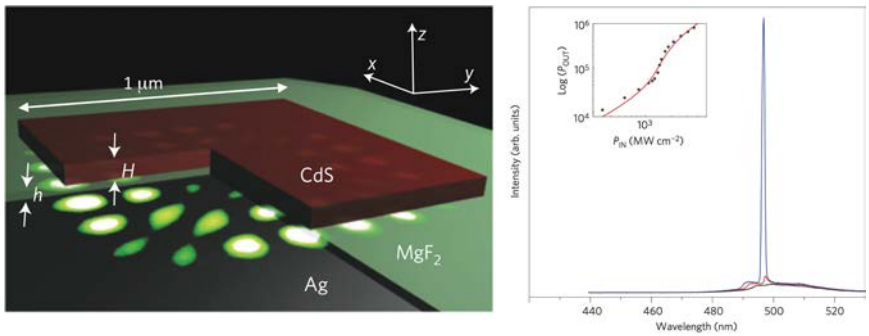


Fig. 5.12. (a) Schematic diagram of the room-temperature plasmon laser showing a thin CdS square atop a silver substrate separated by a 5 nm MgF_2 gap. (b) The room-temperature laser spectra and the input-output curve (inset) showing the transition from spontaneous emission (1.096 $MWcm^{-2}$, black) through amplified spontaneous emission (1.280 $MWcm^{-2}$, smallest line intensity) to full single-mode laser oscillation (1.459 $MWcm^{-2}$, largest line intensity). Adopted from Ref. [53].

The family of plasmonic nanolasers is rapidly growing, utilizing new materials and morphologies. One particularly interesting geometry, based on the total internal reflection of surface plasmons in the structure consisting of 1 μm CdS square "nanopatch" separated from Ag surface by a thin layer of MgF_2 spacer (Fig. 5.12(a)), has been demonstrated in Ref. [53]. This room-temperature semiconductor sub-diffraction-limited plasmonic laser combined strong confinement with low metal loss (as the most intense electric fields resided in lossless MgF_2.) High cavity quality factors, approaching 100, along with strong $\lambda/20$ mode confinement, lead to enhancements of spontaneous emission rate by up to 18-fold. By controlling the structural geometry, the number of cavity modes was reduced to achieve a single-mode lasing.

Surface plasmon polaritons supported by Ag or Au surfaces have relatively large loss in the visible part of the spectrum, making the stimulated emission difficult to achieve. At the same time, SPP loss is much smaller in the near-infrared spectral range. Correspondingly, it has been experimentally demonstrated[54] that the threshold of the SPP stimulated emission can be reduced nearly twenty-fold if

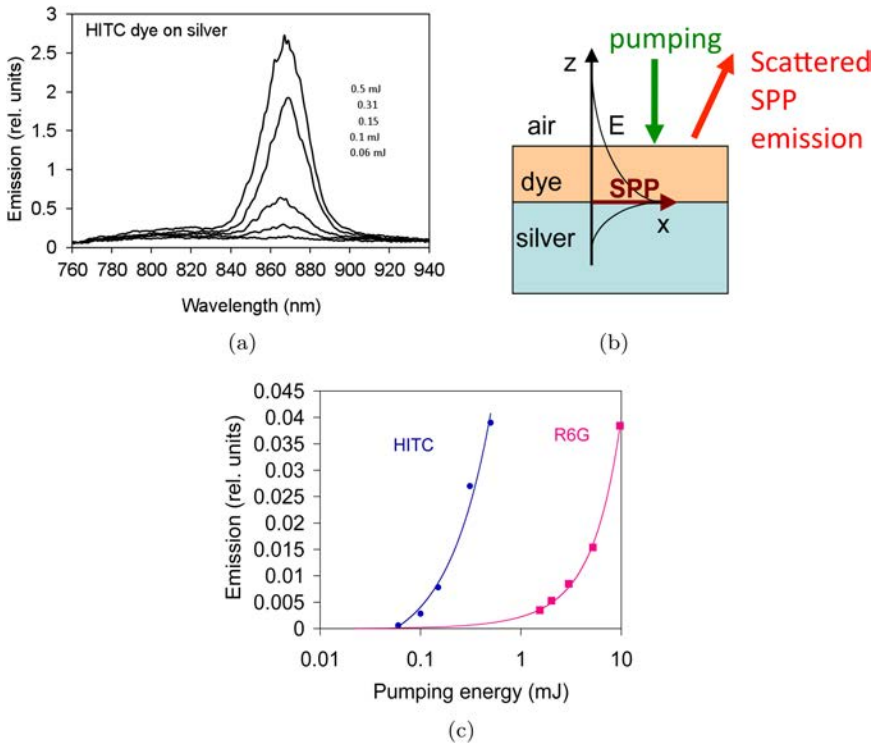


Fig. 5.13. (a) Stimulated emission spectra of the HITC:PMMA film on top of thick (∼200 nm) Ag film. Pumping energies (from bottom to top) range from 0.06 mJ to 0.5 mJ (t_{pulse}∼10 nm, pumped spot $d \sim 1.7$ mm). (b) Schematics of the experimental sample. Stimulated emission, scattered by unintentional scatterers, was collected from the front of the sample. (c) Input/output curves of the SPP stimulated emission at 0.87 μm (HITC dye) and 0.6 μm (R6G dye), showing nearly twentyfold reduction of the lasing threshold in the former.

HITC dye, whose stimulated emission has a maximum at ∼0.87 μm, replaces R6G dye that emits at ∼0.6 μm, Figs. 5.13(a,b,c).

5.5.2. *Spaser based nanolaser (3D sub-diffraction mode confinement)*

A SPASER, theoretically proposed in Ref. [31], provides for stimulated emission of surface plasmons in resonating metallic nanostructures adjacent to the gain medium. It generates coherent surface

plasmons, in the same way as a LASER, whose invention more than half a century ago revolutionized science and technology, generates stimulated emission of coherent photons.[55] In the plasmonic lasers discussed in Section 5.5.1, the mode had sub-diffraction confinement in one dimension but was spread over more than one wavelength in the other two dimensions. At the same time, spaser is the sole means of enabling a true nanolaser, because only a localized SP resonance is capable of squeezing optical frequency oscillations into a nanoscopic cavity in all three dimensions. The grand challenge in the experimental demonstration of a spaser is the absorption loss in metal. This problem is particularly severe when the mode is nanoscopic (sub-diffraction) in all three dimensions.[56–58]

In Ref. [55], the loss of localized SPs has been completely overcome by gain and the stimulated emission achieved in the 44 nm diameter hybrid Au/silica/dye SPASER nanoparticles. In this study, the 14 nm diameter Au nanospheres with a thin sodium silicate shell were covered by a thick (15 nm) silica shell doped with the Oregon Green dye, OG-488, Figs. 5.14(a,b,c). The key to a success was obtaining a sufficiently large number of dye molecules per nanoparticle ($\sim 3 \times 10^3$), which was sufficient to overcome SP absorption loss in the metallic core. The water suspension of nanoparticles was pumped with ∼5 ns pulses from an Optical Parametric Oscillator (OPO) lightly focused into a ∼2.4 mm spot. When the pumping energy exceeded a critical threshold value, a narrow peak appeared at $\lambda = 531$ nm in the emission spectrum, Fig. 5.14(c). The intensity of this peak plotted versus pumping energy resulted in an input-output curve with a pronounced threshold, characteristic of lasers, Fig. 5.14(d). In a separate experiment, oscillating behavior of the emission kinetics, characteristic of relaxation oscillations in lasers, has been observed in the range of the pumping energies exceeding the threshold. The combination of the three phenomena, (i) narrowing of the emission spectrum, (ii) input/output curve with the threshold, and (iii) relaxation oscillations in the emission kinetics, manifested stimulation emission in a spaser-based nanolaser — the first operating at a visible wavelength and the smallest (at that time) reported in the literature.

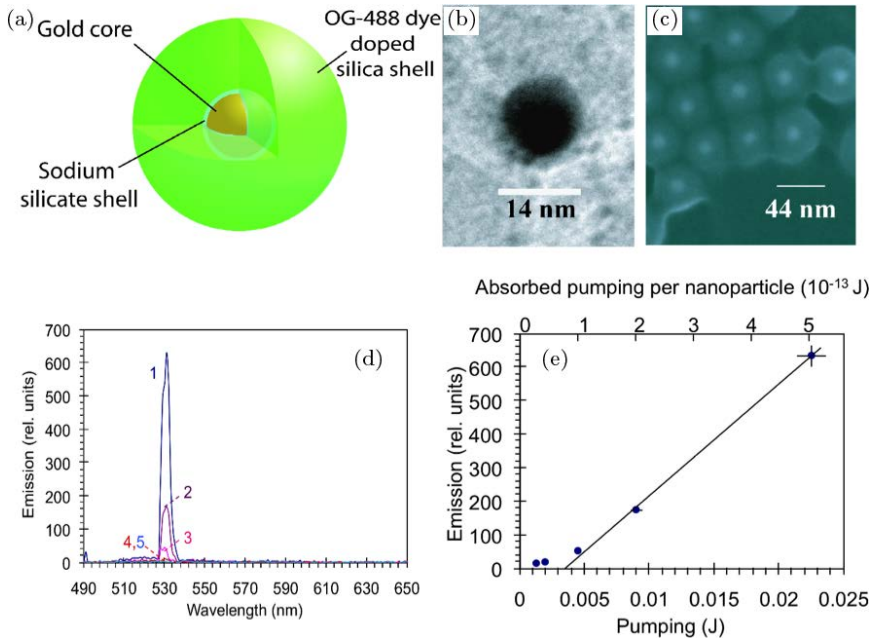


Fig. 5.14. (a) Schematic of the hybrid nanoparticle architecture (not to scale) indicating dyes throughout the silica shell. (b) TEM image of Au core. (c) SEM image of Au/silica/dye core-shell nanoparticles. (d) stimulated emission spectra of the nanoparticle sample pumped with 22.5 mJ (1), 9 mJ, (2) 4.5 mJ (3), 2 mJ (4) and 1.25 mJ (5) 5 ns OPO pulses at $\lambda = 488$ nm. (e) Corresponding input-output curve (lower axis — total launched pumping energy, upper axis — absorbed pumping energy per nanoparticle). Adopted from Ref. [55].

Note that the spaser effect in the core-shell nanoparticles, whose morphology was similar to those in Ref. [55], although with a different dye and multiple sizes of Au cores, has been recently reported in Ref. [59].

5.5.3. *Plasmonic lasers with nanowire gain medium (2D mode confinement)*

Sections 5.5.1 and 5.5.2 above highlighted plasmonic lasers with 1D mode confinement (SPPs on a metallic surface) and 3D mode confinement (nanoparticle based spaser). Plasmonic lasers, whose mode confined in two dimensions, are described below.

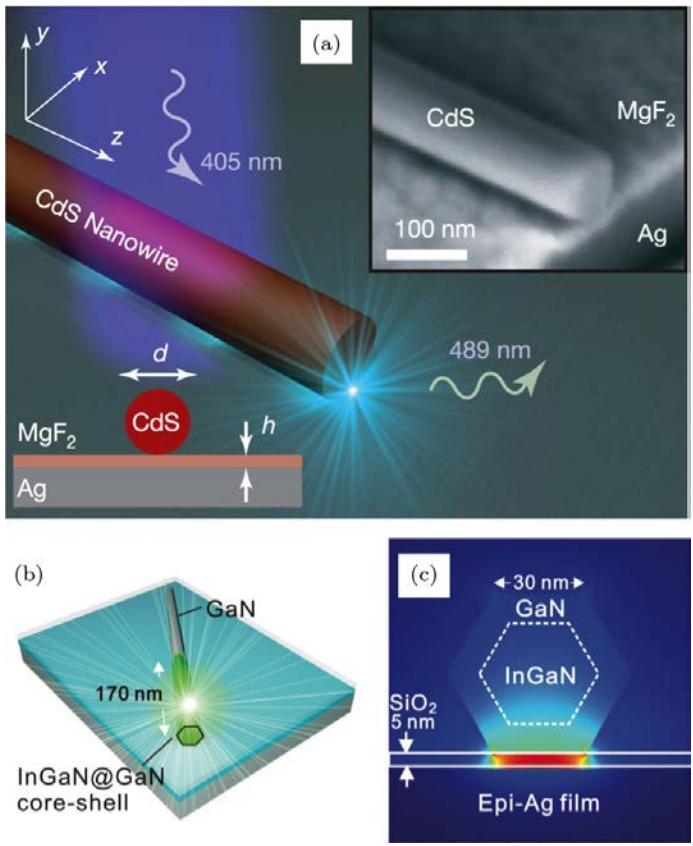


Fig. 5.15. (a) Plasmonic laser based on CdS nanowire: CdS semiconductor nanowire on top of a silver substrate, separated by a nanometre-scale MgF_2 layer. Adopted from Ref. [60]. (b) Plasmonic laser based on InGaN@GaN core-shell nanorod: single InGaN@GaN core-shell nanorod on top of SiO_2-covered epitaxial Ag film (28-nm-thick). (c) The calculated energy-density distribution of the device in Fig. b. Adopted from Ref. [61].

The first nanolaser of this kind, consisting of a high-gain cadmium sulphide semiconductor nanowire, separated from the silver surface by 5-nm-thick MgF_2 spacer, has been demonstrated in Ref. [60], Fig. 5.15(a). Direct measurements of the emission lifetime revealed a broad-band enhancement of the nanowire's exciton spontaneous emission rate (up to six times) owing to the strong mode

confinement. This has allowed one to realize the nanolaser, whose optical mode was ∼5 nm in one transverse direction and ∼38 nm in the other transverse direction — much smaller than the diffraction limit at the 489 nm lasing wavelength. At the same time, the length of the nanowire (and the corresponding longitudinal size of the laser mode) was comparable to or greater than the wavelength. The very tight transversal localization of the laser mode was predicted to enable extreme light-matter interaction, opening up new avenues in the fields of active photonic circuits, bio-sensing and quantum information technology.

Conceptually similar nanolaser, although based on a different semiconductor nanowire serving as the gain medium (with InGaN core and GaN cladding), has been demonstrated in Ref. [61], Figs. 5.15(b,c). The special effort in this study was aimed at growth of atomically smooth epitaxial silver film. The atomic smoothness of the metallic film is crucial for reducing the plasmonic loss, the modal volume and, correspondingly, the lasing threshold. The authors of Ref. [61] have further emphasized that advancement of the technology allowing one to grow atomically smooth epitaxial Ag on Si — a new platform for plasmonics — promotes the development of monolithically integrated plasmonics and Si-based electronics.

Note that more recently, the growth of high-quality epitaxial silver films has been reported in Ref. [62] and, the sputtered Ag films have been shown to support SPPs propagating $\geq$100 μm at $\lambda \geq 650$.[63]

Before concluding this section, we highlight the theoretical proposal[64] of using the spaser as an ultrafast nanoamplifier — an optical counterpart of the MOSFET (metal-oxide-semiconductor field-effect transistor). As it has been predicted in Ref. [64], the spaser can perform functions of an ultrafast nanoamplifier in the two regimes: transient (amplifier) and bistable (dynamic memory). In particular, it has been shown that the spaser can amplify with gain $\geq$50 and the switching time $\leq$100 fs (potentially, ∼10 fs). This prospective spaser technology has a promise of enabling ultrafast microprocessors working at 10–100 THz clock speed. Other potential applications

of spasers include sensing and ultrafast high-density information storage.

5.5.4. *Stimulated emission in metamaterials*

As we have seen in Sections 5.5.1, 5.5.2 and 5.5.3, (i) both localized and propagating surface plasmons can be generated via the stimulated emission process similar to that in lasers and (ii) outcoupling of SP oscillations to photonic modes enables nanoscopic lasers with the mode confined in one, two or three dimensions. Can more complex metamaterial geometries improve the spaser performance and make possible new functionalities, which are not achievable in simple plasmonic structures? The affirmative answer to this question is supported by several examples outlined below.

The stimulated emission of surface plasmons propagating on top of multilayered hyperbolic metamaterials coated with dye-doped polymeric films (HITC:PMMA) was studied in Ref. [65], Figs. 5.16(a,b). It was shown to have a much lower (2.5–7 times) threshold than that in similar dye-doped films on top of thick Ag films, Fig. 5.16(c). Red shift of the stimulated emission spectrum on top of the metamaterial, as compared to that on top of thick Ag film, is consistent with the lower threshold in the metamaterial sample, Fig. 5.16(d). This observation correlates with more than 2-fold shortening of the spontaneous emission kinetics on top of a lamellar metamaterial as compared to that on top of silver. The propagation of surface plasmons on top of the metamaterial and silver substrates was modeled (in the Otto geometry) using COMSOL Multiphysics. It has been shown that at given system parameters, the plasmon propagation loss in a metamaterial is smaller than that in silver by $\sim$14% and cannot explain the dramatic reduction of the lasing threshold observed experimentally in the metamaterial samples. It has been inferred that the stimulated emission of propagating plasmons on top of a metamaterial is enhanced by the nonlocal dielectric environment and high local density of photonic states.

The spaser, originally proposed in Ref. [31], was deigned to produce stimulated emission of surface plasmons supported by *single*

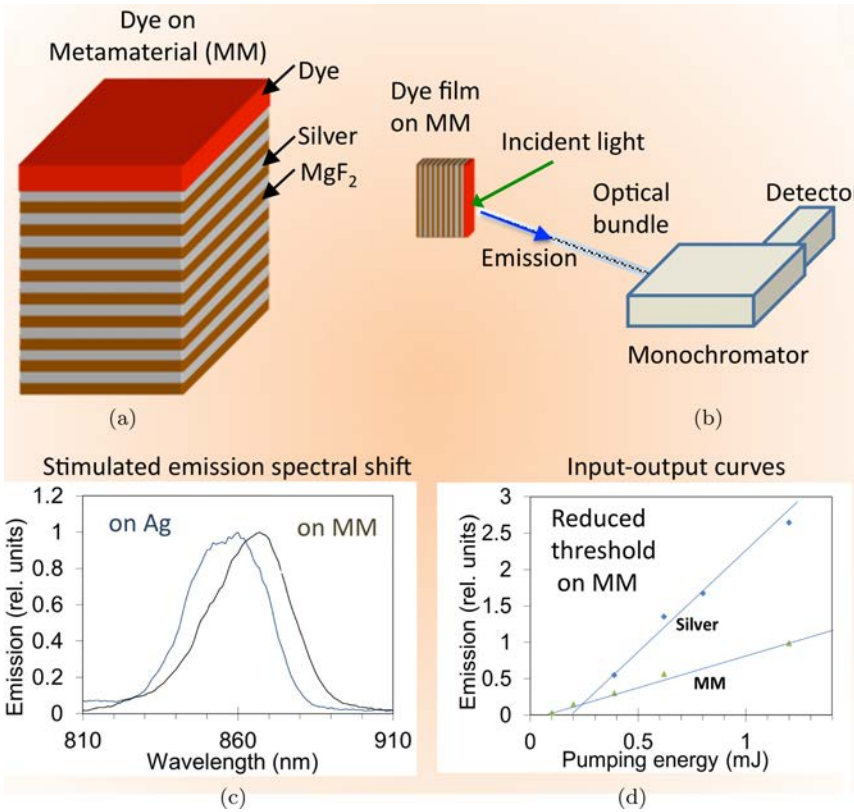


Fig. 5.16. (a) Experimental sample — dye doped polymeric film (HITC:PMMA) on top of the lamellar Ag/MgF_2 metamaterial (25 nm Ag and 35 nm MgF_2). (b) Stimulated emission setup. (c) Stimulated emission spectra of the HITC:PMMA film on top of the 200 nm thick Ag film and metamaterial. Red shift of the spectrum on the metamaterial is consistent with the lower threshold. (d) Input/output stimulated emission curves demonstrating smaller threshold on the metamaterial than on the thick silver film. Adopted from Ref. [65].

metallic nanoparticles (plasmonic oscillators) coupled to the gain medium. The further development of the spaser concept, aligned with that of active metamaterials, has been done in Ref. [66], proposing that one can create a narrow-divergence coherent source of electromagnetic radiation based on multiple "single meta-atom" spasers coupled to each other. The particular design of Ref. [66] involved a

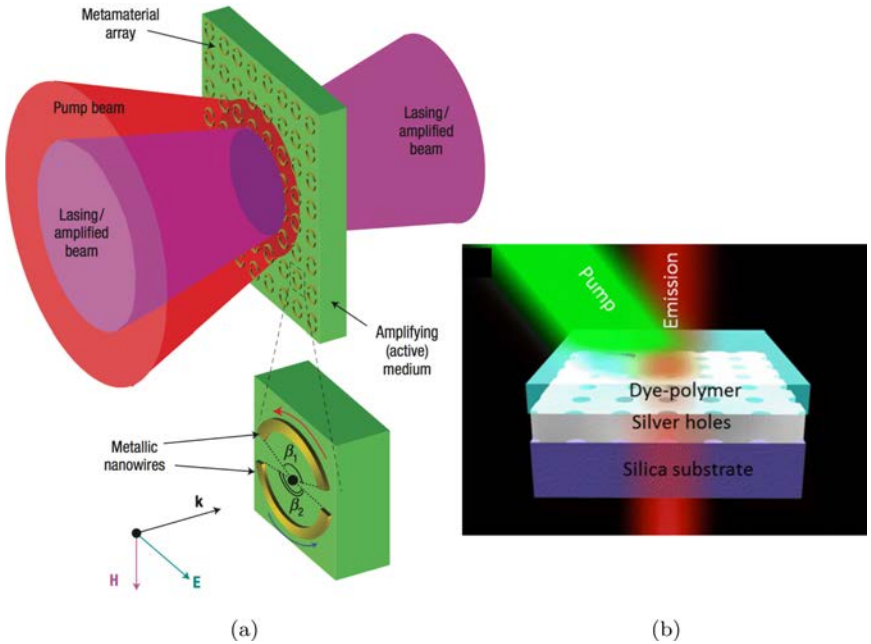


Fig. 5.17. Lasing spaser. (a) The structure consists of a gain medium slab (green) supporting a periodic array of metallic asymmetric split-ring resonators. The dashed box indicates an elementary translation cell of the array, and the arrows along the arcs of the ring illustrate the antisymmetric currents of plasmonic oscillations. In-phase plasmonic oscillation in individual resonators leads to the emission of spatially and temporarily coherent light propagating in the direction normal to the array. Adopted from Ref. [66]. (b) Alternative configuration of the experimentally demonstrated lasing spaser device. The system is composed of a periodic hole array covered by a thin layer of the R101(dye):PVA(polymer) film. Adopted from Ref. [67].

two-dimensional array of asymmetric split-ring resonators supporting coherent current excitations with high quality factors, Fig. 5.17(a). The proposed device, termed a 'lasing spaser', was predicted to act as a planar source of spatially and temporally coherent radiation.

The originally proposed design depicted in Fig. 5.17 is not the only possible realization a lasing spaser, and many other types of spaser meta-atoms, arranged into periodic arrays and coupled to each other, can lead to a qualitatively similar result. The first experimental

demonstration of the lasing spaser has been reported in Ref. [67], in which a highly directional spaser emission in the visible part of the spectrum has been observed in a periodic subwavelength array of holes perforated in the silver film (functioning as plasmonic nanocavities) combined with an organic laser dye providing for a gain. The single mode lasing effect occurred in the red wavelength range. It has been inferred that the optical feedback for spasing is provided by the SPP–Bloch wave. This conclusion is supported by the fact that no spasing was attained in aperiodic holes as well as in periodic holes that do not support the SPP–Bloch wave at the spasing wavelength.

Another very interesting concept of a metamaterial-based laser, theoretically proposed in Ref. [68], rests on the idea that the stimulated emission feedback can be provided by the stopped light (SL), Fig. 5.18(a). When light is brought to a standstill, its interaction with the gain media increases dramatically due to a singularity in the density of optical states.[68] Concurrently, stopped light provides

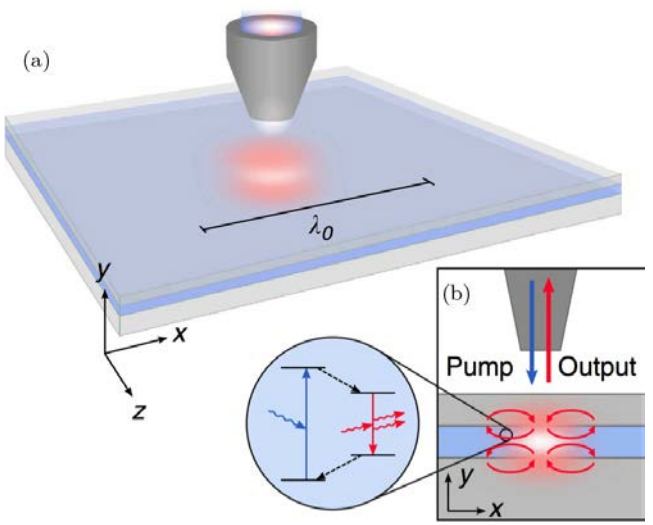


Fig. 5.18. (a) The core layer of the metal-dielectric SL multilayer structure is filled with gain material (blue). (b) Spatially selective excitation of the homogeneous gain layer using a near-field tip leads to the formation of a subwavelength spot of inverted gain, in which the stimulated emission processes take place.

for an inherent cavity-free feedback mechanism, similar to the one demonstrated in large-scale disordered media and random lasers.[49] The fundamental principle behind the stopped light lasing is to trap and amplify photons in the same location where they were emitted. In planar plasmonic heterostructures, the stopping of light arises from balancing of opposing energy flows in layers of materials with positive (dielectric) and negative (metal) permittivities. At the SL point, the overall energy flow cancels, forming a characteristic closed-loop vortex (Fig. 5.18(b)), which results in a strong enhancement of the local density of optical states. Combined with gain, this SL feedback mechanism can lead to coherent amplification of the trapped photons via the stimulated emission processes. The experimental realization of this feedback mechanism would be a major breakthrough in nanolasing and cavity-free active quantum plasmonics.

5.6. Discussion and Summary

To summarize, in multiple examples, we have demonstrated that optical gain adjacent to plasmonic surfaces, nanostructures and meta-atoms (which compose metamaterials) can compensate absorption loss in metal, enable stimulated emission of plasmonic lasers, and provide for ultra-fast signal processing. In the visible part of the spectrum, the smallest critical gain, $\sim 10^3$ cm^{-1}, which is required to conquer optical loss in metal, has been theoretically predicted and experimentally demonstrated for surface plasmon polaritons propagating on top of metallic surfaces. This magnitude of gain is relatively easy to achieve in dye-doped polymeric films and concentrated dye solutions pumped with nanosecond or picosecond laser pulses. Several examples of stimulated emission of SPPs, with and without laser feedback, have been reported in the literature.[46–48,53,54]

Similar values of critical gain (in scattering and stimulated emission experiments) are predicted for localized surface plasmons supported by elongated or aggregated Ag nanoparticles.[26] At the same time, the gain required to compensate loss in isolated metallic and metal-dielectric nanospheres is larger, $\sim 10^4$ cm^{-1} to $\sim 10^5$ cm^{-1}.[33]

This level of gain is still achievable in laser dyes, although the laser damage of dye-doped polymeric films and photobleaching of dye molecules present a serious problem.

Given that the maximal concentration of dye molecules in a solid can be as high as $\sim 10^{21}$ cm^{-3} and the emission cross section can be as high as $\sim 10^{-15}$ cm^2,[69] the maximal gain can, in principle, be of the order of $\sim 10^6$ cm^{-1}. This upper limit is unlikely achievable due to very strong luminescence quenching in highly concentrated molecular ensembles. However, the gain of $\sim 10^5$ cm^{-1} has been demonstrated in Ref. [55].

Overall, the very large gain, which can be achieved in optically pumped laser dyes, makes them the material of choice many for proof-of-principle demonstrations. However, laser damage, photobleaching, and the need for a short-pulsed pumping laser make optically pumped organic dyes unsuitable for many practical applications.

Quantum dots (QDs) have much higher stability than dye molecules and, reportedly, can have optical gain of $\sim 10^5$ cm^{-1}.[70,71] They have been theoretically proposed as a gain medium for loss compensation in the fishnet metamaterial with negative index of refraction[72] and used in the loss compensation experiment employing a different metamaterial design[44] (Fig. 5.7). However, the experimentally demonstrated loss compensation was not very high.[47]

Practical applications of plasmonic lasers and active plasmonics require electrical (rather than optical) pumping. It can be realized in semiconductor heterostructures and multiple quantum wells. (The latter, reportedly, can have optical gain exceeding 10^3 cm^{-1} [73]). Several micrometer- and submicrometer-scale semiconductor lasers with electrical pumping have been reported in the literature,[74–76] Fig. 5.19. However, none of them had the mode size (in all three dimensions) substantially smaller than a half of the lasing wavelength in the medium.

It has been argued[56,58,77] that electrically pumped semiconductor nanolasers, whose mode is smaller than half of the wavelength in the medium in all three dimensions, are nearly impossible to realize (using available plasmonic materials), since the threshold

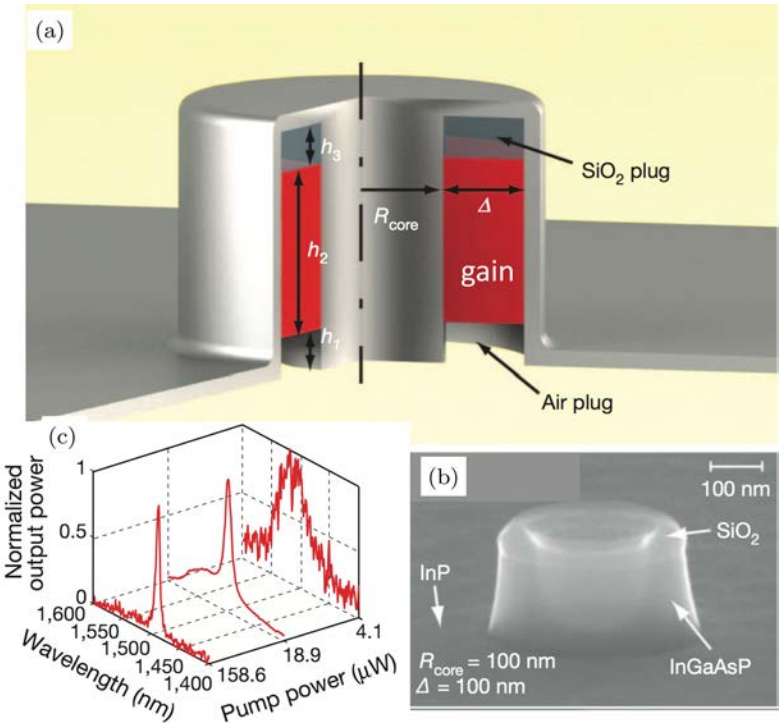


Fig. 5.19. (a) Laser schematics. The coaxial waveguide that supports plasmonic modes is composed of a metallic rod enclosed by a metal-coated semiconductor ring (six quantum wells gain medium). (b) Scanning electron microscope image of the device. (c) Evolution of the emission spectra with increase of the pumping power (room temperature). Adopted from Ref. [76].

current density would significantly exceed the damage threshold. The reduction of loss in plasmonic material was claimed to be the only possible way of reducing the lasing threshold.[77] The loss reduction can, in principle, be achieved by moving from the visible spectral range to the near-infrared range,[54] reduction of the temperature,[78–81] and improving the quality of the metallic (silver) films. The recent advance in the latter direction[62,63] gives a hope of bringing electrically pumped plasmonic lasers, as well as the whole technological area of active plasmonics and metamaterials, to reality.

References

1. Engheta, N. and Ziolkowski, R. W. (editor) (2006). *Electromagnetic Metamaterials: Physics and Engineering Explorations*, Wiley-IEEE Press, 440 p.
2. Cui, T. J., Smith, D. and Liu, R. (editor) (2009). *Metamaterials: Theory, Design and Applications*, Springer US, 367 p.
3. Cai, W. and Shalaev, V. (Author) (2010). *Optical Metamaterials: Fundamentals and Applications*, Springer New York, 200 p.
4. Noginov, M. A. and Podolskiy, V. A. (Eds.) (2011). *Tutorials in Metamaterials*, CRC Press, Taylor & Francis group, 308 p.
5. Li, J. and Chan, C. T. (2004). Double-negative acoustic metamaterial, *Phys. Rev. E* **70**, 055602(R).
6. Zhang, S., Yin, L. and Fang, N. (2009). Focusing ultrasound with an acoustic metamaterial network, *Phys. Rev. Lett.* **102**, 194301.
7. Schittny, R., Kadic, M., Guenneau, S. and Wegener, M. (2013). Experiments on transformation thermodynamics: Molding the flow of heat, *Phys. Rev. Lett.* **110**, 195901.
8. Pendry, J. B. (2000). Negative refraction makes a perfect lens, *Phys. Rev. Lett.* **85**, 3966.
9. Shelby, R. A., Smith, D. R. and Schultz, S. (2001). Experimental verification of a negative index of refraction, *Science* **292**, 77–79.
10. Smith, D. R., Schurig, D., Rosenbluth, M., Schultz, S., Ramakrishna, S. A. and Pendry, J. B. (2003). Limitations on subdiffraction imaging with a negative refractive index slab, *Appl. Phys. Lett.* **82**, 1506.
11. Jacob, Z., Alekseyev, L. V. and Narimanov, E. (2006). Optical hyperlens: Far-field imaging beyond the diffraction limit, Jacbo, Z. Alekseyev, L. V. and Narimanov, E. (Eds.), *Optics Express* **14**, 8247–8256.
12. Liu, Z., Lee, H., Xiong, Y., Sun, C. and Zhang, X. (2007). Far-field optical hyperlens magnifying sub-diffraction-limited objects, *Science* **315**, 1686.
13. Pendry, J. B., Schurig, D. and Smith, D. R. (2006). Controlling electromagnetic fields, *Science* **312**, 1780.
14. Chen, H., Chan, C. T. and Sheng, P. (2010). Transformation optics and metamaterials, *Nature Materials* **9**, 387–396.
15. Vakil, A. and Engheta, N. (2011). Transformation optics using graphene, *Science* **332**, 1291–1294.
16. Cai, W., Chettiar, U. K., Kildishev, A. V. and Shalaev, V. M. (2007). Optical cloaking with metamaterials, *Nature Photonics* **1**, 224–227.
17. Ergin, T., Stenger, N., Brenner, P., Pendry, J. B. and Wegener, M. (2010). Three-dimensional invisibility cloak at optical wavelengths, *Science* **328**, 337–339.
18. Engheta, N., Salandrino, A. and Alù, A. (2005). Circuit elements at optical frequencies: Nanoinductors, nanocapacitors and nanoresistors, *Phys. Rev. Lett.* **95**, 095504.
19. Engheta, N. (2007). Circuits with light at nanoscales: Optical nanocircuits inspired by metamaterials, *Science* **317**, 1698–1702.

20. Ozbay, E. (2006). Plasmonics: Merging photonics and electronics at nanoscale dimensions, *Science* **311**, 189–193.
21. Chettiar, U. K. and Engheta, N. (2012). Optical frequency mixing through nanoantenna enhanced difference frequency generation: Metatronic mixer, *Phys. Rev. B* **86**, 075405.
22. Huang, K. C. Y., Seo, M.-K., Sarmiento, T., Huo, Y., Harris, J. S. and Brongersma, M. L. (2014). Electrically driven subwavelength optical nanocircuits, *Nature Photonics* **8**, 244–249.
23. Stockman, M. I. (2010). The spaser as a nanoscale quantum generator and ultrafast amplifier, *Journal of Optics* **12**, 024004 (13pp).
24. Boltasseva, A. and Atwater, H. A. (2011). Low-loss plasmonic metamaterials, *Science* **331**, 290–291.
25. Noginov, M. A., Zhu, G., Drachev, V. P. and Shalaev, V. M. (2007). Surface plasmons and gain media, in *Nanophotonics with Surface Plasmons*, Shalaev, V. M. and Kawata, S. Eds., Elsevier, New York, pp. 141–169.
26. Noginov, M. A., Zhu, G., Bahoura, M., Adegoke, J., Small, C. E., Ritzo, B. A., Drachev, V. P. and Shalaev, V. M. (2006). Enhancement of surface plasmons in an Ag aggregate by optical gain in a dielectric medium, *Optics Letters* **31**, 3022–3024.
27. Sudarkin, A. N. and Demkovich, P. A. (1989). Excitation of surface electromagnetic waves on the boundary of a metal with an amplifying medium, *Sov. Phys.Tech. Phys.* **34**, 764–766.
28. Avrutsky, I. (2004). Surface plasmons at nanoscale relief gratings between a metal and a dielectric medium with optical gain, *Phys. Rev. B* **70**, 155416.
29. Nezhad, M. P., Tetz, K. and Fainman, Y. (2004). Gain assisted propagation of surface plasmon polaritons on planar metallic waveguides, *Optics Express* **12**, 4072–4079.
30. Lawandy, N. M. (2004). Localized surface plasmon singularities in amplifying media, *Appl. Phys. Lett.* **85**, 540–542.
31. Bergman, D. J. and Stockman, M. I. (2003). Surface plasmon amplification by stimulated emission of radiation: Quantum generation of coherent surface plasmons in nanosystems, *Phys. Rev. Lett.* **90**, 027402.
32. Stockman, M. I. (2008). Spasers explained, *Nature Photonics* **2**, 327–329.
33. Gordon, J. A. and Ziolkowski, R. W. (2007). The design and simulated performance of a coated nano-particle laser, *Opt. Express* **15**, 2622–2653.
34. Raether, H. (1988). *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*, Springer Verlag, Berlin (1988).
35. Noginov, M. A., Podolskiy, V. A., Zhu, G., Mayy, M., Bahoura, M., Adegoke, J. A., Ritzo, B. A. and Reynolds, K. (2008). Compensation of loss in propagating surface plasmon polariton by gain in adjacent dielectric medium, *Opt. Express* **16**, 1385–1392.
36. Noginov, M. A., Zhu, G., Bahoura, M., Adegoke, J., Small, C. E., Ritzo, B. A., Drachev, V. P. and Shalaev, V. M. (2006). Enhancement of surface

plasmons in an Ag aggregate by optical gain in a dielectric medium, *Opt. Lett.* **31**, 3022–3024.
37. Shalaev, V. M., *Nonlinear Optics of Random Media*, Springer (2000).
38. Drachev, V. P., Buin, A. K., Nakotte, H. and Shalaev, V. M. (2004). Size dependent $\chi^{(3)}$ for conduction electrons in Ag nanoparticles, *Nano Letters* **4**, 1535–1539.
39. Johnson, P. B. and Christy, R. W. (1972). Optical constants of the noble metals, *Physical Review B* **6**, 4370–4379.
40. Seidel, J., Grafström, S. and Eng, L. (2005). Stimulated emission of surface plasmons at the interface between a silver film and an optically pumped dye solution, *Phys. Rev. Lett.* **94**, 177401.
41. Ambati, M., Nam, S. H., Ulin-Avila, E., Genov, D. A., Bartal, G. and Zhang, X. (2008). Observation of stimulated emission of surface plasmon polaritons, *Nano Letters* **8**, 3998–4001.
42. Noginov, M. A., Zhu, G., Bahoura, M., Agegoke, J., Small, C., Ritzo, B. A., Drachev, V. P. and Shalaev, V. M. (2007). The effect of gain and absorption on surface plasmons in metallic nanoparticles, *Appl. Phys. B.* **86**, 455–460.
43. Sivan, Y., Xiao, S., Chettiar, U. K., Kildishev, A. V. and Shalaev, V. M. (2009). Frequency-domain simulations of a negative-index material with embedded gain, *Opt. Express* **17**, 24060–24074.
44. Plum, E., Fedotov, V. A., Kuo, P., Tsai, D. P. and Zheludev, N. I. (2009). Towards the lasing spaser: Controlling metamaterial optical response with semiconductor quantum dots, *Opt. Express* **17**, 8548–8551.
45. Xiao, S., Drachev, V. P., Kildishev, A. V., Ni, X., Chettiar, U. K., Yuan, H.-K. and Shalaev, V. M. (2010). Loss-free and active optical negative-index metamaterials, *Nature* **466**, 735–740.
46. Noginov, M. A., Zhu, G., Mayy, M., Ritzo, B. A., Noginova, N. and Podolskiy, V. A. (2008). Stimulated emission of surface plasmon polaritons, *Phys. Rev. Lett.* **101**, 226806.
47. Kitur, J. K., Podolskiy, V. A. and Noginov, M. A. (2011). Stimulated emission of surface plasmon polaritons in microcylinder cavity, *Phys. Rev. Lett.* **106**, 183903.
48. Kitur, J. K., Zhu, G., Barnakov, Yu. A. and Noginov, M. A. (2012). Stimulated emission of surface plasmon polaritons on smooth and corrugated silver surfaces, *J. Opt.* **14**, 114015.
49. Noginov, M. A. (2005). *Solid-State Random Lasers*, Springer, New York, 235 p.
50. Kumar, A., Yu, S. F. and Li, X. F. (2009). Random laser action in dielectric-metal-dielectric surface plasmon waveguides, *Appl. Phys. Lett.* **95**, 231114.
51. Polson, R. C., Huang, J. D. and Vardeny, Z. V. Analysis of random lasers in thin films of p-conjugated polymers, *Photonic Crystals and Light Localization in the 21st Century*, Soukoulis, C. M., ed., NATO Science Series, Vol. 563 (Kluwer, Boston, 2001), p. 405.

52. Polson, R. C., Raikh, E. M. and Vardeny, Z. V. (2002). Universality in unintentional laser resonators in π-conjugated polymer films, *Comptes Rendus Physique* **3**, 509–521.
53. Ma, R.-M., Oulton, R. F., Sorger, V. J., Bartal, G. and Zhang, X. (2010). Room-temperature sub-diffraction-limited plasmon laser by total internal reflection, *Nature Materials* **10**, 110–113.
54. Kitur, J. K., Zhu, G. and Noginov, M. A. (2014). Low-threshold stimulated emission of surface plasmons polaritons, *Journal of Optics* **16**, 114020.
55. Noginov, M. A., Zhu, G., Belgrave, A. M., Bakker, R., Shalaev, V. M., Narimanov, E. E., Stout, S., Herz, E., Suteewong, T. and Wiesner, U. (2009). Demonstration of a spaser-based nanolaser, *Nature* **460**, 1110–1112.
56. Khurgin, J. B. and Sun, G. (2012). How small can "Nano" be in a "Nanolaser"?, *Nanophotonics* **1**, 3–8.
57. Wang, F. and Shen, Y. R. (2006). General properties of local plasmons in metal nanostructures, *Phys. Rev. Lett.* **97**, 206806.
58. Khurgin, J. B. and Sun, G. (2011). Scaling of losses with size and wave-length in nanoplasmonics and metamaterials, *Appl. Phys. Lett.* **99**, 211106.
59. Galanzha, E. I., Weingold, R., Nedosekin, D. A., Sarimollaoglu, M., Kuchyanov, A. S., Parkhomenko, R. G., Plekhanov, A. I., Stockman, M. I. and Zharov, V. P. (2017). Spaser as a biological probe, *Nature Communications* **15528**.
60. Oulton, R. F., Sorger, V. J., Zentgraf, T., Ma, R.-M., Gladden, C., Dai, L., Bartall, G. and Zhang, X. (2009). Plasmon lasers at deep subwavelength scale, *Nature* **461**, 629–632.
61. Lul, Y.-J., Kim, J., Chen, H.-Y., Wu, C., Dabidian, N., Sanders, C. E., Wang, C.-Y., Lu, M.-Y., Li, B.-H., Qiu, X., Chang, W.-H., Chen, L.-J., Shvets, G., Shih, C.-K. and Gwo, S. (2012). Plasmonic nanolaser using epitaxially grown silver film, **6093**, 450–453.
62. Wu, Y., Zhang, C., Estakhri, N. M., Zhao, Y., Kim, J., Zhang, M., Liu, X.-X., Pribil, G. K., Alù, A., Shih, C.-K. and Li, X., Intrinsic (2014). Optical properties and enhanced plasmonic response of epitaxial silver, *Advanced Materials* **26**, 6106–10.
63. High, A. A., Devlin, R. C., Dibos, A., Polking, M., Wild, D. S., Perczel, J., de Leon, N. P., Lukin, M. D. and Park, H. (2015). Visible-frequency hyperbolic metasurface, *Nature* **522**, 192–196.
64. Stockman, M. I. (2010). The spaser as a nanoscale quantum generator and ultrafast amplifier, *Journal of Optics* **12**, 024004.
65. Kitur, J. K., Gu, L., Tumkur, T., Bonner, C., and Noginov, M. A. (2015). Stimulated emission of surface plasmons on top of metamaterials with hyperbolic dispersion, *ACS Photonics* **2**, 1019–1024.
66. Zheludev, N. I., Prosvirning, S. L., Papasimakis, N. and Fedotov, V. A. (2008). Lasing spaser, *Nature Photonics* **2**, 351–354.
67. Meng, X., Liu, J., Kildishev, A. V. and Shalaev, V. M. (2014). Highly directional spaser array for the red wavelength region, *Laser & Photonics Reviews* **8**, 896–903.

68. Pickering, T., Hamm, J. M., Page, A. F., Wuestner, S. and Hess, O. (2014). Cavity-free plasmonic nanolasing enabled by dispersionless stopped light, *Nature Communications* **5**, 4972.
69. Gu, L., Livenere, J., Zhu, G., Narimanov, E. E. and Noginov, M. A. (2013). Quest for organic plasmonics, *Appl. Phys. Lett.* **103**, 021104.
70. Bimberg, D., Kirstaedter, N., Ledentsov, N. N., Alferov, Zh. I., Kop'ev, P. S. and Ustinov, V. M. (1997). In GaAs–GaAs quantum-dot lasers, *IEEE Journal of Selected Topics in Quantum Electronics* **3**, 196–205.
71. Strauf, S., Hennessy, K., Rakher, M. T., Choi, Y.-S., Badolato, A., Andreani, L. C., Hu, E. L., Petroff, P. M. and Bouwmeeste, D. (2006). Self-tuned quantum dot gain in photonic crystal lasers, *Phys. Rev. Lett.* **96**, 127404.
72. Dong, Z.-G., Liu, H., Li, T., Zhu, Z.-H., Wang, S.-M., Cao, J.-X., Zhu, S.-N. and Zhang, X. (2010). Optical loss compensation in a bulk left-handed metamaterial by the gain in quantum dots, *Appl. Phys. Lett.* **96**, 044104.
73. Asada, M., Kameyama, A. and Suematsu, Y. (1984). Gain and intervalence band absorption in quantum-well lasers, *IEEE Journal of Quantum Electronics* **QE-20**, 745–753.
74. Hill, M. T., Oei, Y.-S., Smalbrugge, B., Zhu, Y., de Vries, T., van Veldhoven, P. J., van Otten, F. W. M., Eijkemans, T. J., Turkiewicz, J. P., de Waardt, H., Geluk, E. J., Kwon, S.-H., Lee, Y.-H., Nötzel, R. and Smit, M. K. (2007). Lasing in metallic-coated nanocavities, *Nature Photonics* **1**, 589–594.
75. Hill, M. T., Marell, M., Leong, E. S. P., Smalbrugge, B., Zhu, Y., Sun, M., van Veldhoven, P. J., Geluk, E. J., Karouta, F., Oei, Y.-S., Nötzel, R., Ning, C.-Z. and Smit, M. K. (2008). Lasing in metal-insulator-metal sub-wavelength plasmonic waveguides, *Opt. Express* **17**, 11107.
76. Khajavikhan, M., Simic, A., Katz, M., Lee, J. H., Slutsky, B., Mizrahi, A., Lomakin, V. and Fainman, Y. (2012). Thresholdless nanoscale coaxial lasers, *Nature* **482**, 204–207.
77. Khurgin, J. B. and Sun, G. (2012). Practicality of compensating the loss in the plasmonic waveguides using semiconductor gain medium, *Appl. Phys. Lett.* **100**, 011105.
78. Liu, M., Pelton, M. and Guyot-Sionnest, P. (2009). Reduced damping of surface plasmons at low temperatures, *Phys. Rev. B* **79**(3), 035418.
79. Bouillard, J.-S. G., Dickson, W., O'Connor, D. P., Wurtz, G. A. and Zayats, A. V. (2012). Low-temperature plasmonics of metallic nanostructures, *Nano Lett.* **12**(3), 1561–1565.
80. Mayy, M., Zhu, G., Mayy, E., Webb, A. and Noginov, M. A. (2012). Low temperature studies of surface plasmon polaritons in silver films, *J. Appl. Phys.* **111**(9), 094103.
81. Jayanti, S. V., Park, J. H., Dejneka, A., Chvostova, D., McPeak, K. M., Chen, X., Oh, S.-H. and Norris, D. J. (2015). Low-temperature enhancement of plasmonic performance in silver films, *Opt. Mat. Express* **5**, 1148–1155.

CHAPTER 6

Active Metamaterials

MICHAEL D. SINGH and MAREK S. WARTAK*

Wilfrid Laurier University, Waterloo, Canada

The outstanding properties of metamaterials open the door of opportunity for a number of exciting practical applications. Fascinating applications such as: perfect lenses that break the diffraction limit of conventional lenses, optical quantum storage, and invisibility cloaking. These dramatic new technological advancements are being severely hindered by the significant loss effects of metamaterials. Overcoming loss limitations is currently one of the most important topics in metamaterials research. Presented in the paper are loss reduction schemes regarding: the geometric shape of metamaterial components, parameter manipulation, and parametric amplification. Active gain is then discussed with experimental and theoretical approaches involving loss compensation of active metamaterials. Next, surface plasmons are introduced discussing the intriguing concepts of spasers and metamaterial lasing spasers. Finally, we move to the opposite direction and consider increasing losses in metamaterials to create near perfect metamaterial absorbers that can find applications in the solar cell industry as well as other fields of science.

6.1. Introduction

The properties exhibited by a substance which is found in nature is determined by the chemical elements and bonds that form the

*Corresponding author: mwartak@wlu.ca

structure. By precisely arranging the elements of a structure, an artificial material can be created. This artificially engineered material, a material that is beyond a naturally occurring one, is known as a metamaterial ("meta" — Greek word for beyond).[1] Since a metamaterial is artificially produced and to the best of our knowledge cannot be found in nature, it exhibits exotic and potentially advantageous properties that we can utilize.

The electromagnetic properties of any medium are determined by two fundamental parameters: electric permittivity (ϵ) and magnetic permeability (μ). The square-root of the product of the two parameters defines the refractive index of a material

$$n = \sqrt{\epsilon\mu} \tag{6.1}$$

A material simultaneously having negative electric permittivity and negative magnetic permeability is a metamaterial. Since both ϵ and μ are negative, the product of the two is a positive real number. Interestingly, this means that the refractive index is a real negative number. To illustrate why the refractive index becomes a negative number, see this simple example (for a more complete discussion, see[2])

$$\begin{gathered}\text{Let, } \epsilon = -1 \text{ and } \mu = -1 \\ n = \sqrt{\epsilon\mu} \\ n = \sqrt{(-1)(-1)} \\ n = \sqrt{(-1)}\sqrt{(-1)} \\ n = i \times i \\ n = -1\end{gathered}$$

Thus, metamaterials are characterized by having a negative index of refraction, and are often referred to as negative index materials.

In recent years the field of metamaterials has received remarkable attention with the number of published papers growing exponentially. This is due to unusual properties of such systems (see Ref. 3 for a recent review) and also important practical applications like perfect lenses,[4] invisibility cloaking,[5,6] slow light,[7] and enhanced optical

nonlinearities.[8] Parallel to theoretical developments there has been spectacular experimental progress.[9]

The full exploitation of metamaterials suffers due to the existence of losses which at present are orders of magnitude that are too large for practical applications and are considered as an important factor limiting practical applications of metamaterials. For example, detailed analytical studies show that losses limit the super-resolution of a theoretical super-lens.[10] There was some controversy about the possibilities of eliminating losses[11] but as shown by Webb and Thylen[12] it is possible to completely eliminate losses in metamaterials.

In recent years there has been several reports[13–15] about possible compensation of losses in metamaterials by introducing gain elements. For example, Wegener *et al.*[14] formulated a simple model where gain is represented by a fermionic two-level system which is coupled via a local-field to a single bosonic resonance representing the plasmonic resonance of metamaterial. Also recently, Fang *et al.*[15] described a model where gain system is modelled by generic four-level atomic system. They conducted numerical analysis using FDTD technique and showed that one can compensate the losses in of the dispersive metamaterials. Some other reports of the design and analysis of active metamaterials are by Yuan *et al.*[16] and Sivan *et al.*[17]

6.2. Metamaterial Structures

The idea of left-handed materials was first introduced theoretically by Vesalgo in his theoretical work,[18] but it was not until Pendry *et al.*[19] in 1999 when the scientific world took notice. Their work demonstrated a practical way of fabricating the once theoretical metamaterials by using a split-ring resonator structure. Since then metamaterial research has been a highly popular and significant research area. Currently the two most prominent metamaterial structures which effectively achieve simultaneous negative electric permittivity and magnetic permeability are split-ring resonator (SRR) structures and fishnet structures.[20]

6.2.1. *Split-ring resonator structures*

A split-ring resonator consists of a pair of concentric non-magnetic metallic rings, with slit separations on opposite sides (see Fig. 6.1). One can generate a magnetic resonance by inducing rotational currents within the SRR. As shown in Ref. 19, a negative permeability can be realized in a periodic array of SRRs. A negative electric permittivity can be achieved by using ultra-thin metallic wires, at all frequencies below the plasma frequency.[19] Thus a periodic array of SRRs consisting of non-magnetic metallic nanowires can achieve simultaneous negative electric permittivity and magnetic permeability, given the correct parameters, as shown in Ref. 21.

6.2.2. *Fishnet structures*

A fishnet metamaterial structure is made of two very thin metallic layers, separated by a thin dielectric slab, with periodically arrayed rectangular holes.[22] This can be seen in Fig. 6.2(a).

The goal is to achieve a simultaneously negative: electric permittivity and magnetic permeability, for a given frequency of incident

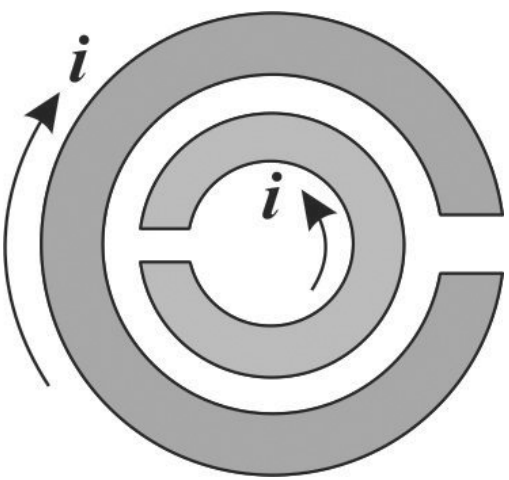


Fig. 6.1. A split-ring resonator structure.

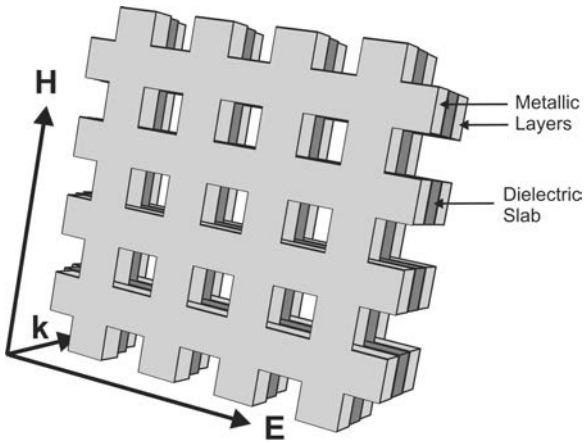


Fig. 6.2. A fishnet structure with metallic layers on either side of a dielectric slab.

light wave. A negative permeability can be achieved if a strong magnetic resonance is formed at a frequency, ω_m.[23] Such a process can be described and analyzed using an inductance-capacitance (LC) circuit where,

$$\omega_m = \frac{1}{\sqrt{LC}} \tag{6.2}$$

For an incident optical signal, an (LC) circuit is formed in the fishnet structure by the wire-dielectric-wire configuration that runs parallel to the incident magnetic field whilst simultaneously inducing anti-parallel currents in the metallic layers on either side of the dielectric slab.[22] These two processes allow for the formation of a magnetic field and ultimately a negative permeability through a strong magnetic resonance between the incident and induced magnetic field.

A negative electric permittivity can be achieved through a system of parallel wires separated by a dielectric slab. These parallel wires essentially exhibit plasmonic behaviour similar to that of Drude metals. [22,23] The permittivity then can be represented by,

$$\epsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2} \tag{6.3}$$

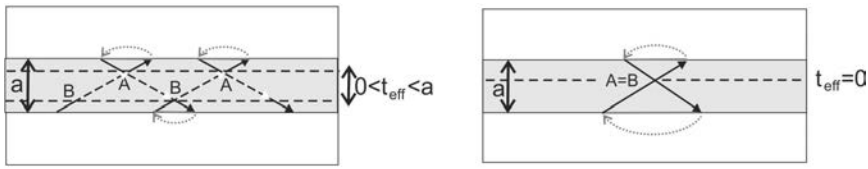
where ω_p is the plasma frequency which can be tuned by changing the distance between the wires (i.e., by changing the thickness of the dielectric slab). Thus from Eq. 6.3, a slab thickness corresponding to an ω_p value such that, $\omega_p > \omega$, will result in a negative electric permittivity ($\epsilon < 0$). By viewing Fig. 6.2, it can be clearly seen that parallel wires are separated by a dielectric slab in a fishnet structure, hence providing a perfect environment for a negative electric permittivity.

6.3. Applications

The growing popularity of metamaterial research can likely be attributed to the fascinating potential applications. Among others, metamaterials have been proposed to be capable of: storing light, breaking the diffraction limit or in other words; achieving super-resolution, making objects invisible, and becoming nearly perfect absorbers (optical black holes).

6.3.1. *Photon storage*

Photon storage is of particular interest for storing qubits in memory for quantum computing, i.e., quantum optical memory. Stopping light or even slowing it down can provide a major technological boost in quantum electronics. Storing light has recently been demonstrated (2007) theoretically by Kosmas L. Tsakmakidis *et al.*, in Ref. 7. The 2007 paper proposes the use of a tapered negative index waveguide core to completely and effectively stop the propagation of light. Analysis of said waveguide reveals the propagation of a light ray which becomes trapped due to the waveguides effective thickness reducing it to zero at a critical point.[7] To illustrate this effect, imagine a cladding of positive refractive index surrounding a waveguide of negative index that gradually decreases in thickness. As the light ray propagates through the negative index core, it will experience a negative refraction (known as a negative Goos-Hänchen lateral displacement[7]) at the interfaces of the mediums. Once the critical waveguide thickness is reached, a double light cone (referred to as



(a) (b)

Fig. 6.3. (a) The negative index waveguide approaching critical thickness. The light ray experiencing negative Goos-Hänchen lateral displacements. (b) A double light cone ('optical clepsydra') forms at critical thickness.

an optical clepsydra in Refs. 7) will be formed, thus permanently trapping the ray of light. This process is shown in Fig. 6.3.

6.3.2. *Perfect lenses*

The sharpness of an image from a conventional lens is limited by the diffraction limit, preventing the fine details from being realized. A conventional lens cannot focus light onto an area smaller than the square of the light's wavelength.[4] To understand this limit, as shown by Pendry in Ref. 4, one can consider an infinitesimal dipole with freedom ω placed in front of a lens. The electric component of the field can be represented by a 2D Fourier expansion,

$$\vec{E}(\vec{r},t) = \sum_{\sigma,k_x,k_y} \vec{E}_\sigma(k_x,k_y) \times exp(ik_z z + ik_x x + ik_y y - i\omega t) \quad (6.4)$$

Using Maxwell's equations and choosing z to be the axis of the lens,

$$k_z = \sqrt{\frac{\omega^2}{c^2} - k_x^2 - k_y^2}, \quad \text{where } \frac{\omega^2}{c^2} > k_x^2 + k_y^2 \quad (6.5)$$

It is shown in Ref. 4 for large values of the transverse wave vector, that propagating waves are limited to: $k_x^2 + k_y^2 < \frac{\omega^2}{c^2}$ such that,

$$k_{max} = \frac{\omega^2}{c^2} \quad (6.6)$$

then the maximum attainable resolution of an image is,

$$R_{max} = \frac{2\pi}{k_{max}} = \frac{2\pi c}{\omega} = \frac{2\pi c}{2\pi f} = \lambda, \quad \text{where } \omega = 2\pi f \tag{6.7}$$

Thus it can clearly be seen that the maximum area one can focus light onto is equal to λ^2.

J. B. Pendry proposed in Ref. 4 that this limit in the optical range of a classical lens can be surpassed using a negative index material (metamaterial).

As shown earlier, the refractive index of a medium will become negative if its electric permittivity and magnetic permeability are simultaneously negative. However the impedance of the medium defined by,

$$Z = \sqrt{\frac{\mu\mu_0}{\epsilon\epsilon_0}} \tag{6.8}$$

will remain positive.[4] Moreover, if one selects values of $\epsilon = -1$ and $\mu = -1$, the medium matches that of a vacuum space allowing for zero reflection at the interface between vacuum and the negative index medium. At both boundaries of the medium (entering and exiting the medium), the impedance match allows the light to stay unchanged. The transmission coefficient of the negative index medium turns out to be,

$$T = e^{-id\sqrt{\omega^2 c^{-2} - k_x^2 - k_y^2}} \tag{6.9}$$

where d is the thickness of the metamaterial slab. This slab can prevent the decay of evanescent waves by amplifying them, as shown in Ref. 4. Sparing the intensive details, after some analysis the transmission coefficient of the S-polarized waves when the limit of ϵ and μ each approach -1 is taken, one obtains,

$$T_s = e^{-ik_z d} \tag{6.10}$$

and similarly taking the same limit for P-polarized waves,

$$T_p = e^{-ik_z d} \tag{6.11}$$

Thus amplification of evanescent waves is achievable using a negative index medium. More importantly, the propagating waves as well as the evanescent waves can contribute to the resolution of the image.[4] This should be a point of emphasis that theoretically speaking, *a metamaterial lens can focus light onto an area smaller than the square of the light's wavelength.*

Quite interestingly, it was proposed that the earlier mentioned limit can even be broken using a thin slab of silver, which alone is not capable of achieving a negative refractive index, only rather a negative electric permittivity given the correct design parameters.

If one considers the dimensions of the silver slab to be smaller than the incident light wavelength, known as the electrostatic limit where,

$$\omega \ll c_0 \sqrt{k_x^2 + k_y^2} \tag{6.12}$$

the transmission coefficient of said slab, as shown in Ref. 4, is completely independent of μ. Following the analysis in,[4] the transmission coefficient in the electrostatic limit turns out to be,

$$\lim_{k_x^2+k_y^2 \to \infty} T = \frac{4\epsilon e^{ik_z d}}{(\epsilon+1)^2 - (\epsilon-1)^2 e^{2ik_z d}} \tag{6.13}$$

It can clearly be seen that the transmission is dependent on the electric permittivity and independent of the magnetic permeability. Thus even a metal such as silver, with a negative electric permittivity, can achieve image focusing beyond the classical limit.

6.3.3. *Invisibility*

The prospect of making objects appear invisible is a fascinating idea that once seemed to be of science fiction. Invisibility now appear to be a real possibility with the advent of the field of metamaterials.

Typically, we are able to manipulate light with homogeneous materials such as lenses to magnify and focus light to produce images.[24] These homogeneous materials generally have a uniform refractive index throughout. However, since metamaterials consist

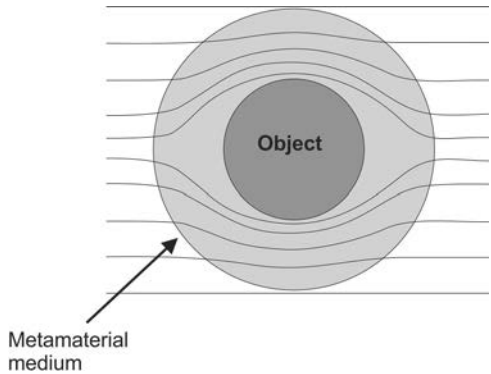


Fig. 6.4. Object surrounded by a metamaterial medium, acting as a cloak, with varying refractive index such that light rays curve around the object, effectively making the object completely invisible.

of sub-wavelength elements to achieve their characteristic negative ϵ and μ, one can arrange those elements in a way to produce a varying ϵ and μ, resulting in a varying refractive index throughout the device. This metamaterial device will then be an inhomogeneous material, which can be tailored to have a gradient of refractive index, capable of manipulating light in ways of our choosing based on its design.[24] With the manipulability of metamaterial structures, one can create a metamaterial with a gradient refractive index designed to curve light around a given object (See Fig. 6.4). Such a design would make the object completely invisible while allowing anything behind the object to be visible as if nothing was in front of it.

Invisibility using metamaterials is now just a sub-field of research within a larger field known as transformation optics.

6.4. Losses in Metamaterials

6.4.1. *Origin of losses*

The current metamaterial structures, namely fishnet and SRR structures, utilize metallics such as copper, gold, or silver to obtain the characteristic properties of a left-handed material. Metals are known

to exhibit high energy dissipation with incident electromagnetic radiation, which is what we would refer to as losses in many practical applications. The fundamental processes which occur that result in losses of metals are rarely discussed in scientific literature; something that should be much more extensively studied since it pertains to losses in metamaterials. Jacob B. Khurgin has some well written work [25–27] who discusses the origin of losses in metallics, offering a unique point of view, as well as making significant connections to metamaterials.

A propagating electromagnetic wave in an optical cavity will exchange energy between two forms, stored in the electric field and stored in the magnetic field. The total energy conservation expression will be,

$$E_E + E_H = E_{total} \tag{6.14}$$

Substituting $E_E = 1/2\epsilon E^2$ and $E_H = 1/2\mu H^2$ to obtain,

$$1/2\epsilon E^2 + 1/2\mu H^2 = E_{total} \tag{6.15}$$

where E and H are the electric and magnetic field magnitudes, ϵ is the electric permittivity, and μ is the magnetic permeability.

If free carriers are introduced (i.e., electrons), such as in a metal, it can be shown that a third form of energy will enter the system in the form of carrier kinetic energy (see Refs. 26 and 27). Equation 6.14 then becomes,

$$E_E + E_H + E_k = E_{total} \tag{6.16}$$

Subsequently, substituting $E_k = 1/2\epsilon_0(\omega_p^2/\omega^2)E^2$, where ϵ_0 is vacuum permittivity, ω and ω_p is the frequency and plasma frequency, respectively, to obtain,

$$1/2\epsilon E^2 + 1/2\mu H^2 + 1/2\epsilon_0(\omega_p^2/\omega^2)E^2 = E_{total} \tag{6.17}$$

It can then be shown that the electric field energy, E_E, is much greater than the magnetic field energy, E_H. In fact, for a confinement scale close to the plasma wavelength, $\lambda_p = 2\pi c/\omega_p$, the magnetic field energy almost vanishes.[25] Thus the majority of the energy will

be exchanged between the electric field and the kinetic motion of electrons.[25] It means that half of the time, nearly all of the energy of the propagating electromagnetic field is stored in the form of carrier kinetic energy. This kinetic energy eventually gets dissipated throughout the metal due to scattering by defects and phonons at a rate of 2γ, where γ is the damping rate.[26]

6.4.2. *Limitations of loss compensation*

So far, reducing losses experienced by metamaterials has proven to be quite a challenging task. However, this has not deterred theorists and experimentalists alike from attempting to achieve low-loss metamaterials which would unleash a plethora of extraordinary technologies. As mentioned earlier, the main problem is the high dissipative loss that occurs due to metallic layers used in the fabrication of a metamaterial device.

One technique to compensate metal loss in metamaterial, and currently probably the most promising, is to introduce an active gain medium with sufficient pumping parameters to achieve full compensation and even lasing.[28] Loss compensation of a metamaterial device through active gain media has been experimentally successful through meticulous fabrication, which is discussed in Section 6.6. However, according to Khurgin *et al.* in Ref. 29, the only successful demonstration of internal miniature lasers in *plasmonic waveguides*, so far, has been the ones which operate with optical pulses rather than continuous signals and are always larger than half of the wavelength in each structure in at least one direction.[29]

Currently, to the best of our knowledge, the only detailed theoretical analysis of a gain integrated plasmonic waveguide was performed by Khurgin *et al.* in Ref. 29. They show that there appears to be general limitations for loss compensation and lasing using active gain media. These apparent limitations should be studied much more extensively and understood before attempting to create a loss compensation or lasing environment. A review of the work by Khurgin *et al.* in Ref. 29 will be presented.

Analysis on two different practical plasmonic waveguides with semiconductor gain materials — one with a silver waveguide and one with a gold waveguide — revealed two interesting findings. The first, pertains to tightly confining plasmonic waveguides such that the confinement is much less than half the wavelength, the modal loss was found to inevitably approach the loss of the metal itself, independent of the geometry of the structure. The second shows that the current densities necessary to achieve complete loss compensation approach unsustainable levels.

If one assumes that the total energy of light is stored in the kinetic energy of carriers for half of a full oscillation, as shown Section 6.4.1. above, then the expression for modal loss can be represented by,

$$\gamma_{eff} = \gamma(1 - n_{eff}^{-2}) \tag{6.18}$$

where γ is the damping rate, and n_{eff} is the effective refractive index.

Calculations can be carried out and then γ and n_{eff} can be plotted as functions of effective waveguide thickness. These plots can be seen in Ref. 29, showing that once n_{eff} reaches a value of approximately 1.5, the modal loss is commensurate with the metal loss i.e., $10^{14}/s$. The calculations assumed no specific geometry and therefore the results are independent of geometry.

To compensate for this modal loss, the modal gain per second was represented by,

$$g(\omega) = B\sqrt{\hbar\omega - E_{gap}}(f_c(\omega) - f_v(\omega))\Gamma \tag{6.19}$$

where f_c and f_v are the Fermi-factors that depend on carrier density in the conduction and valence bands, and B is the stimulated emission coefficient. The confinement factor is expressed as,

$$\Gamma = 2q_s \int_0^{d_a} e^{-2q_s x} dx \tag{6.20}$$

where the thickness of the active layer was chosen to be $d_a = \frac{1}{2}q_s$, allowing for a confinement factor value to be obtained.

The transparency carrier density, when $g(\omega) = \gamma_{eff}(\omega)$, was then plotted as a function of waveguide thickness yielding a reasonable range of: 10^{18}–$10^{19}\,\mathrm{cm}^{-3}$. However, such transparency density range requires extremely high current densities, i.e., $100\,\mathrm{kA\,cm}^{-2}$ for silver waveguides and $300\,\mathrm{kA\,cm}^{-2}$ for gold, at effective waveguide thickness $d_{eff} < 0.25$. These current densities are at minimum two orders of magnitude higher than threshold current densities in high powered double heterostructure semiconductor lasers. The current density expression used in Ref. 29,

$$J_{tr} = \frac{4ed_a\epsilon_s^{3/2}}{c\lambda^2}BF_p\int_{E_{gap}}^{\infty}\sqrt{\hbar\omega - E_{gap}}f_c(\omega)(1 - f_v(\omega))d_\omega \qquad (6.21)$$

includes Purcells factor,

$$F_p = 1 + \pi\Gamma q_s k_x k_s^3 \omega\left(\frac{dk_x}{d\omega}\right) \qquad (6.22)$$

which can get large due to the reduced group velocity and tight confinement which resulted in decreased recombination times by orders of magnitude. See Ref. 29 for specific numbers.

The current densities necessary to sustain full compensation and then lasing seem to be impractically high. However it should be emphasized that though these current densities seem to be unsustainable for *continuous* pumping, they are achievable through optical *pulsed* pumping.

With all this said, it seems that more extensive research should be done in this area to validate the work in Ref. 29. There is much to learn about plasmonics and many more configurations and new materials to explore.

6.5. Loss Reduction Schemes

Techniques for reducing loss in these left-handed materials are imperative in the field of metamaterials. Many proposed techniques attempt to optimize the geometric design of a given metamaterial in order to minimize energy dissipation.

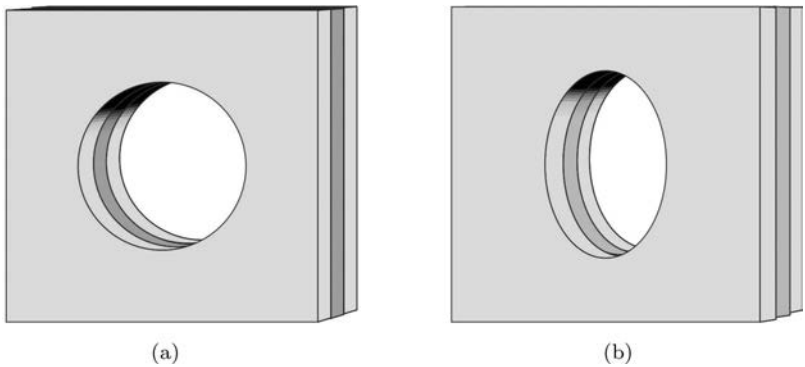


Fig. 6.5. (a) A fishnet structure section with a circular hole instead of a typical rectangular/square hole (b) Fishnet section with an elliptical hole.

6.5.1. *Effects of different hole shapes*

The double-fishnet structure currently seems to be the most effective fabrication of a metamaterial device. As such, exploring various design characteristics of the fishnet structures can provide useful information regarding loss reduction. In Fig. 6.2, the holes of the fishnet structure are shown to be of square or rectangular nature. One can also design a fishnet to have circular or elliptical holes (See Fig. 6.5). Another design characteristic is the dielectric spacer thickness, as alluded to in the previous section. The authors in Ref. 30 numerically analyze the effects on the figure of merit (FOM) of a fishnet structure containing different shaped holes. The main finding was that differing hole shapes can have a large impact on losses.

As a standard of comparison, Dolling *et al.* in Ref. 30, used the highest experimentally verified FOM. At the time of their writing, the highest achieved FOM of a silver-based double-fishnet structure operating at around 1.4 μm was approximately 3, reported in Ref. 31, where this structure contained rectangular holes. Thus, the numerical simulations used the same parameters described in Ref. 31 to ensure consistency. Evidently, the theoretical FOM of a rectangular-hole fishnet structure containing those same parameters turned to be, FOM = 3.72.

To determine the optimum circular-hole structure FOM, they considered the number of free parameters of a rectangular hole structure. Three free parameters are apparent; the width and the length of the rectangle, and the lattice constant. The problem with a circular-hole structure is that there is only one dimensional parameter, being the radius, and the other parameter being the lattice constant; resulting in two free parameters. Thus another parameter must be considered, which was the thickness of the dielectric spacer. The same reasoning can be used for square shaped holes, since the lengths and widths are equal. In principle, since the circular-hole/square-hole structure has the same number of the free parameters as the rectangular-hole structure, one should be able to achieve performance comparable to that of rectangular-holes.

The simulations were done by simultaneously varying the lattice constant and either the width of the square or radius of the circle, depending on which structure was being simulated. The operational wavelength was held fixed at $1.4\,\mu m$, and the highest FOM was searched for as the parameters varied. The retrieved optimum FOM's of the circular-hole and square-hole structures were 1.21 and 1.29, respectively. As they predicted, the FOM's had large differences from the rectangular-hole case (3.72); a difference of 2.51 and 2.43.

These losses were said to occur due to magnetic resonance deterioration leading to the negative real part of the magnetic permeability almost vanishing above resonance. Also, the real part of the electric permittivity becomes overly negative, larger than the rectangular-hole case, due to the larger fraction of metal in the diluted-metal parts. It seems that one needs more free parameters for loss optimization of square/circular holes. From the work presented in Ref. 30, it is apparent that rectangular holes for fishnet structures provide the best configuration for low losses in the metamaterial device.

6.5.2. *Loss reduction via geometric tailoring*

It has been proposed that ohmic losses can be reduced by geometrically tailoring a metamaterial structure. This section will be an overview of the proposed technique in Ref. 32.

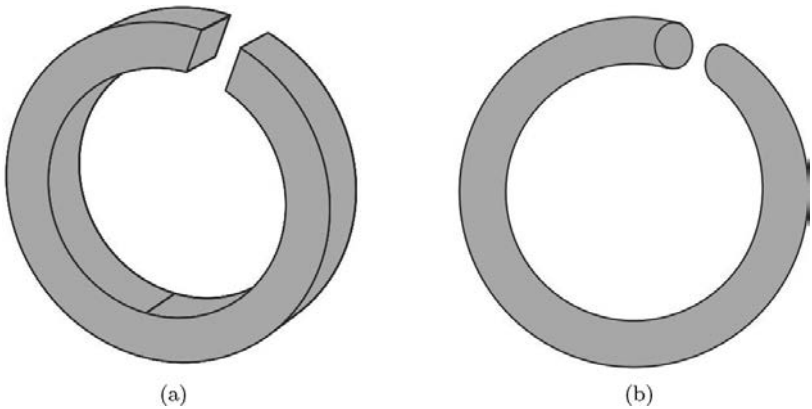


Fig. 6.6. (a) square cross-section (b) SRR with circular cross-section.

6.5.2.1. *Ohmic loss in left-handed materials*

Ohmic loss of a metamaterial is the power dissipation resulting in heat production due to resistance in the metallic layers of the structure.[32] In the high frequency range, the dominant loss component in a metamaterial is ohmic loss. Precisely modifying the geometric details of the structure can significantly change the amount of ohmic loss experienced. In this way, field concentration and current distribution will be affected with different curvatures and precise dimensional changes.[32] A geometric tailoring technique is computationally demonstrated in Ref. 32 for split-ring resonator (SRR) structures, however it should also be applicable to other metamaterials such as fishnet structures.[32] Two types of split-ring resonators are shown in Fig. 6.6.

6.5.2.2. *Low frequency loss*

In general, it was found that sharp corners and edges of a SRRs metallic wires greatly affect current density distributions, as the current tends to concentrate in sharp edged areas of the structure. It can be seen in Fig. 6.7 that the current density is distributed much more uniformly in the SRR with circular cross-section, than in the SRR of

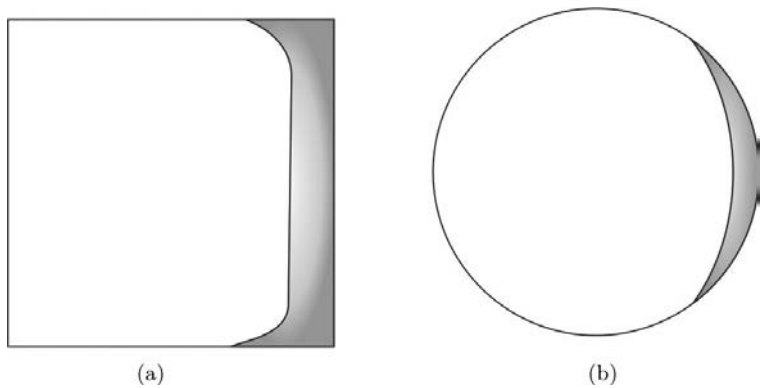


Fig. 6.7. The current density distribution is shown for (a) and (b). Cross-sectional areas are the same for both SRRs. (a) SRR with square cross-section. Current density seems to be increased in the corners. (b) SRR with circular cross-section. The current density seems to be much more evenly distributed.

square cross-section. The square cross-section SRR exhibits increased current concentration in the corners of the wire. Power density goes with the square of current density, thus decreasing the current density concentration in specific areas will decrease the overall ohmic loss.[32] Therefore geometrically tailoring a structure in a metamaterial to have smooth curvatures will, for low frequencies, decrease the power loss experienced by the structure.

6.5.2.3. *High frequency loss*

The smoothening of corners will more uniformly distribute the current density at low frequencies, ultimately resulting in the reduction of power loss. However for high frequencies, the ratio of skin depth to wire thickness is much higher, meaning the current density is already distributed uniformly.[32] The previous technique then is of no use for high frequencies.

To reduce the skin depth ratio, one can increase the wire thickness as can be seen in Fig. 6.8. By increasing the wire thickness toward the bulk limit, the skin depth decreases significantly, effectively confining the current distribution to the surface of the wire. Increasing wire

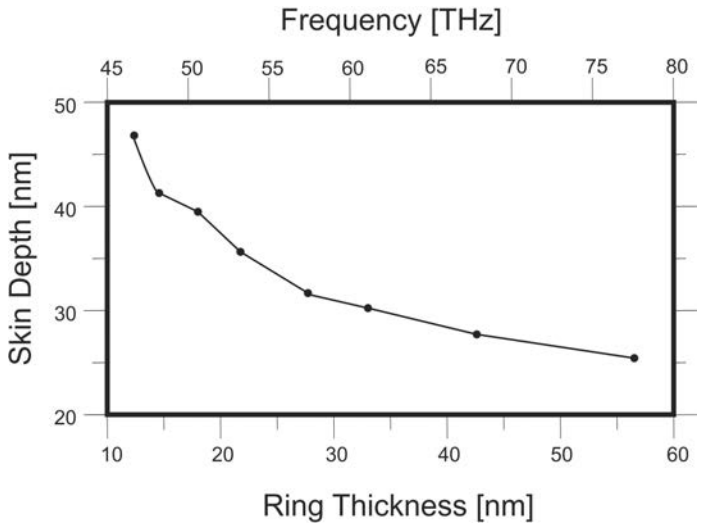


Fig. 6.8. The trend demonstrates that as ring thickness increases, skin depth decreases.

thickness also directly decreases the resistance of a wire. Consider the common resistance equation for a conductive wire,

$$R = \frac{\rho L}{A} \tag{6.23}$$

where R is the total resistance of the wire, ρ is the resistivity, L is the length of the wire, and A is the cross-sectional area of the wire. One can see that by increasing the cross-sectional area (wire thickness), A, the total resistance of the wire, R, will decrease. As alluded earlier, the resistance of the metallic layers is directly involved in the amount of ohmic loss a metamaterial structure experiences; where resistance is proportional to ohmic loss. Thus, reduced skin depth along with increased wire thickness will effectively result in a decrease of ohmic loss in a metamaterial.[32]

However, problems may arise with this technique as one cannot arbitrarily increase wire thickness since wires need to be sufficiently thin to obtain negative electric permittivity in the structure.[19]

6.5.3. *Loss reduction via parameter manipulation*

It can be theoretically shown that increasing the ratio of inductance (L) to capacitance (C) of a fishnet structure, will effectively reduce its losses. This section will be an overview of the findings in Ref. 33.

An important aspect of the findings is that loss reduction can be realized using the proposed technique in the optical frequency range, which so far has been difficult to demonstrate. This method is shown specifically for fishnet metamaterial structures, but it is a generic method that should also be applicable to alternate structures.[33]

6.5.3.1. *Resistance, inductance, and capacitance*

The resonance frequency of a fishnet structure is independent of the structure's separation S. This property can be used to our advantage, as it allows us to change the capacitance and/or the inductance without affecting the resonance frequency.[33] As mentioned earlier, an RLC circuit can be used to model the magnetic element of a fishnet structure unit cell. The magnetic element is the vertical wire that lies along the H axis (see Fig. 6.9). The RLC circuit schematic can be seen in Fig. 6.10.

The expression for total capacitance is given by,

$$C = \frac{1}{2} C_m \tag{6.24}$$

$$C = \frac{1}{4S} \epsilon_r (lw) \tag{6.25}$$

where,

$$C_m = \frac{\epsilon_r (lw)}{2S} \tag{6.26}$$

and total inductance is given by,

$$L = L_m = \frac{\mu_r (lS)}{w} \tag{6.27}$$

Parameters l, S, and w are shown in Fig. 6.9.

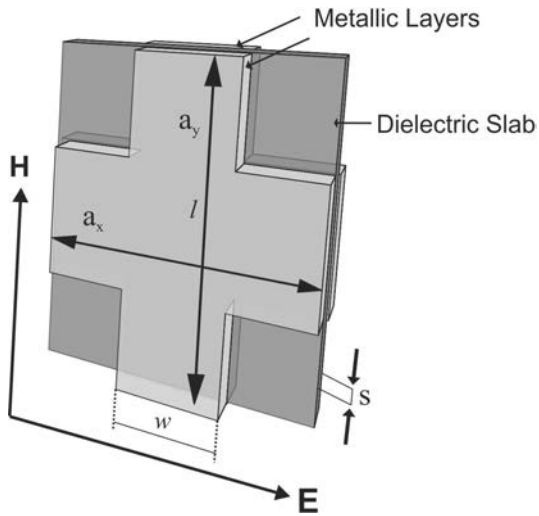


Fig. 6.9. Front facing view of a fishnet structure.

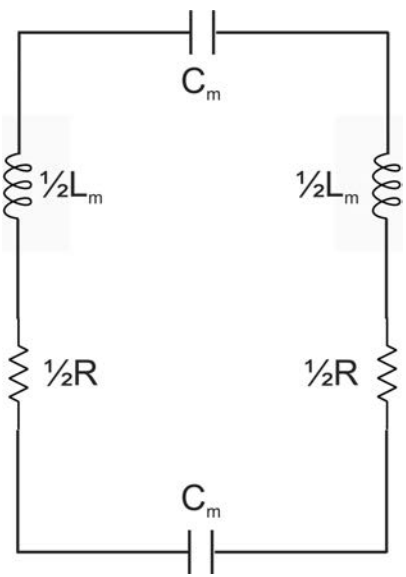


Fig. 6.10. RLC circuit model.

It is known that for RLC circuits, the loss is indirectly proportional to the quality factor,[33]

$$Q = \frac{1}{2R}\sqrt{\frac{L}{C}} \tag{6.28}$$

It can immediately be seen from Eq. 6.28 that by decreasing the resistance R and/or decreasing the capacitance C and/or increasing the inductance L, the quality factor Q can be increased. Due to the relationship of Q and the loss, the losses of the RLC circuit, and ultimately the fishnet structure metamaterial, can effectively be reduced by varying R,L, and C.[33]

Reducing the resistance R can simply be done by choosing a good conductor material such as copper or gold. Inductance L increases linearly with the spacing S (from Eq. 6.27), while the capacitance C decreases linearly with S (Eq. 6.26). These proportionalities work in our favour. One can increase the thickness of the dielectric slab which increases the spacing between the metallic layers, which will increase the inductance while simultaneously decreasing the capacitance.

Increasing the separation S cannot be done arbitrarily, as it is restricted by the unit cell a_z which can be defined from Fig. 6.9 as the perpendicular vector to a_x and a_y. The separation can go no larger than a_z, and a_z must be much smaller than the wavelength of electromagnetic radiation.[33] It must also be kept small enough to not affect the negative electric permittivity of the long wires.[33] Another potential problem occurs according to the computational findings in Ref. 33 that the magnetic resonance disappears after the spacing is increased to a certain value larger than the width of the wires. This apparently occurs because the short wires decouple from each other at that critical value.[33]

6.5.4. *Optical parametric amplification*

Optical parametric amplification was theoretically proposed in Ref. 34 as a means of compensating loss in negative index materials. The scheme utilizes a positive-index auxiliary electromagnetic field

signal as a control field which enables a loss-balancing optical parametric amplification for a negative-index signal wave.[34] This section will present a review of the proposed loss compensation scheme.

The setup considers a negative-index wave signal of frequency ω_1 which has a wave vector $\vec{k_1}$ in the positive direction, let us say *along* the z-axis. This negative-index signal's Poynting vector $\vec{S_1}$ is anti-parallel to $\vec{k_1}$; directed *against* the z-axis. The auxiliary control signal travels in a positive-index domain, and hence is a positive-index signal along the z-axis with frequency ω_3 and wave vector $\vec{k_3}$. This positive-index wave has a Poynting vector $\vec{S_3}$ which is directed parallel to $\vec{k_3}$. The two coupled waves with parallel wave vectors $\vec{k_1}$ and $\vec{k_3}$ will give rise to an idle wave of frequency, $\omega_2 = \omega_3 - \omega_1$, that travels in the positive-index domain. The idle wave then has a wave vector $\vec{k_2}$ and a Poynting vector $\vec{S_2}$ that travel along the z-axis. Figure 6.11 shows the direction of each signal's wave vectors and Poynting vectors. The generated frequency-difference idle wave will enable a phase-matching scheme for optical parametric amplification

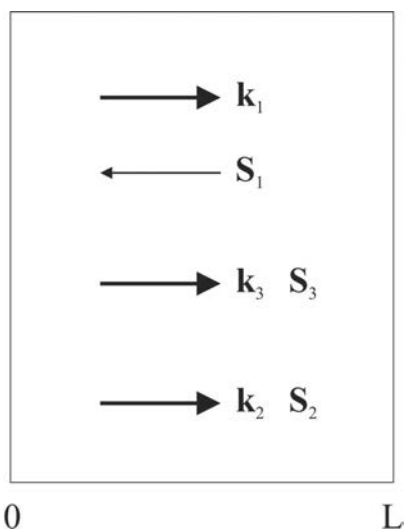


Fig. 6.11. The direction of each signal's wave vectors and Poynting vectors. L is the slab thickness.

at ω_1 through three-wave mixing by converting the pump field energy at ω_3 into a signal.[34]

Following the analysis in Ref. 34, the normalized amplitudes for the signal and idler waves can be expressed as,

$$a_j = \left(\frac{e_j}{\mu_j}\right)^{\frac{1}{4}} \frac{h_j}{\sqrt{\omega_j}} \tag{6.29}$$

Then the rate of change of amplitudes along the z-axis for the signal waves and idler waves can be written as,

$$\frac{da_1}{dz} = -iga_2^* e^{i\Delta kz} + a_1\left(\frac{\alpha_1}{2}\right) \tag{6.30}$$

$$\frac{da_2}{dz} = -iga_1^* e^{i\Delta kz} - a_2\left(\frac{\alpha_2}{2}\right) \tag{6.31}$$

where,

$$g = \sqrt{\omega_1\omega_2}\left(\frac{\epsilon_1\epsilon_2}{\mu_1\mu_2}\right)^{-\frac{1}{4}}\left(\frac{8\pi}{c}\right)\chi^{(2)}h_3 \tag{6.32}$$

and $\Delta k = k_3 - k_2 - k_1$. The pump amplitude h_3 is assumed to be a constant, α_j are the absorption coefficients. The values a_1^2 and a_2^2 are proportional to the number of photons at each corresponding frequency.

Using boundary conditions: $a_1(L) = a_{1L}$ and $a_2(0) = a_{20}$ where L is the slab thickness, one can solve Eqs. 6.30 and 6.31 to obtain solutions,

$$a_1(z) = A_1 e^{\beta_1^+ z} + A_2 e^{\beta_2^+ z} \tag{6.33}$$

$$a_2^*(z) = \kappa_1 A_1 e^{\beta_1^- z} + \kappa_2 A_2 e^{\beta_2^- z} \tag{6.34}$$

where

$$\beta_{1,2}^{\pm} = \beta_{1,2} \pm (i\Delta k/2) \quad \text{and} \quad \beta_{1,2} = \frac{1}{4}(\alpha_1 - \alpha_2) \pm iR$$

$$A_{1,2} = \pm\frac{1}{D}\left[a_{1L}\kappa_{2,1} - a_{20}^{*}e^{\beta_{2,1}^{+}L}\right]$$

$$D = \kappa_2 e^{\beta_1^{+}L} - \kappa_1 e^{\beta_2^{+}L}$$

$$\kappa_{1,2} = \frac{1}{g}(\pm R + is) \quad \text{and} \quad R = \sqrt{g^2 - s^2}$$

$$s = \frac{1}{4}(\alpha_1 + \alpha_2) - i\Delta k/2$$

Then the amplification factor is,

$$\eta_{1a}(z) = \left|\frac{a_1(z)}{a_{1L}}\right|^2 \tag{6.35}$$

and for the case $a_{1L} = 0$,

$$\eta_{1a}(z) = \left|\frac{a_1(z)^*}{a_{20}}\right|^2 \tag{6.36}$$

Upon simulation, resulting graphs show that amplification of several orders of magnitude occurs for phase-matched waves. Figure 6.12 shows a plot of the output amplification factor, η_{1a}, versus gL, where

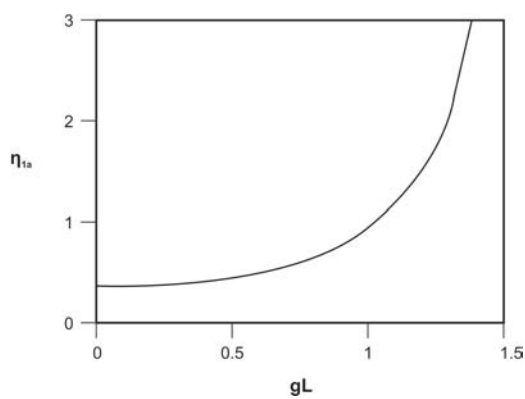


Fig. 6.12. The output amplification factor, η_{1a}, plotted against gL.

g is from equation 6.35. Such amplification can provide complete loss compensation of a metamaterial medium.

6.6. Active Gain

Gain in metamaterials can be incorporated utilizing fluorescent dyes,[35] and quantum dots,[36,37] and quantum wells.[38,39] Several metallic structures were investigated, including SRR, inverted SRR, and fishnets. Gain material is typically modelled as a four-level system.[40,41]

A simple model of plasmonic material with gain known as toy model[14] will now be summarized. The model can provide some general insight about the influence of gain.

6.6.1. *Toy model*

Several years ago a simple model of a metamaterial incorporating optical gain was proposed by Wegener *et al.*[14] It consists of a fermionic two-level system (TLS) which represents gain medium and a single bosonic resonance which represents plasmonic resonance of metamaterial. The systems are connected via a local-field coupling, see Fig. 6.13.

Fermionic TLS is described by the Bloch equations

$$\frac{dp_{TLS}}{dt} + (i\Omega_{TLS} + \gamma_{TLS})p_{TLS} = \frac{i}{\hbar} d_{TLS}(E + L\,p_{pl})(1 - 2f) \tag{6.37}$$

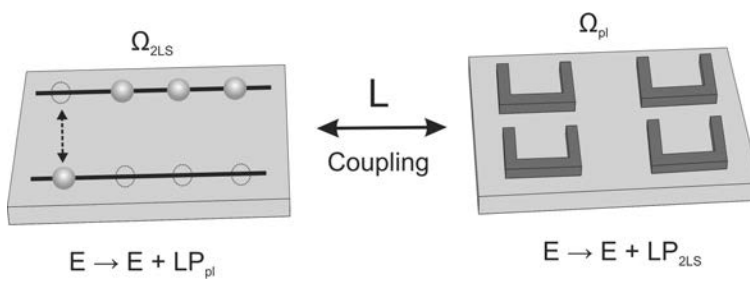


Fig. 6.13.

$$\frac{df}{dt} + \Gamma_{TLS} f = \frac{i}{\hbar}(p^*_{TLS} d_{TLS}(E + L\, p_{pl}) - p_{TLS} d^*_{TLS}(E + L\, p_{pl})^*) \tag{6.38}$$

In the above expressions, p_{TLS} represents (dimensionless) complex transition amplitude, $f = f_{TLS}$ is the occupation probability of the upper level, d_{TLS} is the dipole matrix element, Ω_{TLS} is the transition frequency, γ_{TLS} is the damping or transverse relaxation rate, Γ_{TLS} is the longitudinal relaxation rate, L is the phenomenological coupling constant between fermionic and bosonic systems. Pump rate Γ_{pump} can also be added on the right-hand side of Eq. (6.38). It will account for pumping of TLS via additional energy levels.

The equation of motion for bosonic mode is

$$\frac{dp_{pl}}{dt} + (i\Omega_{pl} + \gamma_{pl}) p_{pl} = \frac{i}{\hbar} d_{pl}(E + L\, p_{TLS}) \tag{6.39}$$

with Ω_{pl} being the plasmonic frequency, γ_{pl} is the damping rate and d_{pl} dipole matrix element. Occupation factor for plasmonic mode is $(1 - 2f) = 1$. E the external electric field of the light. As can be seen from the above equations the coupling between both systems is provided by the local-field effects.

For the toy model, effective parameters can be obtained from the Maxwell-Garrett effective-medium theory. Some of the results are illustrated in Fig. 6.14 where the real part of the refractive index is shown with and without coupling L for an increasing values of occupation probability f from 0 to 1.[14]

Time-dependent analysis with a pump rate $\Gamma_{pump} = \Gamma_0(1 - f)$ reveals relaxation oscillations. Those can be obtained by numerical

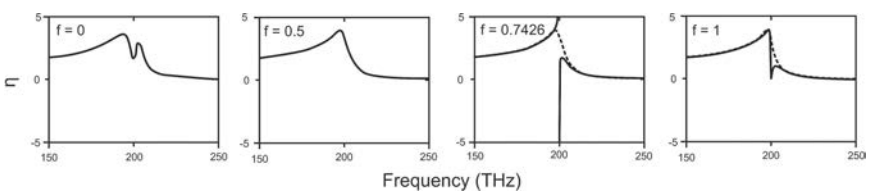


Fig. 6.14. The real part of the refractive index plotted against frequency with the occupation probability increasing from 0 to 1.

integration of the above equations. The results can help in designing practical spaser structures.

6.6.2. *Maxwell-Bloch Theory*

In this Section we summarize the fundamental equations which are used in the study of dynamical processes inside metamaterials. The approach involves combining the finite-difference time-domain (FDTD) method with auxiliary differential equations that self-consistently describe both, the material dispersion of the plasmonic cladding and the nonlinear, spatially resolved polarization response of the gain medium.

Gain medium is modelled as a four-level system with pump operating between level 0 and level 3;[42] the emission transitions take place between levels 2 and 1, see Fig. 6.15. The approach is known as the Maxwell-Bloch (MB) formalism.

The spatio-temporal dynamics of the occupation densities N_0 to N_3 are given by the set of equations[40]

$$\frac{\partial N_3}{\partial t} = \frac{1}{\hbar\omega_{r,a}}\left(\frac{\partial \mathbf{P}_a}{\partial t} + \Gamma_a \mathbf{P}_a\right)\cdot \mathbf{E}_{loc} - \frac{N_3}{\tau_{32}} \tag{6.40}$$

$$\frac{\partial N_2}{\partial t} = \frac{N_3}{\tau_{32}} + \frac{1}{\hbar\omega_{r,e}}\left(\frac{\partial \mathbf{P}_e}{\partial t} + \Gamma_e \mathbf{P}_e\right)\cdot \mathbf{E}_{loc} - \frac{N_2}{\tau_{21}} \tag{6.41}$$

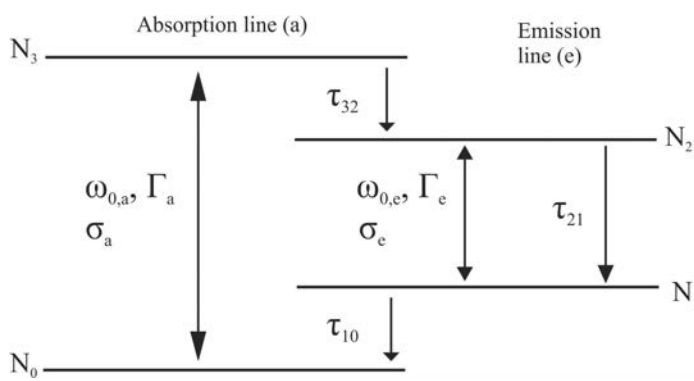


Fig. 6.15. Schematic of the four-level system and its parameters.

$$\frac{\partial N_1}{\partial t} = \frac{N_2}{\tau_{21}} - \frac{1}{\hbar\omega_{r,e}}\left(\frac{\partial \mathbf{P}_e}{\partial t} + \Gamma_e \mathbf{P}_e\right)\cdot \mathbf{E}_{loc} - \frac{N_1}{\tau_{10}} \tag{6.42}$$

$$\frac{\partial N_0}{\partial t} = \frac{N_1}{\tau_{10}} - \frac{1}{\hbar\omega_{r,a}}\left(\frac{\partial \mathbf{P}_a}{\partial t} + \Gamma_a \mathbf{P}_a\right)\cdot \mathbf{E}_{loc} \tag{6.43}$$

The evolution of polarization densities $\mathbf{P}_a = \mathbf{P}_a(\mathbf{r}, t)$ of the transition $0 \leftrightarrow 3$ and $\mathbf{P}_e = \mathbf{P}_e(\mathbf{r}, t)$ of the transition $1 \leftrightarrow 2$ under the local electric field $\mathbf{E}(\mathbf{r}, t)$ is described by the differential equations $(i = a, e)$

$$\frac{\partial^2 \mathbf{P}_i}{\partial t^2} + 2\Gamma_i \frac{\partial \mathbf{P}_i}{\partial t} + \omega_{0,i}^2 \mathbf{P}_i = -\sigma_i \Delta N_i \mathbf{E}_i \tag{6.44}$$

The resonance frequencies are defined as $\omega_{0,i} = (\omega_{r,i}^2 + \Gamma_i^2)^{1/2}$, $\Delta N_a(\mathbf{r}, t) = N_3(\mathbf{r}, t) - N_0(\mathbf{r}, t)$ is the inversion of the pump transition and $\Delta N_e(\mathbf{r}, t) = N_2(\mathbf{r}, t) - N_1(\mathbf{r}, t)$ is the inversion of the probe transition, σ_i is a phenomenological coupling constant.

Details of the numerical implementation are provided in Ref. 43.

6.6.3. *Approach by Wuestner et al. (2010)*

Wuestner et al applied MB approach and considered two configurations, passive and active. In the passive configuration two silver fishnet films were embedded inside a dielectric host with a value of refractive index $n_h = 1.62$. The permittivity of silver was modeled by Drude approach corrected by two Lorentzian resonances to match experimental data at visible wavelengths. In the active configuration it is shown that incorporation of a gain medium in a structure of a double-fishnet metamaterial (see Fig. 6.16) can fully compensate losses in the regime where the real part of the refractive index is negative.

6.6.4. *Experiment by Xiao et al.*

It has been experimentally demonstrated in Ref. 35 by Xiao *et al.* that a fishnet metamaterial structure with an active gain medium spacer can effectively overcome optical losses, resulting in an essentially lossless negative index material.

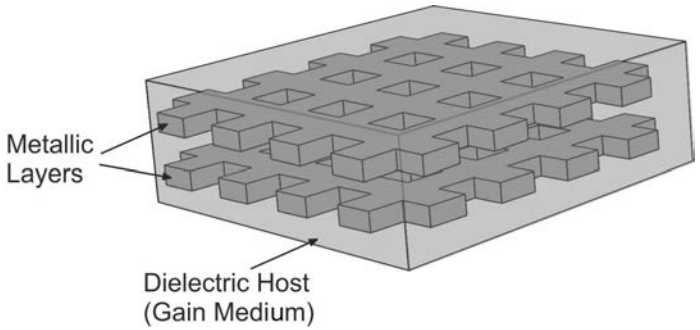


Fig. 6.16. Two silver fishnet films embedded in a dielectric host consisting of gain material (dye molecules).

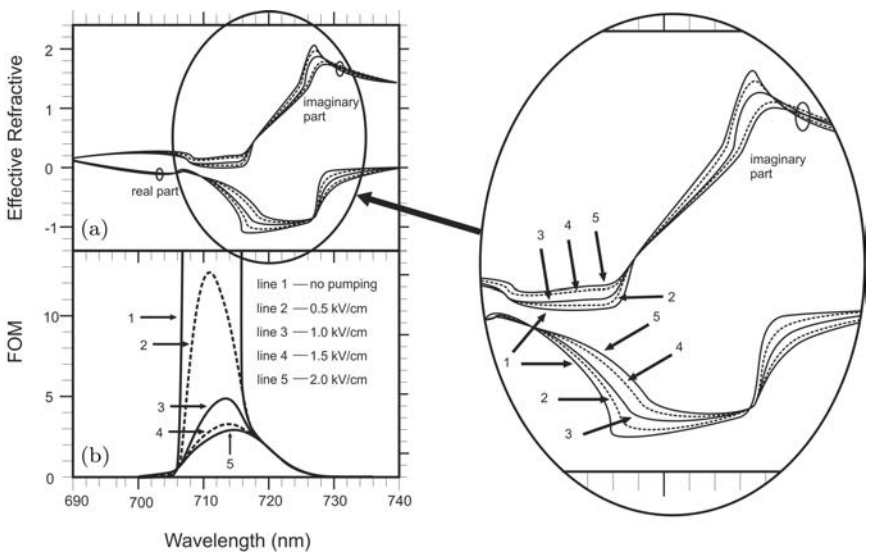


Fig. 6.17. Real and imaginary part of the retrieved effective refractive indices of the double-fishnet structure for different pump amplitudes.

The typical dielectric medium spacer used in fishnet material (see Fig. 6.2) was replaced with a gain medium. A fishnet structure was initially fabricated with an Alumina spacer. This Alumina spacer was then removed and immediately replaced with the gain medium

which was the epoxy medium doped with rhodamine 800 (Rh800) dye molecules.

The active fishnet structure can be optically pumped to form a population inversion within the Rh800 dye molecules. The laser pump must be of frequency, $\omega = (E_3 - E_0)/\hbar$, which corresponds to a wavelength equivalent to the stimulated emission wavelength of the dye molecules. The quantum processes which occur to compensate for loss can be described using the four-level system shown in Fig. 6.15. The device is pumped by a light pulse which will propagate through the device and excite electrons within the Rh800 molecules from the ground state level, N_0, to the third energy level, N_3, (see Fig. 6.15). This is known as a population inversion. After a lifetime, τ_{32}, excited electrons will transition from, N_3, to, N_2, either radiatively or non-radiatively. Then after another lifetime, τ_{10}, electrons in the, N_1, energy level will relax to the ground state level, N_0, and the process will repeat.[28]

Measurements were obtained using a pump-probe setup, in which the delay time between the pump and the probe was optimized for maximal gain. The results are shown in Fig. 6.18 where the black line is the transmission spectrum with pumping power 1 mW, and the gray line is the transmission spectrum with no pump. It can clearly be seen that the losses are significantly decreased when pumping the

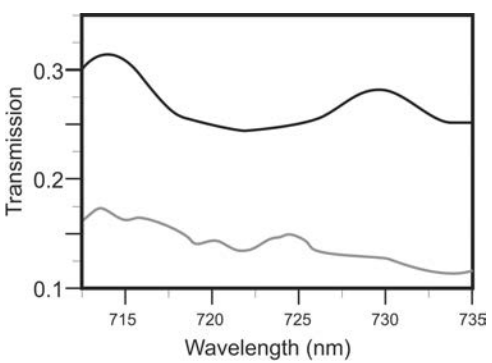


Fig. 6.18. Transmission spectra with no pumping (gray) and of pumping power 1 mW (black line).

gain medium. The radiative transitions from N_3, to, N_2, amplify the probe light, and thereby increase the transmission as it propagates through the device. This is only possible if the correct delay time is chosen. Ideally, the probe light must happen when the radiative transitions are occurring.

This experimental work conclusively shows that, although the metamaterial structure must be meticulously constructed, it is possible to completely overcome losses in a realistic negative index material by replacing the typical dieletric spacer with a gain medium.

6.6.5. *Experiment by Plum et al.*

As mentioned earlier, gain can be incorporated into a metamaterial device using quantum dots. The first experimental demonstration of the use of semiconductor quantum dots to fully compensate for Joule (ohmic) loss in a metallic metamaterial device was done by Plum *et al.* in 2009 in Ref. 36.

The metamaterial structure was a hybrid metamaterial device consisting of an array of asymmetrically split ring resonators made of gold, which sat in between a silica gain substrate (below) and a layer of semiconductor PbS quantum dots (above). Asymmetrical SRR's were chosen since the quality factor of the trapped-mode resonance is controllable by changing the amount of asymmetry in the split.[36] The asymmetrical SRR's can also create arrays of trapped-mode currents which would allow the device to form a lasing spacer, where the gain substrate could be used as an energy source to support the spaser mechanism. See Section. 6.7.1 for a discussion on spasers.

An interesting finding that resulted from the deposited layer of quantum dots was a large red-shift in the transmission spectrum, moving from 860 nm to 1000 nm. It was reasoned in Ref. 36 that the red-shift was due to the shortened excitation wavelength caused by the increased effective permittivity.

The more relevant and significant finding was the reduction in Joule loss due to the optically pumped quantum dots. The transmission of a quantum dot layer atop a glass substrate was

first measured, and found to steadily increase as the pump intensity increased. This transmission increase was attributed to the quantum dot luminescence. When the quantum dots within the hybrid metamaterial device were optically pumped, the transmission spectrum at the trapped mode resonance frequency was significantly altered, in the background of quantum dot luminescence. The transmission spectrum modification was separated from the quantum dot luminescence which depended upon pump intensity, by introducing pump-induced change of the differential transmission signal for two perpendicular polarizations, which was then plotted. It can clearly be seen (see Ref. 36) that the luminescence signal has no dependence on the polarization of the probe light used to measure the transmission. Thus, it can be concluded that any change in the transmission spectrum is solely caused by gain in the metamaterial device. It was argued that the gain must occur due to the strong interaction between the optically pumped quantum dot layer and surface plasmon modes.

Thus, Plum *et al.* have shown that by creating a hybrid metamaterial device with an optically pumped quantum dot layer, one can effectively reduce Joule loss within the device.

6.6.6. *Powered active metamaterial cells*

Experiments have successfully shown loss compensation through powered active cells, however only in unrealistic configurations.[16] Yuan *et al.* in Ref. 16 were able to experimentally demonstrate a realistic tunable powered active-magnetic metamaterial, capable of completely compensating for loss at certain frequencies.

The reasoning in Ref. 16 for their design considered the magnetic moment of a unit cell and the resulting effective magnetic permeability. The relationship between the magnetic field, $\vec{B}$, and the magnetic field strength, $\vec{H}$, where magneto-electro coupling is nonexistent, can be expressed as,

$$\vec{B} = \mu_0(\vec{H} + \vec{M}) \tag{6.45}$$

where $\vec{M}$ is the magnetization vector. Using Eq. 6.45, and recognizing the relation,

$$\vec{B} = \mu_0 \mu_r \vec{H} \tag{6.46}$$

one can determine the relative permeability tensor. Assuming $\vec{B}$, $\vec{H}$, and $\vec{M}$ are colinear, the relavant component of the permeability tensor can be expressed by,

$$\mu_r = 1 + \frac{\vec{M}}{\vec{H}} \tag{6.47}$$

where magnetization $\vec{M}$ is the resulting magnetic moment due to the magnetic strength field, $\vec{H}$, and thus can be given by,

$$\vec{M} = |M| e^{i\phi} \tag{6.48}$$

where ϕ is the phase of $\vec{M}$ with respect to $\vec{H}$.

Using[44] as a reference, the authors in Ref. 16 recognized that control of the phase ϕ from Eq. 6.48 enables control of the real and imaginary parts of the magnetic permeability. The design of the metamaterial device used this fact to enable for high tunability.

The experiment utilized an active-magnetic metamaterial consisting of arrays of unit cells, where each unit cell was powered by a DC supply. The unit cell design was inspired by the work in Ref. 44. Each unit cell comprised of a sensing loop for the incident magnetic field, a radiofrequency (RF) amplifier to amplify the sensed signal, and a driven loop in order to tune the magnetization vector $\vec{M}$. Adjacent to the sensing and driven loops was an SRR, which reduced the amount of active gain required for the device.[16] The RF amplifier inputs a signal into a voltage-controlled phase shifter which controls the phase at the driven loop. The phase shifter turns out to be of high importance as it distinguishes the unit cell from any other previously designed powered active metamaterial. This phase shifter is ultimately responsible for controlling the real and imaginary parts of the effective magnetic permeability. Since the phase shifter is voltage-controlled, the unit cell then has a highly tunable magnetic permeability property. This unique design and capability can be of

great significance when considering loss in the magnetic regime of a metamaterial device.

Upon testing a single unit cell by applying DC bias voltages ranging from 0 V to 12 V, it was found that at each applied voltage there was a strong magnetic response just above the self resonant frequencies of the SRRs. Zero magnetic loss frequencies were found at corresponding points of real ($\mu_r^{'}$) and imaginary ($\mu_r^{''}$) permeability, at specific bias voltages. More interestingly, at certain $\mu_r^{'}$ and $\mu_r^{''}$, it was found that zero magnetic loss is possible for an array of unit cells. However, one cannot choose an arbitrary frequency as it was found that the electric response can become large, resulting in a device that exhibits zero magnetic loss but nonzero overall loss. For certain frequencies, the electric antiresonance is negligible allowing the material to be completely lossless. See Ref. 16 for specific numbers and figures.

The authors in Ref. 16 experimentally demonstrated loss compensation of a realistic metamaterial device, albeit only for certain frequencies. This work however provides a great step forward for low-loss metamaterials.

6.7. Surface Plasmons

Surface plasmons (SPs) can be described as collective and coherent oscillations of electron-fields confined to the surface of a medium. Plasmons are collective oscillations of a free electron gas within a medium, away from the surface. Metallics in particular have surface plasmons that are tightly bound to the surface, which makes them effective scatterers of incident electromagnetic waves.[45] As mentioned in Section 6.4., much of the energy of electromagnetic waves gets dissipated by the kinetic energy of electrons. This kinetic energy is usually in the form of coherent oscillations, i.e., (surface) plasmons.

Surface plasmons can also propagate along the surface of metals, reaching distances of up to several hundreds of microns.[45] Since propagating SPs are moving charges, an electromagnetic wave is generated on the interface of the metal and dielectric. These propagating

electromagnetic field waves which couple with the moving SPS are known as surface plasmon polaritons (SPPs).[45]

6.7.1. *Spaser*

The spaser (acronym for "*surface plasmon amplification* by *stimulated emission* of *radiation*") device was first proposed by Bergman and Stockman in 2003. A spacer is analogous to a laser, the main difference being that, as the name suggests, a spaser emits surface plasmons instead of photons. The resonant cavity of a spaser is a nanoparticle, such as a semiconductor quantum dot. An externally excited gain medium can then be used as the energy source for spasing to occur.[46]

Stockman proposed a particular design of a spacer in Ref. 46 to consist of a dielectric core surrounded by a silver nanoshell, which was then surrounded by a few monolayers of nanocrystal quantum dots (NQD) (see Fig. 6.19). The dielectric core surrounded by the silver nanoshell would then act as the resonant nanoparticle and the NQDs would act as the gain medium.

The idea is to use an external optical radiation source to excite an electron-hole pair from an NQD, which will eventually relax into

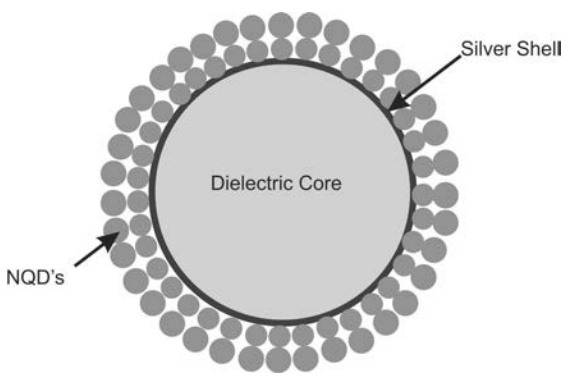


Fig. 6.19. Proposed spaser setup with an interior dielectric core surrounded by a silver shell which is surrounded by nanocrystal quantum dots.

an excitonic level state due to carrier multiplication.[46] Typically the recombination of an exciton would result in the generation of a photon, however since the NQDs are in contact with a resonant nanoparticle, the energy is transferred virtually radiationless to the resonant SPs of the nanoparticle. The excited SPs will then continue to stimulate electron-hole transitions from the gain medium, which will in turn excite more SPs.[46] This is precisely the process of a spaser.

6.7.2. *Metamaterial lasing spaser*

Zheludev *et al.* combined the concepts of spasers along with metamaterials in Ref. 47 to propose a metamaterial lasing device which uses plasmonic oscillations to generate a coherent source of electromagnetic radiation. The proposed device consists of a slab of gain medium containing an array of metallic asymmetric split-ring resonators (See Fig. 6.20).

This array must be a special type of metamaterial array of plasmonic resonators that support a high quality factor for current oscillations in which total emission losses for the in-phase oscillations are at a minimum. These metamaterial arrays were referred to as coherent metamaterials. It was recently shown in Ref. 48 that SRR's

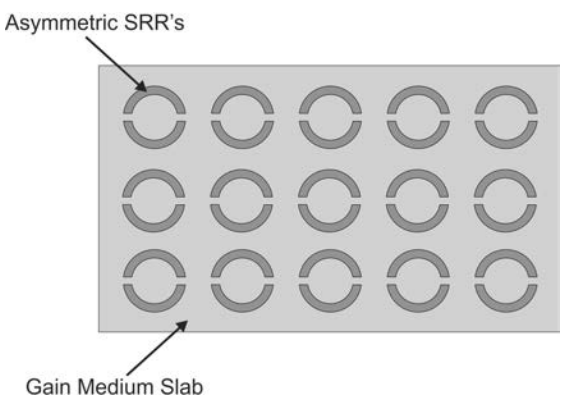


Fig. 6.20. Slab of gain medium containing an array of metallic asymmetric SRR's.

with weak antisymmetry may be excited by a high-quality mode of intense anti-symmetric current oscillations.[47]

The active medium of the lasing spaser is the combination of the electromagnetic resonators where the supporting gain substrate is the energy source. Since the trapped-mode oscillations in a plasmon resonator are typically non-radiative, a small asymmetry can be introduced to make the resonator radiative. Some of the energy that would typically by trapped could then be emitted into free space as electromagnetic radiation.[47] The authors in Ref. 47 drew a comparison to that of the leakage of radiation through the output coupler of a laser resonator. The electromagnetic emission would then be directed perpendicular to the plane of the metamaterial array slab (see Fig. 6.21).

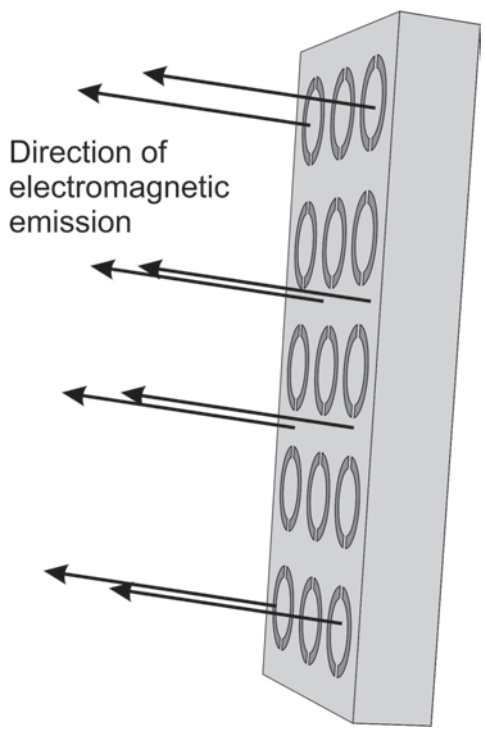


Fig. 6.21. Direction of electromagnetic emission.

6.8. Moving in the Opposite Direction: Increasing Losses to Create Metamaterial Absorbers

The inherent lossy effects of metallics that are used for metamaterial structures severely plague metamaterial practicality in a plethora of ways. However, if one were to look from a different perspective, the high lossy effects could be used to our advantage. Loss in metamaterials is the result of the metallic's ability to absorb incident radiation and convert it into another form of energy (typically heat). One can increase the losses in metamaterials to create near perfect electromagnetic absorbers. Techniques can then be devised to convert the absorbed energy into useful electrical energy rather than heat. In theory, the absorptivity of a metameterial absorber can approach unity, mimicking a miniature black hole.[49] Such an absorber can produce practical technological advancements such as photodetectors, microbolometers, thermal images, and novel solar cell systems.[50]

As it is well known, the electric permittivity and magnetic permeability is characterized by real and imaginary parts such that,

$$\epsilon = \epsilon_{re} + i\epsilon_{im} \tag{6.49}$$

$$\mu = \mu_{re} + i\mu_{im} \tag{6.50}$$

Most of the focus regarding optimizing metamaterials for practical use has been on the real parts of ϵ and μ. However in the case of creating novel applications where loss is required, the imaginary parts (ϵ_{im} and μ_{im}), which are the loss components of the permittivity and permeability, must become points of interest. Independent manipulation of the resonances of ϵ_{im} and μ_{im} can result in a metamaterial that can absorb both incident electric and magnetic fields. By matching the two variables, the metamaterial can be impedance-matched to free space which will in turn minimize reflectivity.[51]

Landy *et al.* in Ref. 51 proposed, fabricated, and then numerically simulated and experimentally tested, a metamaterial absorber structure which demonstrated a near perfect absorbance. The absorber, made exclusively with metallic elements, consisted of a type of a split-ring resonator, referred to as the electric-ring resonator (ERR,

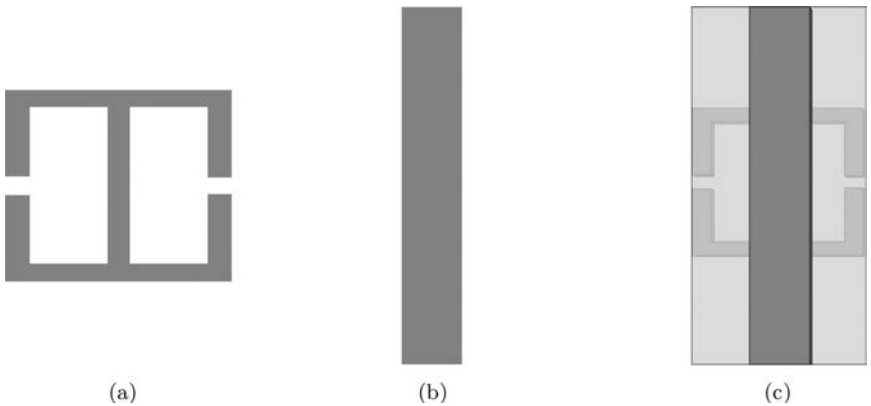


Fig. 6.22. (a) Electric-ring Resonator (ERR) (b) Cut wire (c) The metamaterial absorber unit cell consisting of the ERR and the cut wire separated by a dielectric substrate.

shown in Fig. 6.22(a)), and a cut wire (Fig. 6.22(b)) separated by a dielectric substrate, as shown in Fig. 6.22. The electric coupling was attributed to the ERR and the magnetic coupling was supplied by the final combination of the ERR and the cut wire separated by the substrate (shown in Fig. 6.22(c)). The authors drew a comparison to that of a fishnet structure where two anti-parallel currents are driven in metallic layers that give rise to the magnetic response, as explained in Section. 6.2.2. The magnetic response was then tunable by changing the shape of the cut wire and the spacing between the cut wire and ERR (dielectric slab thickness). Thus the magnetic coupling was manipulable independent of the ERR, allowing ϵ and μ to be decoupled which enabled each response to be individually tunable.[51]

The experimental absorbance data was obtained and plotted along with simulation data for comparison. The simulated data showed an absorbance peak at 96% at 11.48 GHz, while the experimental data showed a peak at approximately 88%. The plots can be seen in Fig. 6.23. Thus the authors in Ref. 51 were able to experimentally demonstrate a near perfect metamaterial absorber that excellently agreed with the simulated data.

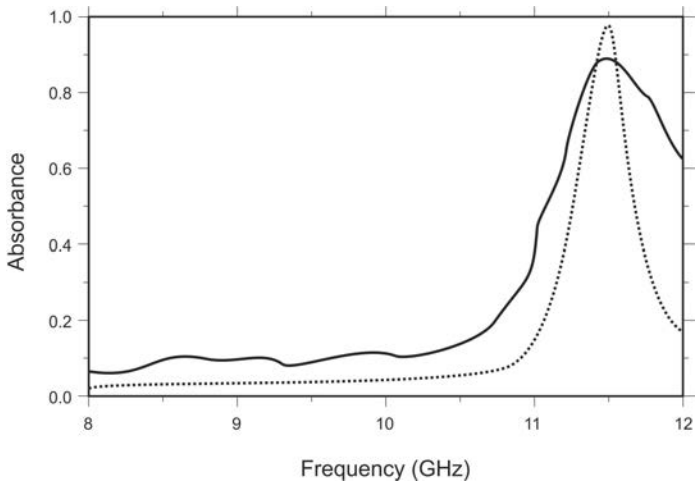


Fig. 6.23. The solid line is the measured data (experimental) and the dotted line is the simulated data.

On one side of the spectrum, multitudes of research is being done towards reducing losses in metamaterials in hopes of creating novel devices, but it seems that much more research should be done on the other end of the spectrum where loss can be significant for equally novel devices. The work done by the authors in Ref. 51 is a great step in that direction.

6.9. The Future for Metamaterials

A central problem hindering the practicality and progress in the development of negative index materials is the substantial dissipative losses which are exhibited. Reducing such losses via geometric tailoring seems to work up to a certain extent but not enough to completely overcome the losses. The future seems to lie in the active metamaterials where experiments such as the ones mentioned in Section 6.6 conclusively show total loss compensation using active gain media. Even though the fishnet structures used in the mentioned experiments are judiciously designed, in ways that may be far from practical designs, they should not be overlooked. Building upon these

foundational experimental and theoretical works is imperative if we are to see metamaterial devices meet their true potential.

One interesting idea to think about is if there are better, less lossy, structures that bring about the simultaneous negative permittivity and permeability. One may think that maybe instead of attacking surface problems such as loss, the root of the problem should be addressed; which is the components of the structures themselves that are causing the loss. What if there are better suited structures that completely differ from SRR's and fishnets that provide the characteristic negative refractive index which exhibit considerably less loss? This may sound questionable but it may be true that the ideal design for a metamaterial structure is yet to be engineered.

However, proposals have already been made towards replacing the typical gold or silver parts that comprise negative index structures with heavily doped semiconductors or by various alloys. Such replacements seem to show some loss decrease but need to be examined further. Another proposal considers an all-dielectric metamaterial that also showed better loss performance.[45]

Metamaterials display tremendous potential with a promise to advance numerous technologies beyond our imagination. Fascination of metamaterials within the scientific community will continue at its fast pace until these technologies are actualized. The remarkable growth of attention this field has received and continues to receive ensures a very bright future for metamaterials.

Acknowledgment

The authors acknowledge financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

1. Liu, Y. and Zhang, X. (2011). Metamaterials: A new frontier of science and technology, *Chemical Society Reviews*. **40**(5), 2494–2507.
2. Kinsler, P. and McCall, M. W. (2008). Criteria for negative refraction in active and passive media, *Microwave and Optical Technology Letters*. **50**, 1804–1807.

3. Ramakrishna, S. and Grzegorczyk, T. (2009). *Physics and Applications of Negative Refractive Index Materials*. (SPIE Press and CRC Press, Bellingham, Washington, USA).
4. Pendry, J. (2000). Negative refraction makes a perfect lens, *Phys. Rev. Lett.* **85**(18), 3966.
5. Leonhardt, U. (2006). Optical conformal mapping, *Science*. **312**(5781), 1777–1780.
6. Schurig, D., Mock, J., Justice, B., Cummer, S., Pendry, J., Starr, A. and Smith, D. (2006). Metamaterial electromagnetic cloak at microwave frequencies, *Science*. **314**(5801), 977–980.
7. Tsakmakidis, K., Boardman, A. and Hess, O. (2007). Trapped rainbow storage of light in metamaterials, *Nature*. **450**(7168), 397–401.
8. O'Brien, S., McPeake, D., Ramakrishna, S. and Pendry, J. (2004). Near-infrared photonic band gaps and nonlinear effects in negative magnetic metamaterials, *Phys. Rev.* **B**. **69**(24), 241101.
9. Shalaev, V. (2007). Optical negative-index metamaterials, *Nature Photonics*. **1**, 41–48.
10. Hafner, C., Xudong, C. and Vahldieck, R. (2006). Resolution of negative-index slabs, *J. Opt. Soc. Am. A*. **23**(7), 1768–1778.
11. Stockman, M. (2007). Criterion for negative refraction with low optical losses from a fundamental principle of causality, *Phys. Rev. Lett.* **98**, 177404.
12. Webb, K. and Thylén, L. (2008). Perfect-lens-material condition from adjacent absorptive and gain resonances, *Optics Letters*. **33**(7), 747–749.
13. Bratkovsky, A., Ponizovskaya, E., Wang, S., Holmstrom, Y. F. P., Thylen, L. and Agren, H. (2008). A metal-wire/quantum-dot composite metamaterial with negative ε and compensated optical loss, *Appl. Phys. Let.* **93**, 193106.
14. Wegener, M., GarcÝa-Pomar, J., Soukoulis, C., Meinzer, N., Ruther, M. and Linden, S. (2008). Toy model for plasmonic metamaterial resonances coupled to two-level system gain, *Optics Express*. **16**(24), 19785–19798.
15. Fang, A., Koschny, T., Wegener, M. and Soukoulis, C. (2009). Self-consistent calculation of metamaterials with gain, *Phys. Rev.* **B**. **79**(24), 241104.
16. Yuan, Y., Popa, B. and Cummer, S. (2009). Zero loss magnetic metamaterials using powered active unit cells, *Optics Express*. **17**(18), 16135–16143.
17. Sivan, Y., Xiao, S., Chettiar, U., Kildishev, A. and Shalaev, V. (2009). Frequency-domain simulations of a negative-index material with embedded gain, *Optics Express*. **17**(26), 24060–24074.
18. Veselago, V. (1968). The electrodynamics of substances, *Sov. Phys. Usp.* **10**(4), 509–514.
19. Pendry, J., Holden, A., Robbins, D. and Stewart, W. (1999). Magnetism from conductors and enhanced nonlinear phenomena, *IEEE Trans. Microw. Theory Tech.* **47**(11), 2075–2084.
20. Zaoui, W., Chen, K., Vogel, W. and Berroth, M. (2012). Low loss broadband polarization independent fishnet negative index metamaterial at 40GHz, *Photonics and Nanostructures: Fundamentals and Applications*. **10**(3), 245–250.

21. Smith, D., Padilla, W., Vier, D., Nemat-Nasser, S. and Schultz, S. (2000). Composite medium with simultaneously negative permeability and permittivity, *Phys. Rev. Lett.* **84**(18), 4184.
22. Dutta, N., Mirza, I., Shi, S. and Prather, D. (2010). Fabrication of large area fishnet optical metamaterial structures operational at near-IR wavelengths, *Materials.* **3**(12), 5283–5292.
23. Kafesaki, M., Tsiapa, I., Katsarakis, N., Koschny, T., Soukoulis, C. and Economou, E. (2007). Left-handed metamaterials: The fishnet structure and its variations, *Phys. Rev.* **B**. **75**(23), 235114.
24. Pendry, J., Schurig, D. and Smith, D. (2006). Controlling electromagnetic fields, *Science.* **312**(5781), 1780–1782.
25. Khurgin, J. and Sun, G. (2010). In search of the elusive lossless metal, *Appl. Phys. Let.* **96**(18), 181102.
26. Khurgin, J. and Sun, G. (2011). Scaling of losses with size and wavelength in nanoplasmonics and metamaterials, *Appl. Phys. Let.* **99**(21), 211106.
27. Khurgin, J. (2015). How to deal with the loss in plasmonics and metamaterials, *Nature Nanotechnology.* **10**(1), 2–6.
28. Fang, A., Koschny, T. and Soukoulis, C. (2010). Lasing in metamaterial nanostructures, *Journal of Optics.* **12**(2), 024013.
29. Khurgin, J. and Sun, G. (2012). Practicality of compensating the loss in the plasmonic waveguides using semiconductor gain medium, *Appl. Phys. Let.* **100**(1), 011105.
30. Dolling, G., Wegener, M., Soukoulis, C. and Linden, S. (2007). Design-related losses of double-fishnet negative-index photonic metamaterials, *Optics Express.* **15**(18), 11536–11541.
31. Dolling, G., Enkrich, C., Wegener, M., Soukoulis, C. and Linden, S. (2006). Low-loss negative-index metamaterial at telecommunication wavelengths, *Optics Letters.* **31**(12), 1800–1802.
32. Güney, D., Koschny, T. and Soukoulis, C. (2009). Reducing ohmic losses in metamaterials by geometric tailoring, *Phys. Rev.* **B**. **80**(12), 125129.
33. Zhou, J., Koschny, T. and Soukoulis, C. (2008). An efficient way to reduce losses of left-handed metamaterials, *Optics Express.* **16**(15), 11147–11152.
34. Popov, A. and Shalaev, V. (2006). Compensating losses in negative-index metamaterials by optical parametric amplification, *Optics Letters.* **31**(14), 2169–2171.
35. Xiao, S., Drachev, V., Kildishev, A., Ni, X., Chettiar, U., Yuan, H. and Shalaev, V. (2010). Loss-free and active optical negative-index metamaterials, *Nature.* **466**(7307), 735–738.
36. Plum, E., Fedotov, V., Kuo, P., Tsai, D. and Zheludev, N. (2009). Towards the lasing spacer controlling metamaterial optical response with semiconductor quantum dots, *Optics Express.* **17**, 8548.
37. Tanaka, K., Plum, E., Ou, J., Uchino, T. and Zheludev, N. (2010). Multifold enhancement of quantum dot luminescence in plasmonic metamaterials, *Phys. Rev. Lett.* **105**, 227403.

38. Meinzer, N., Ruther, M., Linden, S., Soukoulis, C. M., Khitrova, G., Hendrickson, J., Olitzky, J., Gibbs, H. and Wegener, M. (2010). Arrays of ag split-ring resonators coupled to ingaas single-quantum-well gain, *Optics Express.* **18**, 24140.
39. Meinzer, N., Konig, M., Ruther, M., Linden, S., Khitrova, G., Gibbs, H., Busch, K. and Wegener, M. (2011). Distance-dependence of the coupling between split-ring resonators and single-quantum-well gain, *Appl. Phys. Let.* **99**, 111104.
40. Wuestner, S., Pusch, A., Tsakmakidis, K., Hamm, J. and Hess, O. (2010). Overcoming losses with gain in a negative refractive index metamaterial, *Phys. Rev. Lett.* **105**(12), 127401.
41. Huang, Z., Droulias, S., Koschny, T. and Soukoulis, C. M. (2014). Mechanism of the metallic metamaterials coupled to the gain material, *Optics Express.* **22**, 28596.
42. Pusch, A., Wuestner, S., Hamm, J., Tsakmakidis, K. and Hess, O. (2012). Coherent amplification and noise in gain-enhanced nanoplasmonic metamaterials: A Maxwell-Bloch Langevin approach, *ACS NANO.* **6**, 2420–2431.
43. Wuestner, S., Pusch, A., Tsakmakidis, K., Hamm, J. and Hess, O. (2011). Gain and plasmon dynamics in active negative-index metamaterials, *Phil. Trans. R. Soc. A.* **369**, 3523–3550.
44. Popa, B.-I. and Cummer, S. (2007). An architecture for active metamaterial particles and experimental validation at rf, *Microwave and Optical Technology Letters.* **49**(10), 2574–2577.
45. Wuestner, S. and Hess, O. (2014). Active optical metamaterials, *PROGRESS IN OPTICS, VOL 59.* **59**, 1–88.
46. Stockman, M. (2008). Spasers explained, *Nature Photonics.* **2**(6), 327–329.
47. Zheludev, N., Prosvirnin, S., Papasimakis, N. and Fedotov, V. (2008). Lasing spaser, *Nature Photonics.* **2**(6), 351–354.
48. Fedotov, V., Rose, M., Prosvirnin, S., Papasimakis, N. and Zheludev, N. (2007). Sharp trapped-mode resonances in planar metamaterials with a broken structural symmetry, *Phys. Rev. Lett.* **99**(14), 147401.
49. Chen, H., Miao, R. and Li, M. (2010). Transformation optics that mimics the system outside a schwarzschild black hole, *Optics Express.* **18**(14), 15183–15188.
50. Hao, J., Zhou, L. and Qiu, M. (2011). Nearly total absorption of light and heat generation by plasmonic metamaterials, *Phys. Rev.* **B**. **83**(16), 165107.
51. Landy, N., Sajuyigbe, S., Mock, J., Smith, D. and Padilla, W. (2008). Perfect metamaterial absorber, *Phys. Rev. Lett.* **100**(20), 207402.

Index

World Scientific Handbook of

Metamaterials and Plasmonics

Volume 4

Recent Progress in the Field of Nanoplasmonics

World Scientific Series in Nanoscience and Nanotechnology*

ISSN: 2301-301X

The Series aims to cover the new and evolving fields that cover nanoscience and nanotechnology. Each volume will cover completely a subfield, which will span materials, applications, and devices.

Published

Vol. 16 *World Scientific Handbook of Metamaterials and Plasmonics (In 4 Volumes)*
Volume 1: Electromagnetic Metamaterials
Volume 2: Elastic, Acoustic, and Seismic Metamaterials
Volume 3: Active Nanoplasmonics and Metamaterials
Volume 4: Recent Progress in the Field of Nanoplasmonics
edited by Stefan A Maier (Imperial College London, UK)

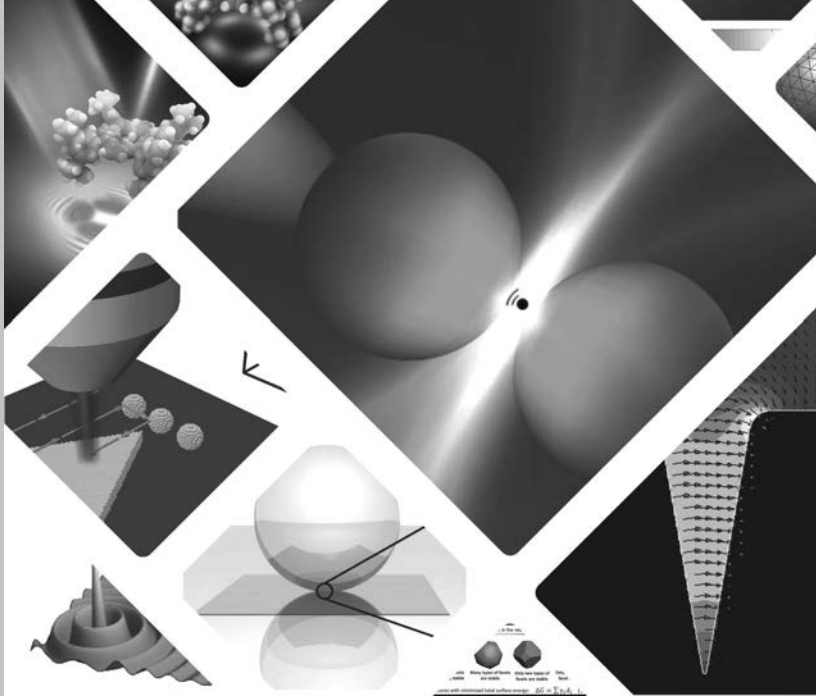
Vol. 15 *Molecular Electronics: An Introduction to Theory and Experiment Second Edition*
by Juan Carlos Cuevas (Universidad Autónoma de Madrid, Spain) and Elke Scheer (Universität Konstanz, Germany)

Vol. 14 *Synthesis and Applications of Optically Active Nanomaterials*
by Feng Bai (Henan University, China) and Hongyou Fan (Sandia National Laboratories, USA)

For further details, please visit: http://www.worldscientific.com/series/wssnn

(Continued at the end of the book)

Volume
16
World Scientific Series in
Nanoscience and Nanotechnology



World Scientific Handbook of

Metamaterials and Plasmonics

Volume 4
Recent Progress in the Field of Nanoplasmonics

Volume Editor
J. Aizpurua
Spanish Council for Scientific Research (CSIC)

Editor
Stefan A. Maier
Imperial College London, UK

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Cataloging-in-Publication Data
Names: Maier, Stefan A., editor.
Title: World Scientific handbook of metamaterials and plasmonics / edited by:
Stefan Maier (Imperial College London, UK).
Description: Singapore ; Hackensack, NJ : World Scientific, [2017] |
Series: World Scientific series in nanoscience and nanotechnology, ISSN 2301-301X ;
volume 16 | Includes bibliographical references and index.
Identifiers: LCCN 2017029450 | ISBN 9789813227613 (set) | ISBN 9789813227620
(v. 1 ; hardcover ; alk. paper) | ISBN 9813227621 (v. 1 ; hardcover ; alk. paper) |
ISBN 9789813227637 (v. 2 ; hardcover ; alk. paper) | ISBN 981322763X
(v. 2 ; hardcover ; alk. paper) | ISBN 9789813227644 (v. 3 ; hardcover ; alk. paper) |
ISBN 9813227648 (v. 3 ; hardcover ; alk. paper) | ISBN 9789813227651
(v. 4 ; hardcover ; alk. paper) | ISBN 9813227656 (v. 4 ; hardcover ; alk. paper)
Subjects: LCSH: Metamaterials. | Plasmons (Physics)
Classification: LCC TK7871.15.M48 W67 2017 | DDC 621.3028/4--dc23
LC record available at https://lccn.loc.gov/2017029450

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Where applicable, figures in this publication will appear in color for the online version.

Desk Editor: Rhaimie Wahap

Typeset by Stallion Press
Email: enquiries@stallionpress.com

Printed in Singapore

Preface by Main Editor

It is our pleasure to present you this *Handbook of Metamaterials and Plasmonics*, charting the tremendous progress that has occurred in this exciting area of research over the last years. What continues to fascinate me about the field above all is its interdisciplinary broadness — we have arrived at a stage where metamaterials make an impact on many arrays of science where control over waves is a prominent ingredient — be they electromagnetic, acoustic, elastic, or even seismic! In these four volumes, we hence attempt to set out the richness of the field, taking metamaterials in the widest sense as artificial media with sub-wavelength structure for control over wave propagation.

Volume 1 focuses on the fundamentals of electromagnetic metamaterials in all their richness, including metasurfaces and hyperbolic metamaterials. Volume 2 widens the picture to include elastic, acoustic, and seismic systems, whereas Volume 3 presents nonlinear and active photonic metamaterials. Finally, Volume 4 includes recent progress in the field of nanoplasmonics, used extensively for the tailoring of the unit cell response of photonic metamaterials.

In its totality, we hope that this handbook will be useful for a wide spectrum of readers, from students to active researchers in

industry, as well as teachers of advanced courses on wave propagation. I want to thank the volume editors Ekaterina Shamonina, Richard Craster, Sébastien Guenneau, Ortwin Hess and Javier Aizpurua, and all the authors for their excellent and sustained work to put these four volumes together.

Stefan A. Maier
Imperial College London, UK
April 2017

Preface by Volume Editor

The revolution of nanotechnology has impacted a broad range of disciplines across science and technology. The fabrication and manipulation of nanostructures has reached levels of control never foreseen, and a variety of spectroscopies and microscopies are now able to routinely probe different physical and chemical properties of matter at the nanoscale. The interaction of light and matter has not escaped off this progress: sophisticated routes of chemical synthesis and a variety of lithographic techniques have managed to construct semiconductor and metallic nanostructures which act as effective blocks that bring light down to the nanoscale, achieving the longstanding dream of beating the diffraction limit of light, pushing optics beyond conventions.

Light can exchange energy and momentum very effectively with electron-hole pair excitations in semiconductor quantum dots, however, it is its interaction with the collective behavior of the conduction electrons in metals what provides one of the best opportunities to establish a direct connection between light and the nanoworld. The oscillation of the electron gas confined in finite metallic nanostructures provides a resonant and effective manner of localizing and enhancing light down to the nanoscale, showing a nanoantenna effect that can be used to allow a variety of physical processes such as enhancing molecular spectroscopy signals, controlling

nanoscale optoelectronics, driving optical forces, producing an enhanced nonlinear response, or interfacing the exchange of quantum states of light, among others.

In this volume, a review of many of these aspects is covered by some of the best experts worldwide on each topic. An introductory chapter by Luis Martín Moreno and Francisco J. García Vidal guides through the basics of nanoplasmonics, where the common model of the dielectric function for metals is presented, together with its application to derive the dispersion relationship in plasmonic waveguides and in localized surface plasmons. This chapter provides a clear picture of the basic concepts that will serve to understand many complex aspects of the optical response in nanostructures which involve plasmonic excitations, thus introducing basic plasmonic nanostructures as canonical building blocks of nanooptics.

In Chapter 2 of the volume, Jean-Jacques Greffet provides very useful insights about the role of plasmonic nanostructures as optical nanoantennas, developing a battery of concepts that connect nanooptics, electrical engineering and quantum optics. Concepts such as impedance of an antenna, gain, and efficiency are introduced in great detail and adapted to the situation of plasmonic antennas, with expressions that allow to describe important processes of light and matter interaction such as the spontaneous emission of an emitter, the modification of its decay rate by the presence of a local density of photonic states, or molecular spectroscopies that can be strongly modified and enhanced by the presence of plasmonic nanoantennas, as in Surface-Enhanced Raman scattering, and Surface-Enhanced fluorescence.

Chapter 3 completes the local description of light and nanostructures within the linear dielectric response theory, typically used to solve Maxwell's equations in inhomogeneous media, by considering the importance of nonlocality in the description of the optical response of plasmonic nanostructures. The dynamical screening of the electron gas can be effectively introduced with the use of a hydrodynamical model which considers an extra pressure term in the dynamics of the electron gas. This model can be extended

with the consideration of a drift-dissipation term, which effectively produces a damping of plasmonic modes. The implementation of hydrodynamical model in powerful computation tools based on Finite Element Methods to solve Maxwell's Equations is an important highlight introduced in this chapter by N. Asger Mortensen, Jacob Khurgin and Martijn Wubs, which allows to adopt the hydrodynamical model in the calculation of the optical response of arbitrarily shaped plasmonic nanostructures.

In Chapter 4, Andrey Borissov and myself introduce the importance of the quantum nature of the electron gas in the optical response of plasmonic nanoantennas. By adopting a jellium model of the electron gas, the optical response of metals in confined nanostructures is developed within the framework of the Time-Dependent Density Functional Theory (TDDFT). The quantum nature of the electron gas is considered in this way, with the excitations in the metal subjected to the boundaries of the different surfaces defining the nanostructure. Particular attention is paid to the location of the centroid of charge of the surface plasmon excitation as a key to understand the spectral evolution of the quantum size effect, as well as the tendencies found for Ångstrom-size plasmon rulers in plasmonic dimers at subnanometric distances. This chapter also addresses tunneling across the gap at optical frequencies, and how this effect modifies the optical response in subnanometric gap-plasmonics.

Coming back to classical descriptions, during the last years a powerful methodology to address the optical response in plasmonic antennas has been developed in the context of transformation optics. Antonio Fernández-Domínguez and Sir John B. Pendry, together with Yu Luo and Rongkuo Zhao introduce the concepts of transformation optics in Chapter 5, and describe their application in the conformal mapping of 2-dimensional plasmonic structures that present singular structures, such as edges and corners, as well as blunt morphologies and hybridized structures. A fascinating aspect of transformation optics is that it allows to obtain analytical expressions for the optical response and local fields in a very elegant way for many plasmonic structures that show singular shapes, and which could

not be tackled otherwise. 3-dimensional structures are also addressed in this chapter with the study of equally challenging morphologies. Important corrections related to radiative effects and nonlocality are also considered in this review, fully completing the picture of transformation optics in nanoplasmonics. Finally, an application to the calculation of Van der Waals forces and comparisons with experiments are nicely illustrated.

Chapters 6 and 7 deal with a modern and vibrant field within nanophotonics which is at the heart of the essence of nanoplasmonics: the ultrafast nature of plasmonic excitations. Otto Muskens, Peter Wiecha and Arnaud Arbouet describe in great detail in chapter 6 the time scales and the processes involved in the dynamics of metallic nanostructures, as well as the optical nonlinearities generated in them. Nonlinear processes in plasmonic nanoparticles such as Second-harmonic generation, third-harmonic generation, or two-photon luminescence are described in the chapter, and both experimental evidence and theoretical calculations of the processes are shown. Electron gain spectroscopy is also addressed in the context of nonlinear response, and finally, the access to the nonlinear response of plasmonic structures via femtosecond time-resolved spectroscopy is introduced in single particles paying special attention to their vibrational modes of excitation. Finally, the applications of plasmonic structures in nonlinear devices ends this chapter.

Chapter 7 retakes the ultrafast response of plasmonic structures by Jue-Min Yi, Petra Groß, and Christoph Lienau, focusing in the access to spectral information of plasmonic nanostructures from time-domain spectroscopy techniques. The spectral response of single particles is introduced by Fourier transform spectroscopy, together with aspects of plasmon propagation in time domain, and the nonlinear response of particles is shown to be accessed via interferometric frequency-resolved autocorrelation techniques. The application of time-resolved techniques to nanoplasmonics allows to study the effect of photoemission in nanotips, and thus to control local fields at the nanoscale. This effect is also covered in this chapter, which ends with a topic of current interest connected with the strong coupling of

emitters and plasmonic antennas: time-domain techniques, as pump-probe spectroscopy allow to access Rabi oscillations, typical of the strong interaction regime, in real time, showing a very nice opportunity to address interactions between quantum units of potential application in quantum nanooptics.

As an alternative to optical excitations in metallic nanostructures, graphene has emerged as a novel material that shows a collective response of the electron gas in the mid-infrared and TeraHertz range of the spectrum. In Chapter 8, Alexey Nikitin introduces the basic concepts of graphene plasmons, and derives its dispersion relation, in connection with the properties of its conductivity. This serves to introduce different aspects of the propagation of these special plasmons. After analysing the different approaches to describe the conductivity in the graphene layer, a chart of methods to optically excite plasmons in graphene is reviewed, giving account of the typical mechanisms of plasmon launching by optical means. Among others, optical excitation of point emitters, of graphene medium interfacing with other media, or the use of resonators on top of the graphene sheet, are among some of the options pointed out for graphene plasmon excitations. In this context, the existence of a modulation in the conductivity, as a grating for optical excitation, is described in great detail. This chapter serves, among other purposes, to reveal the rich possibilities of alternative materials to handle plasmonic excitations.

Another fascinating option to excite surface plasmons in nanostructures is provided by an anternative excitation probe: fast electrons as those commonly used in Scanning Transmission Electron Microscopy. Ulrich Hohenester describes in great detail in Chapter 9, the basic principles of interaction between fast electrons and plasmonic nanoparticles within a classical electrodynamical description of the electromagnetic interactions involved. A variety of methods to tackle the induced fields responsible for the energy losses in Electron Energy Loss Spectroscopy (EELS) is also described in detail, and the usefulness of the quasi-static approach to deal with a modal expansion of the solution to the interaction is brought up. A set of key nanostructures where surface plasmons can be excited and mapped,

including metallic disks, coupled particles, and particles deposited on substrates is also reviewed in the context of EELS. Some effects like the interplay with photons, the possibilities of electron tomography, and quantum effects are quoted in this context. As observed in this chapter, electrons have become a usual tool to explore plasmonic excitation in nanostructures and reveal many properties which are elusive to optics.

Although the localization and enhancement of light is a leitmotiv in plasmonics, plasmons show another important aspect related to their decay in the form of heat. Guillaume Baffou and Romain Quidant elaborate in Chapter 10 a review of thermoplasmonics where they introduce the basic concepts of heat generation, and temperature increase in these nanostructures, focusing on arrays of particles and typical experimental situations of pulsed illumination. After discussing the thermal microscopy techniques used to map temperature in these nanoenvironments, a variety of applications of thermoplasmonics are reviewed ranging from heat-assisted magnetic recording to catalysis, through thermophotovoltaics or chemical vapor deposition. The possibilities in nanomedicine also deserve special attention in this chapter.

Even if the excitation of surface plasmons is often revealed in far-field optical spectroscopy techniques as resonances in the response, the most powerful aspect of plasmonics is the actual nanoscale distribution of the local fields induced around plasmonic nanostructures. The access to this information can be achieved mainly by the use of near-field optical probes. In Chapter 11, Edward Yoxall and Rainer Hillenbrand introduce the principles of Scattering-type Scanning Near-field Optical Microscopy (s-SNOM) that allow to access the local field of plasmonic excitations in nanostructures. Once the technique is introduced, interesting near field patterns in dipolar antennas, focusing antennas, and propagating plasmons are revealed and described, showing the evanescent nature of the plasmonic fields evolving and being distributed in the nanoscale.

It has been mentioned in the beginning of this preface that the control of the systematic fabrication of metallic nanoparticles showing very sophisticated morphologies and compositions on demand has been one of the key drivers of the field of plasmonics. It is thus important to have access to the state of the art in chemical synthesis of plasmonic nanostructures, as Dorleta Jiménez de Aberasturi, Cyrille Hamon and Luis Liz-Marzán report in the last chapter of this volume. In spite of the long tradition of more than a century in this field of science, wet chemistry synthesis of inorganic nanoparticles is now more than ever a vibrant field where many different parameters and effects need to be controlled. The team of Luis Liz-Marzán stresses the role of the seed crystallinity and the effect of adsorbates in the synthesis process, and how its control allows to generate a vast chart of plasmonic nanoparticles. Isotropic particles, nanorods, core-shell particles and other novel possibilities are reviewed and itemized in this chapter providing a clear perspective of the current possibilities of this field of chemistry.

All together, this volume becomes an excellent reference for a reader to approach different aspect of nanoplasmonics which are connected with hot topics currently being developed in the field. Plasmonics is a very broad discipline within nanooptics, and of course any compilation unavoidably leaves relevant things for future completion. It might occur that not everything which is important in plasmonics is covered in this volume, but everything in this volume is very important to plasmonics. I am devoted and tremendously thankful to all the collaborators of the different chapters. All of them have shown their best skills and attitude to build this compilation work. We all hope that every reader can enjoy this volume, making nanoplasmonics feel closer.

Javier Aizpurua
Donostia-San Sebastián
April 2017

This page intentionally left blank

Contents

CHAPTER 1

Basics of Nanoplasmonics

LUIS MARTÍN-MORENO

CSIC-Universidad de Zaragoza, Spain
lmm@unizar.es

FRANCISCO J. GARCÍA-VIDAL

Universidad Autónoma de Madrid, Spain
fj.garcia@uam.es

This chapter is intended to serve both as a brief introduction and a reference frame for the rest of the chapters of this book. First, we analyze the optical response of metals, describing why noble metals are best suited for plasmonics in the optical regime. We also show the theoretical foundations of the surface plasmons supported by two-dimensional (2D) metal surfaces, with special emphasis on analyzing their dispersion relations, confinement and propagation lengths. A historical note on the discovery of these surface electromagnetic (EM) modes in the fields of Optics and Condensed Matter Physics is also presented. Finally, the surface plasmon modes supported by metal particles (0D structures) and 1D metal waveguides and their feasible applications are briefly introduced.

1.1. General Introduction

Nanoplasmonics encompasses the electromagnetic (EM) phenomena that either occurs in the nanoscale proximity to, or are due to the

existence of nanometric features in, metal–dielectric interfaces. These interfaces support EM modes known as surface plasmon polaritons (SPPs), which result from a combination of EM fields with the collective excitation of electrons. SPPs possess several appealing characteristics: ability to concentrate light beyond the diffraction limit, strong modification of the local density of photonic states, high sensitivity to the dielectric environment and ultra-fast response. Furthermore, to a large extent, these properties can be tailored by geometry. These properties are attractive for applications in physics, medicine, security and environmental monitoring.

This chapter presents a succinct description of the basic properties of SPPs that, hopefully, will lay the foundations needed to follow the different aspects of nanoplasmonics covered by this book. The interested reader can find a more extended introduction to nanoplasmonics in the text by Greffet,[1] and several more in the extended monographs on the topic.[2–7] A more general overview on the field of nanophotonics (i.e. not only covering plasmonic systems) can be found in the book by Novotny and Hetch.[8]

In order to study the EM properties of metal structures, we need to know the constitutive relation between the displacement vector $\mathbf{D}$ and the electric field $\mathbf{E}$ inside the metal. Assuming that the external electric fields are much weaker than the interatomic field, this relation is linear. Assuming also that the metal is both homogeneous and isotropic, the constitutive relation can be expressed as a convolution integral:

$$\mathbf{D}(\mathbf{r},t) = \int \mathrm{d}\mathbf{r}' \int \mathrm{d}t' \epsilon(\mathbf{r}-\mathbf{r}', t-t')\, \mathbf{E}(\mathbf{r}', t') \tag{1.1}$$

leading to

$$\mathbf{D}(\mathbf{k},\omega) = \epsilon(\mathbf{k},\omega)\, \mathbf{E}(\mathbf{k},\omega) \tag{1.2}$$

for the Fourier components of the fields, defined as $\mathbf{F}(\mathbf{r},t) = \int \mathrm{d}\mathbf{r}\mathrm{d}t \mathbf{F}(\mathbf{k},\omega) e^{i(\mathbf{k}\mathbf{r}-\omega t)}$. Maxwell's equation $\nabla \cdot \mathbf{D} = 0$ leads to $\epsilon(\mathbf{k},\omega)\, \mathbf{k} \cdot \mathbf{E}(\mathbf{k},\omega) = 0$. Thus, the electric field is transversal

($\mathbf{E}(\mathbf{k},\omega) \perp \mathbf{k}$) whenever $\epsilon(\mathbf{k},\omega) \neq 0$. But, when $\epsilon(\mathbf{k},\omega) = 0$, there are longitudinal solutions with $\mathbf{k} \cdot \mathbf{E}(\mathbf{k},\omega) \neq 0$. These solutions are known as *bulk plasmons*.

It is usually an excellent approximation to ignore the spatial dispersion of the dielectric function and work with $\epsilon(\mathbf{r},\omega)$. The validity of this local description arises from the large difference between the time variation of optical fields and speed of electrons in a metal, as the Fermi velocity is ∼100–1000 times smaller than the speed of light. Thus, the local approximation is valid when the relevant length scales involved are ∼100–1000 times smaller than the free space wavelength. As a rule of thumb, non-local effects are important when the length scales are smaller than 2 nm. Alternatively, non-local effects must be considered when the relevant wavevectors in the problem are ∼100–1000 times larger than the free space wave vector. In what follows in this chapter we will always use the local approximation. The analysis of the linear response of metals beyond the local approximation will be analyzed in Chapter 3, while nonlinear effects will be covered in Chapter 6.

In principle, the dielectric function in metals can be computed with the general expressions for linear response functions provided by the Kubo formalism. As this involves the evaluation of expectation values of operators in the ground state of an *interacting electronic system*, in practice this is a very complex many-body problem that cannot be solved without approximations. The crudest one is assuming that the electrons are independent. The next level of approximation is treating the electron–electron interaction within the mean field, leading to what is called the Random-Phase Approximation (RPA).

Within RPA, and ignoring the existence of a crystalline structure, the dielectric function of a metal can be accurately described by the classical Drude formula:

$$\epsilon_{\mathrm{RPA}}(\omega) = 1 - \frac{\omega_p^2}{\omega\,(\omega + \imath\Gamma)}, \tag{1.3}$$

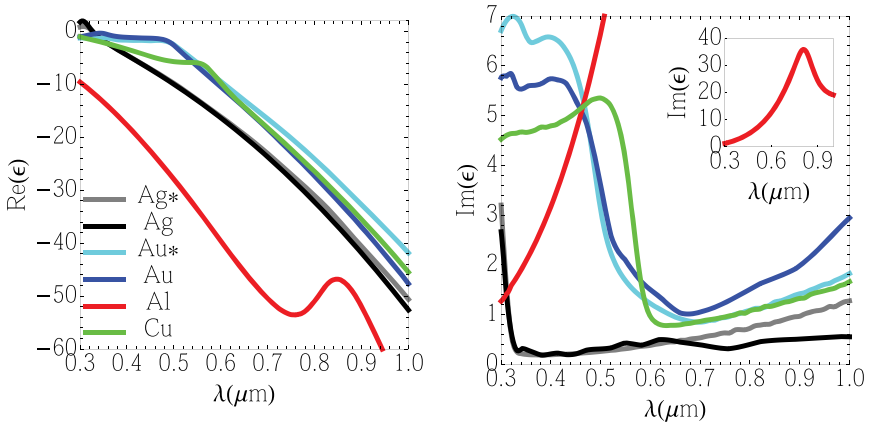


Fig. 1.1. Measured real (left panel) and imaginary (right panel) parts of the dielectric constant of the metals most commonly used in nanoplasmonics: Ag, Au, Al and Cu. The inset in the right panel shows Imϵ for Al. All data are taken from Ref. 9, except the ones marked by *, which are from Ref. 10.

where $\omega_p = \sqrt{4\pi ne^2/m}$ is the plasma frequency, which only depends on the density of free carriers, n and on the electron charge (e) and mass (m), and Γ is the electronic transport mean free rate. Typical values for ω_p in metals are in the 6–15 eV range. The existence of a crystalline structure (and thus a periodic potential seen by the electrons) modifies the electronic band structure in the metal, allowing for interband transitions that contribute to the dielectric function. In practice, the dielectric function of metals is obtained phenomenologically from the fitting of experimental data (like reflectance and absorbance of incoming radiation).

Figure 1.1 renders the tabulated dielectric constant spectra of several metals that, as will be discussed later on, are the most promising for nanoplasmonics. In the case of two best candidates, gold and silver, we include the data from both Johnson and Christie[10] (a classical reference) and those compiled by Norris *et al.*,[9] who recently brought to attention a set of processing rules that lead to high-quality metal surfaces. For later comparison, Fig. 1.2 shows the data for other metals, which are not commonly used in plasmonics.

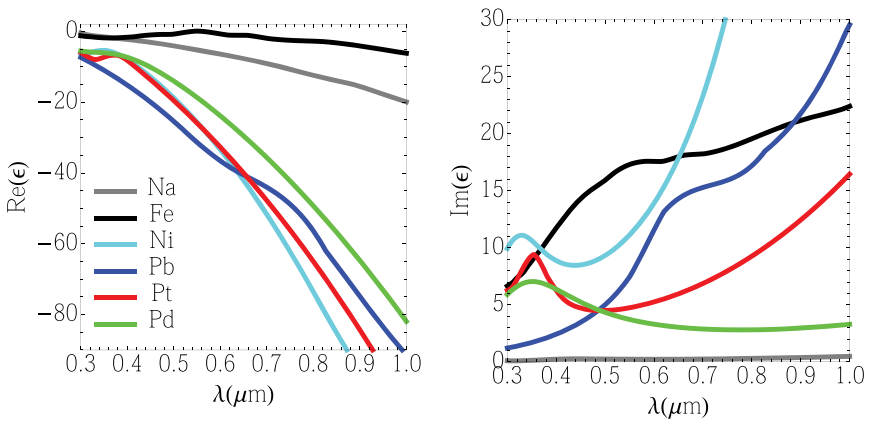


Fig. 1.2. Measured real (left panel) and imaginary (right panel) parts of the dielectric constant of some metals which are not commonly used in nanoplasmonics. The data for Na is taken from Ref. 11, Ref. 12 provided the data for Fe and Ni, and the data for Pb, Pt and Pd are from Ref. 13.

1.2. Surface Plasmon Polaritons

1.2.1. *Characteristics of surface plasmon polaritons*

As mentioned above, SPPs are surface EM modes that appear at metal–dielectric interfaces. The simplest geometry that can be analyzed is a planar interface separating two semi-infinite materials: at $z > 0$, the material is characterized by a dielectric constant ϵ_1, while the material at $z < 0$ is characterized by ϵ_2.

Suppose that we look for a solution to Maxwell's equations that is bounded to the interface. As each semi-infinite medium is translationally invariant, the solutions must propagate as plane waves, with $E_{xi} = Ee^{\imath kx}e^{\pm \imath k_{zi} z}$, where E is the amplitude of the electric field, $i = 1, 2$ labels the media, x is the in-plane direction of propagation, k is the in-plane wavevector (which is the same for both E_{x1} and E_{x2}, in order to satisfy the matching of the parallel E-field component along the interface) and $\pm$ refers to 1 and 2, respectively. The corresponding z-components are $E_{zi} = \mp(q/q_{zi})E_{xi}$, where the dimensionless $\mathbf{q}$ is the modal wavevector in units of the free-space wavenumber ($\mathbf{q} \equiv \mathbf{k}/k_0$, with $k_0 = \omega/c$), so that $q_{zi} = \sqrt{\epsilon_i - q^2}$. An

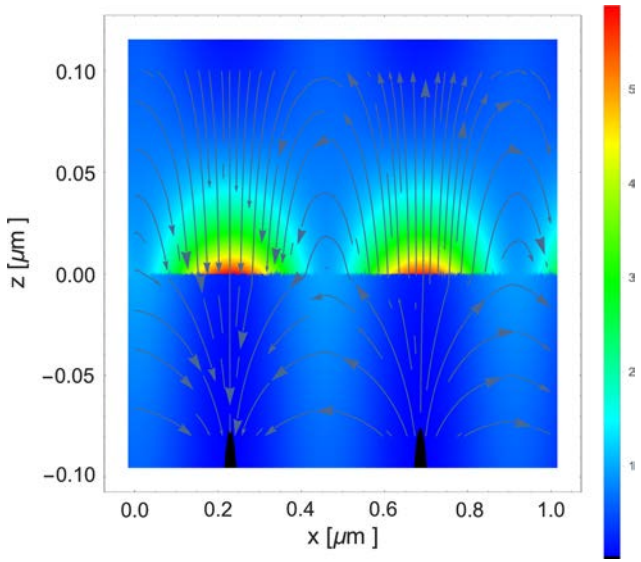


Fig. 1.3. Snapshot of the electric field of an SPP in a silver–air interface. The chosen wavelength is $\lambda = 430\,\text{nm}$ and the metal is characterized by $\epsilon_M = -6 + 0.4\imath$ (as corresponding to Ag[9]). The amplitude of the in-plane component of the electric field at $z = 0$ is taken as 1, and the color code renders the intensity of the electric field $|E(x, z)|^2$.

important technical point is that the sign of the square root must be taken such that $\text{Im}q_z > 0$, as the field must vanish at infinity.

Continuity of the normal component of the displacement vector, $\epsilon_i E_{zi}$, readily leads to the condition for the existence of an EM mode bounded to the interface:

$$\epsilon_1 q_{z2} = -\epsilon_2 q_{z1}. \tag{1.4}$$

For lossless materials ($\text{Im}\epsilon_1 = \text{Im}\epsilon_2 = 0$), this equation does not have solutions when $\epsilon_1\epsilon_2 > 0$. However, when the dielectric constants in both media have different signs, there is an EM mode bounded to the surface with wavevector

$$q_p = \sqrt{\frac{\epsilon_1\epsilon_2}{\epsilon_1 + \epsilon_2}}. \tag{1.5}$$

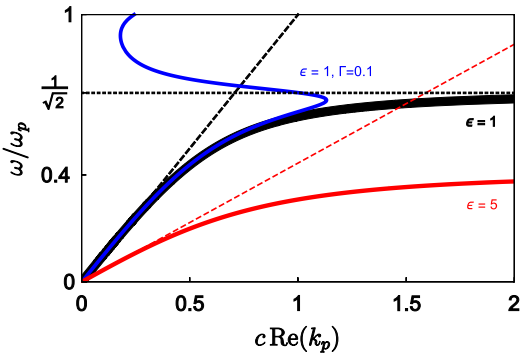


Fig. 1.4. Dispersion relation of SPPs for a metal characterized by a Drude dielectric constant. The black and red curves are for the case of a lossless metal, while the blue curve has been computed for the case $\Gamma = 0.1\omega_p$. The red curve is when the dielectric is characterized by a dielectric constant $\epsilon = 5$, while in the other two cases the dielectric is vacuum. The dashed lines represent the corresponding light cones, while the dotted line shows the asymptote of the dispersion relation when $\epsilon = 1, \Gamma = 0$.

Figure 1.3 renders a representation of the electric field associated with an SPP. It is apparent that the normal component of the electric field (i) is discontinuous across the surface and (ii) changes sign along it. This is related to the existence of a surface polarization wave. On the other hand, Figure 1.4 shows the dispersion relation $\omega(k_p)$ of these SPP modes for the case of a metal and a dielectric, characterized by dielectric functions that are Drude-like for the metal and a frequency-independent ϵ for the dielectric. This figure illustrates several characteristics of the dispersion of SPPs:

- The dispersion relation of SPPs (as for any bounded mode) lies to the right of the *light cone* $\omega = ck/\sqrt{\epsilon}$. In fact, the distance of the SPP dispersion relation to the light cone is directly related to the SPP confinement to the surface: the larger the distance, the more confined the SPP mode becomes.
- The light cone is the asymptote of the SPP dispersion relation at small frequencies ($\omega \ll \omega_p$).
- In a lossless system, no EM bounded modes exist when $\epsilon_M > -\epsilon$, i.e. when $\omega > \omega_{\mathrm{SP}} \equiv \omega_p/\sqrt{1+\epsilon}$. The frequency ω_{SP} is the

surface plasmon frequency, which would be obtained if retardation effects were neglected, i.e., in a model where the electrons interact via the instantaneous Coulomb potential. Therefore, SPPs can be understood as resulting from the coupling of dispersionless electron density surface modes with light waves.

- Without absorption, SPPs would propagate indefinitely. Absorption causes the modes to acquire an $\mathrm{Im}k_p \neq 0$ and thus propagate a typical distance $L_{\mathrm{SPP}} = 1/\mathrm{Im}k_p$. Additionally, absorption induces a back-bending in the dispersion relation close to the ω_{SP} asymptote.
- SPPs exist for a wide frequency range, which implies that they can provide an inherent fast response. This topic will be covered in Chapter 7.

It must be noted that Fig. 1.4 and the previous analysis have been done considering that ω is real, so the solution to Eq. (1.4) is a complex k_p. The dispersion relation for real k_p and complex ω is very similar at small frequencies, but it does not present back-bending.[1] Different experiments may explore different parts of the complex-ω versus complex-k_p dispersion relations, but usually the general picture shown in Fig. 1.4 is a good starting point as, in real metals, SPPs close to the ω_{SP} are strongly damped anyway.

The dispersion relation of SPPs in real metals is represented in Figs. 1.5 and 1.6 in a slightly different way. The left panels show the real part of the dimensionless SPP wavevector as a function of free-space wavelength. The advantage of this representation is that the in-plane plasmon period (λ_{spp}) is readily obtained as $\lambda_{\mathrm{spp}} = \lambda/\mathrm{Re}(q_p)$. The right panels show the spectra of the SPP *figure of merit* $\mathrm{FOM} \equiv \mathrm{Re}(q_p)/\mathrm{Im}(q_p)$. FOM is 2π times the number SPP of periods that propagate before its field amplitude decays by a factor $1/e$.

These figures show that SPPs propagate in the visible regime longer in Ag, Au and Cu than in other metals, with propagation lengths of the order of several microns, while Al is the most promising material for plasmonics in the UV. Of course, the characteristics of the dielectric constant of the metal is not the only parameter

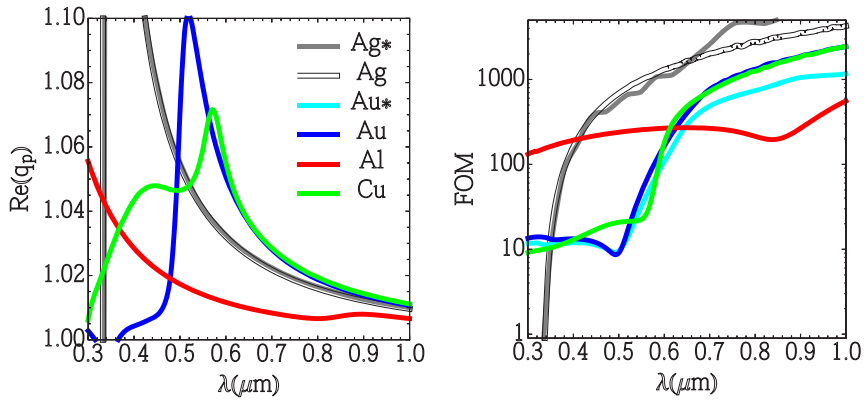


Fig. 1.5. Left panel: Dimensionless in-plane plasmon wavevector $q_p = k_p/(\omega/c)$ for the metals most commonly used in plasmonics. Right panel: corresponding figure of merit FOM $\equiv$ Req_p/Imq_p. dielectric constant of the metals taken from the same references as Fig. 1.1.

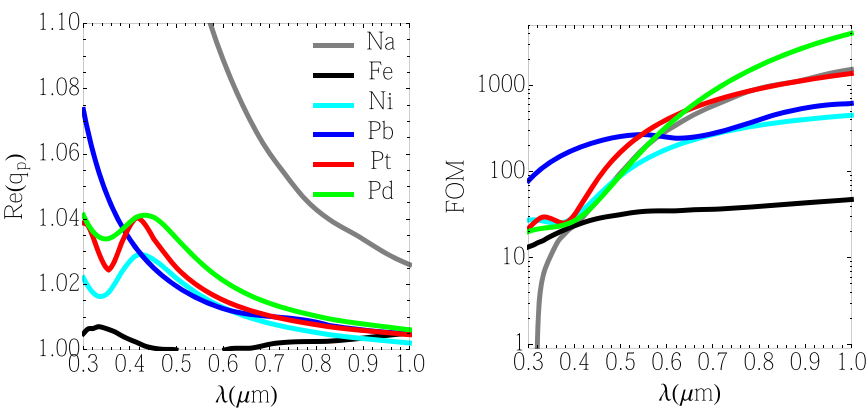


Fig. 1.6. Left panel: Dimensionless in-plane plasmon wavevector $q_p = k_p/(\omega/c)$ for metals not commonly used in plasmonics. Right panel: Corresponding figure of merit. Dielectric constant of the metals taken from the same references as Fig. 1.2.

that determines the performance in plasmonics; the possibility to prepare ultra-smooth surfaces and to keep them uncontaminated is of paramount importance. In this regard, Au is usually preferred to both Ag and Cu, which oxidize fast and require surface protection for applications.

In order to complete the description of the SPP wavefield, Fig. 1.7 shows the confinement length of an SPP. This length is defined as the distance to the interface at which the field amplitude has decayed by a factor $1/e$. In the optical regime, the confinement length inside the metal (commonly referred to skin depth, $\delta = L_z^{\text{metal}}$) is of the order of a few tens of nm and the corresponding confinement in air is smaller than the wavelength.

These properties can be understood by using Eq. (1.5) assuming both $-\epsilon_m \gg \epsilon$ and $|\text{Re}(\epsilon_m)| \gg \text{Im}(\epsilon_m)$, which are good approximations for a metal–air interface and in the optical regime, respectively (see Fig. 1.1). Then, the dimensionless SPP wavevectors in the normal direction in the dielectric (q_{zd}) and in the metal (q_{zm}) are $q_{zd} \approx \imath\epsilon/\sqrt{\epsilon_m}$ and $q_{zm} \approx \imath\sqrt{\epsilon_m}$. Therefore, these vectors satisfy $\text{Im}(q_{zd})\,\text{Im}(q_{zm}) \approx \epsilon$, which is an illustrative relation showing that (i) in adimensional units, the confinement in the metal is inversely proportional to the confinement in the dielectric, and (ii) this confinement increases with the dielectric constant of the dielectric. In real space, these expressions lead to

$$L_z^{\text{metal}} L_z^{\text{dielectric}} \approx \frac{\epsilon}{4\pi^2}\lambda^2. \tag{1.6}$$

In the case of a Drude metal in the optical regime, $\sqrt{-\epsilon_m} \approx \omega_p/\omega \propto \lambda$, so the skin depth does not depend on frequency, while the confinement length in air scales with λ^2. The metals represented in Fig. 1.7 are not exactly Drude metals, but they follow these trends.

Obviously, the considered system of two semi-infinite media is an idealization which is never present in nature. However, it is useful when the metal thickness, W, is much larger than the penetration depth of the field inside the metal, δ, in which case the two interfaces support different SPPs (note that δ is in principle frequency-dependent, so the following argument may apply differently in different frequency regimes). If $W \sim \delta$, the SPP field in one interface will feel the boundary condition imposed by the opposite interface, causing the hybridization of the SPPs of the isolated interfaces, in a process very much like the formation of molecular orbitals from atomic ones. One of the resulting hybridized modes has

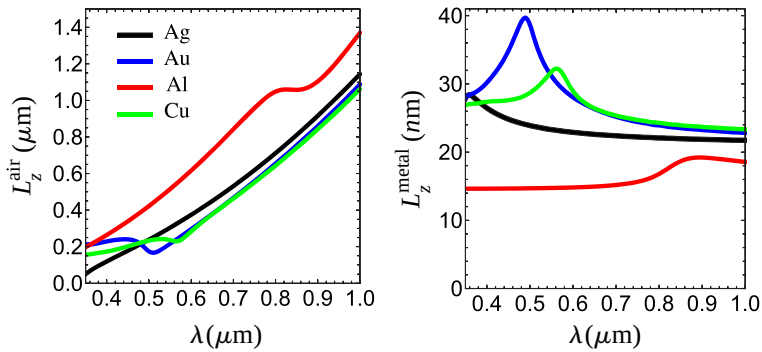


Fig. 1.7. Confinement length of an SPP in the direction perpendicular to the air–metal interface. The left panel shows the decay in air, while the right panel shows the decay in the metal. All dielectric constants have been taken from Ref. 9.

a reduced electric field in the metal (in the case of equal substrate and superstrate, this mode would be the antisymmetric combination of the SPPs of the individual interfaces). Correspondingly, compared to the independent SPPs, this hybridized mode has a larger extension into the dielectric, a dispersion relation closer to the light-cone and a longer propagation length, hence its name long-range SPP (LRSPP). The other hybridized mode, or short-range SPP (SRSPP) has, again compared to the independent SPPs, a larger proportion of the electric field inside the metal, a reduced proportion in the dielectric and a shorter propagation length. Also, it has a smaller in-plane wavelength and thus a dispersion relation that deviates more from the light cone. All these characteristics are progressively more apparent as the film thickness decreases. In fact, nowadays, it is possible to manipulate SRSPPs in the extreme limit of atomic-thin materials. To date, the best-studied case is graphene, which will be covered in Chapter 8. The interested reader may find recent introductions to the plasmonic properties of other families of 2D materials in Refs. 14,15.

1.2.2. *Historical note*

The prediction that the metal–air interface supports a bound EM mode at a frequency $\omega_{SP} = \omega_p\sqrt{2}$ was first made by Ritchie in 1957,[16]

when analyzing the energy loss spectra of fast electrons in thin metal films.

The experimental data showed that electrons were not only exciting the predicted bulk plasmons at frequencies $\approx\omega_p$, but also some other modes at lower frequencies. Ritchie proposed that these modes were surface excitations. Definite experimental confirmation was not straightforward, due to the strong influence of surface preparation and possible contamination on the electron energy loss spectra (EELs), but it appeared 2 years later, when EELS experiments in Al films of high surface and volume purity were conducted.[17] The extreme sensitivity of SPPs to changes in surface environment was highlighted by Stern and Ferrell,[18] who found that even a 2 nm thick oxide layer could appreciably modify the surface plasmon frequency. In fact, this latter work was the one that introduced the term *surface plasmon* which, as the name suggests, was defined as the *quanta* of surface plasma oscillations. This extreme sensitivity of SPPs to the environment is the basis of sensing.

All these theoretical works considered a non-retarded interaction between electrons, and electron losses were ignored, so they could only provide the high frequency "horizontal" asymptote of the full SPP dispersion relation (see Fig. 1.4). Interestingly, the full electrodynamical calculation leading to the condition for the existence of a bound surface mode (Eq. (1.4)) had already been published by Zenneck,[19] 50 years before. Here, the context was the amazing results on long-distance wireless radio propagation that had been reported by Marconi. Zenneck found that the interface between air and a lossy dielectric supports a non-radiative surface EM mode, which could carry radio signals. Sommerfeld analyzed how these modes could be excited with an oscillating electrical dipole (which is why some authors denote these surface modes as Zenneck waves, while others use the name Zenneck–Sommerfeld waves). After the realization that the main mechanism for radio propagation is reflection by the ionosphere (itself a plasma at radio frequencies!), Zenneck waves disappeared from the forefront of research.

We stress that previous paragraphs do not intend in any way to diminish the accomplishments of the scientists previously named. But it is interesting to note that neither Zenneck nor Sommerfeld seem to have considered whether their modes would exist in a metal at optical frequencies, of if they did, to the best of our knowledge, they did not report it. Maybe they did consider the possibility, but the sub-mm distances travelled by SPPs did not appeal to them at the time. The scientific community had to wait for the insight of Ritchie to discover the existence of surface plasma oscillations supported by metal interfaces, using an approach more familiar in Condensed Matter Physics than in Classical Optics.

1.2.3. *Excitation of SPPs*

The SPPs of a flat interface cannot be directly excited by incoming radiation: the dispersion relation of incident radiation (which lies within the light-cone) and the SPP one does not have any common point (see Fig. 1.4), so parallel momentum and energy could not be simultaneously conserved in the process. There are two main strategies to excite SPPs:

- *Breaking the conservation of parallel momentum*: This can be achieved by eliminating the translational symmetry of either the structure or the source. Instances of the first case are the presence of surface roughness and gratings. In fact, uncontrolled coupling to SPPs were a highly detrimental feature of metal surfaces in optics until the influence of surface roughness could be tamed. Gratings provide a good way to excite SSPs with a particular wavevector with the flexibility in their design even allowing for unidirectional launching.[20] Alternatively, the structure can be translationally symmetric, but the source may not be so, for instance by being localized in a region of space Δx (for simplicity, we focus on one spacial dimension in our analysis, but the extension to a source with 2D or 3D confinement is straightforward). In this case, by Fourier transformation, the source can be visualized as

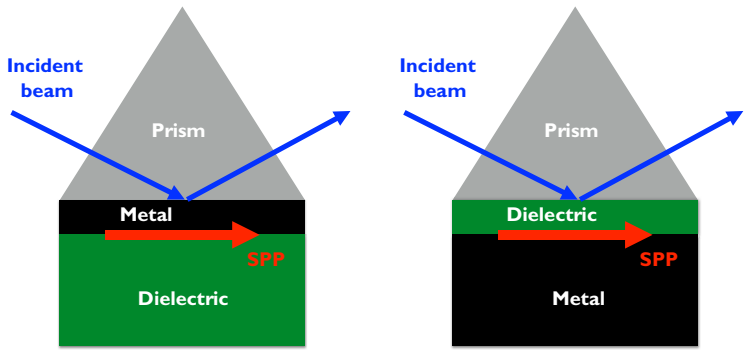


Fig. 1.8. Geometries for evanescent coupling of SPPs. Left panel: Kretschmann configuration. Right panel: Otto configuration.

extended in k-space with components within $\Delta k \sim 1/\Delta x$ that can be outside the light-cone. These components are evanescent from the source, but, if the source is placed close enough of the metal-dielectric interface, they can excite SPPs. Another possibility is using the electric fields confined to electron beams, as was done in the original experiments leading to the discovery of SPPs. Recent developments of EELS in connection with the study of SPPs are covered in this book in Chapter 9.

- *Evanescent coupling*: This technique excites the SPP of a given metal–dielectric interface with light incoming from another dielectric with a higher dielectric constant. To illustrate this concept, consider Fig. 1.4. The spectral region between the red and black discontinuous curves corresponds to waves that propagate in the medium with higher ϵ (as this region is inside the light-cone of that medium), but are evanescent in air. Those propagating waves can excite the SPPs of the metal–air interface, but not those of the metal–high dielectric one. The dielectric should be placed at a distance of the order of the SPP confinement length in air, as otherwise, the coupling to SPPs would be negligible (for much larger distances), or the characteristics of the metal–air SPP would be strongly modified (at much smaller ones). In practice, the high dielectric medium is a prism, illuminated normally at one of the

lateral surfaces, so that there is no diffraction at the air–prism interface, see Fig. 1.8. The coupling to SPPS is then either across the metal film (Kretschmann configuration) or by the evanescent field present in conditions of total internal reflection (Otto configuration).

1.3. 0D Surface Plasmons (Particle Plasmons) and 1D Surface Plasmons (Plasmonic Waveguides)

The existence of surface plasma oscillations is not restricted to the geometry of the plane. On the contrary, virtually any metal–dielectric interface will support them in the optical regime. In fact, SPPs in particles have been used for centuries due to its ability to color surfaces without degradation with time. Renowned examples are the Lycurgus cup (4th century, whose glass contains nanoparticles of gold and silver which give it a different color when looked in reflection than in transmission), the stained glass that can be found in medieval cathedrals from the 10th century, and the lustre technique used in ceramics to provide iridescence (developed also around the 4th century).

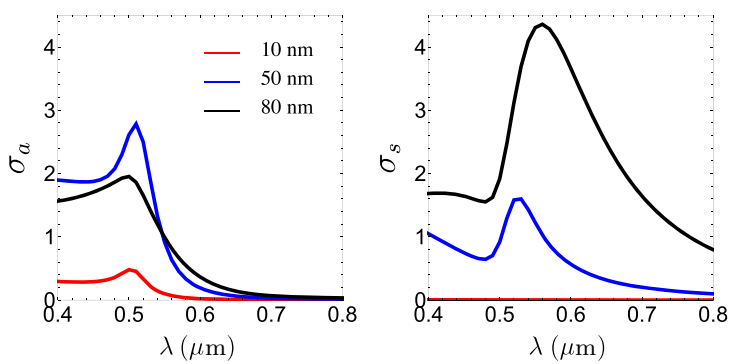
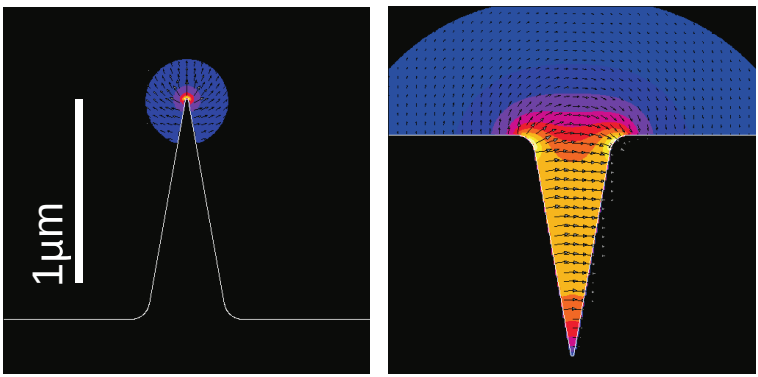


Fig. 1.9. Cross-section spectra (normalized to the geometrical cross-section) of spherical Ag nanoparticles for three different radii: 10, 50 and 80 nm. Left panel: absorption cross-section. Right panel: scattering cross-section. The dielectric constant of Ag is taken from the experimental data in Ref. 9.

As a difference with their 2D counterpart, SPPs in particles can be directly excited by incident radiation. To illustrate this, we consider the scattering of light by a spherical homogeneous particle. This problem was solved analytically by Mie already in 1908[21] and which, despite its apparent simplicity, presents a rich physical behavior.[22] Figure 1.9 summarizes how a spherical silver nanoparticle interacts with light, showing both absorption and scattering cross-sections, normalized to the geometric cross-section, for nanoparticles of three different sizes. The scattering by a spherical particle presents resonances at frequencies that are largely independent on particle size. Depending on the particle size, the particle cross-section is mainly due to absorption (smaller particles) or to scattering (larger ones), the cross-over being at a radius $\sim$50 nm.

Particles with other geometrical shapes also present resonances at frequencies that depend on geometry. Correspondingly, the EM field close to the particle is strongly enhanced. A great deal of work has been devoted to understanding, controlling and optimizing the spectral position and field distribution of SPPs in particles. These efforts are summarized in several chapters in this book. Chapter 12 is devoted to the actual fabrication of the nanoparticles, whereas Chapter 5 studies the field enhancement in a variety of geometries (both in a single particle and in the gaps created between two nearby particles) within the quasistatic approximation. In fact, the present capability to create interparticle gaps in the nm and sub-nm scale requires understanding how both the electronic wavefunction spill-out out of the geometric boundary described by the atoms and electron quantum tunneling across the gap influence the plasmonic modes of the coupled particles. These quantum effects are addressed in Chapter 4. Generally speaking, the coupling between radiation and particle plasmons is of great relevance to many scientific areas, as these plasmons concentrate light in the nanoscale, thus acting as optical antennas, an aspect that is covered in Chapter 2.

Once particle plasmons are excited, they act as localized sources of light, which can help us beat the diffraction limit and be able to "see" objects with nanoscale resolution with optical fields (subject



(a) (b)

Fig. 1.10. Transverse electric field of (a) wedge plasmon and (b) channel plasmon, in gold, defined with the same geometrical parameters by interchanging the metallic and vacuum regions. The panel lateral size is 2 μm and the wavelength is $\lambda = 1.5\,\mu$m. This figure is an adaptation of Ref. 26, courtesy of Esteban Moreno.

covered by Chapter 11). Additionally, the absorption of energy by the plasmonic modes can heat up the particle, which thus becomes a localized source of heat, with promising applications for heat-assisted magnetic recording and, in medicine, in cancer therapy and the controlled release of drugs (see Chapter 10).

Let us finish this section by stating that surface plasmons also exist in one-dimensional (1D) systems, which can be used then as plasmonic waveguides. These waveguides are lossy and present the typical trade-off between confinement and losses: the higher the confinement, which is usually a desired quality, the higher the losses, which are usually to be avoided. However, although this trade-off exists for all geometries, the actual figures may greatly depend on geometry. This is exemplarily represented in Fig. 1.10, which shows the modal field of both a gold wedge and a channel in gold, defined by the same geometrical parameters.

Plasmonic waveguides are also interesting in combination with both classical dipoles and quantum two-level systems, as the small modal volume leads to an enhanced coupling between the emitters and the waveguide, which opens the way to "blind" optical circuits.[23–25]

References

1. Greffet, J.-J. (2012). *Introduction to Surface Plasmon Theory*, In S. Enoch and N. Bonod eds., *Plasmonics: From Basics to Advanced Topics*, pp. 105–148. Springer Berlin Heidelberg, Berlin, Heidelberg.
2. Agranovich, V. M. and Mills, D. L. (1982). *Surface Polaritons*. (North Holland, Amsterdam).
3. Raether, H. (1988). *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*. (Springer-Verlag, Berlin).
4. Zayats, A. V., Smolyaninov, I. I. and Maradudin, A. A. (2005). Nano-optics of surface plasmon polaritons, *Phys. Rep.* **408**, pp. 131–314.
5. Maier, S. A. (2007). *Plasmonics: Fundamentals and Applications*. Springer, Berlin.
6. Pitarke, J. M., Silkin, V. M., Chulkov, E. V. and Echenique, P. M. (2007). Theory of surface plasmons and surface-plasmon polaritons, *Rep. Prog. in Phys.* **70**(1), p. 1.
7. Stockman, M. I. (2011). Nanoplasmonics: Past, present, and glimpse into future, *Opt. Exp.* **19**(22), pp. 22029–22106.
8. Novotny, L. and Hecht, B. (2006). *Principles of Nano-Optics*. Cambridge University Press, Cambridge.
9. McPeak, K. M., Jayanti, S. V., Kress, S. J. P., Meyer, S., Iotti, S., Rossinelli, A. and Norris, D. J. (2015). Plasmonic films can easily be better: Rules and recipes, *ACS Photon.* **2**(3), pp. 326–333.
10. Johnson, P. B. and Christy, R. W. (1972). Optical constants of the noble metals, *Phys. Rev. B.* **6**(12), pp. 4370–4379. doi:10.1103/PhysRevB.6.4370.
11. Inagaki, T., Emerson, L. C., Arakawa, E. T. and Williams, M. W. (1976). Optical properties of solid Na and Li between 0.6 and 3.8 ev, *Phys. Rev. B.* **13**, pp. 2305–2313.
12. Johnson, P. B. and Christy, R. W. (1974). Optical constants of transition metals: Ti, V, Cr, Mn, Fe, Co, Ni, and Pd, *Phys. Rev. B.* **9**, pp. 5056–5070.
13. Werner, W. S. M., Glantschnig, K. and Ambrosch-Draxl, C. (2009). Optical constants and inelastic electron-scattering data for 17 elemental metals, *J. Phys. Chem. Ref. Data.* **38**(4), pp. 1013–1092.
14. Basov, D. N., Fogler, M. M. and García de Abajo, F. J. (2016). Polaritons in van der waals materials, *Science* **354** (6309), doi:10.1126/science.aag1992.
15. Low, T., Chaves, A., Caldwell, J. D., Kumar, A., Fang, N. X., Avouris, P., Heinz, T. F., Guinea, F., Martin-Moreno, L. and Koppens, F. (2017). Polaritons in layered two-dimensional materials, *Nat Mater.* **16**(2), pp. 182–194.
16. Ritchie, R. H. (1957). Plasma losses by fast electrons in thin films, *Phys. Rev.* **106**, pp. 874–881.
17. Powell, C. J. and Swan, J. B. (1959). Origin of the characteristic electron energy losses in aluminum, *Phys. Rev.* **115**, pp. 869–875.
18. Stern, E. A. and Ferrell, R. A. (1960). Surface plasma oscillations of a degenerate electron gas, *Phys. Rev.* **120**, pp. 130–136.

19. Zenneck, J. (1907). Über die fortpflanzung ebener elektromagnetischer wellen längs einer ebenen leiterfläche und ihre beziehung zur drahtlosen telegraphie, *Annalen der Physik.* **328**(10), pp. 846–866. ISSN 1521-3889.
20. López-Tejeira, F., Rodrigo, S. G., Martín-Moreno, L., García-Vidal, F. J., Devaux, E., Dintinger, J., Ebbesen, T. W., Krenn, J. R., Radko, I. P., Bozhevolnyi, S. I., González, M. U., Weeber, J. C. and Dereux, A. (2008). *New J. Phys.* **10**, p. 033035.
21. Mie, G. (1908). Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen, *Annalen der Physik.* **330**, pp. 377–445.
22. Bohren, C. F. and Huffman, D. R. (2007). *Absorption and Scattering of Light by Small Particles.* (Wiley-VCH Verlag GmbH).
23. Falk, A. L., Koppens, F. H. L., Yu, C. L., Kang, K., de Leon Snapp, N., Akimov, A. V., Jo, M.-H., Lukin, M. D. and Park, H. (2009). Near-field electrical detection of optical plasmons and single-plasmon sources, *Nat. Phys.*
24. Gonzalez-Tudela, A., Martin-Cano, D., Moreno, E., Martin-Moreno, L., Tejedor, C. and Garcia-Vidal, F. J. (2011). Entanglement of two qubits mediated by one-dimensional plasmonic waveguides, *Phys. Rev. Lett.* **106**, p. 020501.
25. Kress, S. J. P., Antolinez, F. V., Richner, P., Jayanti, S. V., Kim, D. K., Prins, F., Riedinger, A., Fischer, M. P. C., Meyer, S., McPeak, K. M., Poulikakos, D. and Norris, D. J. (2015). Wedge waveguides and resonators for quantum plasmonics, *Nano Lett.* **15**(9), pp. 6267–6275.
26. Moreno, E., Rodrigo, S. G., Bozhevolnyi, S. I., Martín-Moreno, L. and García-Vidal, F. J. (2008). Guiding and focusing of electromagnetic fields with wedge plasmon polaritons, *Phys. Rev. Lett.* **100**, p. 023901.

CHAPTER 2

Plasmonic Nanoantennas

JEAN-JACQUES GREFFET

Institut d'Optique Graduate School,
Université Paris-Saclay, France
jean-jacques.greffet@institutoptique.fr

Antennas are widely used to couple efficiently electromagnetic waves with subwavelength emitters and receivers. In this chapter we discuss how the concept of antenna can be used in the optics regime to control light-matter interaction processes such as absorption, scattering or spontaneous emission by a quantum system such as an atom, a molecule or a quantum dot. The aim of this basic introduction is to establish connections between the electrical engineer, the optics and the cavity quantum electrodynamics point of view.

2.1. Introduction

Antennas are widely used in radiowaves to efficiently couple waves propagating in open space to electrical lines which are subsequently coupled to detectors or emitters. The wires are subwavelength in the radiowave regime so that antennas confine electromagnetic waves well below the diffraction limit, something that cannot be achieved with lenses and mirrors. In order to efficiently transfer electromagnetic energy between the source/detector and the vacuum, it is required to avoid reflection in the lines. In this context, a key concept is the line impedance which has to be matched with the emitter/receiver

impedance on one hand, with the antenna impedance on the other hand. When comparing radiowaves antennas with optical antennas, some differences appear. The subject of optical antennas was first envisioned in the context of near-field imaging.[1–3] Early discussion of modification of decay rate and emission directivity can be found in Refs. 4–6 and experimental demonstrations were reported in Refs. 7–10. Excellent reviews can be found in Refs. 11–13. In this chapter we focus on the interaction between a nanoantenna and a single emitter such as a molecule, an atom or a quantum dot.[5] The first section reviews briefly the key figures of merit of radiowaves antennas. This will be used as a guide to identify the relevant concepts and functions of an antenna. The second section introduces the corresponding key features in the electromagnetic optics framework. We then specifically study how elementary light-matter interaction processes at the single quantum emitter level are modified by the presence of an antenna. We first compare a nanoantenna with a microcavity and show that the concepts introduced in cavity quantum electrodynamics can be applied to nanoantennas. Finally, we analyse different light-matter interaction processes such as spontaneous emission, Raman Scattering and Fluorescence both in the linear regime and in the saturation regime. Surface Enhanced Infra Red Absorption (SEIRA), non-local effects and quantum effects will not be touched upon here as they will be the subject of specific chapters in this book.

2.2. Antennas: Basic Concepts

2.2.1. *Radiation resistance, efficiency, superradiance*

Let us first consider an antenna used to emit radiowaves. The antenna is fed by an electrical source through a line as shown in Fig. 2.1(a). From the electrical circuit point of view, the antenna is a dipole connected to the line. It is characterized by an impedance Z_{ant} whose real part is the antenna resistance R.

The power dissipated by the antenna is either transformed in heat in the antenna wires or radiated. Two resistances are introduced to

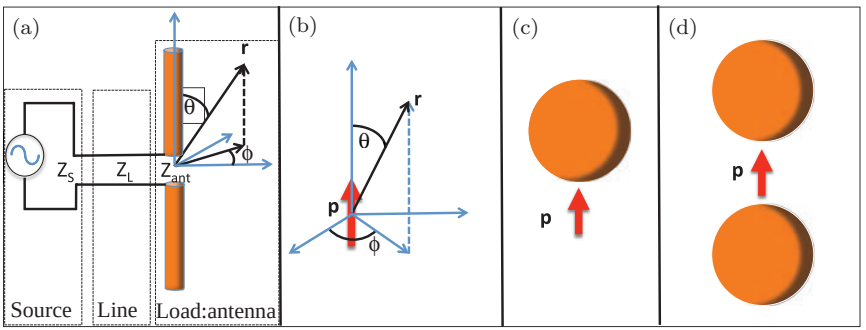


Fig. 2.1. Basic antenna concepts. (a) An electrical source is connected to an antenna through a line. (b) Definition of the angles for a dipolar emitter with dipole moment **p**. (c) Metallic nanosphere antenna, (d) dimer of dipolar antennas.

account for these two processes:

$$R = R_R + R_{NR}, \tag{2.1}$$

where R_R is the radiation resistance (accounting for power radiated only) and R_{NR} accounts for the non-radiative losses (i.e. Joule losses in a metallic antenna). The fraction of dissipated power which is converted into radiation is given by the antenna efficiency

$$\eta_{ant} = \frac{R_R}{R_R + R_{NR}}. \tag{2.2}$$

In order to optimize the emitted power, it is important to avoid reflections between the source and the line and between the line and the antenna. In other words, the line needs to be impedance matched with both the source and the antenna.[14] Here, we want to pinpoint the very basic physical phenomenon responsible for the power radiated by an antenna. To start this qualitative discussion, we consider a short antenna wire with length $L \ll \lambda$ where a current with intensity I is flowing at frequency $\omega = ck$. The emitted power is given by[14]

$$P_R = \frac{\eta_0}{12\pi}(kL)^2 I^2, \tag{2.3}$$

where $\eta_0 = \mu_0 c = 377\Omega$ is the vacuum impedance and $k = 2\pi/\lambda$ is the wavevector modulus. It is clearly seen that the field increases linearly with L so that if the length is multiplied by N while satisfying

the condition $NL \ll \lambda$, the radiated power is multiplied by N^2. An antenna with length NL can be viewed as a collection of N antennas of length L driven by the same current. This simple effect is a direct consequence of *constructive interferences* of emitted electromagnetic fields. It is called collective radiation or *superradiance*[15] in the context of light emission by an ensemble of N atoms in the low excitation regime.[a] This result follows directly from the fact that the field varies linearly with the antenna length L. Assuming that all the sources are in phase, we simply add the fields produced resulting in a factor N for the field amplitude and a factor N^2 for the radiated power.

This simple remark can also be used to analyze the power emitted by a metallic nanosphere. It is proportional to the square of its dipole moment and therefore to the square of its volume or in other words, to the square of the number of electrons. By realizing that a particle with a radius of 50 nm contains approximately 10^7 free electrons, it is clear that such a particle is a much better emitter than a single atom or molecule. This remark explains why the metallic nanoantennas must be chosen to be on the order of a fraction of a tenth of a wavelength to be efficient. If the radius is smaller, the radiated power decays faster than non-radiative losses so that the efficiency drops. Let us make an order of magnitude estimate. The typical decay time due to losses (i.e. generation of a hot electron) is on the order of 5 fs in a metal whereas a typical atomic decay time by spontaneous emission is on the order of 10 ns. If the radius is 5 nm, the number of electrons is on the order of 10^4 so that the radiation decay time becomes 1 ps. Instead, for a radius of 50 nm, we have $N \approx 10^7$ so that the radiative decay time is on the order of 1 fs and becomes faster than non-radiative losses. If the radius is larger than $\lambda/10$, the path differences cannot be neglected anymore. In other words, the fields

[a]This classical superradiance effect corresponds to the spontaneous emission of one excited atom among $N-1$ non-excited atoms. The emitted power scales as N^2, N times more than an ensemble of N incoherent emitters. The spontaneous emission rate is then multiplied by N. This effect should not be confused with the superfluorescence effect observed when the N atoms are all excited. In that case, a burst of light is observed on a time scale also given by τ/N.[15]

emitted by different electrons are no longer all in phase so that the (classical) super radiance effect saturates. In summary, the figure of merit to characterize the amount of radiated power is the radiation resistance. As we have seen in the two examples, the key parameter is the size of the antenna or in more physical terms, the number of electrons radiating coherently.

2.2.2. *Gain*

Another key feature of an antenna is the control of the angular radiation pattern. To characterize the ability of an emitting antenna to radiate light in direction (θ, ϕ) (see Fig. 2.1(a)), the gain $G(\theta, \phi)$ is defined as the flux of the Poynting vector $\mathbf{S}(\mathbf{r})$ at an observation point $\mathbf{r}$ divided by $P/4\pi r^2$:

$$G(\theta, \phi) = \frac{\mathbf{S}(\mathbf{r}) \cdot \hat{\mathbf{r}}}{P/4\pi r^2}, \tag{2.4}$$

where P is the total power delivered to the antenna. Hence, the gain compares the antenna with an isotropic antenna that would radiate a power P. A slightly different quantity called directivity is defined when normalizing with the actual *radiated* power $P_R = \eta_{ant} P$. The directivity accounts for the antenna efficiency:

$$D(\theta, \phi) = \frac{G(\theta, \phi)}{\eta_{ant}}, \tag{2.5}$$

so that it verifies:

$$\int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \sin\theta \mathrm{d}\theta D(\theta, \phi) = 4\pi. \tag{2.6}$$

It is worth discussing briefly the conditions required to obtain a directional antenna. Let us consider the radiation by an antenna described by a monochromatic current density $\mathbf{j}(\mathbf{r}, \omega)$ in a finite volume V. We use the following definition of Fourier transform:

$$\mathbf{j}(\mathbf{r}, t) = \int_{-\infty}^{\infty} \mathbf{j}(\mathbf{r}, \omega) \exp(-i\omega t) \frac{\mathrm{d}\omega}{2\pi}. \tag{2.7}$$

The retarded vector potential is given by:

$$\mathbf{A}(\mathbf{r},\omega) = \frac{\mu_0}{4\pi}\int_V \mathbf{j}(\mathbf{r}',\omega)\frac{\exp(i\frac{\omega|\mathbf{r}-\mathbf{r}'|}{c})}{|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}'. \tag{2.8}$$

This form clearly displays the fact that different elements of the source produce spherical waves. The resulting total potential is due to the interferences between these waves. To discuss directivity, we need to move to the far field. In the far-field approximation, valid provided that $r \gg L^2/\lambda$ where L is a characteristic length of the current distribution, the retarded potential can be simplified:

$$\mathbf{A}(\mathbf{r},\omega) = \frac{\mu_0}{4\pi}\frac{\exp(i\frac{\omega}{c}r)}{r}\int_V \mathbf{j}(\mathbf{r}',\omega)\exp\left(-i\frac{\omega}{c}\frac{\mathbf{r}}{r}\cdot\mathbf{r}'\right)d^3\mathbf{r}'. \tag{2.9}$$

In this approximation, the interference of the fields radiated by different points located at different positions $\mathbf{r}'$ appears clearly: the potential vector results from adding the contributions of all the volume elements accounting for a path difference given by $\frac{\mathbf{r}}{r}\cdot\mathbf{r}'$. This equation shows explicitly how the vector potential depends on the emission direction $\mathbf{r}/r$. It is seen that the dependence is given by the Fourier transform of the current density. Hence, it follows immediately from a very general property of the Fourier transform $\Delta x\Delta k_x \geq 2\pi$ that the radiation in the far field can be confined to small solid angles if and only if the antenna is much larger than a wavelength. Assuming an extension along the x-axis of L, we find that the radiated field has a spatial spectrum with width $\Delta k_x \approx 2\pi/L$ corresponding to an angular width λ/L. We recover the well-known result of diffraction theory. This is not unexpected as diffraction theory is based on the Huygens-Fresnel principle which amounts to transforming the diffraction problem into a radiation problem by introducing secondary sources. As this discussion is based on the retarded potential solution of Maxwell equations and the far-field approximation, the conclusion is very general and can be applied to any radiating system. It follows that any subwavelength antenna cannot be highly directional. Note however that some degree of

directivity can be obtained due to either polarization effects (e.g. a $\sin^2\theta$ dependence for dipolar radiation) or to interferences between dipolar and quadrupolar modes.

2.2.3. *Receiving antenna*

When an antenna is used to collect incident power and funnel it into a detector, an interesting quantity is the *effective antenna area* $A_{eff}(\theta, \phi)$. The effective area is defined by the equation:

$$P_{abs} = S_{inc} A_{eff}, \tag{2.10}$$

where P_{abs} is the absorbed power in the load and S_{inc} is the Poynting vector amplitude of an incident plane wave. In this equation, *it is assumed that the antenna is impedance matched with the line and the line is impedance matched with the load so that there are no reflections*. This effective area is close to the real geometrical cross section of the antenna for very large antennas. By contrast, for small resonant antennas, it can be very different. In particular, for a resonant dipolar electromagnetic antenna, it is given by[14]:

$$A_{eff} = \frac{3\lambda^2}{8\pi} \sin^2\theta, \tag{2.11}$$

even if the dipole is much smaller than the wavelength. A very important result is the connection between the receiving effective area and the gain:

$$A_{eff} = \frac{\lambda^2}{4\pi} G(\theta, \phi), \tag{2.12}$$

a general result valid for any antenna as can be shown using the reciprocity theorem.[14] This is a fundamental property of antennas which shows that emitting and receiving properties are connected. It also clearly shows that the directivity of an antenna is given by its effective area in units of λ^2:

$$G(\theta, \phi) = 4\pi \frac{A_{eff}}{\lambda^2}. \tag{2.13}$$

2.2.4. *A simple model of a nanoantenna*

In this section, we introduce an elementary model to illustrate how a nanoantenna can enhance the radiation of a quantum emitter. In our model, the emitter is characterized by an electric dipole moment $\mathbf{p}$. The plasmonic nanoantenna is simply a metallic nanosphere with radius a and dielectric permittivity $\epsilon(\omega)$. We assume that the permittivity is given by a Drude model $\epsilon(\omega) = 1 - \omega_p^2/(\omega^2 + i\Gamma\omega)$. We choose the dipole to be aligned with the axis connecting the dipole and the nanoantenna (see Fig. 2.1(c)). In the near-field regime ($r \ll \lambda/2\pi$), the leading contribution to the field generated by the dipole at the nanoantenna position is:

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\frac{\mathbf{p}}{r^3}. \tag{2.14}$$

This field polarizes the spherical nanoantenna which has a polarisability:

$$\alpha(\omega) = 4\pi a^3 \frac{\epsilon(\omega) - 1}{\epsilon(\omega) + 2}, \tag{2.15}$$

so that the antenna dipole becomes:

$$\mathbf{p}_{ant} = \alpha(\omega)\epsilon_0\mathbf{E} = \left[\frac{a}{r}\right]^3 \frac{\epsilon(\omega) - 1}{\epsilon(\omega) + 2}\mathbf{p}. \tag{2.16}$$

With this simplified model, we see that the dipole moment of the emitter $\mathbf{p}$ is essentially replaced by the dipole moment of the antenna $\mathbf{p}_{ant}$. Hence the radiated power will be much larger. The enhancement factor contains two terms: a purely geometrical factor $(a/r)^3 < 1$ as $r > a$ and a resonant factor $\frac{\epsilon(\omega)-1}{\epsilon(\omega)+2}$. The first term is specifically a near-field effect similar to the lightning rod effect: it is due to the concentration of surface charges on a conductor. In the case of a sphere, it is smaller than 1 so that there is no enhancement. However, for other geometries such as a dimer (see Fig. 2.1(d)), the enhancement can be much larger than 10. The second term is associated to a plasmon resonance at $\omega_p/\sqrt{3} + i\Gamma/2$ corresponding to $\mathrm{Re}(\epsilon) + 2 = 0$. A more detailed analysis can be found in Ref. 25. This discussion has put forward two key

characteristics of a plasmonic nanoantenna: it must be much larger than the emitter in order to radiate more power, it needs to generate a large electric field in the near field (a hot spot) in order to be efficiently coupled to the dipole source. In summary, a nanoantenna is an intermediate resonator which is well coupled to both the emitter and the propagating waves. With this point of view, the concept of impedance matching to a line is lost. We will discuss in Section 2.9 how it is possible to introduce this concept.

2.3. Spontaneous Emission

The goal of this section is to discuss how the presence of a nanoantenna modifies the spontaneous emission of a two level system characterized by an electric dipolar transition. We will compare the classical antenna point of view and the quantum point of view for radiation emission. We will specifically discuss what is quantum and what is classical in the spontaneous emission process.

2.3.1. *Decay rate of a classical oscillator*

We start by establishing the connection between the power radiated and the decay rate of the energy of the oscillator. We model the emitter by a harmonic oscillator with amplitude $x(t) = x_0(t)\cos(\omega_0 t)$, velocity $v(t) = \dot{x}(t)$ and electric dipole moment $-ex(t)$. The oscillator has a resonant frequency ω_0 and a mass m. The instantaneous energy is given by

$$U = \frac{1}{2}mv^2(t) + \frac{1}{2}m\omega_0^2 x^2(t). \tag{2.17}$$

The amplitude $x_0(t)$ is a slowly decaying function of time due to radiative losses. Assuming a slow amplitude decay $\dot{x_0} \ll \omega_0 x_0$, we can approximate the energy by:

$$U = \frac{1}{2}m\omega_0^2 x_0^2(t). \tag{2.18}$$

We can write an energy conservation equation:

$$\frac{\mathrm{d}U}{\mathrm{d}t} = -P_R(t) = -\frac{\omega_0^4 e^2 x_0^2(t)}{12\pi\epsilon_0 c^3} = -\gamma_{cl} U, \tag{2.19}$$

where the classical decay rate γ_{cl} is given by

$$\gamma_{cl} = \frac{e^2\omega_0^2}{6\pi\epsilon_0 c^3 m}. \tag{2.20}$$

As expected, the larger the dipole moment, the larger the radiated power and the larger the decay rate.

2.3.2. *Quantum spontaneous emission decay rate in vacuum*

The full quantum calculation of the decay rate in vacuum gives a very different result[19] with a different frequency dependence:

$$\gamma_0 = \frac{|d_{12}|^2\omega^3}{3\pi\epsilon_0 c^3\hbar}, \tag{2.21}$$

where d_{12} is the matrix element of the dipole operator and $2\pi\hbar$ is Planck's constant. The presence of $\hbar$ stresses the quantum nature of this result. It can be shown using the Wigner-Weisskopf method that the decay is exponential.[17] The decay rate of this exponential can be found using Fermi golden rule[19]:

$$\gamma_R = \frac{2\pi}{\hbar^2}|\hat{W}_{if}|^2 g(\omega), \tag{2.22}$$

where $\hat{W}_{if}$ is the matrix element of the interaction Hamiltonian and $g(\omega)\mathrm{d}\omega$ is the number of electromagnetic states in the frequency interval $\mathrm{d}\omega$ with the electric field parallel to the dipole moment. Let us stress that an electromagnetic state is nothing but a mode of Maxwell equations (e.g. a mode in a cavity or in a waveguide). In vacuum, the states or modes are plane waves characterized by a frequency and a

wavevector. The density of states is given by[b]:

$$g(\omega) = V\frac{\omega^2}{3\pi^2 c^3}. \tag{2.23}$$

Using the quantized electromagnetic field with amplitude $\sqrt{\hbar\omega/2\epsilon_0 V}$, it is found that the interaction Hamiltonian matrix element $\hat{W}_{if} = \langle f|\hat{d}\cdot\hat{E}|i\rangle$ is given by:

$$|\hat{W}_{if}|^2 = \frac{|d_{12}|^2\ \hbar\omega}{2\epsilon_0 V}. \tag{2.24}$$

Upon inserting Eqs. (2.23) and (2.24) into Eq. (2.22), we obtain the radiative decay rate in vacuum (2.21). Although the quantum result is very different from the classical one, it is interesting to note that it is possible to recover the quantum result by simply replacing two terms in the classical calculation by their quantum counterpart. The first term is the classical energy $U = \frac{1}{2}m\omega_0^2 x_0^2$ which can be replaced by $\hbar\omega$, the second term is the classical dipole moment ex_0 which can be replaced by $2d_{12}$.

2.3.3. *Classical radiation in vacuum: Radiation reaction work*

In this section, we establish a connection between the power radiated and the power done by the radiation reaction, namely, the field radiated by a source on itself. We start by considering the power radiated by a classical dipole in vacuum. We consider a dipole source inside a volume V enclosed in a surface A. Energy conservation in

[b]In k-space, the volume corresponding to electromagnetic modes with frequency in the interval $[\omega, \omega + \mathrm{d}\omega]$ is $4\pi k^2 \mathrm{d}k$ where $k = \omega/c$. Using the periodic boundary conditions for a cubic volume with side L, it is shown that the volume associated to a single mode is given by $(2\pi/L)^3 = 8\pi^3/V$. A factor 2 needs to be included to account for the two polarizations for each plane wave. Finally, a factor 1/3 is included to account for the fact that the electric dipole couples only to the z-component of the electric field, parallel to the dipole moment. This is called the projected density of states.

time domain yields:

$$\int_V \frac{\partial u}{\partial t}\,\mathrm{d}^3\mathbf{r} = -\int_V \mathbf{j}(\mathbf{r})\cdot\mathbf{E}(\mathbf{r})\,\mathrm{d}^3\mathbf{r} - \int_A \mathbf{S}\cdot\mathrm{d}\mathbf{A}, \tag{2.25}$$

where u is the electromagnetic energy density, $\mathbf{j}$ is the current density and $\mathbf{E}$ the electric field. We now consider a monochromatic point-like dipole source with dipole moment $\mathbf{p}$ in stationary regime such that $\mathbf{j}(\mathbf{r}) = -i\omega\mathbf{p}\delta(\mathbf{r}-\mathbf{r}')$. We can now compute the time-averaged energy conservation relation using complex amplitudes:

$$\frac{1}{2}\mathrm{Re}\left[i\omega\mathbf{p}\cdot\mathbf{E}^*\right] = \int_A \langle\mathbf{S}\rangle\cdot\mathrm{d}\mathbf{A} = P_R, \tag{2.26}$$

where we have introduced the radiated power P_R. The term on the left involves the electromagnetic field generated by the dipole on itself. This field is known as radiation reaction. The electric field generated by the dipole source can be cast in the form $\mathbf{E}(\mathbf{r}) = \mu_0\omega^2\overset{\leftrightarrow}{G}^{(0)}(\mathbf{r},\mathbf{r}')\mathbf{p}$, where we have introduced the vacuum Green tensor[16,17]:

$$\overset{\leftrightarrow}{G}^{(0)}(\mathbf{r},\mathbf{r}') = P\left[\overset{\leftrightarrow}{I} + \frac{1}{k_0^2}\nabla\nabla\right]\frac{\exp[ik_0R]}{4\pi R} - \frac{\overset{\leftrightarrow}{I}}{3k_0^2}\delta(\mathbf{r}-\mathbf{r}'), \tag{2.27}$$

and where P stands for principal value of the integral and $k_0 = \omega/c$.

Using Eq. (2.26), the power generated by a dipole $\mathbf{p} = p\mathbf{e}_z$ can be cast in the form:

$$P_R = \frac{\mu_0\omega^3}{2}\mathrm{Im}\left[G^{(0)}_{zz}(\mathbf{r},\mathbf{r})\right]|\mathbf{p}_z|^2. \tag{2.28}$$

Although the real part of $G^{(0)}_{zz}$ diverges, its imaginary part is well defined and takes the value $\omega/(6\pi c)$ so that we recover the usual Larmor formula for the power radiated by a dipole:

$$P_R = \frac{\mu_0\omega^3}{2}\frac{\omega}{6\pi c}|\mathbf{p}_z|^2 = \frac{1}{4\pi\epsilon_0}\frac{\omega^4|\mathbf{p}|^2}{3c^3}. \tag{2.29}$$

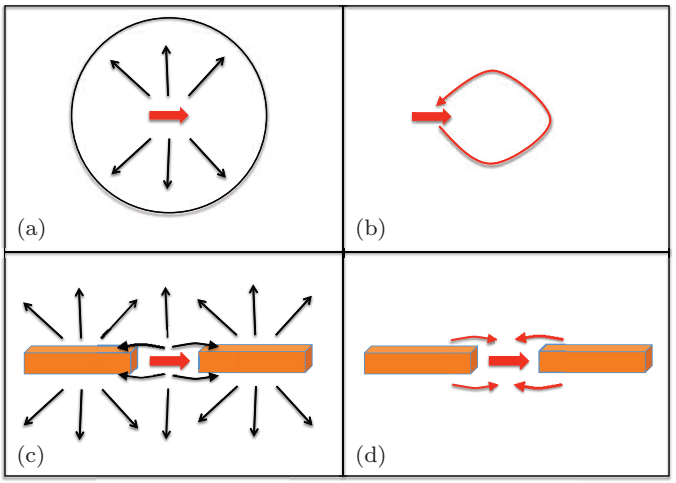


Fig. 2.2. Illustration of the energy conservation: the work done by the radiation reaction (b and d) is equal to the power radiated (a and c). Figure (a) shows a dipole radiating in vacuum. The black lines represent symbolically the lines of the Poynting vector. (b): The red line represents symbolically the field produced by the dipole acting on itself. (c): the black lines represent the flow of energy radiated directly or through the antenna. (d): the red line represents symbolically the radiation reaction produced by the antenna on the dipole.

2.3.4. *Classical radiation in vacuum: Damped oscillator*

In the previous section, we have found that the emitted power can be viewed as the work done by the radiation reaction. We adopt in what follows an alternative point of view to illustrate the role of the radiation reaction. Here, we write a dynamical equation for the dipole moment in vacuum. Let us consider for the sake of simplicity that the dipole consists of a charge $-e$ oscillating along the axis $0z$. We assume that the forces due to the atom are modelled by a parabolic potential $m\omega_0^{*2}z^2/2$. Furthermore, there is an electric field produced by the charge itself (namely, the radiation reaction) given by $\mu_0\omega^2 G_{zz}(\mathbf{r},\mathbf{r})(ez)$. Newton's equation yields:

$$-m\omega^2 z = -m\omega_0^{*2}z + e\mu_0\omega^2 G_{zz}(\mathbf{r},\mathbf{r})(ez). \tag{2.30}$$

It follows that the eigenfrequency is a solution of the equation:

$$\omega^2 - \omega_0^2 + i\frac{e^2\mu_0\omega^2}{m}\mathrm{Im}[G_{zz}] = 0, \tag{2.31}$$

where we have introduced the notation $\omega_0^2 = \omega_0^{*2} - \frac{e^2\mu_0\omega^2}{m}\mathrm{Re}[G_{zz}]$. Hence, it is seen that accounting for the radiation reaction introduces both a frequency shift and an imaginary part describing the radiation damping. The frequency shift is the Lamb shift.[18] The eigenfrequency can be approximated by

$$\omega \approx \omega_0 - i\frac{\gamma_{cl}}{2}, \tag{2.32}$$

where we recover the classical decay rate

$$\gamma_{cl} = \frac{e^2\mu_0\omega_0}{m}\mathrm{Im}[G_{zz}] = \frac{e^2\omega_0^2}{6\pi m\epsilon_0 c^3}. \tag{2.33}$$

2.3.5. *Classical radiation in the presence of an antenna*

In this subsection, we consider the power radiated by a classical dipole source with dipole moment $\mathbf{p}_S$ *in the presence of an antenna*. What is called antenna here can be any object made of any arbitrary material. The dipole and the antenna are located in a volume V. The dipole is outside the object. It drives a current density $\mathbf{j}_{ant}(\mathbf{r}')$ in the environment which leads to an ohmic dissipation rate $\mathbf{j}_{ant}(\mathbf{r}') \cdot \mathbf{E}(\mathbf{r}')$ where $\mathbf{E}(\mathbf{r}')$ is the electric field induced in the antenna. The energy conservation can now be cast in the form:

$$\frac{1}{2}\mathrm{Re}\left[i\omega\mathbf{p}_S \cdot \mathbf{E}^*\right] = \int_V \mathbf{j}_{ant}(\mathbf{r}) \cdot \mathbf{E}(\mathbf{r})\, \mathrm{d}^3\mathbf{r} + \int_A \mathbf{S} \cdot \mathrm{d}\mathbf{A}. \tag{2.34}$$

This equation shows that the power delivered by the dipole is equal to the sum of the radiated power and the power dissipated in the antenna by losses. It is thus seen that the energy conservation can

now be written in the form:

$$\frac{\mu_0\omega^3}{2}\mathrm{Im}\left[G_{zz}(\mathbf{r},\mathbf{r})\right]|\mathbf{p}|^2 = P_{NR} + P_R \tag{2.35}$$

It is important to note that the Green tensor is no longer the vacuum Green tensor. It provides the field produced by the dipole in the presence of the antenna.

$$\mathbf{E}(\mathbf{r}) = \mu_0\omega^2 \overleftrightarrow{G}(\mathbf{r},\mathbf{r}')\mathbf{p}, \tag{2.36}$$

In particular, the radiation reaction includes a term which is the field radiated by the antenna back onto the dipole source. Hence, the only formal modification to the previous classical calculation is to replace $G^{(0)}$ by G. The decay rate is thus proportional to $\mathrm{Im}[G_{zz}]$. The decay rate γ_{ant} in the presence of the antenna is given by:

$$\frac{\gamma_{ant}}{\gamma_0} = \frac{\mathrm{Im}[G_{zz}]}{\mathrm{Im}[G^{(0)}_{zz}]}. \tag{2.37}$$

Although the spontaneous emission decay rate depends on $\hbar$ and on the quantum electric dipole moment matrix element, the spontaneous emission decay rate γ_{ant} in an arbitrary environment is simply the vacuum decay rate γ_0 in vacuum multiplied by a correction factor accounting for the modification of the radiation reaction. This is simply given by the ratio of the imaginary part of Green tensor.

Using the quantum point of view, we can rephrase this statement by invoking the modification of the density of states. However, our previous discussion based on the radiation reaction clearly pointed out that the decay rate depends on the exact position of the emitter. In terms of density of modes, this requires to introduce the concept of *local* density of states $\rho(\mathbf{r},\omega)$ which is a density of states per unit volume. To understand the origin of the spatial dependence of the density of states, it suffices to think of the spatial distribution of a mode in a gap between two mirrors forming a cavity. There are modes with nodes and antinodes. A dipole will couple efficiently to a mode if it is located at an antinode.

Finally, we stress that the decay rate can be increased either by the increase of the radiated power or by the additional decay channels

due to losses in the antenna as shown in Eq. (2.35). In other words, the calculation of the local density of states based on the imaginary part of the Green tensor takes into account the non-radiative decay.

2.3.6. *When quantum physics meets classical physics*

We now compare the classical and quantum approches. The connection between the two points of view can be established in a simple manner. In vacuum we have:

$$\frac{\omega^2}{\pi^2 c^3} = \frac{6\omega}{\pi c^2}\mathrm{Im}[G_{zz}^{(0)}(\mathbf{r},\mathbf{r}')] \tag{2.38}$$

where the left hand side term was introduced as a density of electromagnetic states and the right hand side term was introduced as the radiation reaction. If we now consider a more general situation than vacuum by introducing nanoantennas, the radiation reaction is still well defined. It turns out that it can be shown[20] that the local density of states $\rho(\mathbf{r},\omega)$ is proportional to the imaginary part of the Green tensor:

$$\rho(\mathbf{r},\omega) = \frac{2\omega}{\pi c^2}\mathrm{Im}[\mathrm{tr}\overleftrightarrow{G}(\mathbf{r},\mathbf{r})]. \tag{2.39}$$

This formula is a general formula for the local density of states in presence of lossy objects.[c,20] When dealing with spontaneous emission, we restrict the LDOS to the axis parallel to the dipole moment to obtain the projected LDOS. Inserting this form in the spontaneous emission decay rate given in Eq. (2.22), we obtain:

$$\gamma_R = \frac{2\pi}{\hbar^2}|\hat{W}_{if}|^2\frac{2\omega}{\pi c^2}\mathrm{Im}[G_{zz}(\mathbf{r},\mathbf{r})]. \tag{2.40}$$

It is seen that the structure of the decay rate is given by the product of a quantum term ($\frac{2\pi}{\hbar^2}|\hat{W}_{if}|^2$) and the local density of states

[c]The formula is valid in a non-lossy medium close to lossy media but is not valid inside a lossy local medium. The formula predicts an unphysical divergence of the density of states. The origin of the divergence is a consequence of the absence of cutoff frequencies for wavevectors when using a local model of the permittivity. This is similar to the divergence of the number of phonon modes in a crystal which is avoided by introducing the Debye cutoff frequency.

which is a classical quantity. Upon comparing the decay rate with antenna and in vacuum, we recover again the simple result:

$$\gamma_{ant} = \gamma_0 \frac{\mathrm{Im}[G_{zz}]}{\mathrm{Im}[G_{zz}^{(0)}]}. \tag{2.41}$$

To summarize, the classical point of view identifies the power radiated with the power due to the radiation reaction which is proportional to $\mathrm{Im}[G_{zz}]$. The quantum point of view shows that the decay rate is proportional to the density of states which is proportional to $\mathrm{Im}[G_{zz}]$. At this point, we emphasize that both the classical and the quantum approach show that the spontaneous decay rate is proportional to $\mathrm{Im}[G_{zz}]$. The key remark is that modifying the environment of an emitter amounts modifying the local density of states and also the radiation reaction. Both modifications are described by the imaginary part of the Green tensor. It follows that a classical calculation of the modification of the spontaneous decay rate allows predicting the quantum result.

2.4. Controlling the Spontaneous Decay Rate by Modifying the Environment

The purpose of this section is to explain how the lifetime of an emitter can be modified by changing the environment. We will discuss several cases: modification of the refractive index, the case of lossy materials, the presence of a cavity and the presence of an antenna.

2.4.1. *Modification of the spontaneous decay rate in a dielectric*

We first consider the modification of the decay rate γ_R of an emitter located in a homogeneous transparent dielectric with a real refractive index n as compared to the decay rate in vacuum γ_0. The first obvious consequence of the presence of a refractive index is that light speed is changed from c to $v = c/n$. It follows directly from Eq. (2.23) that the local density of states is multiplied by a factor n^3. However, this does not mean that the spontaneous emission rate has been multiplied by

the same factor. Indeed, the amplitude of the quantized electric field in a dielectric is given by $\sqrt{\hbar\omega/\epsilon_0\epsilon V}$. Hence, the matrix element $|\hat{W}_{if}|^2$ is divided by n^2. Finally, it is seen that the density of states term in Eq. (2.22) is multiplied by n^3 while $|\hat{W}_{if}|^2$ is divided by $\epsilon = n^2$. The spontaneous emission rate is thus $\gamma_R = n\gamma_0$.

We now use the radiation reaction point of view to address the same issue. We have seen in Eq. (2.28) that the power radiated by a classical dipole with a fixed dipole moment is proportional to the imaginary part of the Green tensor. In a homogeneous dielectric, the imaginary part of the Green tensor is given by $\frac{n\omega}{6\pi c}$. Hence, we see from Eq. (2.28), that the power radiated by a classical dipole in a lossless dielectric with refractive index n is multiplied by a factor n.

2.4.2. *Modification of the density of states close to an interface*

We now discuss the behaviour of the local density of states (LDOS) in the vicinity of an interface separating two homogeneous isotropic materials. This effect was first observed experimentally by Drexhage[21] for europium ions located above a silver mirror. The distance was varied by inserting an adjustable number of organic monolayers. A seminal contribution by Chance, Prock and Silbey[22] showed that this result could be explained using a classical dipole model. The reason for the success of this classical model was just discussed in the previous sections. Radiation reaction and density of states are classical concepts. Here, we will discuss the case of an interface between two dielectrics. We start by displaying the result in Fig. 2.3 for a dielectric interface. This question was studied extensively in Ref. 23.

It is seen that the LDOS oscillates as the distance to the interface changes. This behaviour can be understood in terms of interferences. As the radiated field is the sum of the field emitted upwards and the field emitted downwards and reflected on the substrate, it is seen that the total field radiated depends on the distance to the interface. An alternative and equivalent interpretation is to consider the structure of the modes in the presence of an interface. The modes

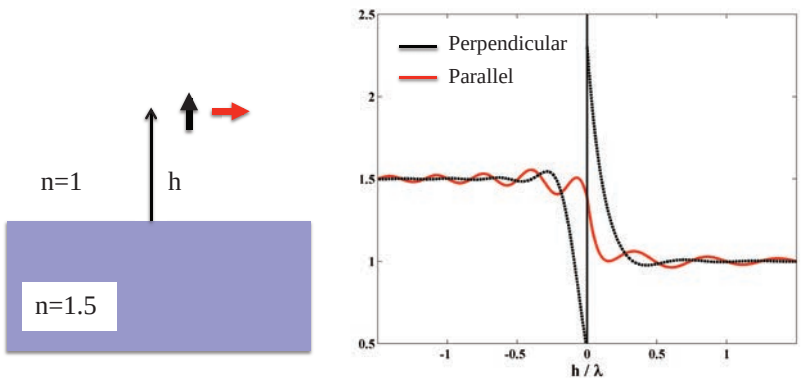


Fig. 2.3. **Local density of states close to an interface**. The LDOS normalized by the LDOS in vacuum has an asymptotic value of 1.5 in the dielectric away from the interface. Note the very different behaviour depending on the field polarization. For both dipole orientations, the presence of the interface is responsible for oscillations.

in vacuum are plane waves. In the presence of interfaces, the modes are the sum of an incident and a reflected wave on one side of the interface and a transmitted wave on the other side. The interference between the incident and reflected waves produces an oscillating field with a period that depends on the z-component of the wave vector. By adding all possible angles of incidence, the visibility is smeared out except close to the interface where the path difference is null. Indeed, at the interface, the boundary condition imposes the phase of the interference pattern. For instance, for a perfect conductor, the field has to be zero on the interface whatever the angle of incidence so that all the interference patterns have a node at the interface.

We observe in Fig. 2.3 that the LDOS close to a dielectric interface presents a clear maximum in the vacuum for p-polarization (for a dipole normal to the interface). This variation is observed close to the interface. It seems that the emitters in vacuum benefit from the larger LDOS of the dielectric material while the LDOS in the dielectric decays close to the interface. The decay of the LDOS in the dielectric can be analysed as a consequence of the interferences

between an incident wave and the reflected wave. In total internal reflection, the z-component of the reflected field changes sign so that the LDOS decreases as compared to the bulk in dielectric. We also observe that the LDOS increases in vacuum. This is due to the modes existing in the dielectric and totally reflected at the interface. These modes have an evanescent tail which extends in vacuum providing an additional contribution to the LDOS in vacuum. This mechanism also explains why the enhancement is confined close to the interface at a distance roughly given by $\lambda/2\pi n$.

To explain why the LDOS is larger in vacuum than in the dielectric, we note that the normal component of $\mathbf{D}$ is continuous across the interface so that the normal electric field is larger by a factor ϵ in vacuum. The ratio of LDOS is proportional to the ratio of intensity and is thus given by $\epsilon^2 = n^4$. In the example shown in Fig. 2.3, a factor 5.06 is expected in good agreement with the exact numerical calculation. By contrast, for parallel polarization, the electric field is continuous so that no discontinuity of the LDOS is expected.

We now consider the case of an interface between a metal and vacuum. The result is shown in Fig. 2.4(a) in vacuum above the interface. We observe again the oscillatory behaviour due to interferences. It is seen that for the field polarized perpendicular to the interface, the LDOS diverges when the distance to the interface goes to zero. This large LDOS in the near field is associated to the excitation of non-radiative modes in the metal. These non-radiative modes decay into heat. We discuss this topic in the following section.

2.4.3. *Radiation quenching*

One of the major limitations of nanoantennas is the competition between losses and radiation. The goal of this section is to provide a physical picture of the mechanisms at work in this process. We start with an analysis in the framework of classical electrodynamics. We consider a monochromatic electric dipole located above a metallic interface at a distance d. The electric field produced by the electric dipole at a distance r varies as $1/r^3$ when $r \ll \lambda$. If a

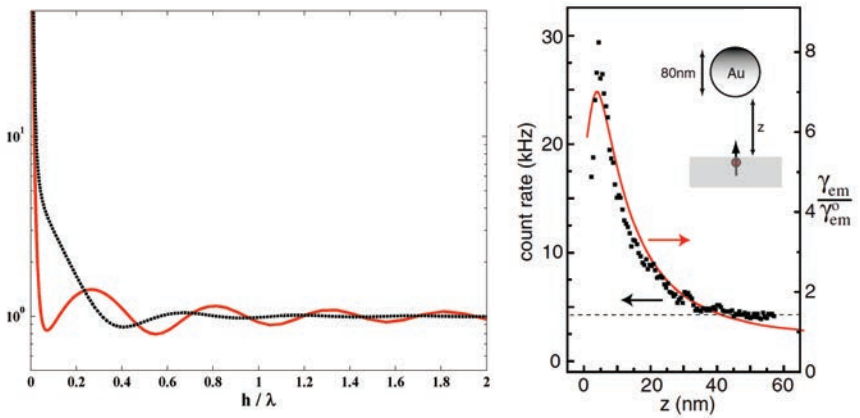


Fig. 2.4. Influence of a metal on the spontaneous emission. (a) Decay rate of an electric dipole normalized by the decay rate in vacuum. h is the distance in μm to a gold film. It is seen that the decay rate increases at small distances due to the non-radiative energy transfer to the surface. (b) Fluorescence signal of a molecule at a distance d from a gold nanosphere. It is seen that the signal decays for a distance smaller than 5 nm indicating non-radiative energy transfer to the metal nanoparticle. The fluorescence is quenched. Figure reproduced from Ref. 17

subwavelength electric dipole is brought at distances smaller than λ from an interface separating vacuum from a lossy material, it induces large current densities which results in large losses. These losses may become much larger than the power radiated so that the radiative efficiency drops to zero. This phenomenon is called quenching and has been known for many years. Although the previous discussion is made in the framework of classical electrodynamics, it is valid for quantum emitters such as atoms or quantum dots.[22,24] Non-radiative losses in the metal are due to interband absorption, electron-hole excitation and electron-scattering losses. The trade-off between quenching and fluorescence is well illustrated in Fig. 2.4(b) where the fluorescence enhancement of a molecule as a function of the distance to a gold nanosphere is plotted. This result suggests that fluorophores should be at distances larger than 5 nm from the surface to prevent quenching.

It is possible to analyse quenching in terms of local density of non-radiative states. To this aim, we simply derive the asymptotic behaviour of the LDOS for small distances[20]:

$$\frac{\mathrm{Im}[G_{zz}]}{\mathrm{Im}[G_{zz}^{(0)}]} = \frac{\mathrm{Im}(\epsilon)}{|\epsilon+1|^2}\frac{3}{4(k_0 z)^3}. \tag{2.42}$$

This result shows clearly that the divergence of the LDOS is proportional to $\mathrm{Im}(\epsilon)$ and therefore associated to losses in the metal. The presence of the resonant term $\frac{1}{|\epsilon+1|}$ indicates a surface plasmon contribution at the frequency such that $\mathrm{Re}(\epsilon) = -1$. For a distance of 1 nm and gold, the increase of the decay rate is about 5 orders of magnitude. In other words, all the energy is transfered to the metal losses. This is referred to as non-radiative energy transfer. As this process becomes very efficient, the radiative decay becomes very unlikely. This is called quenching. This suppression of radiation due to fast non-radiative decay has been studied in detail in.[24] Of particular importance is the review of non-local effects which become important for length scales smaller than a 1 nm.

The physical origin of the non-radiative additional states in a metal is the presence of electrons. More generally, the increased density of states close to matter is due to the presence of electronic excitations which are coupled to the electromagnetic field. The simple idea is as follows. Since each atom contributes to one electron in the conduction band, the number of decay channels per unit volume is several orders of magnitude larger than the number of electromagnetic modes per unit volume. The vast majority of these modes have a large wavevector and are therefore non-radiative. A simple way to avoid the non-radiative decay is to increase the distance between the emitter and the metal taking advantage of the spatial decay $1/(k_0 z)^3$. It is only recently that it has been recognized that the contribution to the LDOS of a few modes properly engineered can compete with the non-radiative states. In the next section, we introduce the idea that a single mode can produce a very large *density* of states per unit volume and a given frequency interval if it is properly confined in space.

2.4.4. *Controlling* $\mathbf{Im}[G_{zz}]$ *with a cavity: Modifying the LDOS*

The first proposal to control spontaneous emission by modifying the environment was introduced in the context of nuclear magnetic resonance by Purcell in the microwave regime. We now give a back of the envelope derivation of the modification of the spontaneous emission rate. Assume that there is an emitter in a single mode cavity with mode volume V and a decay rate κ. Although the cavity has a single mode, the presence of radiative and non-radiative losses of the cavity introduces a broadening of the cavity spectrum. If the cavity width κ is much larger than the two-level width γ_R, the cavity can be considered to be a quasi continuum to which the two-level system is coupled so that the Fermi golden rule can be used. We further assume that the resonance is Lorentzian so that the spectral density is given by the normalized function:

$$g(\omega) = \frac{1}{2\pi} \frac{\kappa}{(\omega - \omega_0)^2 + \kappa^2/4}. \tag{2.43}$$

At resonance, the spectral density is $g(\omega_0) = 2/\pi\kappa$. The local density of states is thus easily found. We have one mode in a volume V with a spectral density $g(\omega_0)$. We also assume that the mode is uniform in the cavity for the sake of simplicity. It follows that the local density of electromagnetic states is given by $g(\omega_0)/V$. Equation (2.37) shows that the acceleration factor of the decay rate, also called Purcell factor F_P, γ_{cav}/γ_R is given by the ratio of LDOS:

$$F_P = \frac{\gamma_{cav}}{\gamma_R} = \frac{2}{\pi\kappa V} \frac{3\pi^2 c^3}{\omega^2 n^3} = \frac{3Q}{4\pi^2} \frac{\lambda^3}{V}, \tag{2.44}$$

where $2\pi/\lambda = n\omega/c$ and the cavity quality factor is defined by $Q = \frac{\omega}{\kappa}$. This is the form originally introduced by Purcell. It is seen that the LDOS can be increased by confining the mode spectrally (large Q factor) and spatially (small mode volume). For a dielectric microcavity, the quality factor can be higher than 10^5. On the other hand, the ratio λ^3/V is typically on the order of 10 or less.

2.4.5. *Controlling* $\mathbf{Im}[G_{zz}]$ *with a cavity: The cavity quantum electrodynamics (CQED) point of view*

In this section, we adopt the notations of the CQED community to describe the modification of spontaneous emission by the cavity. We aim at establishing a link between different notations and points of view used by different communities. Losses are responsible for a broadening so that, if $\kappa \gg \gamma_R$, the cavity can be modelled as a continuum of states and Fermi golden rule can be used. The density of states is given by $g(\omega_0) = 2/\pi\kappa$. In CQED notations, the interaction energy $|\hat{W}_{if}|$ is denoted $\hbar g$. It follows from Eq. (2.22) that:

$$\gamma_{cav} = \frac{2\pi}{\hbar^2}(\hbar g)^2 \frac{2}{\pi\kappa} = \frac{4g^2}{\kappa}. \tag{2.45}$$

The Purcell factor in the cavity is thus given by:

$$F_P = \frac{\gamma_{cav}}{\gamma_R} = \frac{4g^2}{\gamma_R \kappa}. \tag{2.46}$$

With these notations, the Purcell factor appears as a dimensionless quantity. It is worth mentioning that the quantity $g^2/(\kappa\gamma)$ is often called cooperativity or single-photon cooperativity and denoted C_1 in a different context. The reason for this name is that in the case of N emitters in a cavity, the quantity that describes the collective (or cooperative) interaction between the N emitters and the cavity is $C_N = Ng^2/(\kappa\gamma_R)$. Obviously, with a single emitter, there are no cooperative or collective effects but the comparison shows that $F_P = 4C_1$.

2.4.6. *Controlling* $\mathbf{Im}[G_{zz}]$ *with a nanoantenna*

If we now consider a nanoantenna, it is usually difficult to estimate a mode volume associated with an antenna. The field confinement is mostly associated to evanescent waves whereas the field of a microcavity is associated to stationary waves in a cavity. On the other hand, it is possible to derive the Green tensor. An elementary derivation is made for a toy model in the case of a metallic sphere modelled

by a dipole in Ref. 25. A much more powerful technique has been introduced recently by Sauvan *et al.*[26] to deal with lossy antennas.

The key difference between a microcavity and a nanoantenna is that the antenna has usually a quality factor on the order of 10 and a mode volume significantly subwavelength. In other words, the electric field is extremely confined in space while the spectrum is poorly confined in frequency domain. This makes nanoantennas very interesting for broadband emitters. The bandwidth of a radio antenna is of course an important figure of merit. It is desirable to use the same antenna while scanning different transmission channels with different carrier frequencies. In summary, a nanoantenna or a cavity allows engineering the field confinement in direct space and in frequency space.

2.4.7. *The Rabi oscillation regime*

In this section, we give a brief introduction to the strong coupling regime. The reader is refered to Ref. 27 for an extensive discussion. Let us consider a single atom in the cavity depicted in Fig. 2.5. Here, we ask the question: will a photon trapped in the cavity be absorbed by the atom? To answer that question, we simply use the Beer-Lambert law: the transmission through a scattering medium with thickness L is given by $\exp(-L/l_{abs})$. The absorption length is given by $l_{abs} = 1/\rho\sigma$ where ρ is the number of atoms per unit volume and σ is the absorption cross section. A photon in the cavity spends on average a time $\tau = 1/\kappa$ in the cavity so that it travels over a distance $L = (c/n)(1/\kappa)$. The distance normalized by the absorption length $L/l_{abs} = c/(n\kappa l_{abs})$ is called optical thickness. The absorption probability increases with the optical thickness. With one atom in the cavity with volume V, we have $\rho = 1/V$ and the absorption cross-section at resonance is given by $\sigma = 3\lambda^2/2\pi$.[d] It follows

[d]Here a vocabulary note may be useful. In the context of atoms, absorption means that the atom is excited. The absorbed energy will be subsequently radiated by spontaneous emission. In the case of a molecule with a more complex energy spectrum, there will be fluorescence and the emission frequency may be smaller. In the case of a particle, absorption means that the electromagnetic incident energy

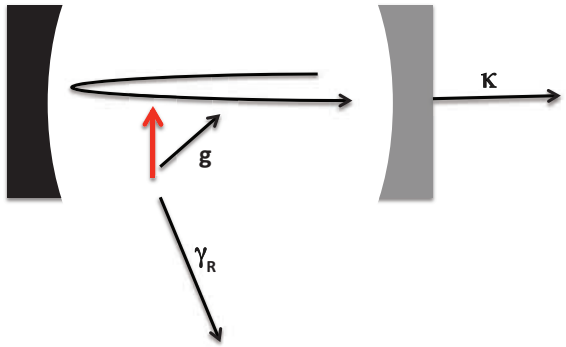


Fig. 2.5. Increasing the spontaneous emission rate with a cavity. The spontaneous emission decay rate is denoted γ_R, the coupling rate between the two-level system and the cavity is denoted g. The cavity decay rate is denoted κ, it is due to the transmission of the partially reflecting mirros for a Fabry-Perot cavity. The Purcell regime corresponds to the situation where the coupling to the cavity rate $4g^2/\kappa$ is faster than the coupling to the vacuum γ_R. In the bad cavity regime $2g \ll \kappa$, the cavity energy is released. Noe that the same picture applies to a nanoantenna.

that

$$\frac{c}{n\kappa}\rho\sigma = \frac{c}{n\kappa}\frac{1}{V}\frac{3\lambda^2}{2\pi} = \frac{c}{n\omega}\left(\frac{\omega}{\kappa}\right)\frac{1}{V}\frac{3\lambda^2}{2\pi} = \frac{3Q}{4\pi^2}\frac{\lambda^3}{V} = F_P = 4C_1. \tag{2.47}$$

It is seen that the Purcell factor appears to be the optical thickness of the cavity with one two-level system at resonance. With this argument, it is seen that a cooperativity C_1 larger than 1 means that a photon in the cavity will be absorbed by the atom before escaping the cavity. This suggests that the energy can be periodically exchanged between the atom and the cavity. The regime where the energy is

will be transformed into heat in the particle. While the excitation of an atom can relax only through spontaneous emission, a plasmon excitation of a particle can relax either through radiation or through internal losses. The consequence is that the absorption cross-section at resonance is given by $\sigma = 3\lambda^2/8\pi$ for a resonant dipolar particle and $\sigma = 3\lambda^2/2\pi$ for a resonant atom.

exchanged periodically between the cavity and the atom is the Rabi oscillation regime. It can be shown that the Rabi oscillation frequency is given by g. Hence, the oscillations are observed if their period is smaller than $1/\kappa$. This condition defines two regimes: either $g \ll \kappa$ or $\kappa \ll g$. In the regime

$$\gamma_R \ll g \ll \kappa, \tag{2.48}$$

the coupling between the emitter and the cavity takes place on a time scale $1/g$ much shorter than the spontaneous decay time $1/\gamma_R$. Once the cavity is excited, the photon is released as the cavity decay time is even shorter $g \ll \kappa$. This is the so-called "bad-cavity" limit or Purcell regime. Antennas operate in this regime. Conversely, if the cavity has a large quality factor ω/κ such that

$$\gamma_R \ll \kappa \ll g, \tag{2.49}$$

the photon is reabsorbed before leaving the cavity. In the context of the Rabi oscillation regime, the quantity $g^2/(\kappa\gamma)$ is called cooperativity instead of Purcell factor. The system oscillates between a photonic state and an atomic state, it is the Rabi oscillation regime. In this regime, the eigenstates of the system are linear superposition of the emitter state and the cavity state. The spectrum is characterized by a Rabi splitting given by $\Omega_R = 2g$. Note that the terminology differs between authors. For some of them, the regime of strong coupling covers only the case of Rabi oscillation $\gamma_R \ll \kappa \ll g$. Others use the term strong coupling if the cooperativity is large ($g^2 \gg \kappa\gamma_R$), i.e. for both the case of Rabi oscillation and the bad-cavity regime with large Purcell factor. Finally, we conclude by noting that the key difference between cavities and antennas is the quality factor. The usual situation for antennas is to operate in the bad cavity limit. The small quality factor is compensated by a very small mode volume. The observation of the Rabi oscillation regime with plasmonic nanoantennas is extremly challenging because the cavity decay rate is extremly fast due to metallic losses. On the other hand, it is possible to reduce the mode volume to extremly small values and therefore to enhance the coupling parameter g. Recent experiments have been

reported with spectral scattering signatures showing the Rabi splitting.[28,29]

2.5. Raman Scattering

We now address the question of the control of the Raman scattering process using a nanoantenna.[30] Before discussing the phenomena, we need to define the set-up. We consider a confocal microscope set-up as depicted in Fig. 2.6 with two pinholes in front of the source and the detector. The light coming from the source S is collimated and focussed on the sample by an objective. The antenna is located in the focus of the objective. The molecules are thus illuminated and scatter the light which is collected by the same objective and then reflected by the dichroic mirror towards the detector. Note that the detector is placed behind a pinhole in a confocal setup so that its position is the symmetric of the source position. We are interested in the analysis of the light scattered at frequency ω_2 by the molecules illuminated with a laser light at frequency ω_1. To analyse the system,

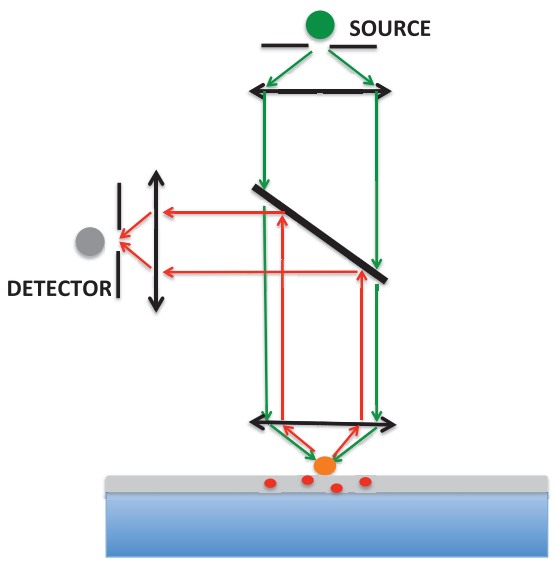


Fig. 2.6. Confocal set-up for Raman scattering. The antenna is represented by an orange nanosphere, the fluorophores are represented by red dots.

we consider that the incident light is produced by a polarized point-like source with dipole moment $\mathbf{p}_S$ located at the pinhole position $\mathbf{r}_S$. It follows that the field at $\mathbf{r}$, close to the focus of the objective can be cast in the form:

$$\mathbf{E}(\mathbf{r}) = \mu_0 \omega^2 \mathbf{G}(\mathbf{r}, \mathbf{r}_S, \omega_1) \mathbf{p}_S. \tag{2.50}$$

In the presence of the antenna, the field is modified and we denote it as:

$$\mathbf{E}_{ant}(\mathbf{r}) = \mu_0 \omega^2 \mathbf{G}_{ant}(\mathbf{r}, \mathbf{r}_S, \omega_1) \mathbf{p}_S. \tag{2.51}$$

The effect of the antenna is to enhance the electric field amplitude by a factor denoted $K(\omega_1)$:

$$K(\omega_1) = \frac{|\mathbf{G}_{ant}(\mathbf{r}, \mathbf{r}_S, \omega_1) \mathbf{p}_S|}{|\mathbf{G}(\mathbf{r}, \mathbf{r}_S, \omega_1) \mathbf{p}_S|}. \tag{2.52}$$

We stress that this field enhancement differs from the LDOS enhancement. The former deals with the value of the field in a specific illumination condition whereas the latter is an intrinsic quantity. The former is related to $|\mathbf{G}_{ant}(\mathbf{r}, \mathbf{r}_S, \omega_1) \mathbf{p}_S|$ whereas the latter is proportional to $\mathrm{Im}[G_{ant}(\mathbf{r}, \mathbf{r}, \omega_1)]$. The field produced by the induced dipole at frequency ω_2 due to the Raman shift is proportional to the Green tensor $\mathbf{G}_{ant}(\mathbf{r}_D, \mathbf{r}, \omega_2)$. By noting that D and S are symmetrical, we can replace the Green tensor by $\mathbf{G}_{ant}(\mathbf{r}_S, \mathbf{r}, \omega_2)$. We now make use of the reciprocity of the Green tensor $G_{nm}(\mathbf{r}, \mathbf{r}', \omega) = G_{mn}(\mathbf{r}', \mathbf{r}, \omega)$.[e] It follows that $G_{nm,ant}(\mathbf{r}_S, \mathbf{r}, \omega_2) = G_{mn,ant}(\mathbf{r}, \mathbf{r}_S, \omega_2)$. According to Eq. (2.52), the latter is simply proportional to the amplitude enhancement at frequency ω_2. This result demonstrates that the antenna serves both to increase the amplitude of the field illuminating the dipole, and the field radiated by the dipole towards the detector. The reciprocity theorem shows that these are two sides of

[e]This equality follows from the equality $\mathbf{p}_1 \cdot \mathbf{E}_2 = \mathbf{p}_2 \cdot \mathbf{E}_1$ derived from Maxwell equations where the electric field $\mathbf{E}_i$ is produced by the dipole $\mathbf{p}_i$ in an arbitrary environment with the only restriction that the materials have a symmetric permittivity tensor $\epsilon_{ij} = \epsilon_{ji}$. Inserting $\mathbf{E}_n(\mathbf{r}_1) = \mu_0 \omega^2 G_{nm}(\mathbf{r}_1, \mathbf{r}_2, \omega) p_{2m}$, we find that reciprocity implies $G_{nm}(\mathbf{r}_2, \mathbf{r}_1, \omega) = G_{mn}(\mathbf{r}_1, \mathbf{r}_2, \omega)$.

the same coin. Thus, we find that the field amplitude at the detector is multiplied by $K(\omega_1)K(\omega_2)$ and the intensity is multiplied by the square of this quantity. Finally, if we assume that the antenna and the microscope do not change between the two frequencies, we conclude that the presence of the antenna amounts to multiplying the Raman signal by a factor of $K^4(\omega_1)$. It is worth emphasizing that the field enhancement can be two orders of magnitude so that the signal enhancement can be eight orders of magnitude. We also emphasize that scattering is not limited by saturation, a key difference with fluorescence as we now discuss.

2.6. Fluorescence in Stationary Regime

We now consider the effect of an antenna on fluorescence.[30] The effect of the antenna cannot be summarized in a few words. It depends on the illumination: pulsed or stationary, intensity in the linear or in the saturated regime. It also depends on the nature of the emitter: the effect on a two level system or a dye molecule is not the same. This effect has been studied by many groups using either tips[4] or controlled nanospheres.[8,9] The effect of the antenna is more subtle than for Raman scattering because fluorescence is not a pure scattering process, it depends on the population of the excited emitters. It follows that there may be saturation effects. The aim of this section is to give a simple analysis of the dependence of the fluorescence signal on the parameters of the antenna.

2.6.1. *Fluorescence signal*

In the absence of an antenna, the fluorescence signal is proportional to the radiative rate of the molecule and the number of molecules N_2 in the excited state. We denote γ_R the radiative rate of the molecule without antenna and Γ_R the radiative rate in the presence of the antenna. The signals with and without antenna S_{ant} and S_0 are thus given by:

$$S_{ant} = C f_{coll} \Gamma_R N_2; \quad S_0 = C \gamma_R N_2, \tag{2.53}$$

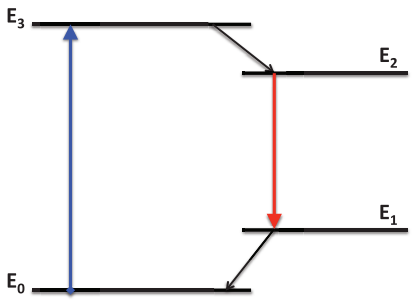


Fig. 2.7. Scheme of the fluorophore energy levels.

where C is a constant that accounts for the detector response and f_{coll} is a number smaller than 1 that accounts for the fraction of collected radiation (collection efficiency). We now need to write populations equations to find the number N_2 of molecules in the excited state.

2.6.2. *Population equations*

Let us begin by introducing a simplified level structure for the fluorophore. The energy scale is depicted in Fig. 2.7. The fundamental level is denoted 0. A laser pump at frequency ω_p illuminates the system and excites the fluorophore to level 3. A fast non-radiative decay process then populates level 2. The fluorescent radiative decay takes place between level 2 and level 1. Finally, level 1 has a fast non-radiative relaxation towards level 0. If we consider that the non-radiative decay of level 3 and 1 are very fast, it is possible to consider that their populations are negligible. The molecule has a radiative decay rate γ_R and a non-radiative decay rate γ_{NR} leading to an intrinsic quantum yield η_{int}:

$$\eta_{int} = \frac{\gamma_R}{\gamma_R + \gamma_{NR}} \tag{2.54}$$

With obvious notations, we can then write that the total number of fluorophores N satisfies:

$$N = N_0 + N_2. \tag{2.55}$$

In stationary regime, the rate of variation of N_0 is given by:

$$\frac{\mathrm{d}N_0}{\mathrm{d}t} = \Gamma_{tot} N_2 + B\, W_f N_2 - B'\, W_p N_0 = 0, \tag{2.56}$$

where W_f is the power density at the fluorescence frequency, $W_p = K^2 W_{inc}$ is the power density at the pump frequency W_{inc} enhanced by the antenna, $\Gamma_{tot} = \Gamma_R + \Gamma_{NR} + \gamma_{NR}$ is the total decay rate of the molecule accounting respectively for the radiative decay in the presence of the antenna, the non-radiative decay in the presence of the antenna and the internal non-radiative decay rate. B and B' are the Einstein coefficients for stimulated emission for the transition between levels 1 and 2. The solution yields:

$$N_0 = N_2 \frac{\Gamma_{tot} + BW_f}{B'W_p}. \tag{2.57}$$

Combining Eqs. (2.55) and (2.57) yields:

$$N_0 = N\frac{\Gamma_{tot} + BW_f}{\Gamma_{tot} + BW_f + B'W_p}; \quad N_2 = N\frac{B'W_p}{\Gamma_{tot} + BW_f + B'W_p} \tag{2.58}$$

2.6.3. *Saturation regime*

Here, we consider the case of large pumping intensity $B'\,W_p \gg \Gamma_{tot} + BW_f$ so that $N_2 \approx N$. In this so-called saturation regime, all molecules are excited. The signal is then given by:

$$S_{ant}^{(sat)} = C\, f_{coll}^{ant}\, \Gamma_R N. \tag{2.59}$$

In the saturation regime, the signal is limited by the spontaneous emission rate Γ_R. In simple words, the pumping efficiency is so large that all the molecules are excited. However, as opposed to the scattering signal which cannot be saturated, the fluorescence intensity is limited by the spontaneous emission rate. On average, an excited molecule needs a time $1/\Gamma_R$ to decay radiatively and this does not depend on the incident intensity. Without antenna, the saturated signal S_0 is proportional to γ_R so that the effect of the antenna is

given by:

$$S_{ant}^{(sat)} = \frac{\Gamma_R}{\gamma_R} \frac{f_{coll}^{ant}}{f_{coll}} S_0^{(sat)} \tag{2.60}$$

In this regime, the Purcell effect can increase the signal but the incident field enhancement does not play any role.

2.6.4. *Linear regime*

We now consider the low-pumping regime where we can approximate the population of the excited state by:

$$N_2 = N \frac{B' W_p}{\Gamma_{tot}}, \tag{2.61}$$

so that the population is proportional to the incident intensity. This is the linear regime. The fluorescence signal is then given by:

$$S_{ant}^{(lin)} = C\, f_{coll}^{ant} \Gamma_R \frac{B' W_p}{\Gamma_R + \Gamma_{NR} + \gamma_{NR}}. \tag{2.62}$$

In the case $\Gamma_R + \Gamma_{NR} \gg \gamma_{NR}$, the result can be approximated by

$$S_{ant}^{(lin)} \approx C\, f_{coll}^{ant} \Gamma_R \frac{B' W_p}{\Gamma_R + \Gamma_{NR}} = C\, f_{coll}^{ant} \eta_{ant} B' K^2 W_{inc}. \tag{2.63}$$

Without antenna, the signal is given by

$$S_0^{(lin)} = C\, f_{coll} \eta_{int} B' W_{inc}, \tag{2.64}$$

so that the antenna signal is given by:

$$S_{ant}^{(lin)} = \frac{\eta_{ant}}{\eta_{int}} \frac{f_{coll}^{ant}}{f_{coll}} K^2 S_0^{(lin)} \tag{2.65}$$

The effect of the antenna is therefore manyfold. The intensity is multiplied by the local intensity enhancement factor K^2. The antenna can concentrate the emission pattern into the detection solid angle of the receiving system thereby increasing f_{coll}. Finally, the quantum yield of the molecule η_{int} is replaced by the radiative efficiency of the antenna η_{ant} if $\Gamma_R + \Gamma_{NR} \gg \gamma_{NR}$. It is seen that the antennas are of particular interest for molecules with low intrinsic quantum

yield.[32] The physical reason is simple: a low quantum yield means that the non-radiative decay channels are faster than the radiative decay channel ($\gamma_{NR} \gg \gamma_R$) so that light cannot be emitted efficiently. If the antenna introduces a new decay channel faster than the non-radiative decay channel ($\Gamma_R > \gamma_{NR}$), the molecule energy can be funneled to the antenna. The intrinsic quantum efficiency is then replaced by the antenna radiation efficiency.

2.6.5. *Fluorescence for broad spectrum emitters*

The previous discussion considered the case of emitters with a single fluorescence transition between level 2 and level 1. The effect of the antenna is markedly different[31] if the emission has a broad spectrum as it is the case of dye molecules with a distribution of sublevels in the ground state. In particular, the signal is no longer proportional to the quantum yield in the linear regime. To understand qualitatively this behaviour, let us analyse the signal given by Eq. (2.62). While the numerator is proportional to the decay rate $\Gamma_R(\omega)$ at frequency ω, the denominator results from all the decay channels towards all the sublevels of the ground level at all different frequencies $\Gamma_{tot} = \int_0^\infty \mathrm{d}\omega \Gamma_{tot}(\omega)$. This term is no longer a function of frequency. It follows that the fluorescence signal of a dye close to an antenna in the linear regime is proportional to $\Gamma_R(\omega)$. A consequence of this result is that the emission spectrum can be modified. Indeed, if the antenna has a spectral width similar or smaller than the width of the emission spectrum, the spectrum of $\Gamma_R(\omega)$ will be different from the emitter spectrum.

2.7. Fluorescence in the Impulse Regime

We derive in this section a model for the fluorescence signal when using a periodic pulse excitation. Considering an ensemble of N emitters, we first derive the number of emitters in the excited state as a function of time $N_2(t)$. We consider periodic pulses with a repetition rate R of small duration δt and irradiance ϕ_{inc}. After the pulse, the emitters in the ground state are excited with a

probability Π. Let $N_2(0)$ be the number of excited emitters before the pulse; the number of excited emitters after the pulse is thus $N_2(0^+) = N_2(0^-) + \Pi[N - N_2(0^-)]$. This population decays with a rate Γ and the number of excited emitters at the end of a cycle is $N_2(t = 1/R) = N_2(0^+)\exp(-\Gamma/R)$. Because the excitation is periodic, we have $N_2(t = 1/R) = N_2(0^-)$. We consequently derive the number of excited emitters as a function of time:

$$N_2(t) = N_2(0^+)\exp(-\Gamma t); \quad N_2(0^+) = N\frac{\Pi}{1-(1-\Pi)e^{-\Gamma/R}}. \tag{2.66}$$

We can now use the population to derive the emitted intensity. The radiative decay rate of the emitter is Γ_R and the collection efficiency of the system is f_{coll}. During an excitation cycle, the signal consisting of the mean intensity collected (in photons.s^{-1}) per pulse is:

$$S = R\int_{0^+}^{1/R} \Gamma_R\, f_{coll}\, N_2(t)\mathrm{d}t, \tag{2.67}$$

which yields:

$$S = \frac{\Gamma_R}{\Gamma}\, R\, f_{coll}\, N\frac{\Pi}{1-(1-\Pi)e^{-\Gamma/R}}(1-e^{-\Gamma/R}) \tag{2.68}$$

Possible saturation effects are included in the dependence of the excitation probability Π on the pump irradiance. We introduce this effect using rate equations for populations of an ensemble of emitters. The excitation is assumed to be proportional to the product of the absorption cross-section σ_{abs}, the pump irradiance ϕ_{inc} and the population of emitters in the ground state. In this simple model, we have assumed that internal relaxation processes after excitation are faster than the radiative decay. During the excitation (duration δt), the population equation is:

$$\frac{\mathrm{d}N_2(t)}{\mathrm{d}t} = \frac{\sigma_{abs}\phi_{inc}}{\hbar\omega_{inc}}[N - N_2(t)] - \Gamma_{tot}N_2(t). \tag{2.69}$$

We solve this equation using the initial condition $N_2(0) = 0$. The transition probability Π for an emitter initially in its ground state can then be cast in the form:

$$\Pi(\phi_{inc}) = \frac{N_2(\delta t)}{N} = \frac{1}{1 + \frac{\Gamma \hbar \omega_{inc}}{\sigma_{abs} \phi_{inc}}} \left(1 - e^{-\frac{\delta t \sigma_{abs}}{\hbar \omega_{inc}} \phi_{inc}} e^{-\Gamma \delta t}\right) \quad (2.70)$$

These results give a model for the fluorescence collected per pulse regardless of the relative values of the repetition rate R and the decay rate Γ and for any intensity regime. We now consider the particular case of a low-excitation regime $\phi_{inc} \ll \frac{\Gamma \hbar \omega_{inc}}{\sigma_{abs}}$, so that $\Pi \ll 1$ and a low repetition rate $\Gamma/R \ll 1$. We get the following form of the fluorescence signal:

$$S = \frac{\Gamma_R}{\Gamma} \, R \, f_{coll} \, N \, \frac{\sigma_{abs} \phi_{inc} \delta t}{\hbar \omega_{inc}}. \quad (2.71)$$

This result shows that under these conditions, the fluorescence signal per pulse has a similar structure as in the stationary regime. It is proportional to the antenna radiative efficiency, to the intensity enhancement and to the collection efficiency. However, there is an additional factor that introduces a difference. The total fluorescence signal is proportional to the repetition rate R. As we have made the approximation that the pulse repetition rate R satisfies $\Gamma/R \gg 1$, the signal is bounded by Γ. Hence, in the pulse regime, the signal can be increased by increasing the decay rate.

2.8. Antenna Design Rules

The purpose of this section is to briefly analyse the rules to design an efficient nanoantenna operating in the Purcell regime. More specifically, we address the issue of an efficient coupling between a plane wave and a quantum emitter while preserving a good radiation efficiency. We are thus interested in discussing the trade-off between local field enhancement and quenching. On one hand, the coupling to the antenna mode is increased by reducing the mode volume. Typical examples are a dimer of nanorods or a bow-tie antenna.[33] The

field is increased as the gap between the two parts of the antenna is reduced. However, as the gap is reduced below 10 nm, an emitter placed at the center is at a distance of 5 nm to the metal surfaces so that quenching is expected to start playing a negative role. Hence, enhancing the LDOS by reducing gaps to values below 10 nm has been considered to be plagued by the onset of non-radiative losses. Several results reported after 2012 have challenged this point of view. It has been shown both theoretically[34–37] and experimentally [38] that large Purcell factors with radiative efficiencies on the order of 50% are possible. In what follows, we want to provide some physical arguments providing some hope for future improvement of the antennas efficiencies with very large Purcell factors. Finally, we note that when the distance becomes smaller than 1 nm, tunnel effect becomes possible, considerably modifying the enhancement.[39–41]

2.8.1. *Radiative and non-radiative modes*

A key issue is the competition between radiative and non-radiative modes. As discussed in Section 4, the LDOS can be 4 or 5 orders of magnitude larger than the LDOS in vacuum. This is usually mostly due to the very large numbers of non-radiative modes in metallic structures. However, it has also been shown that the LDOS due to a single mode can be very large provided that the mode volume is very small and its quality factor is good. The key to a 50% efficiency is simply to ensure that the antenna or cavity contribution to the LDOS matches the non-radiative contribution. This can be achieved by two means: (i) using a hot spot to take advantage of the light confinement, (ii) using a plasmonic resonance to further enhance the mode field. The light confinement can be obtained by engineering the antenna. For instance, a tip produces a singular behaviour of the field taking advantage of the so-called lightning rod effect, a dimer of nanospheres produces a highly localized field between the spheres, etc. By combining light confinement and plasmonic resonance, it is possible to design a hot spot characterized by a LDOS which can be five orders of magnitude larger than in vacuum.

2.8.2. *Radiative efficiency of the antenna mode*

Generating a hot spot is by no means sufficient to obtain a good antenna. A hot spot merely ensures that the emitter will excite the antenna mode. A second issue is the radiative efficiency of the antenna mode. Indeed, the antenna mode can either relax by emitting radiation or through losses in the material. It is thus necessary to reduce non-radiative losses and to increase the radiated power. Non-radiative losses can be reduced by designing metallo-dielectric antennas.[34,36] Radiated power can be increased by properly designing the size of the antenna and its structure. This requires a minimum size. As a rule of thumb, it is useful to keep in mind that the absorption cross section of a metallic sphere is equal to its scattering cross section for a diameter on the order of 40 nm. For smaller radii, absorption dominates while radiative losses dominates for larger radii. In summary, an efficient nanoantenna needs to be large to radiate efficiently and to have some localized hot spots to couple efficiently to the emitter. This is the case of all the recently proposed nanoantennas such as nanocones, nanocubes, nanopatches.

2.9. Nanoantenna and Two-Level System Impedances

In this section, we introduce the concept of impedance for nanoantennas following Ref. 42. An extended discussion can be found in Ref. 43. The motivation for the definition of the concept of impedance is based on the fact that there are concepts such as impedance matching, reactive energy, radiation resistance which are widely used in antennas and can be useful in optics. Of particular importance is the idea of impedance matching: how can we optimize the transfer of the power collected by the antenna to a load? We note that this issue can be addressed without using the concept of impedance. By computing the Green tensor of the structure, the field is known and therefore, any quantity of interest can be computed. Yet, it is of practical interest to reformulate the multiple scattering equations of optics using the vocabulary of electrical circuits. When adressing this question, the first difficulty is the definition of the feed points. In

electrical engineering, there is a line connecting the antenna to the source or detector. It is possible to define the line impedance. The line is connected to the antenna at two feed points and it is possible to define the voltage difference between these two points. If we now consider an atom radiating light (the source) in proximity to a metallic nanoparticle (the antenna), it is not possible to define the line nor the feed points. In this section, we first introduce an impedance for the antenna and for a two-level system. We then apply this concept to discuss the maximal power absorbed by a nanoparticle close to an antenna. This is for example the goal of an antenna designed to funnel energy into a subwavelength detector.

2.9.1. *Nanoantenna impedance*

Since impedance is the concept that allows to discuss the power radiated by an antenna, let us try to introduce heuristically the impedance by analyzing the power emitted by an optical antenna. We start from the power P_0 delivered by the dipole to the optical field. The time averaged value is given by $P_0 = \langle -\frac{d\mathbf{p}}{dt} \cdot \mathbf{E} \rangle = \frac{1}{2}\mathrm{Re}(i\omega \mathbf{p} \cdot \mathbf{E}^*)$. There is a clear similarity between the structure of this equation and the familiar form of the electrical power P dissipated in a load $P = \frac{1}{2}\mathrm{Re}(IU^*)$. This suggests to introduce a linear relation between the dipole moment and the field. This is analogous to the relation $U = Z_L I$ where Z_L is the impedance of the load in an electrical circuit. Such a relation can be written using the Green tensor[f] that yields the field radiated at $\mathbf{r}_0$ by a dipole located at $\mathbf{r}_0$:

$$\mathbf{E}(\mathbf{r}_0) = \overleftrightarrow{\mathbf{G}}(\mathbf{r}_0, \mathbf{r}_0, \omega) \cdot \mathbf{p} = \frac{\overleftrightarrow{\mathbf{G}}(\mathbf{r}_0, \mathbf{r}_0, \omega)}{-i\omega} \cdot [-i\omega \mathbf{p}]. \tag{2.72}$$

Figure 2.8 shows the dipole near a nanoantenna. The dipole is assumed to be oriented along the z axis so that $\mathbf{p}\overleftrightarrow{\mathbf{G}}^*(\mathbf{r}_0, \mathbf{r}_0, \omega)\mathbf{p}^* = |\mathbf{p}|^2 G^*_{zz}(\mathbf{r}_0, \mathbf{r}_0, \omega)$. The energy transfered by the dipole to the field can be written in the form $P_o = \frac{1}{2}\mathrm{Im}(\frac{G_{zz}(\mathbf{r}_0, \mathbf{r}_0, \omega)}{\omega})|\omega \mathbf{p}|^2$. This has

[f]Note that the definition used here follows Ref. 42 and differs by a factor $\mu_0\omega^2$ from the definition used before in the chapter.

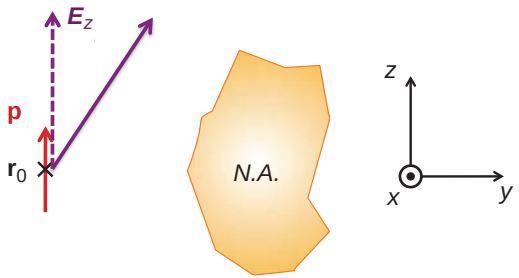


Fig. 2.8. **Sketch of the system**. N.A. stands for Nanoantenna. A dipole is located at $\mathbf{r}_0$ near the nanoantenna and the dipole moment is oriented along the z-axis. E_z is the z-component of the electric field $\mathbf{E}$ at $\mathbf{r}_0$.

the structure of the power $\frac{1}{2}\mathrm{Re}(Z_L)|I|^2$ dissipated in a load Z_L. A comparison between the two forms of the power delivered by a source suggests the following identification:

$$\begin{aligned} I &\leftrightarrow -i\omega p_z, \\ U &\leftrightarrow -E_z(\mathbf{r}_0), \\ Z &\leftrightarrow \frac{-iG_{zz}(\mathbf{r}_0, \mathbf{r}_0, \omega)}{\omega}, \end{aligned} \tag{2.73}$$

where we have $U = ZI$.[g] As for lumped elements, losses are given by the real part R of the impedance $Z = R + iY$. The resistive part of the impedance is thus $\mathrm{Im}(G_{zz})/\omega$ and accounts for both radiative losses and non-radiative losses. It is important to emphasize that this impedance has the dimension of $\Omega \cdot \mathrm{m}^{-2}$. It is thus an impedance per unit area. The difference stems from the fact that we use $-i\omega p_z$ (in Am) instead of I (in A) and E_z (in V/m) instead of U (in V). We shall therefore use the term *specific* impedance. We note that this approach is similar to the so-called emf method introduced by Brillouin[44] in 1922 to derive the input impedance in the case of a wire antenna.

[g]Note that the signs — are due to the relations $\mathbf{j} = \frac{\partial p}{\partial t} = -i\omega\mathbf{p}$ and $\mathbf{E} = -\nabla\phi$ where ϕ is a scalar potential.

2.9.2. *Two-level system impedance*

Defining an internal specific impedance for the source amounts to finding a linear relation between the induced dipole moment $\mathbf{p}^{\text{ind}}$ and the electric field. Far from the saturation regime, such a linear relation is simply given by the polarisability of a two-level system: $\mathbf{p}^{\text{ind}} = \alpha\varepsilon_0\mathbf{E}$. We can thus write $\mathbf{E} = Z_{\text{int}}(-i\omega\mathbf{p}^{\text{ind}})$, where we have introduced the source scalar specific impedance.

$$Z_{\text{int}} = \frac{i}{\omega\alpha\epsilon_0}. \tag{2.74}$$

Let us consider a two-level system for the sake of illustration. The density matrix formalism allows deriving the particular form of the polarisability for such a system. The transition frequency ω_0 is defined by $\hbar\omega_0 = E_2 - E_1$ (from a classical point of view, ω_0 is the bare frequency of the dipole decoupled from all fields) where the atomic Hamiltonian eigenvalues are denoted E_1 and E_2 respectively. The polarisability of the system can be written in the form[19]:

$$\alpha(\omega) = \frac{\alpha_0}{\omega_0^2 - \omega^2 - i\gamma_0\omega} \tag{2.75}$$

where α_0 can be written using the oscillator strength f as $\alpha_0\epsilon_0 = (e^2/m)f$ or $\alpha_0\epsilon_0 = 2d_{12}^2\omega_0/\hbar$, with $d_{12} = |p|$ the dipole moment of the transition. The term γ_0 accounts for the broadening of the resonance. It consists of several contributions. First of all, the radiative emission is characterized by a contribution γ_R. It corresponds to a population decay of level 2. Coupling to the environment can also provide inelastic interactions leading to a population decay of the excited state. It corresponds to a non-radiative decay and is included in the term $\gamma_{NR} = 1/T_1$. Finally, elastic collisions produce a dephasing of the wave function without modifying the population of the excited state. This contribution to the resonance broadening is called dephasing and characterized by a contribution $\gamma_2^* = 2/T_2^*$ so that we have

$\gamma_0 = \gamma_R + \gamma_{NR} + \gamma_2^*$. Finally, the impedance can be cast in the form:

$$Z_{\text{int}} = \frac{\gamma_0}{\alpha_0\epsilon_0} + i\frac{\omega_0^2}{\alpha_0\epsilon_0}\frac{1}{\omega} - i\omega\frac{1}{\alpha_0\epsilon_0} \tag{2.76}$$

$$= R - \frac{1}{iC\omega} - iL\omega \tag{2.77}$$

This impedance has the structure of the impedance of a RLC series circuit where $R = \gamma_0/\alpha_0\epsilon_0$, $L = 1/\alpha_0\epsilon_0$ and $C = \alpha_0\epsilon_0/\omega_0^2$. Let us consider the resistance here. It is proportional to the sum $\gamma_R + \gamma_{NR} + \gamma_2^*$. The first term corresponds to spontaneous emission in vacuum. The corresponding resistance is denoted R_0. The other terms account for non-radiative interactions with the environment. From the impedance point of view, it is easier to consider the radiative resistance as an external load, and to write the impedance of the source using only $\gamma = \gamma_{NR} + \gamma_2^*$ instead of γ_0. We will now write $R_{\text{int}} = \gamma/\alpha_0\epsilon_0$, L_{int} and C_{int} the lumped elements of the source.

2.9.3. *Conjugate impedance matching condition*

We consider here the absorption of energy coming from an external source. In electricity, it is well known that a conjugate impedance matching condition must be fulfilled to optimize absorption in a load impedance. Here, we will apply the concept of impedance to analyse the absorption by a nanoparticle or an atom when located in a given environment. This analysis will provide a guideline to design a structure that enhances the absorption of a nanoparticle or a two-level system.

The nanoparticle (or the atom) has a specific impedance Z_1 and the environment acts as a load Z_2. The system is illuminated by an external field $\mathbf{E}_{\text{ext}}$. The amplitude of the dipole moment of the system, illuminated by a field that is the sum of the incident field and the field scattered by the environment, is given by $p_z = \alpha\epsilon_0[E_{z,\text{ext}} + G_{zz}p_z]$ where α is the nanoparticle (or the atom) polarisability and G_{zz} is the Green tensor that accounts only for the environment. This equation can be reformulated as $E_{z,\text{ext}} = \left[\frac{i}{\omega\alpha\epsilon_0} + \frac{-iG_{zz}}{\omega}\right](-i\omega p_z) = [Z_1 + Z_2](-i\omega p_z)$. We now sketch the equivalent circuit in Fig. 2.9.

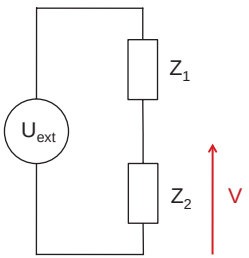


Fig. 2.9. **Equivalent circuit of a nanoparticle in a given environment.** An external field is represented by a source voltage. The potential difference for the load Z_2 is V.

The illuminating field is represented by a voltage U_{ext} applied to Z_1 and Z_2 in series. The potentiel difference for the load is V. We recognize a potential divider so that $V = Z_2/(Z_2 + Z_1)U_{\text{ext}}$. It is now a simple matter to derive the condition for maximum power dissipated in the load. Let us write $Z_1 = R_1 + iY_1$ and $Z_2 = R_2 + iY_2$. The power dissipated in the environment (nanoantenna for instance), due to both radiative and ohmic losses, is given by:

$$P_0 = \frac{1}{2}\frac{R_2|U_{\text{ext}}|^2}{(R_1 + R_2)^2 + (Y_1 + Y_2)^2} \tag{2.78}$$

where R_1 and R_2 are always positive, respectively accounting for the losses of the source and the local density of states at the position of the source. Y_1 and Y_2 can be either positive or negative depending on the frequency. The maximum power dissipated by the nanoantenna is obtained when $Y_1 + Y_2 = 0$ and $R_1 = R_2$. This condition is the usual conjugate impedance matching condition $Z_1 = Z_2^*$ that ensures an optimized power transfer between a source and a load.

References

1. Wessel, J. (1985). Surface-enhanced optical microscopy, *J. Opt. Soc. Am. B* **2**, p. 1538.
2. Pohl, D. W. (1999). Near field optics seen as an antenna problem in *Near-Field Optics: Principles and Applications*, (World Scientific, Singapore).

3. Grober, R. D., Schoelkopf, R. J. and Prober, D. E. (1997). Optical antenna: Towards a unity efficiency near-field optical probe, *Appl. Phys. Lett.* **70**, p. 1354.
4. Azoulay, J., Débarre, A., Richard, A. and Tchénio, P. (2000). Quenching and enhancement of single-molecule fluorescence under metallic and dielectric tips, *Europhys. Lett.* **51**, p. 374.
5. Greffet, J. J. (2005). Nanoantennas for light emission, *Science* **308**, p. 1561.
6. Thomas, M., Greffet, J.-J., Carminati, R. and Arias-Gonzalez, J. R. (2004). Single molecule spontaneous emission close to absorbing nanostructures, *Appl. Phys. Lett.* **85**, p. 3863.
7. Mühlschlegel, P., Eisler, H. J., Martin, O. J., Hecht, B. and Pohl, D. W. (2005). Resonant Optical Antennas, *Science* **308**, p. 1607.
8. Anger, P., Bharadwaj, P. and Novotny, L. (2006). Enhancement and quenching of single-molecule fluorescence, *Phys. Rev. Lett.* **96**, p. 113002.
9. Khn, S., Håkanson, U., Rogobete, L. and Sandoghdar, V. (2006). Enhancement of single-molecule fluorescence using a gold nanoparticle as an optical nanoantenna, *Phys. Rev. Lett.* **97**, p. 017402.
10. Curto, A., Volpe, G., Taminiau, T. H., Kreuzer, M. P., Quidant van, R. and van Hulst, N. F. (2010). Unidirectional emission of a quantum dot coupled to a nanoantenna, *Science* **329**, p. 930.
11. Bharadwaj, P., Deutsch, B. and Novotny, L. (2009). Optical Antennas, *Advances in Optics and Photonics* **1**, p. 438.
12. Biagioni, P., Huang, J. S. and Hecht, B. (2012). Nanoantennas for visible and infrared radiation, *Rep. Prog. Phys.* **75**, p. 024402.
13. Alú, A., Agio, M. Ed. (2013). *Optical Antennas*, (Cambridge University Press, Cambridge).
14. Staelin, D. H., Morgenthaler, A. W. and Kong, J. (1994). *Electromagnetic Waves*. (Prentice Hall International Editions, Englewood Cliffs).
15. Mandel, L. and Wolf, E. (1995). *Optical Coherence and Quantum Optics*. (Cambridge University Press, Cambridge).
16. Van Bladel, J. (1991). *Singular Electromagnetic Fields and Sources*. (IEEE Press, Piscataway).
17. Novotny, L. and Hecht, B. (2006). *Principles of Nano-Optics*. (Cambridge University Press, Cambridge).
18. Hoeppe, U., Wolff, C., Küchenmeister, J., Niegemann, J., Drescher, M., Benner, H. and Busch, K. (2012). Direct observation of non-markovian radiation dynamics in 3D bulk photonic crystals, *Phys. Rev. Lett.* **108**, p. 043603.
19. Loudon, R. (2000). *The Quantum Theory of Light*. (Oxford University Press, Oxford).
20. Joulain, K., Carminati, R., Mulet, J. P. and Greffet, J. J. (2003). Definition and measurement of the local density of electromagnetic states close to an interface, *Phys. Rev. B.* **68**, p. 245405.
21. Drexhage, K. H., Kuhn, H. and Schäfer, F. P. (1968). Ber. Bunsenges. *Phys. Chem.* **72**, p. 329.

22. Chance, R. R., Prock, A. and Silbey, R. (1974). *J. Chem. Phys.* **60**, p. 2744.
23. Lukosz, W. and Kunz, R. E. (1977). Fluorescence lifetime of magnetic and electric dipoles near a dielectric interface, *Opt. Commun.* **20**, p. 195.
24. Ford, G. W. and Weber, W. H. (1984). *Phys. Rep.* **113**, p. 195.
25. Carminati, R., Greffet, J. J., Henkel, C. and Vigoureux, J. M. (2006). Radiative and Non-radiative decay of a single molecule close to a metallic nanoparticle *Opt. Commun.* **261**, p. 368.
26. Sauvan, C., Hugonin, J. P., Maksymov, I. S. and Lalanne, P. (2013). Theory of the spontaneous optical emission of nanosize photonic and plasmon resonators, *Phys. Rev. Lett.* **110**, p. 237401.
27. Törmä, P. and Barnes, W. L. (2015). Strong coupling between surface plasmon polaritons and emitters: A review, *Rep. Prog. Phys.* **78**, p. 013901.
28. Santosh, K., Bitton, O., Chuntonov, L. and Haran, G. (2016). Vacuum Rabi splitting in a plasmonic cavity at the single quantum emitter limit, *Nat. Commun.* **7**, p. 11823.
29. Chikkaraddy, R., de Nijs, B., Benz, F., Barrow, S. J., Scherman, O. A., Rosta, E., Demetriadou, A., Fox, P., Hess, O. and Baumberg, J. J. (2016). Single-molecule strong coupling at room temperature in plasmonic nanocavities, *Nature* **535**, p. 127.
30. Esteban, R., Laroche, M. and Greffet, J. J. (2009). Influence of metallic nanoparticles on upconversion processes, *J. Appl. Phys.* **105**, p. 033107.
31. Ringler, M., Schwemer, A., Wunderlich, M., Nichtl, A., Krzinger, K., Klar, T. A. and Feldmann, J. (2008). Shaping emission spectra of fluorescent molecules with single plasmonic nanoresonators, *Phys. Rev. Lett.* **100**, p. 203002.
32. Khurgin, J. B., Sun, G. and Soref, R. A. (2008). Electroluminescence efficiency enhancement using metal nanoparticles, *Appl. Phys. Lett.* **93**, p. 021120.
33. Crozier, K. B., Sundaramurthy, A., Kino, G. S. and Quate, C. F. (2003). Optical antennas: Resonators for local field enhancement, *J. Appl. Phys.* **94**, p. 4632.
34. Chen, X. W., Agio, M. and Sandoghdar, V. (2012). metallodielectric hybrid antennas for ultrastrong enhancement of spontaneous emission, *Phys. Rev. Lett.* **108**, p. 233001.
35. Yang, J., Faggiani, R. and Lalanne, P. (2016). Light emission in nanogaps: Overcoming quenching, *Nanoscale Horizons* **1**, p. 11.
36. Bigourdan, F., Marquier, F., Hugonin, J. P. and Greffet, J. J. (2014). Design of Highly efficient metallo-delectric patch antennas for single photon emission, *Opt. Express* **22**, p. 2337.
37. Bigourdan, F., Hugonin, J. P., Marquier, F., Sauvan, C. and Greffet, J. J. (2016). Nanoantenna for electrical generation of surface plasmon polaritons, *Phys. Rev. Lett.* **116**, p. 108803.
38. Akselrod, G. M., Argyropoulos, C., Hoang, T. B., Ciraci, C., Fang, C., Huang, J., Smith, D. R. and Mikkelsen, M. H. (2014). Probing the mechanisms of

large Purcell enhancement in plasmonic nanoantennas, *Nature Photon.* **8**, p. 840.
39. Savage, K., Hawkeye, M. M., Esteban, R., Borisov, A. G., Aizpurua, J. and Baumberg, J. (2012). Revealing the quantum regime in tunneling plasmonics, *Nature* **491**, p. 574.
40. Esteban, R., Borisov, A., Nordlander, P. and Aizpurua, J. (2012). Bridging quantum and classical plasmonics with a quantum-corrected model, *Nature Com.* **3**, p. 825.
41. Scholl, J. A., García-Etxarri, A., Leen, A. and Koh, J. A. (2012). Dionne, Observation of quantum tunneling between two plasmonics nanoparticles *Nano Lett.* **13**, p. 564.
42. Greffet, J. J., Laroche, M. and Marquier, F. (2010). Impedance of a nanoantenna and a single quantum emitter, *Phys. Rev. Lett.* **105**, p. 117701.
43. Marquier, F. and Greffet, J. J. (2013). Impedance of a nanoantenna, in *Optical Antennas*. Ed. Alú, A. and Agio, M. (Cambridge University Press, Cambridge).
44. Brillouin, L. (1922). Sur l'origine de la résistance de rayonnement, *Radioélectricité* **3**, p. 147.

CHAPTER 3

Nonlocality in Plasmonics

N. ASGER MORTENSEN*

University of Southern Denmark, Denmark

JACOB B. KHURGIN

Johns Hopkins University, Baltimore, USA

MARTIJN WUBS

Technical University of Denmark, Denmark

3.1. Introduction

The field of plasmonics has developed with a solid foundation in classical electrodynamics employing semiclassical descriptions of the interactions of light with matter.[1–3] In particular, our conceptual understanding of plasmons — the collective oscillations of the free conduction electrons subject to driving optical fields — largely relies on analysis of problems within Drude theory and the local-response approximation (LRA). By assumption, the material response occurs only in the point of space of the perturbation, while there is no response at even short distances. For dielectrics, this is of course a well-established and accurate approach all the way down to atomic length scales [see panel (c) in Fig. 3.1], while it is traditionally

*Corresponding author: namo@mci.sdu.dk

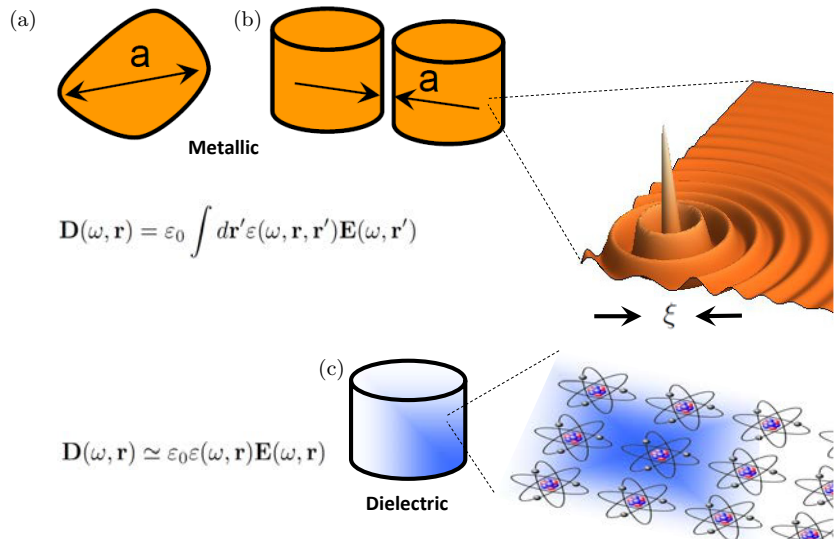


Fig. 3.1. Examples of metallic and dielectic nanostructures, indicating characteristic length scales a for (a) isolated metal particles and (b) metal-particle dimers with gaps. While the local-response approximation is typically adequate for dielectric structures as in (c), the nonlocal correction becomes important for metallic nanostructures in the mesoscopic regime where $a \longrightarrow \xi$, with ξ being an intrinsic length scale associated with the finite range of nonlocal dynamics of the electron gas. In dielectrics, there are no free conduction electrons and the local-response approximation remains accurate all the way down to the atomic length scale.

being adapted for metals too, including for plasmonic nanostructures [see panels (a) and (b) in Fig. 3.1]. Admitting its simplifications, the LRA framework has nevertheless fostered both striking predictions and novel experimental confirmations of a zoo of plasmonic phenomena, such as the squeezing of light beyond the diffraction limit,[4,5] the tunability of the optical properties of metallic structures with size and shape,[6] and large enhancement of the electric field in metal nanoparticles of close proximity[7] as well as in metal geometries with sharp surface corrugation.[8] Basically, all the fascinating aspects of plasmonics have been driven forward by a basic understanding deeply rooted in the LRA approach!

In dealing with light-matter interactions we in a first place commonly rely on linear-response theory. For insulating materials the

further simplification associated with the LRA is usually excellent, while the success of the LRA in nanoplasmonics is perhaps more intriguing. As an example, novel experimental explorations of gold bow-tie dimers with few-nanometer gaps were found to be *classical down to the nanometer scale*.[9] On the other hand, such structures are in general considered potential candidates for the observation of nonlocal response associated with spatial dispersion, which is potentially manifesting when light interacts with the free conduction electrons in metals.[10] So, one could rightfully ask what makes the LRA so reasonable for many nanoplasmonic problems and in which situations can we expect to find appreciable corrections?

As a first reflection on this, we note that the underlying quantum dynamics of the electron gas manifests itself when approaching intrinsic length scales of the metal. Here, the Fermi wavelength λ_F is a key parameter of the electron gas and for most metals of interest to plasmonics it is in the nanometer-to-Ångström regime.[11] This largely explains the success of the LRA and the application of Drude theory to plasmonics even in nanoscale metallic structures. However, this also anticipates the existence of an intriguing mesoscopic size regime hosting electrodynamics with a need for corrections of the LRA description. In fact, there is now a general understanding in the community that the neglect of nonlocal effects is the foundation for field-singularities in the LRA response, e.g. for arbitrarily sharp changes in the metal-surface topography or in dimers with vanishing gaps.[12]

Nonlocal response and spatial dispersion has a long history,[13–15] while developments in the exploration of yet smaller plasmonic structures have stimulated a growing interest and more recent revival of the field.[16–19] Most importantly, theory developments are now stimulated by amazing experimental explorations of plasmons in true nanoscale structures.[20–29] Figure 3.1 shows schematic representations of archetype mesoscopic geometries with competing extrinsic and intrinsic length scales, such as characteristic geometrical dimensions a and the finite range ξ of nonlocal response. In order to address such problems, we will discuss developments of a real-space

formulation[19,30] and numerical implementations[31] of nonlocal hydrodynamic theory. We will also in more detail discuss a more recent extension to a generalized account of drift-diffusion dynamics (the so-called GNOR theory, Generalized Nonlocal Optical Response)[32] along with discussions of Landau damping[33] in connection to nonlocal response.[34–36]

3.2. Linear-Response Nonlocal Theory

In order to simplify our discussion of nonlocal response, we will limit ourselves to the linear-response regime with scalar fields, while later offering a vectorial description. We here consider a generic system that we perturb by a field $\mathcal{F}$, which is in turn causing a response $\mathcal{Z}$. As an example of mechanical system, we recall Hooke's law where $\mathcal{F}$ would be a classical force field, while $\mathcal{Z}$ would then be the associated displacement in position of the mechanical system. When we later turn our discussion to plasmonics, the electrons constituting the plasma are driven out of equilibrium by an electrical field $\mathbf{E}$ and the associated response of the electron gas is given in terms of an induced current density $\mathbf{J}$. Note that alternatively the response could be formulated in terms of the displacement field $\mathbf{D}$. Naturally, $\mathcal{Z}$ should be considered a functional of $\mathcal{F}$, written as $\mathcal{Z}[\mathcal{F}]$. We emphasize that in general the exact functional could be of a complicated nonlinear nature, obviously depending on the detailed dynamics of the underlying microscopic system. However, being interested in only the linear response, we may proceed without such detailed insight in the dynamics. To support a more phenomenological path, we utilize a series expansion of $\mathcal{Z}[\mathcal{F}]$. For our case where fields depend on both space and time, this is known as the Volterra series. Thus, to linear order in the perturbing field we get

$$\mathcal{Z}(t,\mathbf{r}) = \int_{-\infty}^{t} dt' \int d\mathbf{r}'\, \chi(t-t',\mathbf{r},\mathbf{r}')\mathcal{F}(t',\mathbf{r}') \tag{3.1}$$

where the expansion coefficient χ is our linear-response susceptibility of the system. Spatial coordinates are given by $\mathbf{r}$ and $\mathbf{r}'$, while t and t' denote different time instances. Note that the expansion

itself does not help us to determine χ. For this, we need some other physical input, e.g. from microscopic or semiclassical theory or from linear-response experiments. The principle of causality implies that there is only a response at times following the action, i.e. $t > t'$. Note that the linear response function $\chi(t-t', \mathbf{r}, \mathbf{r}')$ contains both temporal memory effects and spatial nonlocal effects. We may conveniently turn to Fourier space, in which case the temporal response is associated with frequency dispersion (ω-dependence) while nonlocal response gives rise to spatial dispersion (k-dependence) in the response function. Furthermore, for a homogeneous medium, the translational invariance implies that $\chi(t-t', \mathbf{r}, \mathbf{r}') = \chi(t-t', \mathbf{r}-\mathbf{r}')$. This consequently turns both the temporal and the spatial integrals in Eq. (3.1) into convolutions, with the immediate consequence that in frequency-momentum space the linear response becomes a simple product, i.e.

$$\mathcal{Z}(\omega, k) = \chi(\omega, k)\mathcal{F}(\omega, k). \tag{3.2}$$

In this way, the synonymous reference to *nonlocal response* and *spatial dispersion* is apparent. In this chapter we focus on finite plasmonic nanostructures, i.e. with broken translational invariance (see Fig. 3.1). Consequently we will focus on the real-space notation, and emphasize the terminology of *nonlocal response*. We will pursue a phenomenological approach[37] that explores

$$\mathcal{Z}(\omega, \mathbf{r}) = \int d\mathbf{r}'\, \chi(\omega, \mathbf{r}, \mathbf{r}')\mathcal{F}(\omega, \mathbf{r}') \tag{3.3}$$

in the situation where nonlocal response is considered a perturbative correction to the common local-response approximation associated with a homogeneous system. Mathematically, this means that

$$\chi(\omega, \mathbf{r}, \mathbf{r}') = \chi_{\mathrm{LRA}}(\omega)\delta(\mathbf{r}-\mathbf{r}') + \delta\chi(\omega, \mathbf{r}-\mathbf{r}'). \tag{3.4}$$

where $\delta\chi$ is a nonlocal response function, being of a short range nature. If the system is furthermore isotropic (as it is the case for the free-electron gas), then $\delta\chi$ is symmetric, i.e. $\delta\chi(\omega, \mathbf{r}-\mathbf{r}') = \delta\chi(\omega, |\mathbf{r}-\mathbf{r}'|)$.

If the perturbing field $\mathcal{F}$ varies slowly on the scale of $\delta\chi$, we may conveniently proceed with a Taylor expansion of $\mathcal{F}$ around the point $\mathbf{r}$,

$$\mathcal{F}(\mathbf{r}') \simeq \mathcal{F}(\mathbf{r}) + [\nabla\mathcal{F}(\mathbf{r})] \cdot (\mathbf{r}' - \mathbf{r}) + \frac{1}{2}(\mathbf{r}' - \mathbf{r})^T \cdot [H\mathcal{F}(\mathbf{r})] \cdot (\mathbf{r}' - \mathbf{r}) + \cdots \tag{3.5}$$

Here, H is the Hessian matrix with elements $H_{ij} = \partial^2/(\partial_i\partial_j)$ where $i, j = x, y, z$. We now substitute this expansion into Eq. (3.3), thus allowing us to perform the integration over $\mathbf{r}'$ term by term to any desired order in our Taylor expansion. In fact, all we now need to do is to calculate different moments of the function $\delta\chi$. By assumption, the zero-order moment satisfies

$$\int dr\, \delta\chi(\omega, r) \ll \chi_{\text{LRA}}(\omega), \tag{3.6}$$

while the first moment vanishes for symmetry reasons (the integral is over the product of an odd and an even function),

$$\int dr\, r\delta\chi(\omega, r) = 0. \tag{3.7}$$

On the other hand, the second moment is finite (the integral is over an even function),

$$\int dr\, r^2\delta\chi(\omega, r) \equiv 2\xi^2, \tag{3.8}$$

where we have introduced the phenomenological parameter ξ as the characteristic range of the nonlocal response function. Note that the factor of 2 in this definition of ξ is only introduced for the later convenience of canceling the factor 1/2! appearing in the second-order term of the Taylor expansion in Eq. (3.5).

To appreciate the importance of these moments, we now return to Eq. (3.3). For simplicity, we absorb the zeroth-order moment into our definition of the local term $\chi_{\text{LRA}}(\omega)$. As the first-order terms vanish, the leading order-correction to the constitutive relation is

now written as

$$\mathcal{Z}(\omega, \mathbf{r}) \simeq \left[\chi_{\mathrm{LRA}}(\omega) + \xi^2 \nabla^2\right] \mathcal{F}(\omega, \mathbf{r}). \tag{3.9}$$

We emphasize that our consideration based on a series expansion does not in itself help us to determine the nonlocal length scale ξ. For that we will need further physical input. As we shall see, for plasmonics this ξ is roughly the distance that an electron can travel during the time of an optical cycle. Also, note that the occurrence of the Laplacian is a consequence of our scalar treatment. In Section 3.3 we relax this assumption and offer a vectorial description, as a result of which contributions from other second-order spatial derivatives will also arise.

At this stage we briefly discuss the immediate implications of the nonlocal correction to the constitutive relation. First of all, Eq. (3.9) is entirely generic for any scalar problem, with the nonlocal response always manifesting itself through a Laplacian correction term. Quite notably, this Laplacian form of the correction holds irrespectively of the microscopic or semiclassical origin of nonlocality. In turn, this also implies the possibility of several coexisting nonlocal mechanisms that would simply add up to form an effective nonlocal length, i.e.

$$\xi^2 = \sum_i \xi_i^2. \tag{3.10}$$

This is a point already hinted to by Landau and Lifshitz[10] and we will provide an example of this below. We also highlight a most remarkable property of Eq. (3.9) when compared to Eq. (3.3): Stating only very few generic assumptions we have conveniently transformed the challenging nonlocal two-point integral relation into a more attractive constitutive law that is mathematically of a local-response form, i.e. only involving a single spatial coordinate. The practical consequences are immediate: if we can handle an additional Laplacian term (indeed, such differential operators are already present in Maxwell's wave equation), then we have conceptually returned to a formulation that is already routinely explored in local-response computational photonics and plasmonics.[38] Thus, with few almost pedestrian

manipulations we have eased both our conceptual understanding and computational explorations of nonlocal effects in nanoplasmonic systems.

Let us next turn to a discussion of the additional physics brought about by the Laplacian term. Laplacian terms are well known from diffusion problems and here the Laplacian term likewise causes a spatial smearing of fields (but not only due to diffusion, as shown below). Thus, we anticipate its importance for field-enhancement phenomena, where it will serve to regularize field singularities associated with the common local-response contribution in Eq. (3.9). This motivates us to also emphasize a more practical aspect of the Laplacian smearing. In computational hydrodynamics based on numerical solutions of the Navier–Stokes equation it is a common pragmatic practice to introduce a small artificial diffusion as a useful and even essential move to stabilize numerics.[39] For computational plasmonics, the Laplacian smearing actually benefits numerical convergence in a similar way. The presence of the smearing term in the wave equation facilitates electrodynamic simulations of metal structures with even arbitrarily sharp features in the surface topography, without suffering the usual convergence issues due to the underlying singular response inherent to the local-response approximation.[31]

Above, we have interchangeably used the terms *nonlocal response* and *spatial dispersion* and we now for completeness provide an explicit link between the two terminologies. When Fourier transforming Eq. (3.9) and relating it to Eq. (3.2), we see immediately that $\chi(\omega, k) \simeq \chi_{\mathrm{LRA}}(\omega) - (k\xi)^2$, i.e. spatial dispersion with a leading quadratic k-correction for $k\xi \ll 1$.

Having now treated the linear-response nonlocal theory in a most general way, we next turn to the electrodynamics associated with plasmonics.

3.3. Linear-Response Electrodynamics

Theoretical modeling of light-matter interactions in plasmonic nanostructures is commonly relying on the macroscopic Maxwell

equations[1] with the optical response of metals described through constitutive relations, which relate the response of the material to the applied field. The common approach is

$$\mathbf{D}(\omega, \mathbf{r}) \simeq \varepsilon_0 \varepsilon(\omega, \mathbf{r}) \mathbf{E}(\omega, \mathbf{r}) \tag{3.11}$$

where the spatially local relationship between $\mathbf{D}$ and $\mathbf{E}$ is almost taken for granted. Of course, this LRA is well motivated for dielectric structures with the polarization of electronic states highly localized to the individual atoms, see Fig. 3.1(c), while the approximation is perhaps less obvious for materials such as metals with also delocalized conduction electrons. Given our more general discussion above, see Eq. (3.3), the displacement field $\mathbf{D}$ caused in response to a perturbing electric field $\mathbf{E}$ is given by

$$\mathbf{D}(\omega, \mathbf{r}) = \varepsilon_0 \int d\mathbf{r}' \varepsilon(\omega, \mathbf{r}, \mathbf{r}') \mathbf{E}(\omega, \mathbf{r}'). \tag{3.12}$$

In passing, we note that the LRA in Eq. (3.11) follows from Eq. (3.12) by assuming $\varepsilon(\omega, \mathbf{r}, \mathbf{r}') \simeq \varepsilon(\omega, \mathbf{r}) \delta(\mathbf{r} - \mathbf{r}')$. Combining Eq. (3.12) with Maxwell's equations we arrive at an integro-differential wave equation

$$\nabla \times \nabla \times \mathbf{E}(\omega, \mathbf{r}) = \left(\frac{\omega}{c}\right)^2 \int d\mathbf{r}' \varepsilon(\omega, \mathbf{r}, \mathbf{r}') \mathbf{E}(\omega, \mathbf{r}') \tag{3.13}$$

which constitutes an eigenvalue problem. While this formulation clearly conveys the nonlocal aspect of the problem, the equation is perhaps not too appealing and this goes for both numerical implementations as well as for further analytical efforts. In order to allow any progress we would as a first prerequisite need the actual response function $\varepsilon(\omega, \mathbf{r}, \mathbf{r}')$. Thus, there is a call for some kind of microscopic or semiclassical description. A more phenomenological avenue is to assume a simple short-range function as a modification to the usual Drude delta-function response.[40] Instead, adapting the approach that lead us to Eq. (3.9), we here transform the integro-differential equation into a more attractive regular partial-differential equation (PDE). Obviously, we should now turn to a vectorial treatment and

with only few additional notational complications of vector-field calculus, we arrive at the conclusion as for the scalar considerations given above. In fact, within the nonlocal hydrodynamic model we get[30,32,41]

$$\nabla \times \nabla \times \mathbf{E}(\omega, \mathbf{r}) = \left(\frac{\omega}{c}\right)^2 \left[\varepsilon_{\mathrm{LRA}}(\omega) + \xi^2 \nabla(\nabla\cdot)\right] \mathbf{E}(\omega, \mathbf{r}). \qquad (3.14)$$

In comparison to Eq. (3.9), we note that the nonlocal correction term is qualitatively of a Laplacian form, as discussed above for scalar fields. To appreciate this, we first quote the relation $\nabla \times \nabla \times \mathcal{F} = \nabla(\nabla\cdot)\mathcal{F} - \nabla^2\mathcal{F}$ that holds for any vectorial field $\mathcal{F}$. Conceptually, we may thus turn the gradient-of-divergence term into a Laplacian term by simultaneously grouping the remaining double-curl operator with the already-existing double-curl operator on the left-hand side of Eq. (3.14). Interestingly, this would only serve to re-normalize the equation and the effect of this re-normalization could even be neglected if $\xi^2\omega^2/c^2 \ll 1$. In fact, this condition is equivalent to $k\xi \ll 1$ or $\xi \ll \lambda$, with $k = 2\pi/\lambda$ being the free-space wave vector. Intuitively, this condition is commonly fulfilled in noble-metal plasmonics. As we shall see in the hydrodynamic model, this is commensurate to the condition $v_F \ll c$. Nevertheless, we will proceed with the explicit structure in Eq. (3.14) since the explicit gradient-of-divergence form will respect the vectorial nature of the problem. In particular, we hereby address transverse and longitudinal field components differently[30] which is key to accurately account for both transverse resonances having a frequency below ω_p and longitudinal resonances occurring above ω_p, with ω_p being the plasma frequency.[19,42]

3.4. Hydrodynamic Equation of Motion: Drift-diffusion Theory

With Eq. (3.10) we have already speculated on the possibility of several underlying physical mechanisms co-existing and contributing to the overall nonlocal response. In this section we offer a semiclassical theory which underlines this possibility, by including two distinct

contributions to ξ from celebrated classical transport mechanisms, i.e. convection and diffusion of charge carriers. The detailed derivation starts from a classical hydrodynamic-diffusion problem of an electron responding to an external electric field. In line with common treatments in classical condensed-matter physics and plasmonics text books[1,11] we apply Newton's second law to the description of an electron subjected to an electrical field. This is the essence of the linearized hydrodynamic equation-of-motion (for simplicity, we suppress spatial and temporal variables)[14,19]

$$\frac{\partial}{\partial t}\mathbf{v} = -\gamma\mathbf{v} + \frac{(-e)}{m}\mathbf{E} - \frac{\beta^2}{n_0}\nabla n_1. \tag{3.15}$$

Here, v is the non-equilibrium velocity correction to the static electron problem, while γ is the phenomenological Drude damping parameter also appearing within LRA. The right-hand side contains an additional semiclassical correction. This pressure term ∇n_1 is classical in spirit, while β is a quantum parameter. In fact, β is really not that mysterious: by dimensional analysis we see that it has dimensions of speed, being thus a characteristic velocity for pressure waves associated with the finite compressibility of the electron gas. In the high-frequency limit ($\omega \gg \gamma$), Thomas–Fermi theory gives $\beta^2 = 3/5v_F^2$, with v_F being the Fermi velocity of the electron gas.[30]

To facilitate the linearization, we may conveniently write the electron density as $n(\mathbf{r},t) = n_0 + n_1(\mathbf{r},t)$. Here, n_0 is the equilibrium density, while n_1 is assumed to be the small ($n_1 \ll n_0$) induced density variation associated with the perturbing E field that drives the system away from its equilibrium density.

Of course, Eq. (3.15) is complemented by the principle of charge conservation and extending our considerations to include both convective and diffusive transport of charge, we invoke the linearized convection-diffusion equation

$$\frac{\partial}{\partial t}\{(-e)n_1\} = D\nabla^2(-e)n_1 - \nabla\cdot\{(-e)n_0\mathbf{v}\} = -\nabla\cdot\mathbf{J} \tag{3.16}$$

with D being the diffusion constant (note that this is not to be confused with the previously discussed displacement field $\mathbf{D}$).

By a straightforward manipulation of Eqs. (3.15) and (3.16), followed by a temporal Fourier transformation, we arrive at a governing equation that explicitly connects the current density $\mathbf{J}$ with the driving electrical field $\mathbf{E}$,

$$\xi^2 \nabla(\nabla \cdot \mathbf{J}) + \mathbf{J} = \sigma_D \mathbf{E}. \tag{3.17}$$

Here, we have introduced the nonlocal parameter

$$\xi^2 = \frac{\beta^2}{\omega(\omega + i\gamma)} + \frac{D}{i\omega} \tag{3.18}$$

and recalling the local-response Ohm's law ($\mathbf{J} \simeq \sigma_D \mathbf{E}$), we have also introduced the usual frequency-dependent Drude conductivity

$$\sigma_D = \frac{i\frac{e^2 n_0}{m}}{\omega + i\gamma} = \frac{i\varepsilon_0 \omega_p^2}{\omega + i\gamma} \tag{3.19}$$

with $\omega_p = \sqrt{\frac{e^2 n_0}{\varepsilon_0 m}}$ being the plasma frequency. Naturally, we immediately recover the LRA in the limit $\xi \longrightarrow 0$ where Eq. (3.17) reduces to Ohm's law.

The generalized constitutive equation, Eq. (3.17), can now be used with the Maxwell equation

$$\nabla \times \nabla \times \mathbf{E} = \left(\frac{\omega}{c}\right)^2 \mathbf{E} + i\omega\mu_0 \mathbf{J} \tag{3.20}$$

where we are for simplicity neglecting interband transitions. We will continue to do so throughout this chapter, while a contribution ε_b due to bound electrons can be included straightforwardly. Obviously, the remaining step is to illustrate that this indeed leads to Eq. (3.14). We might be tempted to substitute Eq. (3.17) into Eq. (3.20) in order to eliminate the current density from the latter. However, the perhaps less obvious procedure is to eliminate the current density from Eq. (3.17) and then use that $\nabla \cdot (\nabla \times \mathcal{F}) = 0$ for any vector

field $\mathcal{F}$. In this way, we finally arrive at[41]

$$\nabla \times \nabla \times \mathbf{E}(\omega, \mathbf{r}) = \left(\frac{\omega}{c}\right)^2 \left[\varepsilon_D + \xi^2 \nabla(\nabla \cdot)\right] \mathbf{E}(\omega, \mathbf{r}) \tag{3.21}$$

where

$$\varepsilon_D = 1 + \frac{i\sigma_D}{\varepsilon_0 \omega} = 1 - \frac{\omega_p^2}{\omega(\omega + i\gamma)} \tag{3.22}$$

is the Drude dielectric function. Eq. (3.21) clearly has the same structure as Eq. (3.14). Consequently, Eq. (3.18) offers the desired semiclassical estimate of the length scale ξ that was first introduced phenomenologically. From Eq. (3.18), we also immediately see how convection and diffusion are playing in concert and how they are adding up to give an effective nonlocal length scale in accordance with Eq. (3.10). Note, that the end result is a complex-valued nonlocal parameter, with convection predominantly contributing to the real part, while diffusion shows up as an imaginary-valued addition. This has immediate consequences for the physics: Diffusion of induced charge over time will degrade the polarization of charge, thereby representing a nonlocal damping mechanism. On the other hand, convection is associated with non-dissipative propagating pressure waves, thereby to a first approximation not causing any additional broadening.

Note that in our derivation of Eq. (3.21) we have deliberately followed a route close in spirit to the common LRA, while carefully maintaining a lowest-order account for nonlocal corrections. Nevertheless, the correction term proportional to ξ^2 can near surfaces (where field derivatives are significant) give rise to pronounced corrections. In fact, the diffusive damping can locally exceed the standard Drude damping proportional to γ. We will later discuss how this diffusive damping amounts to enhanced Landau damping occurring in the near vicinity of any metal surface.

3.5. Boundary Conditions

Naturally, we need to also discuss boundary conditions when solving a PDE in a finite geometry as suggested in Fig. 3.1. For the LRA we have to solve Eq. (3.21) in its $\xi \longrightarrow 0$ version. The boundary condition for $\mathbf{E}$ of course follows as a consequence of Maxwell's equations themselves, i.e. the derivation of the boundary conditions only involves Maxwell's equations and the use of Gauss' and Stokes' theorems. On the other hand, solving the coupled equations for the nonlocal problem, Eqs. (3.20) and (3.17), there are now extra field components to be determined. We clearly need an *additional boundary condition* (ABC) that accounts for the behavior of $\mathbf{J}$ on the boundary of the metal.

We emphasize there is no ambiguity or freedom in the choice of this ABC. Perhaps no need to say, but the ABC should reflect underlying physical assumptions, and not just convenient mathematical or numerical choices. In other words, we should first formulate the physical assumptions and then boundary conditions simply follow from our governing equations. For more detailed accounts, we refer to Refs. 19, 43, 41, 32.

The linearized equation-of-motion in Eq. (3.15) is resting on an inherent assumption of a spatially homogeneous equilibrium electron density n_0, while the induced charge n_1 exhibits the temporal and spatial dynamics, i.e. $n(\mathbf{r}, t) = n_0 + n_1(\mathbf{r}, t)$. In this way, n_0 is assumed constant throughout the metal while it turns abruptly to zero outside the metal surface. This step-like variation unambiguously leads to exactly one required ABC, namely the continuity of the normal component of the free-electron current density $\mathbf{J}$.[19,44] Our implicit assumption of an infinite work function does not allow electrons outside the metal. Consequently, we have $\mathbf{J} = 0$ outside the metal, while being still finite inside the metal. Thus, no electrons move across the metal surface, while they are still free to move parallel to the surface (in this way our hydrodynamics of electrons is different from the more common treatment of hydrodynamic flow assuming a no-slip boundary condition for the fluid velocity). In other words, the normal component of $\mathbf{J}$ is zero at the surface,

which we mathematically express as $\mathbf{n} \cdot \mathbf{J} = 0$, with $\mathbf{n}$ being the normal vector. Here, we have emphasized the physical arguments, but this can also be derived rigorously by applying Gauss' theorem to Eq. (3.16).

We note that in the context of Eq. (3.21), the condition $\mathbf{n} \cdot \mathbf{J} = 0$ implies that $\mathbf{n} \cdot \mathbf{E}$ is only continuous across the boundary in the absence of interband transitions and for vacuum surroundings. However, in general there is a normal-component electric-field discontinuity if the interband contribution on the metal side is not fully compensated by dielectric contributions on the dielectric side of the interface.[43]

Let us now briefly reflect on our physical assumptions. The assumption for the ground state of the electron gas corresponds to an infinite work function, thus suppressing quantum-spill out of electrons beyond the surface of the metal, while we are clearly also neglecting any density variations inside the metal, i.e. confinement effects such as Friedel oscillations. Our pragmatic defense is that despite the limitations, this picture is close in spirit to the LRA that involves an implicit assumption of a spatially uniform equilibrium electron density. We emphasize that this is a reasonable description of noble metals that are commonly employed in plasmonics, while spill-out effects are mainly important for less common cases like sodium.[36,45–48] In passing, we emphasize that there have been attempts of relaxing the assumption of a homogeneous equilibrium density[49–51] and it was recently demonstrated how to include density-gradient corrections to also account for quantum-spill out in a hydrodynamic model.[52–55] We will return to problems with quantum spill-out in our later *ab initio* considerations of Landau scattering near the surface of jellium metals.

3.6. Numerical Solutions

While Eq. (3.21) can be solved analytically in a few special cases, more complex and arbitrarily shaped geometries call for numerical solutions. These solutions can either invoke the coupled-wave formulation [Eqs. (3.20) and (3.17)] or they can rely on the generalized wave

equation [Eq. (3.21)]. Of course, the end result will be the same, while details numerical procedures can be different. In either case, the field of computational nanophotonics has witnessed great progress and there is access to a wide range of numerical methods.[38] In principle the nonlocal problem can be addressed by many approaches ranging from Mie-scattering descriptions applied to cylindrical and spherical geometries[19,57,58] and Fourier-modal methods to periodic systems[59] to finite-element implementations[31,60] and boundary-element methods[61] applicable to arbitrarily shaped metallic geometries.

Mie-scattering descriptions are highly efficient for problems employing nanostructures with e.g. cylindrical or spherical symmetry, and Chulhai and Jensen have released a Mie-based tool to simulate the fields for spherical dimers using the discussed nonlocal model.[62]

The finite-element method constitutes another powerful approach that can handle more arbitrarily shaped geometries. Here, we briefly discuss our finite-element implementation.[31] For our numerical solution of the system of equations [Eqs. (3.20) and (3.17)] we take advantage of a commercially available code dedicated to solving partial differential equations based on the finite-element method (FEM). We rely on a weak-form implementation[63] that allows us to draw on built-in routines for electromagnetic scattering and the code is also offering built-in meshing and mesh-refinement routines for arbitrarily shaped geometries. Our code is made freely available[64] and works as an add-on to COMSOL Multiphysics 4.1. The performance of the implementation has been supported by numerous examples and it has been documented in detail. In particular, convergence tests and rigorous benchmarking have been performed for geometries where semi-analytical accounts, such as Mie expansions, allow for solutions with arbitrary numerical accuracy.[56]

Figure 3.2 illustrates an example of the scattering of an incident plane wave from a metallic nanowire where the cross section has a nontrivial triangular shape. Despite the three-fold cylindrical rotational symmetry of the equilateral triangular cross section, the structure remains too complex for any significant analytical progress.

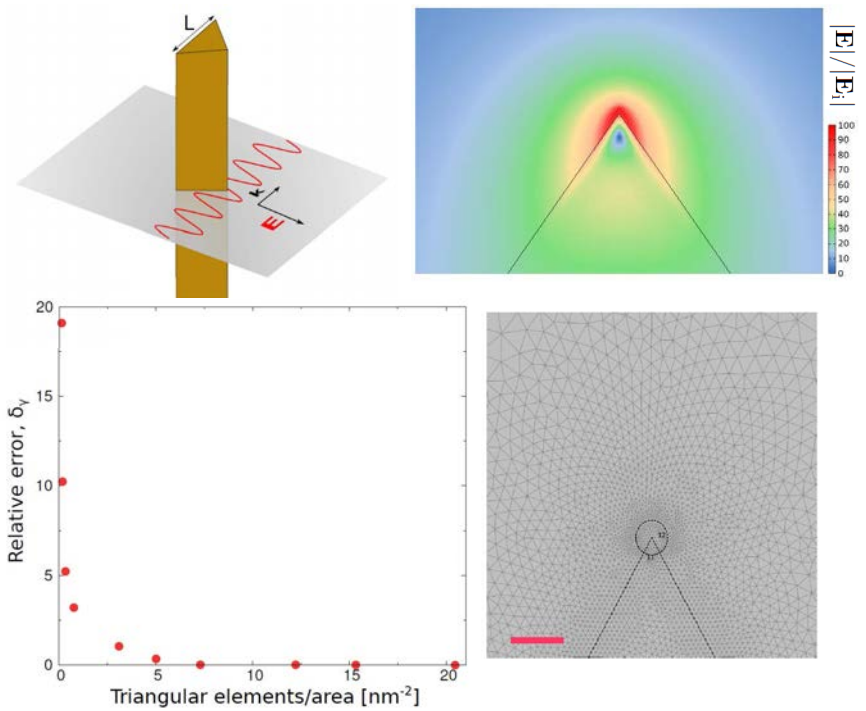


Fig. 3.2. Example of the use of the finite-element method for a numerical solution of the nonlocal wave equation, Eq. (3.21), in a complex geometry with arbitrarily sharp corners. The use of a finite-element mesh allows to discretize a triangular cross with a refinement of the mesh density in regions with abrupt changes in surface topography. While the LRA treatment of such a problem will not converge, the nonlocal smearing facilitates numerical convergence upon appropriate mesh refinement near the corners. Courtesy of Giuseppe Toscano.[56]

As such, the geometry serves our purpose of illustrating the capabilities of our finite-element based numerical approach. Rather than introducing some arbitrary rounding of corners, which is critical to circumvent field divergence in any LRA numerical treatment of this geometry, we straightforwardly account for the arbitrarily sharp corners. No matter the mesh refinement, LRA treatments would exhibit no convergence in the fields due to the singular nature of the problem. However, for the nonlocal treatment, the presence of the nonlocal

length scale ξ changes the situation completely and the problem converges with a mesh refinement that allows spatially resolving variations on the scale of ξ, see lower right panel of Fig. 3.2. This is illustrated in the upper right panel of Fig. 3.2 where the electrical field varies smoothly in space while attaining large but finite values, despite the underlying arbitrarily sharp change in the surface topography of the metal.

As a final comment on the numerics, we emphasize that our finite-element implementation was initially targeting the common hydrodynamic model where ξ^2 reflects only convection dynamics.[19,31] However, in the light of Eq. (3.10) the code is sufficiently general to also include any other effects that are of a short-range and isotropic nature. As such, the inclusion of diffusion is straightforward, since this only renders the length scale in Eq. (3.18) complex valued, while leaving all other governing equations and boundary conditions unchanged.[32] Most importantly, this conclusion also applies to any other method that relies on the original hydrodynamic equations, including other methods and approaches[19,31,57,58,60,61,65] that we have briefly discussed above.

3.7. Estimates of Characteristic Material Parameters

So far, we have with general arguments derived the nonlocal corrections to the LRA and with the hydrodynamic drift-diffusion model we have provided a semiclassical account for the nonlocal length scale ξ, see Eq. (3.10). Obviously, the real part reflects the Fermi velocity of the metal, which is a property resulting from a quantum treatment of the electrons, and as such the correction term is clearly of a semiclassical nature. The imaginary part contains the diffusion constant, which is our classical approach to phenomenologically account for complex many-body interactions and scattering events at the metal surface. One first approach is the semiclassical estimate of D, being related to the mean-free path ℓ and the relaxation rate $\gamma = 1/\tau$ already present in the usual Drude description, i.e. $\ell = v_F\tau$ and likewise, the diffusion constant is given by $D = \ell v_F = v_F^2\tau$.

Table 3.1. Characteristic length scales and parameters for Au, Ag, and Na.

	Fermi wavelength λ_F	Mean-free path ℓ	Convec. length v_F/ω_p	Dif. length $\sqrt{D/\omega_p}$	Scat. time $\omega_p\tau$
Au	0.52 nm	103 nm	0.10 nm	1.9 nm	1000[a]
	0.52 nm	50 nm	0.11 nm	1.3 nm	465[b]
Ag	0.52 nm	103 nm	0.10 nm	1.9 nm	1000[a]
	0.52 nm	40 nm	0.10 nm	1.1 nm	421[b]
Na	0.68 nm	2.6 nm	0.12 nm	0.32 nm	21[b]
	0.68 nm	4.4 nm	0.12 nm	0.42 nm	37[c]
	0.68 nm	2.3 nm	0.12 nm	0.31 nm	20[d]

[a]Data for crystalline materials taken from Ashcroft and Mermin.[11] [b]Data tabulated by Blaber *et al.*[68] [c]Data used in simulations by Teperik *et al.*[47] [c]Data used in simulations by Stella *et al.*[66]

In Table 3.1, we summarize characteristic length scales and parameters for the common plasmonic metals Au and Ag. We have also included data for Na which has recently received considerable attention in the *ab initio* quantum plasmonic community.[47,66,67] The entries for different metals are based on Fermi wavelengths λ_F, Fermi velocities v_F, and plasma frequencies ω_p taken from standard tables,[11] while the values for τ originate from various references as indicated in the right-most column of the table.

Inspecting the numbers, we note that all length scales are nanometric. In fact, in most cases they are even approaching atomic dimensions. This underlines the success of the LRA when applied to larger plasmonic structures; the nonlocal correction to the delta-function response of the LRA is indeed negligible on longer lengths scales due to the vanishing surface-to-volume ratio.

In their textbook discussion of spatial dispersion, Landau and Lifshitz[10] already mentioned the simultaneous occurrence between convection and diffusion length scales, and that spatial dispersion would be dominated by one of these two transport mechanisms, i.e. the one

that manifests itself on the largest scale. Often, we are led to consider diffusion a slow process, which would imply that $D \ll v_F^2/\omega_p$. However, given our classical estimate for D, Table 3.1 indicates that diffusion and convection are indeed playing in concert when considering plasmons in metals at optical frequencies. The two length scales appear comparable and if so, the mechanisms should naturally be treated on an equal footing. This is indeed made possible by our Eq. (3.18)!

3.8. The Relation of GNOR to Higher-level Descriptions

Above, we have used the classically widely accepted drift-diffusion model to account for diffusive effects in the hydrodynamic description of plasmons in metallic nanostructures. Here, we briefly mention more recent discussions of this. In particular, it is long known how the hydrodynamic model with the convective contribution can be derived from higher-level descriptions such as the Boltzmann transport equation or Kubo formalism.[69] *Can our more recent diffusive contribution be derived in a similar manner?* Here, we recall that for convection and diffusion sharing the same underlying microscopic mechanism, their high and low-frequency limits should be intimately connected by the fluctuation-dissipation theorem.[70] Focusing on the convective contribution as it occurs in the standard hydrodynamic model of a bulk metal, we see that it actually already contains both high-frequency convective and low-frequency diffusive contributions, i.e.

$$\frac{v_F^2}{\omega(\omega+i\gamma)} \simeq \begin{cases} \frac{v_F^2}{\omega^2}, & \omega \gg \gamma \\ \frac{v_F^2}{i\omega\gamma} = \frac{D}{i\omega}, & \omega \ll \gamma \end{cases} \tag{3.23}$$

where $D = v_F^2/\gamma$ is the classical estimate for the diffusion constant introduced above. Now, this also indicates that a high-frequency diffusive contribution term would not emerge out of Boltzmann considerations applied to the homogeneous bulk system and indeed this can also be shown explicitly.[71] Simply stated, bulk diffusion

at optical frequencies is negligible. *So, what is then the physical mechanism behind the observed diffusion at optical frequencies?* Here, surface-enhanced Landau damping represents a surface-scattering effect beyond the bulk description of the metal. As we shall see, estimated numbers for D remain in the same ballpark of Table 3.1.

3.9. A Test Case for GNOR Theory: Monomers and Dimers

Nonlocal hydrodynamics is a long established theory with the account for both longitudinal plasmons at frequencies above the plasma frequency and with frequency shifts of transverse plasmon resonance below the plasma frequency. *However, what is the test of the GNOR theory with the inclusion of diffusive effects?*

Here, monomers and their dimer counterparts represent an interesting and important test case. In the context of size-dependent broadening, the Kreibig model[72] only addresses monomers while the quantum-corrected model[73] exclusively applies to dimers, and we turn to both below. Obviously, any versatile theory could rightfully be expected to offer new insights for such archetypical plasmonic structures and of course there should be some unifying aspect.

The phenomenological theory of Kreibig and co-workers successfully mimics the size-dependent spectral broadening observed for few-nanometer sized and close-to-spherical monomers.[72] Although this theory satisfactorily explains size-dependent broadening in spherical nanoparticles quite well, it is not immediately clear how to extend it to particles of other morphology, let alone to dimers. In particular, the theory does not address the question whether diffusive surface scattering would be of importance to small-gap dimers, e.g. dimers composed of monomers that are not necessarily small themselves.

On the other hand, the gap-dependent broadening in such nano-gap dimers has been explained by the phenomenological quantum-corrected model[73] that invokes additional dissipation attributed to relaxation of tunneling currents. Resting on a tunneling picture where charge tunnels across the classically prohibited nano-gap from the one nanoparticle to the other, it is clear that this theory

cannot explain the observed size-dependent broadening of individual nanoparticles, the broadening that the Kreibig theory does capture.

As we shall see, this is where the GNOR proves versatile and with a predictive power reaching far beyond the case of spherical monomers. In fact, despite GNOR's simplicity and semiclassical nature, it unifies the phenomena observed for monomers and dimers, while also suggesting a close connection between size-dependent damping in monomers and gap-dependent broadening in dimers. Here, diffusion associated with surface scattering is the key to understand both monomers and dimers.

We start with a brief comparison of models for a spherical particle of radius R. The approach of Kreibig is to boldly add a Av_F/R size correction to the bulk damping parameter to phenomenologically account for surface-scattering effects obviously not exposed in the bulk of the particle. We note that this size dependence is generic for any scattering effect that reflects the surface-to-volume ratio. In the context of GNOR, the complex-valued nonlocal length scale ξ in Eq. (3.18) leads to $1/R$ corrections too.[32] Furthermore, this observation immediately allows us to link the diffusion constant D to Kreibig's A-coefficient, where A is experimentally found to be of the order of unity for most considered metals. Interestingly, this way of estimating D agrees fairly well with the estimates in Table 3.1, suggesting that surface-scattering rates are not too different from bulk-damping rates. In Fig. 3.3 we illustrate the main differences between the GNOR and the LRA models when applied to a $R = 1.5\,\mathrm{nm}$ spherical metal particle. In particular, the figure is highlighting the LDOS and extinction spectra with the GNOR model exhibiting a blueshift and a further broadening of the dipole resonance as compared to the behaviour within the LRA.

Next, the crucial check is to explore whether the very same choice of D-value would also serve to explain spectra of dimers. In Fig. 3.4 we illustrate simulations of extinction spectra of a Na dimer, contrasting the LRA and the GNOR. For simplicity, the dimer consists of cylindrical nanowires with a radius of $R = 10\,\mathrm{nm}$, while the gap size g is varied between 50 Å and -50 Å in steps of 5 Å. We emphasize that

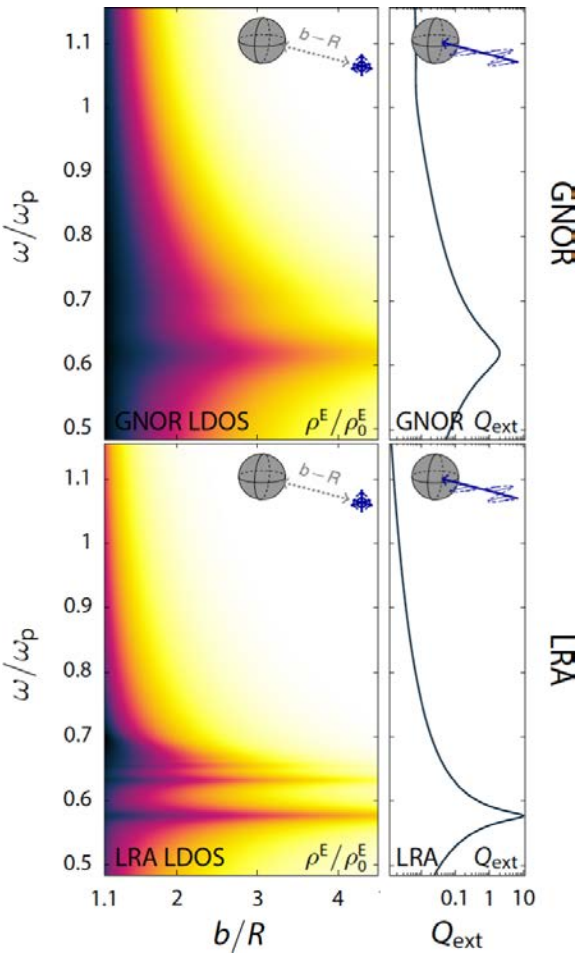


Fig. 3.3. GNOR and LRA calculations for the LDOS and extinction spectra for a $R = 1.5\,\text{nm}$ metal particle. Courtesy of Thomas Christensen.[74]

the negative gap values ($g < 0$) correspond to overlapping wires, in which case conducting nano-junctions are formed, where charge can flow back and forth in a classical manner rather than being transferred in quantum tunneling fashion.

If we now focus on the spectra in Fig. 3.4 for decreasing but positive gap sizes, then we notice that the GNOR dimer spectra become increasingly broadened while the LRA dimer spectra do not.

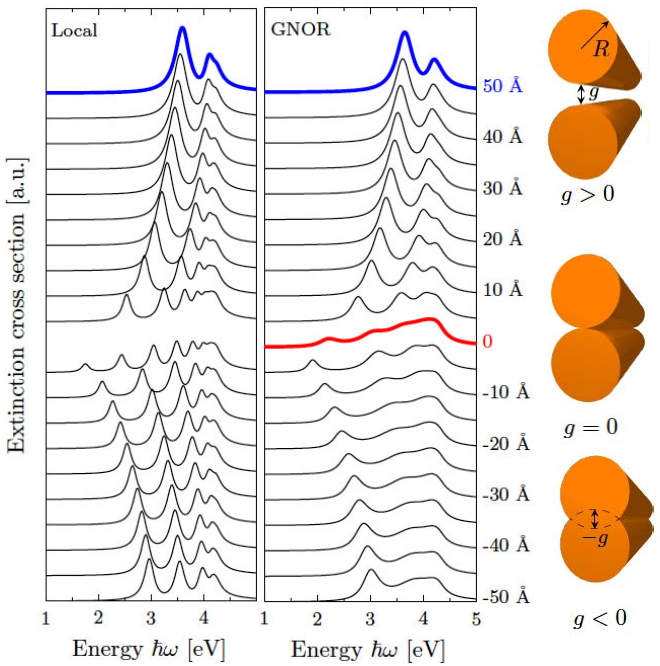


Fig. 3.4. Extinction cross section as a function of energy for a Na dimer with a radius $R = 10$ nm and a gap size g varying from 50 Å to -50 Å in steps of 5 Å. The values next to the spectra denote the corresponding values for the gap size g. Courtesy of Søren Raza.[76]

Such gap-size dependent broadening has been observed experimentally, and sometimes been interpreted solely as due to quantum tunneling. It is therefore important to stress that both the LRA and GNOR spectra in Fig. 3.4 are obtained with an entirely semiclassical model where the simplifications and the resulting boundary conditions (see Sec. 3.5) are incompatible with a quantum-tunneling interpretation. Most critically, the work function is considered infinite and the metal surfaces serve as hard walls or perfectly reflecting mirrors for the electrons, so that inter-nanoparticle charge transport is entirely prohibited for $g > 0$. Instead, the gap-dependent broadening is due to surface scattering as mimicked by the diffusion contribution to the nonlocal response. In this way, size-dependent broadening

becomes pronounced as the gap approaches the magnitude of the nonlocal length scale. Note that broadening also occurs for weakly overlapping wires; a case were quantum-tunneling is suppressed, with the charge-transport being of a classical nature. To summarize, gap-dependent broadening sets in for nanometric dimer gaps and GNOR spectra agree qualitatively with both experiments[22,75] and *ab initio* studies.[47]

So, what is the important common feature of the monomers of Fig. 3.3 and the dimers of Fig. 3.4? Clearly, tunneling has no meaning for monomers! On the other hand, if surface scattering appears near any metal surface, then this effect could explain both the enhanced scattering at the surfaces of monomers as well as the scattering near the surfaces of almost touching dimers. (This being said, although quantum tunneling is systematically excluded from our semiclassical theory, we do not rule out the possibility that quantum tunneling may lead to additional gap-size-dependent broadening in dimers, as an additional physical mechanism of spectral broadening.[36]) In the following, we turn to our *ab initio* studies where more rigorous insight consolidates the importance of surface-enhanced Landau damping (mimicked as diffusive broadening). In particular, we show how this indeed acts as a unifying feature of both monomers and dimers.

3.10. Surface-Enhanced Landau Scattering: *ab initio* Insight

The phenomenological Kreibig model for size-dependent broadening has already been linked to quantum mechanical calculations of Landau damping associated with electron-hole pair generation in the metal near the surface, see recent work.[34] Likewise, *ab initio* studies[36,48] and electron spectroscopy[28,77] have established strong plasmon damping at the very surface of metals. Most importantly, the damping may exceed the attenuation occurring in the bulk. Here, plasmons decay into electron-hole pairs and at the surface they do so more effectively than in the bulk. This is what the GNOR model seeks to capture with the diffusion parameter D, i.e. by using

this single parameter in a classical equation. Since induced charge associated with the plasmon always resides in the near vicinity of the surfaces, density gradients make diffusion effective only near the surface. In other words, our nonlocal term in Eq. (3.21) is most pronounced at the surface where $\mathbf{E}$ varies the most. As a result, the drift-diffusion model successfully mimics both longitudinal pressure waves and surface-enhanced Landau damping. The enhanced damping near the surface is illustrated in Fig. 3.5 where we show the local effective permittivity $\varepsilon_{\text{eff}}(\mathbf{r},\omega)$ at a planar metal-air interface as extracted from GNOR simulations via

$$\mathbf{D}(\mathbf{r},\omega) \equiv \varepsilon_0 \varepsilon_{\text{eff}}(\mathbf{r},\omega)\mathbf{E}(\mathbf{r},\omega). \tag{3.24}$$

The $\varepsilon_{\text{eff}}(\mathbf{r},\omega)$ so defined is called 'effective' because both the displacement and the electric fields are determined using a nonlocal-response theory (being GNOR theory in this case, other theories below), and it is the effectively local quantity $\varepsilon_{\text{eff}}(\mathbf{r},\omega)$ that by definition connects the two electromagnetic fields, i.e. close in spirit to the LRA of Eq. (3.11). Note in Fig. 3.5 that due to the abrupt termination of the surface and the associated hard-wall boundary condition for the normal component of the current, the additional damping is 'forced' to occur slightly inside the surface. This links up to the Feibelman parameter and the importance of the *actual position of the surface of the electron plasma* was recently excellently discussed by Teperik and co-workers[47,78] in the context of plasmonic ruler effects of subnanometer gap dimers. Here, we note that the hard-wall boundary condition can be relaxed to also include density-gradient and spill-out effects in a hydrodynamic model.[52]

The common importance of surface-enhanced Landau damping for monomers and dimers can be conveniently explored with *ab initio* solutions of the optical response of the electron gas near the surface of the metal.[36,79] For the ease of the *ab initio* simulations we consider the jellium approximation applicable to simple metals, such as sodium (Na). Using time-dependent density-functional theory we seek the response to a time-dependent electrical field. The calculation provides us both with the equilibrium density $n_0(\mathbf{r})$ (now with a

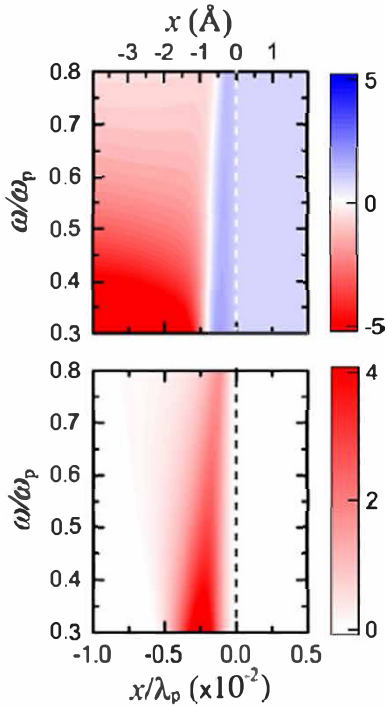


Fig. 3.5. GNOR results for the real part (top panel) and imaginary part (lower panel) of $\varepsilon_{\text{eff}}(x,\omega)$ as defined in Eq. (3.24), for a planar metal-air interface, as a function of position and of frequency, for frequencies throughout the optical range. Reproduced from Ref. 79.

space dependence that we neglected in our GNOR model) and with the induced charge density $n_1(\mathbf{r})$. In the context of Eq. (3.24), we also obtain the displacement field $\mathbf{D}$ that occurs in response to the perturbing $\mathbf{E}$ field, thus allowing us to infer an effective relative dielectric function $\varepsilon_{\text{eff}}(\mathbf{r},\omega)$, see Eq. (3.24). In this way, we can now explore and visualize the enhanced damping associated with the more efficient electron-hole pair generation near the surface.

Figure 3.6(a) shows the *ab initio* counterpart of the GNOR results in Fig. 3.5. In addition, the top panel illustrates the equilibrium density (exhibiting both Friedel oscillations and quantum spill-out) which is assumed step-like in the GNOR treatment. Inspecting

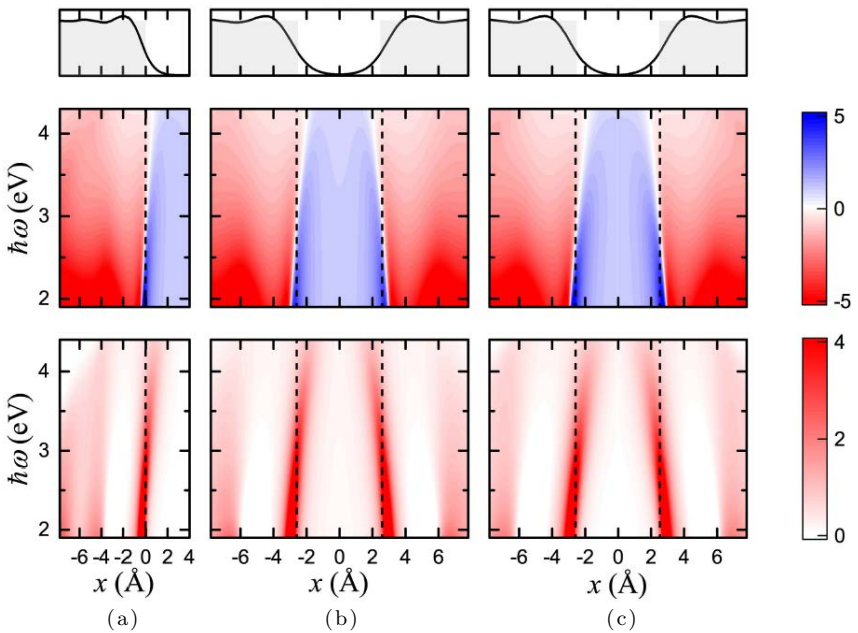


Fig. 3.6. Equilibrium density in the jellium model (top panels), exhibiting both Friedel oscillations and quantum spill-out, along with the TD-DFT results for the real part (middle panels) and imaginary part (lower panels) of ε_{eff} as defined in Eq. (3.24), for different frequencies throughout the optical range. Panel (a) is for a single interface. Panel (b) is a full calculation for interacting surfaces. Finally, panel (c) is for a dimer with a 0.5 nm gap separating two independent interfaces, and is based on the results in (a). Reproduced from Ref. 79.

the imaginary part of the response (lower panel), we immediately note that near the surface the Landau damping is exceeding its bulk limit. Although electron spill-out occurs at the surface, the surface-enhanced imaginary part of the response still occurs mostly inside the classical geometric interface, at least for energies below 3 eV. This may serve as a justification of adopting the simpler GNOR model in that energy range.

We also note that the enhanced imaginary part of $\varepsilon_{\text{eff}}(\mathbf{r}, \omega)$ near the surface is accompanied by a vanishing real part occurring right where the decaying tail of the equilibrium electron density $n_0(r)$ is small enough for the dilute plasma to exactly balance the

vacuum background polarization.[50] As such, the enhanced Landau damping could also be seen as a manifestation of damping inherently associated with epsilon-near-zero phenomena.[80]

We now repeat the calculation for a dimer, i.e. two opposing jellium surfaces separated only by a sub-nanometric vacuum gap. As an example, Fig. 3.6(b) shows the results for a 0.5 nm gap. Quite remarkably, we here recover the physics found for a single interface, but this time with significant Landau damping occurring at both interfaces. On the other hand, there is no pronounced damping occurring inside the gap. The close similarity with the single interface is perhaps best illustrated by comparison with panel (c) which shows a simple superposition of the response of two opposing independent surfaces. The similarity of panels (b) and (c) illustrates how the surface scattering characteristic of a single isolated interface also dominates the dissipation exhibited in a dimer. The truly surprising aspect is that this seems to be the case even for a gap of only 0.5 nm. It is this dominant role of Landau damping that the diffusive term in the GNOR model captures so well.

3.11. Surface-enhanced Landau Scattering: Microscopic Insight

All nonlocal-response phenomena of the previous sections are described by the spatial dispersion of the optical susceptibility [see Eq. (3.2)], both by its real and imaginary parts. In principle, different microscopic theories may give rise to the same nonlocal optical response function. In the phenomenological theory of prior sections, the spatial dispersion of the susceptibility is represented through the nonlocal parameter ξ of Eq. (3.18), whose real part defines the range over which the perturbation at point $\mathbf{r}$ affects the dielectric response at point $\mathbf{r}$. As shown above, this range is given by $\mathrm{Re}\{\xi\} \approx \beta/\omega = \sqrt{3/5}\, v_F/\omega$ i.e. roughly the distance traveled by the electron near the Fermi surface over one optical cycle. In our phenomenological theory this term is interpreted as convection and its inclusion in the model of plasmonic structures leads mostly to the size-dependent changes in resonant frequencies of plasmonic modes.

On the other hand, the imaginary part of the nonlocal term mostly originates from the diffusion of the electrons, i.e. $\mathrm{Im}\{\xi\} \approx \sqrt{D/2\omega}$. When introduced into the plasmonic model this leads to the size-dependent damping of the plasmonic modes, and to changes in the plasmonic spectra that are more drastic than the ones caused by the changes in the real part of ξ. These size-dependent damping phenomena have been studied and described in depth by Kreibig and co-workers[72] and they are associated with surface collisions. According to Kreibig, the electrons moving with Fermi velocity inside the metal particle of size a [see example in panel (a) of Fig. 3.1] are expected to experience a collision with the surface after the time $\tau_c \sim a/v_F$. This collision alters the momentum of electrons and thus leads to additional damping of the Drude dielectric function of the metal. For a spherical particle of radius R, the dielectric function would then appear as

$$\varepsilon_{\mathrm{eff}}(\mathbf{r},\omega) = \begin{cases} 1 - \frac{\omega_p^2}{\omega[\omega - i\gamma - i\gamma_s(R)]}, & r < R \\ 1, & r > R \end{cases} \tag{3.25}$$

where γ is the already introduced *bulk* damping rate, while γ_s denotes the surface origin of the size-dependent surface damping contribution $\gamma_s(R) = Av_F/R$. As we have already discussed briefly in Sec. 3.9, the dimensionless constant A is of the order of unity. Note that the dielectric function in Kreibigs interpretation is entirely within the spirit of the LRA, i.e. it is a stepwise constant local dielectric function that depends only on the size of the nanoparticle. Hence, even though γ_s is referred to as a surface damping, it is included into the Drude expression for the entire volume of the nanoparticle and not just for the points close to the surface.

A simple qualitative interpretation of the Kreibig damping can be made by first considering the conductivity of a lossless Drude metal, $\sigma_D = i\varepsilon_0\omega_p^2/\omega$ [see Eq. (3.19)]. In the lossless metal electrons do not experience any momentum-altering collisions, hence the conductivity is purely imaginary [Eq. (3.19)]. If collisions do occur, the time over which interaction between the electrons and the electric field takes

place is limited by the mean time between collisions $\tau = \ell/v_F$, where ℓ is the mean-free path that we also discussed in Sec. 3.7. Then, according to the Heisenberg Principle the frequency of the electromagnetic wave has an uncertainty $\gamma = 1/\tau$ and its inclusion as $\omega \to \omega + i\gamma$ immediately leads to the Drude formula in Eq. (3.19) and thereby to the expression for the dielectric function (3.25). Note that any event that shortens the interaction time between the electromagnetic wave and the electrons, will contribute to the damping rate, be it a phonon absorption/emission process, scattering on a defect, collisions with a wall, or, simply an electron exiting the region where the electromagnetic wave is concentrated. In that respect, a more descriptive name for γ_s could perhaps be *confinement damping* which states that as long as the region where the electrons and electro-magnetic field interact is restricted to a characteristic size a, the interaction time is limited to a/v_F which causes damping with the rate $\gamma_s \sim v_F/a$. We will return to this question of *confinement* versus *surface-collisions damping* shortly.

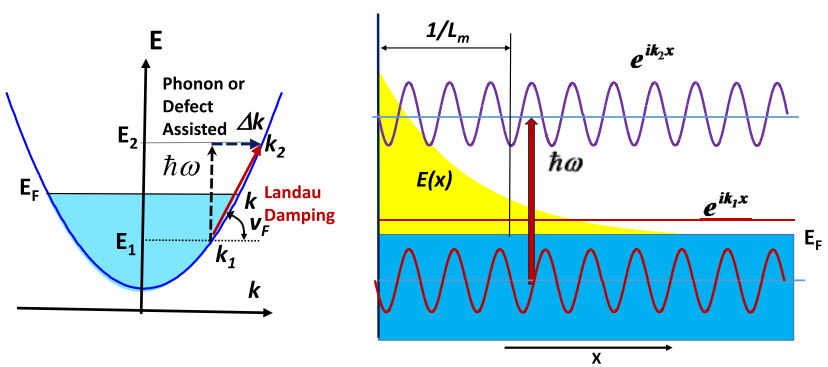
A full quantum origin of the confinement damping can best be described by considering two free-electron states with wave functions $\psi_1(\mathbf{r},t) = \exp[i(\mathbf{k_1}\cdot\mathbf{r} - \omega_\mathbf{1}\mathbf{t})]$ and $\psi_2(\mathbf{r},t) = \exp[i(\mathbf{k_2}\cdot\mathbf{r} - \omega_\mathbf{2}\mathbf{t})]$, with energies $\mathcal{E}_1 = \hbar\omega_1$ and $\mathcal{E}_2 = \hbar\omega_2$ below and above the Fermi level respectively, as shown in panels (a) and (b) of Fig. 3.7 in $\mathbf{k}$-space and real space, respectively. In the presence of the electric field $\mathbf{E}(\mathbf{r},t) = \mathbf{E}_0 \exp[i(\mathbf{k}\cdot\mathbf{r} - \omega\mathbf{t})]$, where $\omega = \omega_2 - \omega_1$, the transition involves a matrix element of the interaction Hamiltonian $H_{12} = -e\mathbf{E}\cdot\mathbf{r}_{12} = -e\mathbf{E}\cdot\mathbf{p}_{12}/m\omega$ which becomes

$$H_{12} = -\frac{e\hbar}{m\omega}(\mathbf{E}_0\cdot\mathbf{k}_2)\int d\mathbf{r}\,\exp[\mathbf{i}(\mathbf{k_2}-\mathbf{k_1}-\mathbf{k})\cdot\mathbf{r}] \qquad (3.26)$$

where m is the free electron mass, and we have used the standard relation between matrix elements of coordinate $\mathbf{r_{12}}$ and momentum $\mathbf{p_{12}}$. The matrix element is finite if $\mathbf{k} = \delta\mathbf{k} = \mathbf{k_2} - \mathbf{k_1}$, or $k > k_2 - k_1 \approx k_{\rm LD}$ where $k_{\rm LD} = \omega/v_F$ is the wave-vector at which the onset of Landau damping occurs. We note that in relation to the hydrodynamic theory and ξ in Eq. (3.18), the wave vector $k_{\rm LD}$ is also a characteristic cutoff for the otherwise unbound wave vector in the LRA.[41,43,81]

Landau damping[33] occurs when the phase velocity for the electromagnetic wave $v_{\rm ph} = \omega/k$ is matched by the velocity of an electron, which allows them to exchange energy. This process is shown as a straight diagonal line in Fig. 3.7(a). For noble metals the Fermi velocity has a value of $v_F \approx 1.4 \times 10^6$ m/s, which is about two orders of magnitude less than the phase velocity of propagating electromagnetic waves. Hence Landau damping is not too effective in the bulk of a metal, even if the phase velocity would be reduced as for polaritons and other surface waves. Because of the small wave-vector of the propagating waves $k \ll \Delta k$, many absorption processes (including free-carrier absorption) involve some other process (phonon, carrier-carrier interaction defect) to provide momentum Δk. This interaction is represented in Fig. 3.7(a) as a two-step process of first absorbing a photon and then subsequently scattering into the final state.

The situation of Landau damping changes, however, if the region in which electrons and the electric field interact is restricted. This occurs either because the field itself is spatially dependent as $\mathbf{E}_0(\mathbf{r})\exp(-i\omega t)$, or because the electrons are confined within the finite region (or, most often, both). Basically, the finiteness of the structure and the abruptness of the surface termination will in a



(a) (b)

Fig. 3.7. Free-carrier absorption in the conduction band between the states k_1 and k_2. (a) k-space diagram of Landau damping (direct diagonal line) and phonon/defect (dashed lines) transitions. (b) Real-space diagram illustrating the wave-functions and confined electric field near the surface.

Fourier representation promote wave-vector components exceeding the cut-off $k_{\mathrm{LD}} = \omega/v_F$ associated with Landau damping. Returning to Eq. (3.26), the integral in the matrix element of the interaction Hamiltonian becomes $\int_{\mathrm{metal}} d\mathbf{r}\, \mathbf{E}_0(\mathbf{r}) \exp[i(\mathbf{k}_2 - \mathbf{k}_1) \cdot \mathbf{r}]$ [see Fig. 3.7(b)], where the integration is taken over the now finite volume of the metal. While the integral is zero for a translationally invariant system, it now clearly remains finite due to the finite size of the structure, and the transition between the states ψ_1 and ψ_2 can take place with a finite rate. This is a quantum picture of Landau damping.

As a first example, consider the most simple case of the surface-plasmon polariton (SPP) propagating on the boundary between a metal and a dielectric,[35] as shown in Fig. 3.7(b). The normal component of the electric field inside the metal can be written as $E_x = E_0 \exp(-x/L_m) \cos(k_z z - \omega t)$ where k_z is the propagation constant and L_m is the decay length in the metal. The matrix element of the interaction Hamiltonian is then

$$\begin{aligned} H_{12} &\sim -\frac{e\hbar E_0 k_x}{2m\omega L} \int_0^\infty dx\, \exp(-x/L_m) \sin(\Delta k_x x) \\ &= -\frac{e\hbar k_x}{2m\omega L} \frac{E_0 \Delta k_x}{L_m^{-2} + \Delta k_x^2}, \end{aligned} \tag{3.27}$$

where L is a quantization length. The sine function appears in Eq. (3.27) due to the boundary conditions for the electron wave function. (For the field polarized along the z-direction one ends up with a much smaller value of integral and that is why we consider only the normal component of electric field here.) One can then in a next step use Fermi's Golden rule in order to estimate the transition rate from a given state k_x in the conduction band as $R(k_x) = 2\pi\hbar^{-1} |H_{12}|^2 L \rho_x(\mathcal{E}_F)$ where $\rho_x(\mathcal{E}_F) = (2\pi\hbar|v_{F,x}|)^{-1}$ is the one-dimensional density of states near the Fermi level and $v_{F,x}$ is the projection of the Fermi velocity onto the x-direction. Using the relation $\Delta k_x \approx v_{F,x}/\omega$ and $\Delta k_x L_m \gg 1$ one quickly arrives at the transition rate $R(k_x) = e^2 |v_{F,x}^3| E_0^2/(4\hbar^2\omega^4 L)$. Performing the summation over all wave vectors $\mathbf{k}$ and subsequent multiplication by $\hbar\omega$

results in the expression for the energy loss per unit of surface area

$$\frac{dU}{dt} = -\hbar\omega L \sum_{\mathbf{k}} R(k_x) = -\frac{e^2 m^2 v_F^4 E_0^2}{16\pi^2 \hbar^3 \omega^2} = -2\gamma_s U \tag{3.28}$$

where $2\gamma_s$, the effective rate at which U decays, is equal to twice the momentum decay rate γ_s. The energy density U experiencing the decay is the time-averaged density of the kinetic energy of free electrons

$$U(x) = \int_0^\infty dx \, \frac{e^2}{4m\omega^2} n_0 E_x^2(x) = \frac{L_m}{24\pi^2} \frac{e^2 m^2 v_F^3 E_0^2}{\hbar^3 \omega^2} \tag{3.29}$$

where $n_0 = k_F^3/3\pi^2 = m^3 v_F^3/3\pi^2\hbar^3$ is the density of free electrons also appearing in the Drude conductivity [Eq. (3.19)]. From Eqs. (3.28) and (3.29) we immediately obtain the relation $\gamma_s = (3/4)v_F/L_m$, which would be remarkably close to Kreibig's phenomenological estimate $\gamma_s \sim Av_F/a$ if in the latter we would have used the decay length L_m in place of the particle size a. Therefore, one may generalize the phenomenological estimate, by stating that the surface damping is determined by the effective size of the region in which both the electrons and the electric field are contained. Obviously, if the size of the metal structure a is smaller than decay length L_m of a given surface plasmon mode, then it is a that determines the surface damping. Alternatively, for the large metal structures it is the decay length that defines γ_s.

To generalize these results we should note that the evaluation of the interaction Hamiltonian for the transition between two free-electron states was in its essence nothing but an evaluation of the spatial Fourier transform of the electric field $\mathbf{E}(\mathbf{r})$. This observation can lead us in the direction of using the Lindhard expression for the dielectric function of the free electron gas to evaluate the effective dielectric function seen by the confined electromagnetic wave. According to Lindhard,[82] the k-dependent longitudinal dielectric function is

$$\varepsilon(\omega, k) = 1 + \frac{3\omega_p^2}{v_F^2 k^2} \left[1 - \frac{\omega}{2v_F k} \ln \frac{\omega + v_F k}{\omega - v_F k} \right]. \tag{3.30}$$

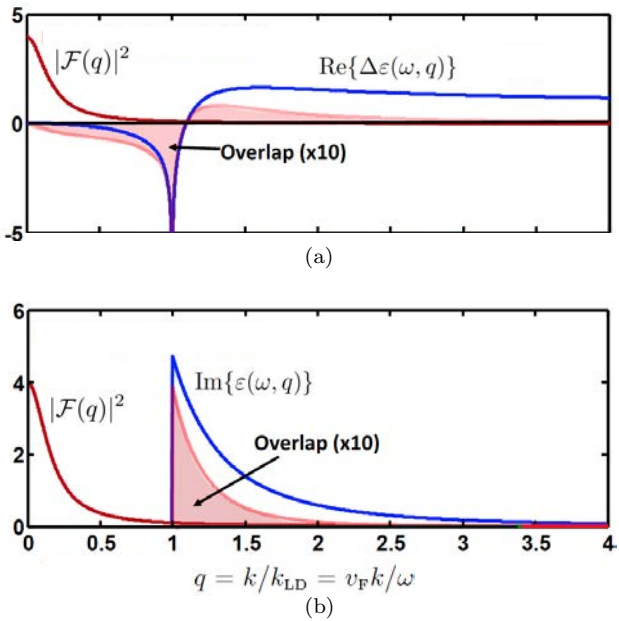


Fig. 3.8. Spatial dispersion of (a) real and (b) imaginary parts of Lindhard metal dielectric function and their overlaps with the spatial power spectrum of the electric field in the metal.

For convenience, we next introduce a normalized wave vector $q = k/k_{\mathrm{LD}} = v_F k/\omega$, i.e. being normalized to the onset of Landau Damping, with the aim to obtain separate expressions for the real and imaginary parts. For the real part one can write $\mathrm{Re}\{\varepsilon(\omega,q)\} = \mathrm{Re}\{\varepsilon(\omega,q\to 0)\}+\mathrm{Re}\{\Delta\varepsilon(\omega,q)\}$, where $\mathrm{Re}\{\varepsilon(\omega,q\to 0)\} = 1-\omega_p^2/\omega^2$ is a long-wavelength Drude dielectric function [see Eq. (3.22)], while

$$\mathrm{Re}\{\Delta\varepsilon(\omega,q)\} = \frac{\omega_p^2}{\omega^2}\left[1+\frac{3}{q^2}-\frac{3}{2q^3}\ln\left|\frac{1+q}{1-q}\right|\right] \tag{3.31}$$

is the nonlocal correction also being illustrated in Fig. 3.8(a).

For a discussion of the imaginary part of the dielectric function we focus on the limit without scattering: For small wave vectors $k < \omega/v_F$ (corresponding to $q < 1$) it remains zero, while for larger wave vectors Landau damping engenders a non-vanishing imaginary part

of the dielectric function. Finally, since the logarithm of a negative number has an imaginary part of $i\pi$, one obtains

$$\mathrm{Im}\{\varepsilon(\omega, |q| > 1)\} = \frac{\omega_p^2}{\omega^2}\frac{3\pi}{2q^3}, \tag{3.32}$$

as depicted in Fig. 3.8(b). One can now evaluate the Fourier transform $\mathcal{F}(\mathbf{k})$ of the Lindhard dielectric function and a power density spectrum $|\mathcal{F}(\mathbf{k})|^2$ of the electric field $E(\mathbf{r})$ inside the metal. It is only logical then to assume that all the longitudinal field components with spatial frequencies with $k_L > \omega/v_F$ get Landau damped and thus contribute to the imaginary part of the dielectric function. For each value of the wave vector $\mathbf{k}$ the power density of the longitudinal field is $|\mathcal{F}_L(\mathbf{k})|^2 = |\mathcal{F}(\mathbf{k}) \cdot \mathbf{k}|^2/k^2$. One can then evaluate the effective longitudinal dielectric function $\varepsilon_{\text{eff}}(\omega)$, which is a function of the size and shape of an SPP mode, by evaluating the overlap of $|\mathcal{F}_L(\mathbf{k})|^2$ with $\varepsilon(\omega, k)$ as seen in Fig. 3.8(a) & (b) for the real and imaginary parts of the dielectric function, respectively. The imaginary part of this dielectric function is then

$$\mathrm{Im}\{\varepsilon_{\text{eff}}\} = \frac{3\pi\omega_p^2}{2\omega^2}\frac{\int_{q>1}^{\infty} d^3q\, q^{-3}|\mathcal{F}_L(q)|^2}{\int_0^{\infty} d^3q\, |\mathcal{F}(q)|^2}. \tag{3.33}$$

Now, according to the Drude formula for bulk metals [Eq. (3.25)] one has $\mathrm{Im}\{\varepsilon_{\text{eff}}\} = \omega_p^2\gamma/\omega^3$ in terms of the bulk damping parameter γ. Likewise we can now use the above analysis to introduce the Landau damping rate as $\mathrm{Im}\{\varepsilon_{\text{eff}}\} = \omega_p^2\gamma_{\text{LD}}/\omega^3$, so we eventually find that

$$\gamma_{\text{LD}} = \frac{3\pi\omega}{2}\frac{\int_{q>1}^{\infty} d^3q\, q^{-3}|\mathcal{F}_L(q)|^2}{\int_0^{\infty} d^3q\, |\mathcal{F}(q)|^2} = \frac{\omega}{Q_s}, \tag{3.34}$$

where Q_s is the structural quality factor of the mode which represents the broadening associated with nonlocal dynamics.[83] The total damping is then $\gamma \to \gamma + \gamma_{\text{LD}}$ as previously expressed in Eq. (3.25). The novelty here is that we have a theoretical estimate for the surface damping rate based on the Landau damping mechanism.

Let us now again evaluate the Landau damping rate for the propagating SPP that we considered above, but this time based on the

power density spectrum analysis that we just developed. The normalized power spectrum has a Lorentzian shape for positive wave vectors, shown in Fig. 3.8,

$$|\mathcal{F}(k)|^2 = \frac{2}{\pi L_m}\frac{1}{k^2 + L_m^{-2}} = \frac{2v_F}{\pi L_m \omega}\frac{1}{q^2 + \left(\frac{v_F}{\omega L_m}\right)^2}. \tag{3.35}$$

Since the second term in the denominator is much smaller than unity, it can be neglected for $q > 1$, and then the evaluation of a simple integral of q^{-5} in Eq. (3.34) yields $\gamma_{\rm LD} = (3/4)v_F/L_m$, which is precisely the value of surface collision damping γ_s that we obtained below Eq. (3.29) using Fermi's Golden rule.

Interestingly enough, if instead the electron rather than the electric field is confined, then a similar result is obtained. For instance, if one considers a very thin slab of metal of thickness $a < L_m$, then the normalized power spectrum of the mode in the metal is

$$|\mathcal{F}(q)|^2 = \frac{a\omega}{\pi v_F}\frac{\sin^2(qa\omega/2v_F)}{(qa\omega/2v_F)^2}. \tag{3.36}$$

For $q > 1$ the sine function oscillates fast and its square can be replaced by its average value 1/2 with the result $\gamma_{\rm LD} = (3/4)v_F/a$, which is no different from the previous one. In fact, similar results can be obtained for just about any mode shape as long as there is a sharp boundary separating the metal from the dielectric, because at large wave vectors the power spectrum would always have the same wave-vector dependence of $\sim q^{-2}$, which of course happens to be the square of the Fourier transform of the Heaviside step function describing the boundary.

On the other hand, if the confinement does not involve a sharp boundary, then the situation is strikingly different. For instance, if, hypothetically, the mode confinement had been Gaussian, $E_x = E_0 \exp(-x^2/2a^2)$ and the metal spread out over both positive and negative x, then from the normalized power spectrum $|\mathcal{F}(q)|^2 = (v_F/a\omega\pi^{1/2})\exp[-(qa\omega a/v_F)^2]$ one would have obtained an entirely different result $\gamma_{\rm LD} \approx (3/4)(v_F/a) \times \pi^{1/2}(v_F/\omega a)^2 \exp[-(\omega a/v_F)^2]$, which decays very quickly with effective confinement width a. For

example even for $a = 1\,\text{nm}$ and a wavelength $\lambda = 750\,\text{nm}$, one would have had $\gamma_{\text{LD}} \approx 0.02 \times v_F/a$, in other words the Kreibig constant A would have come out with a value of only 0.02 rather than of order unity. If, on the other hand a single sharp boundary was restored into our model by restricting the metal to positive x, then the Fourier transform would no longer be Gaussian, but rather a complex error function. The corresponding power spectrum would then be $|\mathcal{F}(q)|^2 \sim (4v_F/a\omega\pi^2)\exp[-(qa\omega/v_F)^2]\left|\text{erfc}(ia\omega/\sqrt{2}v_F)\right|^2$. This function at large values of argument again decays inversely proportional to q^2. As a result, once again one would obtain the relation $\gamma_{\text{LD}} \approx (3/2\pi)v_F/a = Av_F/a$ with $A = 3/2\pi \approx 1/2$.

The above exercise of comparing sharp, loose, and semi-loose electron confinement confirms that the phenomenological "surface collision" model does reflect a correct physical picture — only a sharp discontinuity is capable of providing momentum matching that facilitates significant absorption of photons by free electrons. In passing, we note that this discussion can also be directly linked to Feibelman parameters.[48,71] Our analysis here also justifies the use of sharp boundaries in the GNOR model in combination with the interpretation of the diffusion (and hence size-dependent damping) in the GNOR model as due to Landau damping, while leaving open the possibility for other microscopic mechanisms, such as surface imperfections that produce similar phenomena.

Let us now turn our attention to the real part of ε. In a confined geometry, its change into an effective response can be found as

$$\text{Re}\{\Delta\varepsilon_{\text{eff}}(\omega)\} = \text{Re}\{\varepsilon_{\text{eff}}(\omega)\} - \varepsilon(\omega, q=0) = \frac{\omega_p^2}{\omega^2}\int_0^\infty g_r(q)|\mathcal{F}(q)|^2 dq \tag{3.37}$$

where $g_r(q) = 1 + 3/q^2 - (3/2q^3)\log|1+q|/|1-q|$, as also shown in Fig. 3.8. Numerical calculations show that $\text{Re}\{\Delta\varepsilon_{\text{eff}}(\omega)\} \approx \omega_p^2\gamma_{\text{LD}}^2/\omega^4$, which is in agreement with the amount predicted by the phenomenological theory of Kreibig.[72]

This result predicts a small positive change in the real part of the effective dielectric constant, which then leads to a small redshift of the resonance, while in reality the nonlocal response leads to a

blueshift in noble metals. One can account for this fact by noticing that the integral in Eq. (3.37) is taken all the way to infinity and is small because g_r changes sign as seen in Fig. 3.8(a). It is more reasonable to cut off the upper integration limit when the wave vector approaches the maximum value that can be supported by the Fermi gas, which is of the order of the Thomas–Fermi screening wave vector (inverse convection length), ω_p/v_F. That corresponds to cutting off the integration in Eq. (3.37) at $q_{\max} \approx \omega_p/\omega$, which has a value not much larger than unity. The net result of this cutoff will be a negative change in the real part of ϵ and indeed a resulting blueshift of the resonance. At the same time, cutting off the integration at $q_{\max}$ will not change the result for the imaginary part of ϵ, since the integrand shown in Fig. 3.8(b) does not change sign and decays very fast.

Thus changes of both real and imaginary parts of the dielectric function can lead to the manifestations of nonlocality described in earlier sections. Of those the increased loss and broadening associated with the imaginary part of ϵ becomes important roughly when $v_F/a \approx \gamma$ when Landau (surface) damping becomes commensurate with the bulk damping, i.e. when all the electrons located within the mode are no more than a mean free path away from the sharp interface. Diffusion theory gives essentially the same result. Since the scattering time is related to the bulk broadening as $\tau = \gamma^{-1}$, one can estimate the diffusion term as $D\nabla^2 \approx v_F^2\gamma^{-1}a^{-2}$ and this term is equal to bulk dissipation rate γ when $v_F/a \approx \gamma$. For a typical noble metal Landau (surface damping) thus becomes important for the mode size that is as large as 10 nm.

The corresponding change associated with change in the real part of the dielectric constant (or convection) is manifested as a blueshift. There is no direct analogous estimate for what sizes the nonlocal blueshift becomes significant, since there is no bulk energy shift to compare with. Qualitatively one can estimate that the blueshift becomes important on a significantly smaller scale, when the smallest dimensions, such as gap size, become comparable to the Thomas–Fermi screening length, which is on the order of a few Ångstrom. In experiments this turns out to be too conservative, as blueshifts

in few-nanometer particles are well observed,[23,24] for the smallest particles even larger than the theory predicts.[24,28]

Obviously, the tighter is confinement of electric fields, the larger is γ_{LD}, hence all the resonances broaden. This effect is most noticeable in the study of dimers (Fig. 3.3). But increase in γ_{LD} causes not just broadening but also a change in the shape of the mode itself, as high spatial frequencies of the electric field are attenuated and the resulting change in the shape of the mode can be described as diffusion. For example, as the size of the nanogap in the two cylinder dimer decreases, higher-order modes in the cylinders get excited.[84] But since these higher-order modes are close to the surface, they get strongly damped and suppressed[85] which causes spreading (diffusion) of the gap mode in real space.

3.12. Concluding Remarks

In this chapter we have summarized recent developments in nonlocal plasmonics. In particular, two descriptions of nonlocality have been developed. One is semiclassical or "hydrodynamic" for which we used a real-space formulation, while the other is microscopic and reciprocal space was emphasized.

According to the hydrodynamic description, as explained in earlier sections of this work, electrons are localized in real space and the local effective dielectric function of Eq. (3.24) changes in the vicinity of the surface as a result of diffusion. We showed that the hydrodynamic "GNOR" theory quite accurately describes Landau damping at the surface of individual nanoparticles, reproducing well-known phenomenological size-dependent damping as a nonlocal diffusion phenomenon. Moreover, by comparison of GNOR theory with microscopic TD-DFT calculations, we also find a good agreement for larger dimer structures with small gaps. In particular, even for gap sizes down to 0.5 nm, the agreement is quite good since quantum tunneling, that is neglected in GNOR theory, only sets in for even smaller gaps. It follows that spectral broadening for dimer gaps below 0.5 nm is due to a combination of surface-enhanced damping and quantum tunneling.

According to the microscopic "reciprocal space theory" of Sec. 3.11, free electrons are delocalized in space and different spectral components of the electric field experience different dielectric function. Still, even in this picture, the enhancement of the imaginary part of the dielectric function is caused not simply by the confinement of optical fields, but rather by the presence of sharp discontinuities of the electric field, or the electron density, or both. Kreibig's phenomenological theory for size-dependent damping is again reproduced, now based on quantum calculations. It is also generalized by the quantum calculations, with the prediction that the size-dependent damping will depend on the smallest of two sizes, either the size of the structure, as is commonly assumed, or the typical length of the optical mode in the metal.

Both theories predict damping of large-k-vector components of the field distribution, with diffusion in real space as a result. Not unlike the diffraction limit, this diffusion sets an ultimate limit on the degree of electromagnetic-field confinement achievable in nanoplasmonic structures.

Acknowledgments

We acknowledge funding from the Danish National Research Foundation (CoE Project DNRF103) and the Danish Council for Independent Research–Natural Sciences (Project 1323-00087). J.K. acknowledges support US Army Research Office (W911NF-15-1-0629). N.A.M. is a Villum Investigator supported by Villum Fonden. Over the course of years we have benefited tremendously from numerous interactions with now former PhD students and postdocs as well as with our peers in the plasmonics community. In particular, many of the figures are by courtesy of PhD theses by Giuseppe Toscano,[56] Søren Raza,[76] and Thomas Christensen.[74] We also thank Wei Yan, Christos Tserkezis, Mark Kamper, Christin David, Nicolas Stenger, Sanshui Xiao, Yu Luo, Alexander Uskov, Greg Sun, Uriel Levy, Lasse Jensen, Javier García de Abajo, Antti-Pekka Jauho, and Sergey Bozhevolnyi for long-term interactions and collaborations.

References

1. Maier, S. A. (2007). *Plasmonics: Fundamentals and Applications.* (Springer, New York).
2. Brongersma, M. L. (2015). Introductory lecture: nanoplasmonics, *Faraday Discuss* **178**, pp. 9–36.
3. Baev, A., Prasad, P. N., Ågren, H., Samoć, M. and Wegener, M. (2015). Metaphotonics: An emerging field with opportunities and challenges, *Phys. Rep.* **594**, pp. 1–60.
4. Gramotnev, D. K. and Bozhevolnyi, S. I. (2010). Plasmonics beyond the diffraction limit, *Nat. Photonics* **4**(2), pp. 83–91.
5. Gramotnev, D. K. and Bozhevolnyi, S. I. (2014). Nanofocusing of electromagnetic radiation, *Nat. Photonics* **8**(1), pp. 14–23.
6. Liz-Marzán, L. M. (2006). Tailoring surface plasmons through the morphology and assembly of metal nanoparticles, *Langmuir.* **22**(1), pp. 32–41.
7. Hao, E. and Schatz, G. C. (2004). Electromagnetic fields around silver nanoparticles and dimers, *J. Chem. Phys.* **120**(1), pp. 357–366.
8. Stockman, M. I. (2004). Nanofocusing of optical energy in tapered plasmonic waveguides, *Phys. Rev. Lett.* **93**(13), p. 137404.
9. Duan, H., Fernández-Domínguez, A. I., Bosman, M., Maier, S. A. and Yang, J. K. W. (2012). Nanoplasmonics: Classical down to the nanometer scale, *Nano Lett.* **12**(3), pp. 1683–1689.
10. Landau, L. D., Lifshitz, E. M. and Pitaevskii, L. P. (1984). *Electrodynamics of Continuous Media.* vol. 8, *Landau and Lifshitz Course on Theoretical Physics*, (Butterworth Heinemann, Oxford), 2 edition.
11. Ashcroft, N. W. and Mermin, N. D. (1976). *Solid State Physics.* (Saunders College Publishing, Fort Worth).
12. Romero, I., Aizpurua, J., Bryant, G. W. and García de Abajo, F. J. (2006). Plasmons in nearly touching metallic nanoparticles: Singular response in the limit of touching dimers, *Opt. Express* **14**(21), pp. 9988–9999.
13. Bloch, F. (1933). Bremsvermögen von Atomen mit mehreren Elektronen, *Z. Phys. A.* **81**(5–6), pp. 363–376.
14. Boardman, A. (1982). *Electromagnetic surface modes. Hydrodynamic theory of plasmon-polaritons on plane surfaces.* (John Wiley and Sons, Chichester).
15. Pitarke, J. M., Silkin, V. M., Chulkov, E. V. and Echenique, P. M. (2007). Theory of surface plasmons and surface plasmon polaritons, *Rep. Prog. Phys.* **70**, pp. 1–87.
16. García de Abajo, F. J. (2008). Nonlocal effects in the plasmons of strongly interacting nanoparticles, dimers, and waveguides, *J. Phys. Chem. C.* **112**(46), pp. 17983–17987.
17. McMahon, J. M., Gray, S. K. and Schatz, G. C. (2009). Nonlocal optical response of metal nanostructures with arbitrary shape, *Phys. Rev. Lett.* **103**, p. 097403.

18. David, C. and García de Abajo, F. J. (2011). Spatial nonlocality in the optical response of metal nanoparticles, *J. Phys. Chem. C.* **115**(40), pp. 19470–19475.
19. Raza, S., Toscano, G., Jauho, A.-P., Wubs, M. and Mortensen, N. A. (2011). Unusual resonances in nanoplasmonic structures due to nonlocal response, *Phys. Rev. B.* **84**, p. 121412(R).
20. Kern, J., Grossmann, S., Tarakina, N. V., Häckel, T., Emmerling, M., Kamp, M., Huang, J.-S., Biagioni, P., Prangsma, J. and Hecht, B. (2012). Atomic-scale confinement of resonant optical fields, *Nano Lett.* **12**(11), pp. 5504–5509.
21. Ciracì, C., Hill, R. T., Mock, J. J., Urzhumov, Y., Fernández-Domínguez, A. I., Maier, S. A., Pendry, J. B., Chilkoti, A. and Smith, D. R. (2012). Probing the ultimate limits of plasmonic enhancement, *Science* **337**(6098), pp. 1072–1074.
22. Savage, K. J., Hawkeye, M. M., Esteban, R., Borisov, A. G., Aizpurua, J. and Baumberg, J. J. (2012). Revealing the quantum regime in tunnelling plasmonics, *Nature* **491**(7425), pp. 574–577.
23. Scholl, J. A., Koh, A. L. and Dionne, J. A. (2012). Quantum plasmon resonances of individual metallic nanoparticles, *Nature* **483**(7390), p. 421.
24. Raza, S., Stenger, N., Kadkhodazadeh, S., Fischer, S. V., Kostesha, N., Jauho, A.-P., Burrows, A., Wubs, M. and Mortensen, N. A. (2013). Blueshift of the surface plasmon resonance in silver nanoparticles studied with EELS, *Nanophotonics* **2**(2), pp. 131–138.
25. Raza, S., Stenger, N., Pors, A., Holmgaard, T., Kadkhodazadeh, S., Wagner, J. B., Pedersen, K., Wubs, M., Bozhevolnyi, S. I. and Mortensen, N. A. (2014). Extremely confined gap surface-plasmon modes excited by electrons, *Nat. Commun.* **5**, p. 4125.
26. Tan, S. F., Wu, L., Yang, J. K., Bai, P., Bosman, M. and Nijhuis, C. A. (2014). Quantum plasmon resonances controlled by molecular tunnel junctions, *Science* **343**(6178), pp. 1496–1499.
27. Paria, D., Roy, K., Singh, H. J., Kumar, S., Raghavan, S., Ghosh, A. and Ghosh, A. (2015). Ultrahigh field enhancement and photoresponse in atomically separated arrays of plasmonic dimers, *Adv. Matter.* **27**(10), p. 1751.
28. Raza, S., Kadkhodazadeh, S., Christensen, T., Di Vece, M., Wubs, M., Mortensen, N. A. and Stenger, N. (2015). Multipole plasmons and their disappearance in few-nanometer silver nanoparticles, *Nat. Commun.* **6**, p. 8788.
29. Shen, H., Chen, L., Ferrari, L., Lin, M.-H., Mortensen, N. A., Gwo, S. and Liu, Z. (2017). Optical observation of plasmonic nonlocal effects in a 2d superlattice of ultrasmall gold nanoparticles, *Nano Lett.* **17**(4), pp. 2234–2239.
30. Raza, S., Bozhevolnyi, S. I., Wubs, M. and Mortensen, N. A. (2015). Nonlocal optical response in metallic nanostructures, *J. Phys.: Condens. Matter.* **27**, p. 183204.

31. Toscano, G., Raza, S., Jauho, A.-P., Mortensen, N. A. and Wubs, M. (2012). Modified field enhancement and extinction in plasmonic nanowire dimers due to nonlocal response, *Opt. Express* **20**(4), pp. 4176–4188.
32. Mortensen, N. A., Raza, S., Wubs, M., Søndergaard, T. and Bozhevolnyi, S. I. (2014). A generalized nonlocal optical response theory for plasmonic nanostructures, *Nat. Commun.* **5**, p. 3809.
33. Landau, L. (1946). On the vibrations of the electronic plasma, *Zhurnal Eksperimentalnoi Teor. Fiz.* **16**(7), pp. 574–586.
34. Uskov, A. V., Protsenko, I. E., Mortensen, N. A. and O'Reilly, E. P. (2014). Broadening of plasmonic resonances due to electron collisions with nanoparticle boundary: Quantum-mechanical consideration, *Plasmonics* **9**, pp. 185–192.
35. Khurgin, J. B. (2015). Ultimate limit of field confinement by surface plasmon polaritons, *Faraday Discuss.* **178**, pp. 109–122.
36. Yan, W., Wubs, M. and Mortensen, N. A. (2015). Projected dipole model for quantum plasmonics, *Phys. Rev. Lett.* **115**, p. 137403.
37. Mortensen, N. A. (2013). Nonlocal formalism for nanoplasmonics: phenomenological and semi-classical considerations, *Phot. Nanostr.* **11**(4), pp. 303–308.
38. Gallinet, B., Butet, J. and Martin, O. J. F. (2015). Numerical methods for nanophotonics: Standard problems and future challenges, *Laser Phot. Rev.* **9**(6), pp. 577–603.
39. Monaghan, J. (2005). Smoothed particle hydrodynamics, *Rep. Prog. Phys.* **68**(8), pp. 1703–1759.
40. Ginzburg, P. and Zayats, A. (2013). Localized surface plasmon resonances in spatially dispersive nano-objects: phenomenological treatise, *ACS Nano.* **7**(5), p. 4334.
41. Toscano, G., Raza, S., Yan, W., Jeppesen, C., Xiao, S., Wubs, M., Jauho, A.-P., Bozhevolnyi, S. I. and Mortensen, N. A. (2013). Nonlocal response in plasmonic waveguiding with extreme light confinement, *Nanophotonics.* **2**(3), pp. 161–166.
42. Wubs, M. (2015). Classification of scalar and dyadic nonlocal optical response models, *Opt. Express* **23**(24), pp. 31296–31312.
43. Yan, W., Wubs, M. and Mortensen, N. A. (2012). Hyperbolic metamaterials: Nonlocal response regularizes broadband supersingularity, *Phys. Rev. B.* **86**(20), p. 205429.
44. Jewsbury, P. (1981). Electrodynamic boundary conditions at metal interfaces, *J. Phys. F: Met. Phys.* **11**, p. 195.
45. Liebsch, A. (1993). Surface-plasmon dispersion and size dependence of Mie resonance: silver versus simple metals, *Phys. Rev. B.* **48**(15), pp. 11317–11328.
46. Monreal, R. C., Antosiewicz, T. J. and Apell, S. P. (2013). Competition between surface screening and size quantization for surface plasmons in nanoparticles, *New J. Phys.* **15**(8), p. 083044.

47. Teperik, T. V., Nordlander, P., Aizpurua, J. and Borisov, A. G. (2013). Robust subnanometric plasmon ruler by rescaling of the nonlocal optical response, *Phys. Rev. Lett.* **110**, p. 263901.
48. Christensen, T., Yan, W., Jauho, A.-P., Soljačić, M. and Mortensen, N. A. (2017). Quantum corrections in nanoplasmonics: Shape, scale, and material, *Phys. Rev. Lett.* **118**, p. 157402.
49. Ichikawa, M. (2011). Theory of localized plasmons for metal nanostructures in random-phase approximation, *J. Phys. Soc. Jpn.* **80**(4), p. 044606.
50. Öztürk, Z. F., Xiao, S., Yan, M., Wubs, M., Jauho, A.-P. and Mortensen, N. A. (2011). Field enhancement at metallic interfaces due to quantum confinement, *J. Nanophot.* **5**, p. 051602.
51. David, C. and García de Abajo, F. J. (2014). Surface plasmon dependence on the electron density profile at metal surfaces, *ACS Nano.* **8**(9), pp. 9558–9566.
52. Toscano, G., Straubel, J., Kwiatkowski, A., Rockstuhl, C., Evers, F., Xu, H., Mortensen, N. A. and Wubs, M. (2015). Resonance shifts and spill-out effects in self-consistent hydrodynamic nanoplasmonics, *Nat. Commun.* **6**, p. 7132.
53. Yan, W. (2015). Hydrodynamic theory for quantum plasmonics: Linear-response dynamics of the inhomogeneous electron gas, *Phys. Rev. B.* **91**, p. 115416.
54. Li, X., Fang, H., Weng, X., Zhang, L., Dou, X., Yang, A. and Yuan, X. (2015). Electronic spill-out induced spectral broadening in quantum hydrodynamic nanoplasmonics, *Opt. Express* **23**(23), pp. 29738–29745.
55. Ciracì, C. and Della Sala, F. (2016). Quantum hydrodynamic theory for plasmonics: Impact of the electron density tail, *Phys. Rev. B.* **93**, p. 205405.
56. Toscano, G. (2013). *Semiclassical theory of nonlocal plasmonic excitation in metallic nanostructures*. PhD thesis, Technical University of Denmark.
57. Raza, S., Toscano, G., Jauho, A.-P., Mortensen, N. A. and Wubs, M. (2013). Refractive-index sensing with ultrathin plasmonic nanotubes, *Plasmonics* **8**(2), p. 193.
58. Christensen, T., Yan, W., Raza, S., Jauho, A.-P., Mortensen, N. A. and Wubs, M. (2014). Nonlocal response of metallic nanospheres probed by light, electrons, and atoms, *ACS Nano.* **8**(2), pp. 1745–1758.
59. Yanai, A., Mortensen, N. A. and Levy, U. (2013). Absorption and eigenmode calculation for one-dimensional periodic metallic structures using the hydrodynamic approximation, *Phys. Rev. B.* **88**, p. 205120.
60. Hiremath, K. R., Zschiedrich, L. and Schmidt, F. (2012). Numerical solution of nonlocal hydrodynamic Drude model for arbitrary shaped nano-plasmonic structures using Nédélec finite elements, *J. Comp. Phys.* **231**(17), pp. 5890–5896.
61. Yan, W., Mortensen, N. A. and Wubs, M. (2013). Green's function surface-integral method for nonlocal response of plasmonic nanowires in arbitrary dielectric environments, *Phys. Rev. B.* **88**, p. 155414.
62. Chulhai, D. and Jensen, L. (2016), Generalized nonlocal optical response (Jun). www.nanoHUB.org, DOI: 10.4231/D3V97ZS7X.

63. Hughes, T. J. R. (1987). *The Finite Element Method — Linear Static and Dynamic Finite Element Analysis*. (Dower Publications, Inc., Mineola, New York).
64. NanoPlasmonics Lab, (2012). http://www.nanopl.org.
65. Li, L., Lanteri, S., Mortensen, N. A. and Wubs, M. (2017). A hybridizable discontinuous Galerkin method for solving nonlocal optical response models, *Comp. Phys. Commun.* **219**, p. 99. DOI: 10.1016/j.cpc.2017.05.012.
66. Stella, L., Zhang, P., García-Vidal, F. J., Rubio, A. and García-González, P. (2013). Performance of nonlocal optics when applied to plasmonic nanostructures, *J. Phys. Chem. C.* **117**(17), pp. 8941–8949.
67. Andersen, K., Jensen, K. L., Mortensen, N. A. and Thygesen, K. S. (2013). Visualizing hybridized quantum plasmons in coupled nanowires: From classical to tunneling regime, *Phys. Rev. B.* **87**, p. 235433.
68. Blaber, M. G., Arnold, M. D. and Ford, M. J. (2009). Search for the ideal plasmonic nanoshell: The effects of surface scattering and alternatives to gold and silver, *J. Phys. Chem. C.* **113**(8), pp. 3041–3045.
69. Rammer, J. and Smith, H. (1986). Quantum field-theoretical methods in transport theory of metals, *Rev. Mod. Phys.* **58**, pp. 323–359.
70. Kubo, R. (1966). The fluctuation-dissipation theorem, *Rep. Prog. Phys.* **29**(1), p. 255.
71. Kamper, M., Jauho, A.-P., Mortensen, N. A. and Tserkezis, C. (2017). Role of diffusive surface scattering in nonlocal plasmonics, *preprint*.
72. Kreibig, U. and Vollmer, M. (1995). *Optical Properties of Metal Clusters*. (Springer-Verlag, Berlin).
73. Esteban, R., Borisov, A. G., Nordlander, P. and Aizpurua, J. (2012). Bridging quantum and classical plasmonics with a quantum-corrected model, *Nat. Commun.* **3**, p. 825.
74. Christensen, T. (2017). *From classical to quantum plasmonics in three and two dimensions*. Springer Theses, (Springer International Publishing). doi: 10.1007/978-3-319-48562-1.
75. Scholl, J. A., Garcia-Etxarri, A., Koh, A. L. and Dionne, J. A. (2013). Observation of quantum tunneling between two plasmonic nanoparticles, *Nano Lett.* **13**(2), pp. 564–569.
76. Raza, S. (2014). *Probing plasmonic nanostructures with electron energy-loss spectroscopy*. PhD thesis, Technical University of Denmark.
77. Jin, D., Hu, Q., Neuhauser, D., von Cube, F., Yang, Y. Sachan, R., Luk, T. S., Bell, D. C. and Fang, N. X. (2015). Quantum-spillover-enhanced surface-plasmonic absorption at the interface of silver and high-index dielectrics, *Phys. Rev. Lett.* **115**, p. 193901.
78. Teperik, T. V., Nordlander, P., Aizpurua, J. and Borisov, A. G. (2013). Quantum effects and nonlocality in strongly coupled plasmonic nanowire dimers, *Opt. Express* **21**(22), pp. 27306–27325.
79. Tserkezis, C., Yan, W., Hsieh, W., Sun, G., Khurgin, J., Wubs, M. and Mortensen, N. A. (2017). On the origin of nonlocal damping in plasmonic monomers and dimers, *Int. J. Mod. Phys. B.* **31**, p. 1740005.

80. Javani, M. H. and Stockman, M. I. (2016). Real and imaginary properties of epsilon-near-zero materials, *Phys. Rev. Lett.* **117**, p. 107404.
81. David, C., Mortensen, N. A. and Christensen, J. (2013). Perfect imaging, epsilon-near-zero phenomena and waveguiding in the scope of nonlocal effects, *Sci. Rep.* **3**, p. 2526.
82. Lindhard, J. (1954). On the properties of a gas of charged particles, *Dan. Mat. Fys. Medd.* **28**(8), pp. 1–57.
83. Yan, W. and Mortensen, N. A. (2016). Nonclassical effects in plasmonics: An energy perspective to quantify nonclassical effects, *Phys. Rev. B.* **93**, p. 115439.
84. Sun, G., Khurgin, J. B. and Bratkovsky, A. (2011). Coupled-mode theory of field enhancement in complex metal nanostructures, *Phys. Rev. B.* **84**, p. 045415.
85. Khurgin, J. B. and Sun, G. (2015). Impact of surface collisions on enhancement and quenching of the luminescence near the metal nanoparticles, *Opt. Express* **23**(24), pp. 30739–30748.

CHAPTER 4

Quantum Effects in the Plasmonic Response

JAVIER AIZPURUA

Center for Materials Physics in San Sebastian (CSIC-UPV/EHU), Spain
aizpurua@ehu.es

ANDREI G. BORISOV

L'Institut des Sciences Moléculaires d'Orsay,
CNRS-Université Paris Sud, France
andrei.borisov@u-psud.fr

Excitation of localized surface plasmon resonances in metal nanoparticles and nanoparticle assemblies allows to manipulate light at nanometer scales well beyond the wavelength and opens the possibility of many applications. The complexity of the variety of methods to fabricate plasmonic structures and the sophistication of optical characterization techniques invites for completing the plasmonics landscape with a comprehensive theoretical understanding of the optical response. Classical approaches based on the solution of Maxwell's equations using local (model or empirical) bulk dielectric constant are extensively used to describe plasmonic systems and gain the necessary intuition. At the same time, recent experimental and theoretical developments have clearly demonstrated that as soon as nanometer and subnanometer scales are approached, quantum effects such as tunneling, non-local screening and the atomistic structure of the metal nanoparticles become important. In this chapter, we discuss the major manifestations of quantum effects

in the properties of localized plasmons, with particular emphasis on the underlying physics.

4.1. Introduction

Interaction of light with metallic nanoparticles can be greatly enhanced if the incident electromagnetic wave is at resonance with collective electron excitations, so called plasmons. Known for a long time in surface science, where intense research on plasmons has been developed in the context of the interpretation of electron energy loss spectroscopy (EELS) experiments,[1] and studied in detail in the context of cluster physics,[2] nowadays plasmonic properties of nanostructured materials are receiving renewed attention as a key building block in nanophotonics.[3–5] Driven by technological progress, plasmonics is now a flourishing field of fundamental and applied research as it allows "to manipulate light at the nanoscale".[3] Indeed, advances in nanofabrication[6,7] allow to engineer the plasmonic response of systems and fully exploit strongly enhanced near fields and scattering resonances in far fields produced by plasmonic nanoparticles. [3–5,8–12] While noble metals such as gold and silver are the most common materials in infrared plasmonics, recent reports address the possibility of using Al nanoparticles with the perspective to exploit their plasmon resonances in the ultraviolet.[13]

The design of the optimal nanostructure with a specific optical response in the near or far fields requires accurate theoretical predictions. Thus, the availability of theoretical tools capable of modelling, predicting, describing, and interpreting experimental data at formal and numerical levels is of paramount importance. Most of the formal approaches and numerical techniques developed to obtain the optical response of a nanoobject root back to the classical macroscopic Maxwell's equations.[5,7,14] The non-dispersive local dielectric permittivity is routinely used to describe the optical response of metallic nanoparticles delimited by sharp boundaries. It is only recently that technology has allowed to engineer plasmonic nanostructures where quantum aspects of the motion of conduction electrons significantly affect the optical response.[15–19] Basically, observing the quantum regime of response implies that the characteristic size(s) of the system

is at the nm scale, comparable to the Fermi electron wavelength in the metal. Thus, several studies have been performed on quantum effects in the energy and lifetime of the collective plasmon excitations in small metal clusters,[2,20–23] as well as on photon-assisted electron tunneling through narrow gaps in semiconductor and metal nanostructures in the THz range.[24,25] Current research on quantum effects in plasmonics at optical frequencies is mainly focused on two aspects:

(1) Non-locality of the screening, where the spatial profile of the induced electron density cannot be infinitely sharp, as assumed in classical theories.
(2) The coupling between the electromagnetic field and electron tunneling across narrow junctions between metal nanoparticles, where the conductive contact can be established prior to the direct geometrical overlap.

The above-mentioned quantum effects are not captured within standard classical approaches, and so the description of the corresponding plasmonic structures requires new developments. The non-local screening can be modelled within a classical electromagnetic theory framework within the non-local hydrodynamical model[26] (NLHD) allowing to partially lift such non-physical results as e.g. the divergence of near fields upon narrowing the gap between plasmonic nanoparticles. This theory has recently been greatly improved to include the quantum spill out of the electron density outside metal boundaries.[27] For a detailed account of the NLHD theory, we refer to Chapter 3 by N. A. Mortensen *et al.* The description of electron tunneling requires a quantum treatment beyond classical electrodynamics. Such approaches become available nowadays, and the optical response of plasmonic nanostructures can be calculated either extending the classical theory to incorporate electron tunneling, or by adopting a full quantum treatment using e.g. the time dependent density functional theory (TDDFT) combined with *ab initio* atomistic or model free-electron descriptions of metal nanoparticles. [28–39] Besides properly accounting for electron tunneling, the TDDFT calculations reveal a sensitivity of the near and far fields to the fine details of the atomistic structure inside the gap between plasmonic

nanoparticles, such as vertices and edges at contact of different crystallographic planes.

Later in this chapter, we present a detailed account of the theoretical developments based on the full quantum TDDFT calculations and discuss related recent experiments. It is noteworthy that owing to the extremely strong field enhancement in the narrow gap between plasmonic nanoparticles, the plasmonic dimer appears as the system of choice to address quantum effects and will be often quoted in this chapter.

4.2. Non-local Screening

Non-locality of the screening results from the collective quantum behavior of the interacting many-electron system and reflects the fact that the response of the conduction electrons at a particular point in space depends not only on the field applied at that point but also on the fields applied at other different points. For a monochromatic incident electromagnetic wave with frequency ω, one can write $\mathbf{P}(\mathbf{r},\omega)$, the polarization of the medium at point $\mathbf{r}$, as:

$$\mathbf{P}(\mathbf{r},\omega) = \int \mathrm{d}\mathbf{r}'\chi(\mathbf{r},\mathbf{r}',\omega)\mathbf{E}(\mathbf{r}',\omega), \tag{4.1}$$

where $\mathbf{E}(\mathbf{r}',\omega)$ is the electric field, and $\chi(\mathbf{r},\mathbf{r}',\omega)$ is the susceptibility of the medium.

Contrary to local classical descriptions, the non-local effect and the spill out of the valence electron density outside the metal prevents a sharp localization of the screening charges at the boundaries of the interfaces. The screening charges thus have finite spatial extension, as we illustrate in Fig. 4.1, where the results of the TDDFT calculation of the electron density dynamics triggered by an external laser pulse incident at a sodium nanowire dimer in vacuum are shown.[33,34] Sodium is a typical free-electron metal for which the jellium metal (JM) approximation applies well. Within the JM model, the ionic cores are not explicitly introduced into the calculation, but they are replaced by a uniform positive background charge with density $n_+ = [4\pi r_\mathrm{s}^3/3]^{-1}$. The Wigner–Seitz radius is $r_\mathrm{s} = 2.12$ Å for Na. The

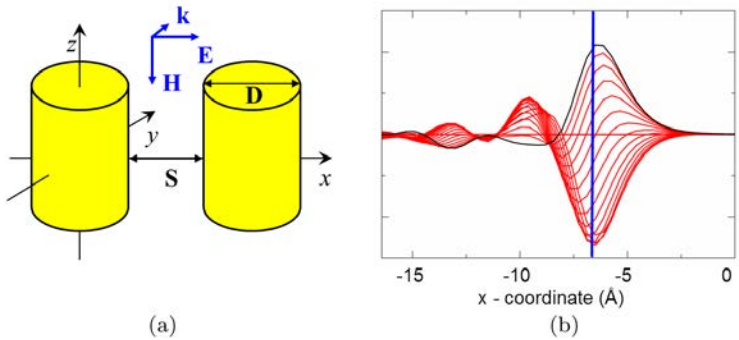


Fig. 4.1. Dynamical screening. (a) Sketch of the geometry of a nanowire dimer. Two identical cylindrical nanowires are infinite along z-axis and have a diameter $D = 2R_{\rm cl}$. The nanowires are separated by a gap of width S. The incident radiation is linearly polarized with the electric field along the dimer axis (x-axis). (b) Evolution of the density induced by an incident field at the surface of the left cylinder facing the junction. The data are shown as a function of the x-coordinate for $D = 9.8\,\text{nm}$, and $S = 1.3\,\text{nm}$. $x = 0$ is at the center of the junction. Different curves correspond to instants of time spanning 1/2 optical period starting from t_0 (black curve). The frequency of the incoming x-polarized electromagnetic wave is in resonance with the bonding dipole plasmon mode at $\omega_{\rm dp} = 3.157\,\text{eV}$. The vertical blue line denotes the position of the jellium edge.

boundary between the background charge density n_+ and surrounding medium is set at half of the spacing between atomic planes in the direction of the surface. This boundary, also called "jellium edge", will be considered here as the geometrical surface of the metal.

While in the local classical approach the screening charges reside at the geometrical surfaces of the cylinders separated by the junction of width S, the shape of the induced charge density obtained in TDDFT calculations shows a pronounced time-dependence in the finite spatial profile. In particular, because of the spill-out effect, the induced charge density extends outside the metal boundaries by some Å. One can then define the real part of the centroid of the induced charge density (Feibelman parameter) $\text{Re}[d(\omega)]$ as a measure of the actual position of the plasmon-induced surface charges with respect to the jellium edge.[20,21,40–42] The Feibelman parameter plays a central role in the theory of dynamic screening at surfaces and allows to describe the dispersion of propagating surface plasmons

and finite size effects in localized nanoparticle plasmons, as we will show below. For the present case of sodium nanowires, $\mathrm{Re}[d(\omega)]$ is positive $\approx$0.9 Å in the frequency range of interest.[42] This means that the screening charge is shifted by 0.9 Å into the vacuum and located at $\approx$3 Å outside the surface atomic plane of the nanoparticle.

One of the most prominent manifestations of the non-local screening is the nanoparticle size-dependent shift of the frequency ω_{dp} of the localized nanoparticle plasmon with dipolar character, as has been thoroughly studied in the case of extinction resonances in cluster physics. [2,20,23,41,43] For nanoparticles characterized by a radius R_{cl}, comparable with the wavelength of the incident radiation λ, retardation effects lead to a redshift of the localized dipolar plasmon frequency. As the size of the nanoparticles decreases, the solution of classical Maxwell's equations predicts that ω_{dp} converges to a constant value ω_{Mie} (Mie frequency) as $\omega_{\mathrm{dp}} = \omega_{\mathrm{Mie}} \left(1 - A[R_{\mathrm{cl}}/\lambda]^2\right)$,[44] where A is a constant. However, for nanoparticles of nanometer scale size, the classical description is no longer applicable, and ω_{dp} is determined by the electron many-body screening at the nanoparticle surface.[20,41]

$$\omega_{\mathrm{dp}}/\omega_{\mathrm{Mie}} = 1 - \mathrm{Re}[d(\omega_{\mathrm{dp}})]/R_{\mathrm{cl}} + O(R_{\mathrm{cl}}^{-2}). \tag{4.2}$$

For alkali metals and aluminum, $\mathrm{Re}[d(\omega_{\mathrm{dp}})]$ is positive. Thus, with decreasing cluster size, the plasmon resonance redshifts from the Mie value ω_{Mie} for small size clusters.[45–49] For noble metals such as Au and Ag, the final size effects lead to a *blue shift* of the dipolar plasmon resonance with decreasing cluster size. [16,43,50–54] The sheer contrast between alkali and noble metals is due to the localized *d*-electron contribution to the total screening in the latter case.[41] The *d*-electrons reside close to the atomic planes beyond the jellium edge determined by the extension of the wave functions of the delocalized *s-p* electrons. When the *d*-electron contribution to the screening is accounted for, $\mathrm{Re}[d(\omega)]$ turns negative indicating that the screening charge is predominantly induced inside the metal.

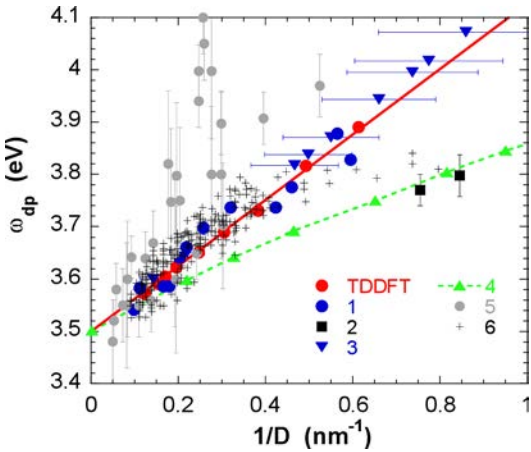


Fig. 4.2. Plasmon energy of Ag clusters as a function of the inverse of the cluster diameter $1/D$. The dipolar bulk limit (Mie plasmon) is at 3.5 eV. The experimental data (1–5) have been scaled to obtain the resonance energies in vacuum.[43] The horizontal error bars on the data (3) give the widths of the size distributions of the corresponding Ag particles. Red line with red circles shows the results of the TDDFT calculations[55] using the jellium model with a polarizable background allowing to account for the *d*-electron contribution. Blue circles (1) — experimental data from Ref. [50]; Black squares (2) — experimental data from Ref. [51]; inverted blue triangles (3) — experimental data from Ref. [52]; green triangles (4) — experimental data from Ref. [53]; gray circles (5) — experimental data from Ref. [16]; (6) — data obtained in the STM induced light emission experiments.[54]

The difference between free electron (alkali and Al) and noble metals is illustrated with Figs. 4.2 and 4.3. Measured and calculated cluster size dependence of the dipolar plasmon resonance frequency is shown for silver and sodium clusters, respectively. For the theoretical calculations, the jellium model of the metal has been used in both cases. In the TDDFT study of the optical response of Ag clusters, the polarizable background with radius $R_{\rm d}$ has been introduced inside the cluster to model the contribution of the localized *d*-electrons to the screening.[36,56] The polarizable background is described by the non-dispersive dielectric function $\varepsilon_d = 4.58$, and it has a radius defined as $R_{\rm d} = R_{\rm cl} - 1.2\,\text{Å}$. As follows from the Figs. 4.2 and 4.3, and in agreement with the general discussion based on the Feibelman parameter, Ag and Na clusters display opposite finite size effect. It is

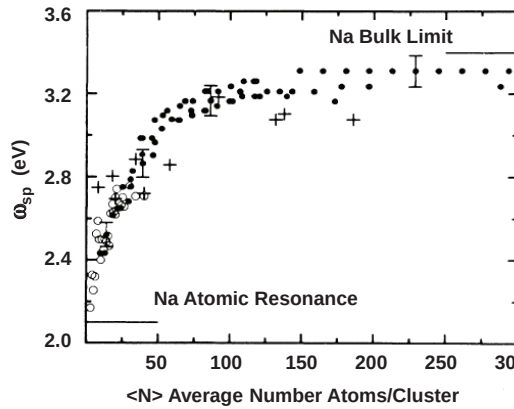


Fig. 4.3. From Ref. 47, energy of the dipolar resonance of Na clusters as a function of the average number of atoms $< N >$ forming the cluster. Frequencies are derived from experimental data (solid and open circles) and compared with theoretical calculations[45] (circles) based on the free-electron description of Na.

worthwhile to stress the good agreement between the experimental data and the theoretical results based on the jellium description of the metal. This indicates that the JM captures the main physics of the screening as recently confirmed by quantum studies that adopt a full atomistic description of the metal nanoparticles.[37,57–61] As appears from the discussion above, the dynamical surface screening is the main mechanism determining the size dependence of the energy of the localized plasmons in small metallic spherical nanoparticles. In a recent work, R. C. Monreal, T. J. Antosiewicz and S. P. Apell show that for quantum confinement effects linked with the quantization of the electronic states to become noticeable the nanoparticle size has to be in the 1 nm range and below.[23]

Dynamic screening not only affects the energies of the plasmon modes in individual nanoparticles, but it also does affect plasmon hybridization in adjacent metallic nanostructures. The plasmonic dimer,[33,34] a canonical coupled structure, allows for a clear cut physical interpretation of quantum effects in the optical response. Fig. 4.4 presents the analysis of the role played by dynamic screening in determining the frequency of the bonding dipole plasmon,

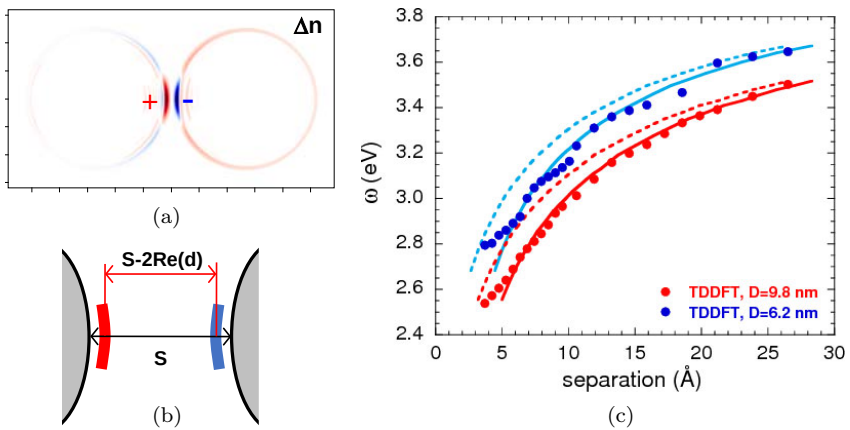


Fig. 4.4. Dynamic screening in a cylindrical nanowire dimer with $D = 9.8\,\text{nm}$ and $S = 0.8\,\text{nm}$, as sketched in Fig. 4.1. (a) The TDDFT result for the charge density Δn induced by the incident x-polarized laser pulse at resonance with the lowest bonding dipole plasmon mode of the system ($\omega_{\text{dp}} = 3.1\,\text{eV}$). Positive (negative) values correspond to red (blue) color code. Δn is shown at the instant of time corresponding to the maximum dipole moment of the dimer. (b) Zoom into the junction with schematic representation of the location of the plasmon-induced screening charges. Within the local classical approach, the screening charges are exactly located at the geometrical surfaces of the cylinders (here equivalent to the jellium edges) separated by the junction of width S. As calculated with TDDFT, the centroids of the screening charges (red area for positive charge and blue area for negative charge) are located at $\text{Re}[d(\omega)]$ in front of the jellium edges, and therefore separated by $S - 2\text{Re}[d(\omega)]$. (b) Energy of the dipole plasmon resonance as a function of the junction width S. Dots: TDDFT results obtained for nanowire dimers formed by $D = 6.2\,\text{nm}$ (blue) and $D = 9.8\,\text{nm}$ (red) individual nanowires. The results of classical Drude calculations for $D = 6.2\,\text{nm}$ and $D = 9.8\,\text{nm}$ dimers are shown, respectively, with blue and red lines. Dashed/solid line type: separation S is measured between the jellium edges/centroids of the induced charges. For more details, see the main text.

also known as gap plasmon. Panel (a) shows the TDDFT result for the charge density induced by a x-polarized laser pulse incident on the $D = 9.8\,\text{nm}$, $S = 0.8\,\text{nm}$ nanowire dimer sketched in Fig. 4.1. The incident radiation is at resonance with the bonding dipolar plasmon of the dimer formed by the hybridization of the plasmon dipolar modes of the individual nanowires. The systems with narrow gaps are characterized by a strong field enhancement in the junction because

of the large charge densities of opposite sign induced by the plasmon at the facing metal surfaces across the gap [see panel (a)]. In turn, this results in an attractive interaction leading to the redshift of the energies of the bonding plasmon modes of the dimer as compared to those of the individual particles. In the classical theory, the field enhancement and the redshift diverge for a vanishing width of the junction $S \to 0$.[62] The "zero width catastrophe" is removed when the electron tunneling is accounted for as we discuss in the next section. At present, we limit the discussion to $S > 0.5\,\text{nm}$, where the tunneling effect is negligible, and the induced charge densities can be well assigned to the individual cylinders forming the junction.

We recall that within TDDFT, the real part of the centroid of the induced charge density $\text{Re}[d(\omega)]$ corresponds to the position of the plasmon-induced surface charges with respect to the jellium edge of each cylinder. For the present case of sodium nanowires, $\text{Re}[d(\omega)] \approx 0.9\,\text{Å}$ with screening charge shifted from the jellium edge into the vacuum. As we schematically show in panel (b) of Fig. 4.4, for a junction with width S measured between the jellium edges, the actual distance between the plasmon-induced charges across the junction is $\Sigma = S - 2\text{Re}[d(\omega)]$. This insight provides an explanation of the dependence of the energies of the bonding dipolar plasmon as a function of gap width, as shown in Fig. 4.4(c) for two different nanowire dimers.[34] While classical electromagnetic calculations reproduce the TDDFT data overall, the latter is systematically redshifted with respect to the former. Indeed, in the classical theory, the screening charges are confined at the metal surfaces separated by the junction of width S, while in the quantum theory, the actual distance between the screening charges Σ is smaller leading to a stronger interaction (dots versus dashed lines in Fig. 4.4(c)). Thus, the TDDFT data obtained for a junction with a specific width S (dots) should be compared with the classical results calculated for a gap of effective width $\Sigma = S - 2\text{Re}[d(\omega)]$ (solid lines). In doing so, the agreement between TDDFT and the classical simulations becomes excellent indicating that in this case the effect of the non-local screening can be accounted for within the classical theory by simple distance

rescaling.[33,34] The distance rescaling provides an estimate of the characteristic gap sizes S, for which the non-local screening is important. Indeed, the shift of the plasmon energy from the classical value starts to be noticeable for $2\mathrm{Re}[d(\omega)]/S \sim 0.1$. Provided that the Feibelman parameter $\mathrm{Re}[d(\omega)]$ is in the Å range for typical metals, we obtain $S \leq 2-3\,\mathrm{nm}$. Note that for small widths of the gap $S < 0.5\,\mathrm{nm}$, the electron densities of the individual nanowires overlap indicating the onset of efficient tunneling which then dominates the optical response (see next section).

In addition to an explanation of the change of energies of the localized nanoparticle plasmons, the dynamic surface screening theory can be also applied to describe the cluster size dependence of the excited electron dynamics in the system. Indeed, via many-body interactions, the collective electron excitation (plasmon) decays into an incoherent excitation of electron–hole pairs. [2,21,22,56,63–66] The surface contribution to this Landau damping mechanism can be described via the imaginary part, $\mathrm{Im}[d(\omega)]$, of the Feibelman parameter.[21] For the sake of clarity, we will use much simpler phenomenological description here.[65] The width of the localized plasmon resonance Γ (inverse of the lifetime of the collective mode $\tau = 1/\Gamma$) is given by several contributions:

$$\Gamma = \Gamma_{\mathrm{rd}} + \Gamma_0 + g_{\mathrm{s}}\, v_{\mathrm{f}}/R_{\mathrm{cl}}. \tag{4.3}$$

The first term, $\Gamma_{\mathrm{rd}} \propto R_{\mathrm{cl}}^3$, stands for the radiative decay rate.[66,67] The second term gives the bulk contribution to the Landau damping and the last term describes the surface contribution. The latter mainly introduces a decay rate which is proportional to the frequency of the collisions with the surface $v_{\mathrm{f}}/R_{\mathrm{cl}}$ of a classical electron moving with Fermi velocity v_{f}. As follows from Eq. (4.3), for large nanoparticle sizes, radiative decay dominates; however, for small nanoparticles the incoherent electron–hole excitation becomes a leading mechanism with the surface contribution increasing as $1/R_{\mathrm{cl}}$. This is further illustrated in Fig. 4.5, where it can be observed that as a silver nanoparticle size decreases, the spectral width, Γ, of the plasmon resonance shows a transition from a radiative decay regime to a regime of decay via electron hole-pair excitation.

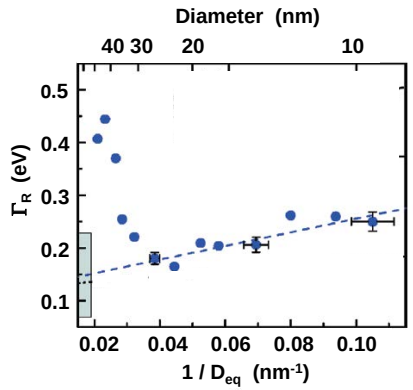


Fig. 4.5. From Ref. 65, spectral width Γ of the localized surface plasmon resonance measured in different single Ag/SiO_2 nanoparticles as a function of the inverse of their equivalent diameter D_{eq}, optically determined by fitting the extinction spectra. Error bars on size determination and width are indicated for three typical sizes. The blue dashed line is a linear fit using Eq. (4.3) and assuming negligible radiative decay $\Gamma_{rd} = 0$, with $g_s = 0.7$ and $\Gamma_0 = 0.125\,eV$. The shaded area indicates the estimated Γ_0 value in bulk silver.

While transfer of the electromagnetic energy trapped by the localized plasmon to the electron–hole excitations in the metal is one of the main bottlenecks in using plasmonic materials for information technology, this effect is extremely beneficial for many applications induced by hot electrons such as local heating, photovoltaics, nanoscale detectors and photochemistry at surfaces.[10,68–71]

4.3. Electron Tunneling

Electron tunneling through the potential barrier separating classically allowed regions of motion is one of the prominent quantum features that cannot be captured by the classical theory. This effect forms the basis of a powerful method of surface analysis, namely scanning tunneling microscopy (STM),[72] where a tunneling current across the junction formed by a STM tip and a surface is triggered by an applied external DC bias. In the AC regime, photon-assisted tunneling in metal–dielectric–metal or semiconductor junctions has been studied in the THz range.[24,25] However, it is only recently that

full quantum theories and experimental studies performed with well-controlled junctions addressed tunneling phenomena in plasmonics at optical frequencies.[15–19,28,30–38,73–79,82–84] Electron tunneling, or, in broader terms, electron transfer between nanoparticles and its interaction with photons coupled to plasmonic excitations is at the origin of many physical phenomena in narrow gaps. To access the tunneling regime in plasmonics allows (i) to understand the transition from separated nanoparticles to touching particles, when the gap sizes are of the order of the distances between atomic planes in the metal, (ii) to design nanoscale electro-optical devices capable of either generating electronic currents with incident photons or producing electroluminescence due to inelastic tunneling of electrons, and (iii) to develop strategies for controlling the plasmon response such as those offered by functionalization of the junction with self-assembled molecular layers (SAMs).

Because of the difficulty to reach reproducible sub-nanometer gaps in experimental conditions, the tunneling effect in plasmonics was first studied in nanoparticle dimers theoretically.[28,30,31] These theoretical predictions have been followed by experimental verification of the tunneling regime.[15–19,73–76] In Fig. 4.6 we show the optical response of a $D = 9.8\,\text{nm}$ jellium Na nanowire dimer in vacuum (Fig. 4.1). The results are obtained using different, classical, quantum and mixed, approaches not allowing or allowing to account for the electron transport through plasmonic gap.[34] Similar results have been reported for a dimer formed by two spherical nanoparticles represented using the free-electron JM,[28,30,31] or described at full atomistic level.[37] Thus, one can identify general and robust features of the modification of the optical response due to the electron transfer between nanoparticles.

A comparison between the quantum TDDFT results and the results from the classical electromagnetic calculations reveals several important differences as the junction width S is decreased:

- The redshift of the energies of the plasmon modes as calculated with TDDFT is finite, not diverging as in the classical theory.

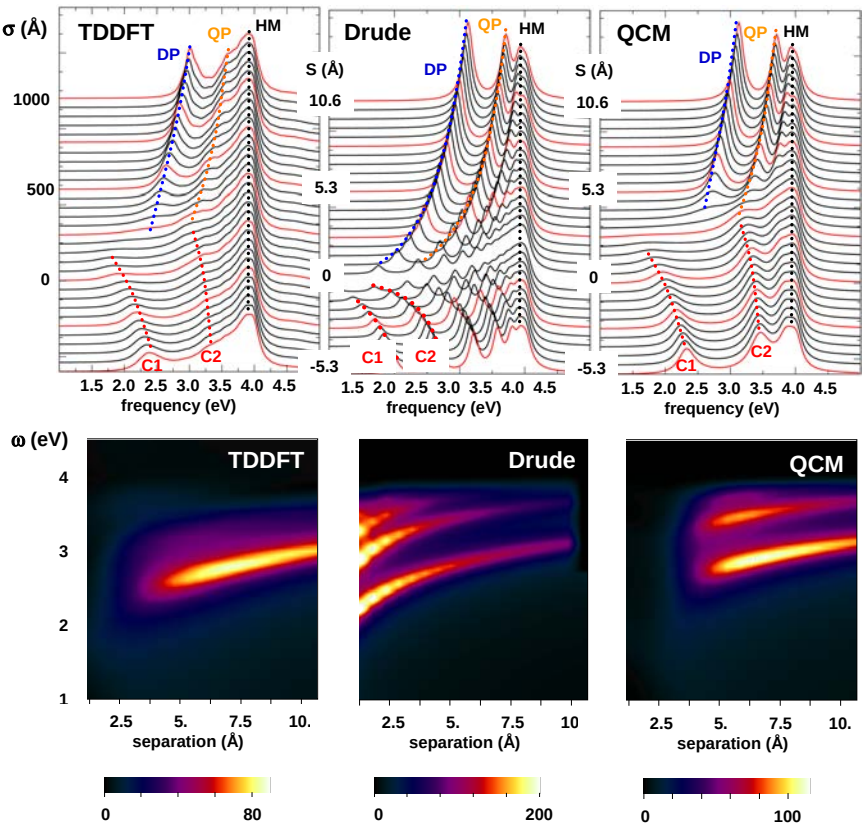


Fig. 4.6. Optical response of a Na nanowire dimer in vacuum. The diameter of each nanowire is $D = 9.8\,\text{nm}$, as depicted in Fig. 4.1. Results are obtained with the TDDFT calculations (TDDFT), with classical Drude electromagnetic calculations (Drude), and with the quantum corrected model (QCM). The latter allows to account for electron transport across the junction within the framework of classical electrodynamics. The incoming plane wave is polarized along the dimer axis x. Upper panels: Waterfall plots of the dipole absorption cross-section per unit length, σ, as a function of the width of the junction S and frequency ω of the incident radiation. The centers of the wires are at $x = \pm(D+S)/(2)$, and S is negative for overlapping cylinders. For clarity, a vertical shift proportional to S is introduced for each absorption spectrum. The red curves are used for a change of S every 2.65 Å. These curves are labeled with the corresponding S-values every 5.3 Å. The plasmonic modes responsible for the peaks in the absorption cross-section are labeled: Bonding Dipole Plasmon (DP), Bonding Quadrupole Plasmon

Fig. 4.6. (*caption continued*) (QP), high-order hybridized mode (HM), lowest (dipole) Charge Transfer Plasmon (C1), and higher energy Charge Transfer Plasmon (C2). Lower panels: Color plots of the local field enhancement at the center of the junction for positive separations. Results are shown as a function of the frequency ω of the incident radiation and junction width S. The color code is displayed at the bottom of the corresponding panels.

- The number of plasmon resonances is much smaller in the TDDFT results.
- The bonding plasmon modes of the separated dimer disappear, and the charge transfer plasmon modes[89] emerge in the TDDFT results prior to direct geometrical contact between the nanowire surfaces at $S = 0$.
- The field enhancement in the middle of the junction, as obtained with TDDFT, stays finite and it is quenched for small S in sheer contrast with the divergence of the fields observed from the classical description.

All over, the transition from a capacitively ($S > 0$) to a conductively ($S \leq 0$) coupled system is not discontinuous as in classical calculations, but continuous. This is because the electron tunneling between metal surfaces separated by the narrow gaps leads to the establishment of a conductive contact across the junction prior to the direct geometrical touch between the surfaces. Indeed, in the local classical calculations the divergence of the shifts of the plasmon modes energies and the field enhancement at vanishing gap width results from the divergent coulomb interaction between the screening charges of opposite sign localized at the geometrical surfaces across the junction. In the quantum calculations, however, the tunneling current neutralizes the plasmon-induced charges and thus reduces the field enhancement and the shift of the plasmon modes energies. This tunneling current also leads to the progressive attenuation of the bonding plasmon modes of the dimer formed by the hybridization of the plasmon modes of the individual nanoparticles. From the point

of view of the underlying physics, accounting for the ability of conduction electrons to tunnel through the potential barrier separating two nanoobjects allows to correctly capture the transition from the the situation of disconnected nanoparticles to the situation with the continuous metal bridge linking the junction.

Recently, quantum effects in the optical response of metallic nanoparticles have been addressed with computationally involved calculations based on a full atomistic description of the nanoparticles.[37–39] Undoubtedly, the atomistic structure of the nanoparticles leads to a quantitative dependence of the near fields and optical cross-sections on particular morphological aspects of the surfaces, e.g. specific surface facets forming the junction. The formation of atomic-size hot spots at the vertices and edges between crystallographic planes has been also reported.[39,85] Nevertheless, the general trends of the plasmon modes evolution in narrow gaps derived from TDDFT calculations within the JM description of the nanoparticles hold.

More insights into the effect of electron tunneling across the junction can be gained by analyzing the dynamics of currents and charges in the nanowire dimer induced by an incident x-polarized laser pulse resonant with the lowest (DP or C1) plasmonic modes, as displayed in Fig. 4.7. For a wide junction with $S = 7.95\,\text{Å}$, the total induced dipole is formed by the in-phase dipolar polarization of each nanocylinder, as expected for the bonding DP mode. The maximum of the dipole is reached at the instant of time when the maximum charge separation has occurred and the currents inside the nanoparticles are minimal. In that situation, high charge densities are induced at the surfaces facing the junction, resulting in large electric field enhancement. The structure of the induced charges and fields is similar to previous classical results for coupled cylinders[86,87] or metal spherical dimers.[62] For such large separations, the probability of tunneling between nanowires is negligible. When the junction width is reduced to $S = 2.65\,\text{Å}$, basically the same profiles for induced densities and fields are obtained. However, in this situation, optically-assisted AC tunneling current between the nanoparticles can now be

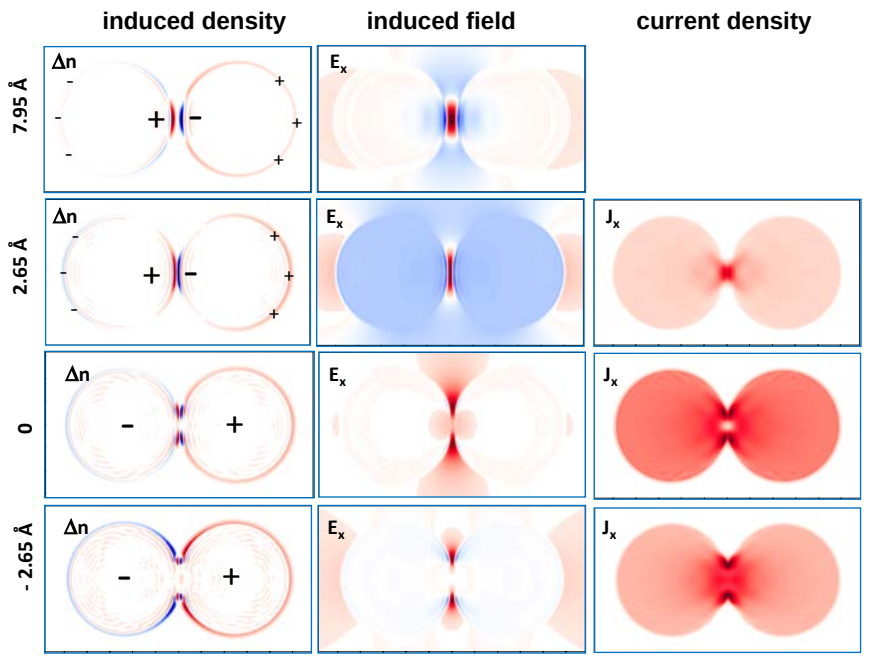


Fig. 4.7. Plasmon dynamics in a $D = 9.8$ nm nanowire dimer system. The incident x-polarized laser pulse is at resonance with the lowest bonding dipolar plasmon DP for $S > 0$, or with the lowest charge transfer plasmon C1 for $S \leq 0$. Panels in the left column: snapshots of the induced charge density Δn; middle column: snapshots of the electric field along the x-axis E_x; right column: current density J_x along the x-axis. Positive (negative) values correspond to the red (blue) color code. Results within each row correspond to a fixed junction width S, as indicated to the left of each row. The induced densities are shown at the instant of time corresponding to the maximum dipole moment of the dimer. The induced currents and fields are shown at the instants of time when the induced fields in the junction reach the maximum. J_x is not shown for the $S = 7.95$ Å because the tunneling current is negligible in that case.

observed.[24,25,80,81] The junction shows a *resistive* character[24] with the maximum tunneling current between nanowires reached at the maximum field and consequently at the maximum induced dipole. Because of the tunneling, the field enhancement in the junction is reduced (see Fig. 4.6).

Further reduction of the junction width S increases the tunneling probability and short-circuits the junction. When the conductive contact is formed, the fields are expelled from the junction,[32,88] the DP mode disappears and the charge transfer modes emerge in the absorption spectrum. In this respect, the results obtained for $S = 0$ and $S = -2.65\,\text{Å}$ are very similar. The maximum dipole moment of the dimer now corresponds to oppositely charged nanowires. The maximum currents and fields are reached when the total dipole moment of the system is minimum. Indeed, the current flowing trough the junction builds the dipole polarization with opposite charges at the left and right nanowires.

The theory of electron tunneling in the presence of a DC electromagnetic field[24,25,80,81] provides a formal framework that allows to describe the evolution of the plasmon response in nanoparticle assemblies with narrow gaps. Assuming linear response regime, the optical field E_ω at frequency ω induces a tunneling current at the same frequency in the junction between two flat metal surfaces. The dissipative component (in phase with the driving field) of this AC tunneling current density is given (in atomic units) by:[24,83]

$$J_\omega = \frac{V_\omega}{2\omega}\left[J_{\text{DC}}(U+\omega) - J_{\text{DC}}(U-\omega)\right], \tag{4.4}$$

where U is the applied DC bias, $V_\omega \simeq \text{SE}_\omega$ is the (small) optical bias across the junction, and $J_{\text{DC}}(U \pm \omega)$ is the DC tunneling current at the constant bias $U \pm \omega$. Here, we assume that E_ω does not vary across the junction. Equation (4.4) establishes a connection between the AC tunneling characteristics of the junction and the static DC tunneling characteristics, commonly used in the description of STM.[81] For slow variation of J_{DC} with U, we obtain the classical limit $J_\omega = \text{SE}_\omega\, \text{d}J_{\text{DC}}(U)/\text{d}U$, which can be written as

$$\begin{aligned} J_\omega &= G(U,S)V_\omega = \sigma(U,S)E_\omega, \\ G(U,S) &= \text{d}J_{\text{DC}}(U)/\text{d}U, \\ \sigma(U,S) &= SG(U,S), \end{aligned} \tag{4.5}$$

where $G(U,S)$ is the DC conductance of the junction at bias U, and $\sigma(U,S)$ is the DC condutivity of the junction. One expects that the tunneling current starts affecting the plasmon modes when it allows partial neutralization of the plasmon-induced screening charge densities ϱ at the metal surfaces across the junction.[89] This implies that $J_\omega \simeq \mathrm{d}\varrho/\mathrm{d}t$. Assuming $J_\omega = G(U,S)\mathrm{SE}_\omega$, and $E_\omega = 4\pi\varrho$ (flat capacitor), the harmonic time dependence of the fields results in $G(U,S)\mathrm{SE}_\omega \simeq \omega E_\omega/4\pi$. Thus, the value of the conductance of the junction at which tunneling effects are noticeable on the plasmon modes can be estimated from:

$$G(U,S) = \omega/4\pi S. \tag{4.6}$$

For a junction of width of 0.5 nm, and $\omega \approx 3\,\mathrm{eV}$, Eq. (4.6) results in $G(U,S) \approx 2\times10^{-3}G_0$, where $G_0 = 7.75\times10^{-5}\,\mathrm{S}$ is the quantum of conductance. This estimate is consistent with the range of conductance values reported for the onset of quantum effects on the bonding plasmon modes of dimer structures within the TDDFT calculations in the absence of applied bias ($U = 0$). [30–34,38,39,90] Since $G(U,S)$ increases exponentially with decreasing junction width, a variation of S is the most efficient way to enter into the quantum regime. However, a modification of the externally applied bias might provide an additional degree of freedom to tune the optical response of the nanostructure, together with the functionalization of the plasmonic gaps with molecular layers, as discussed below.

The approach expressed by Eqs. (4.4) and (4.5) is based on the theory of photon-assisted tunneling, it describes the resistive AC tunneling currents, and thus it can be applied in the case of the junctions with low tunneling probabilities and resistive character. This approach can thus explain the attenuation of the bonding plasmon modes of the dimer (DP, QP) (see Fig. 4.6). However, it does not allow to describe the evolution of the junction from resistive to conductive upon further decreasing S, as shown in Fig. 4.8. Panel (a) of the figure illustrates the evolution of the electron tunneling barrier across the junction between two metal surfaces when the junction width is reduced from S_1 to S_2, where $S_2 < S_1$. Panel (b) displays

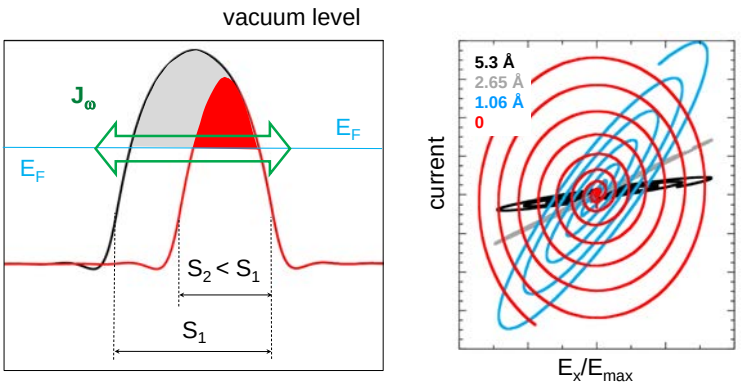
the time-evolution of the current density $J_x(t)$ across the middle of the junction between two $D = 9.8\,\text{nm}$ metal cylinders. $J_x(t)$ is shown as a function of the electric field $E_x(t)$ at the same position. The subscript refers to the projection on the x-axis of the dimer (see Fig. 4.1(a)). The time dependence of the currents and fields is calculated within TDDFT for an incident plane wave resonant with the corresponding plasmon mode of the system (see caption). If the frequency-dependent conductivity of the junction is defined as

$$\sigma(S,\omega) = J_x(\omega)/E_x(\omega), \tag{4.7}$$

then the linear proportionality between $J_x(t)$ and $E_x(t)$ with no relative phase, as shown in panel (b) of Fig. 4.8 for large S (black lines), implies that $\sigma(S,\omega)$ is real valued (resistive junction). This conductivity can be obtained from the TDDFT results, as shown above, or it can be approximated by its DC value defined by Eq. (4.5), since the approach developed to describe tunneling conductances G in STM junctions[31,91] can be applied. At small S, the relative phase between $J_x(\omega)$ and $E_x(\omega)$ implies that $\sigma(S,\omega)$ is complex. Thus, the junction character evolved from resistive [large S] to conductive [small S]. The tunneling barrier is low in that situation, and electrons are efficiently transferred between the nanoparticles. The conductance of the junction is in the $0.1 - 1\ G_0$ range, which is the threshold for observing the emergence of charge transfer plasmon modes.[38,39,89,92] Obviously, at $S = 0$ the potential barrier vanishes and the metal-to-metal contact is formed.

The analysis of the quantum regime in tunneling junctions and its consequences can be summarized as follows:

- In the absence of tunneling, the non-local screening affects the energies of plasmon modes in junctions with S in the range of some nanometers. This effect can be accounted for by rescaling classical results.
- For junction widths below 0.5 nm, electron transfer between nanoparticles is a dominating process that determines the optical response. The range of the electron tunneling effect as given here,



(a) (b)

Fig. 4.8. (a). Schematic representation of the evolution of the electron tunneling barrier across a nanoparticle junction when the junction width is reduced from S_1 to S_2, where $S_2 < S_1$. No bias is applied. The shaded areas represent the tunneling barrier for the electrons at the Fermi energy before (gray) and after (red) changing S. Green arrow shows the AC current J_ω due to the incident electromagnetic field at frequency ω. (b) Conductivity analysis for the $D = 9.8$ nm nanowire dimer system. The current density J_x calculated with TDDFT on the x-axis in the middle of the junction is plotted as a function of the normalized electric field component E_x at the same position. Different colors correspond to different junction widths S as labeled in the insert. The incident x-polarized laser pulse is at resonance with the lowest (DP) plasmon mode for junctions of width $S = 2.65$ Å and $S = 5.3$ Å. For the vanishing junction, $S = 0$ and $S = 1.06$ Å, the incident pulse is at resonance with the C1 charge transfer plasmon mode. For further details see Ref. 34.

is only an estimate. It depends on the specific conditions and can be modified through the change of the potential barrier.

- Standard classical electromagnetic approaches do not allow to describe the electron tunneling effect.
- Quantum calculations naturally include tunneling, but being computationally demanding, are limited to systems too small as compared to those of actual interest in plasmonics.
- For not-too-narrow junctions (low G) much of the understanding of the modal description can be gained from the theory of photon-assisted tunneling. However, this theory does not allow for a simple description of the transition from the tunneling to full contact regime.

A recent proposal to cope with the issues pointed above consists in describing the electron transport properties of the junction via an effective dielectric constant. In this way, the so-called QCM[31] allows to account for the tunneling between plasmonic nanoparticles within the framework of classical Maxwell's equations so that the standard efficient electromagnetic solvers can be applied in practical plasmonic morphologies of considerable size. [15,31,34,90,93] The QCM approach is grounded on two observations. First, the widths of nanogaps where tunneling effects are important are much smaller than the typical curvature radii of the plasmonic nanoparticles forming the junction. The quantum effects in the nanogap can be then modelled using the properties of the electron transfer as calculated for planar metal surfaces separated by S, and then continuously varying S according to the actual geometry of the system. Second, since $S < \lambda$ (λ is the wavelength at optical frequencies), Eq. (4.7) can be considered as a local one. We can thus assume that the local permittivity of the gap can be defined as

$$\varepsilon(S,\omega) = 1 + i\ 4\pi\sigma(S,\omega)/\omega. \tag{4.8}$$

Applying Eq. (4.7), the conductivity $\sigma(S,\omega)$ can be obtained from quantum calculations as shown in Fig. 4.8(b) or reported in Ref. 94. Then, if this $\varepsilon(S,\omega)$ is used to describe the response of the junction in the classical Maxwell's equations, the relationship between the current density across the gap and the optical field inside the gap as obtained from these classical calculations will be the same as the one obtained from quantum calculations. Alternatively, for resistive junctions (large enough gaps), $\sigma(S,\omega)$ can be approximated by its DC value calculated from the STM theory.[91]

For narrow gaps at the transition from the resistive to conductive junction, the STM theories do not apply. However, the TDDFT calculations of $\sigma(S,\omega)$ can be also avoided in this case, by enforcing the transition to a continuous metal for $S = 0$ with the use of a frequency-dependent permittivity of the junction described with the

Drude model:

$$\varepsilon_D(S,\omega) = 1 - \frac{\omega_p^2}{\omega(\omega + i\eta(S))}, \tag{4.9}$$

where ω_p is the plasma frequency characteristic of the metal nanoparticles. $\eta(S)$ is set by an exponential function such that for $S \to 0$, one retrieves the attenuation of the bulk metal, and for large widths of the gap, $\varepsilon_D(S,\omega) \to \varepsilon(S,\omega)$, given by Equation (4.8).[90] Observe that for large S, the dielectric function becomes independent of the particular model parametrization. The QCM is shown to correctly reproduce the full quantum results of the response for various systems such as the spherical nanoparticle dimer,[31] the cylindrical nanowire dimer[34] (see Fig. 4.6) and core-shell plasmonic nanoparticles.[93] The model has been also applied to explain a number of recent experimental observations, as we describe in Fig. 4.9. Screening of the gap plasmon signal in Electron Energy Loss Spectra, in dark-field scattering, in nonlinear signals from gaps, or in gaps functionalized with molecular layers are some of the relevant examples which make quantum effects in plasmonics evident.

4.4. Controlling Tunneling with Molecular Layers

The recognition that the conductance across a gap between metal nanoparticles is a key parameter determining the tunneling effect in the plasmonic response of the system allows one to elaborate a variety of strategies for controling the latter. While varying the width of the gap S seems a natural choice to control the conductance across the junction, using self-assembled molecular layers (SAMs) also provides the possibility to change this conductance, and thus to decrease/increase quantum effects. Junctions functionalized with SAMs offer very promising perspectives as they allow to impact tunneling in two ways:[95] via variation of the overall tunneling barrier,[17] and/or via resonant transfer involving molecular orbitals.[92,96–99]

Examples of molecular control over the plasmon response in metal nanostructures characterized by narrow plasmonic gaps are shown in Fig. 4.10. The nanometer-thick junction between silver

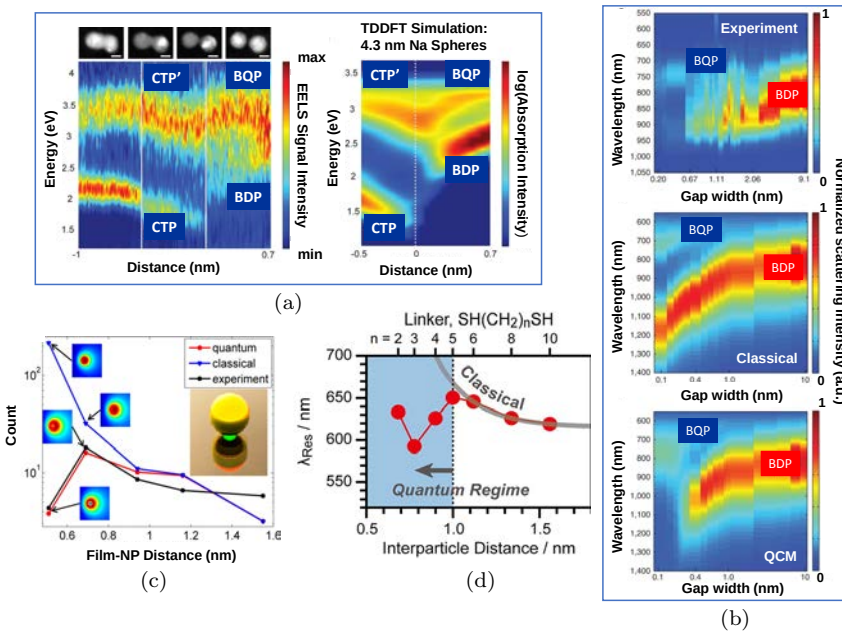


Fig. 4.9. (a) Continuous collection of EELS measured in 9-nm-diameter silver homodimer as a function of the separation distance to observe quantum tunneling effects at small gap sizes (left).[16] Absorption cross-section at different separation distances calculated with TDDFT for the dimer formed by 4.3-nm-diameter sodium spheres in vacuum[30] (right). BDP, BQP: bonding dipole and quadrupole plasmons, CTP, CTP': charge transfer plasmons. (b) Experiment: measured dark-field scattering spectra from gold disk dimers with various gap-widths.[16] Horizontal axis represents measured gap-widths from 9.1 nm to 2 Å in log scale. Classical: Results of the classical electromagnetic simulations of the same situation as above for gap-widths from 1 Å to 10 nm. QCM: Scattering intensities simulated with the QCM. For small gap-widths <4 Å, screening of spectral features related to electron tunneling is observed and reproduced with QCM. The measured and calculated scattering intensities are normalized by their corresponding largest values. (c) Third harmonic signal as a function of the thickness of a self-assembled molecular layer separating 60 nm spherical gold nanoparticles from a gold surface.[18] The experimental data an compared with calculations using the QCM, and the classical electromagnetic approach. Insets show the calculated electric field intensity distribution in the gold film just below the nanoparticle for a wavelength of 1570 nm. The decrease of the nonlinear signal for small film-nanoparticle distance is attributed to the reduction of the field enhancement because of electron tunneling. (d) Evolution of the resonance wavelength of the plasmon band (red dots)

Fig. 4.9. (*caption continued*) in gold nanoparticle dimers as the interparticle distance is reduced by varying the size of the molecular linker.[75] The gray curve is the best fit to an exponential function, representing the classical electromagnetic model for plasmon coupling. The vertical dotted line marks an interparticle distance of 1 nm, where quantum plasmon coupling apparently begins.

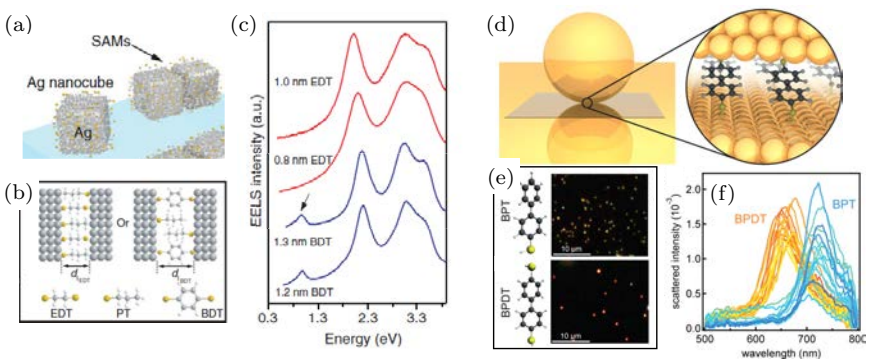


Fig. 4.10. (a)–(c) from Ref. 17. (a): Schematic illustration of the molecular tunnel junctions between two silver nanoparticles bridged by a SAM on a silicon nitride membrane. (b): The distance between two adjacent nanoparticles is determined by the thickness of the SAMs of either saturated, aliphatic 1,2-ethanedithiolates (EDT) or aromatic 1,4-benzenedithiolates (BDT). (c): Measured EELS spectra for double SAMs of EDT (red) and BDT (blue). Tunneling was observed for the double-layer BDT but not in the double-layer EDT junctions. (d)–(f) from Ref. 99. (d): Conductive and non-conductive SAMs in plasmonic junctions. Schematic of a gold nanoparticle separated from a gold film by a thin molecular spacer layer. (e): Dark-field images of 60 nm gold nanoparticles on biphenyl-4-thiol (BPT) and biphenyl-4,4′-dithiol (BPDT), respectively. (f): Normalized scattered intensity from individual 60 nm gold nanoparticles on BPT and BPDT. For nearly the same dielectric constant of BPT and BPDT polymers linking the junction, tunneling leads to a blue shift of the bonding plasmon mode in the case of the BPDT film (slightly conductive) as compared to the BPD (non-conductive).

nanoparticles in one case (panels a, b and c), and between a gold nanoparticle and a gold surface in the other case (panels d, e and f) is functionalized with SAMs. Two types of molecular layers are used in the experiments allowing to compare the results obtained with SAMs of relatively low and high conductance in order to

reveal tunneling effects. For silver nanoparticles separated by a double molecular layer, the low-energy tunneling plasmon associated with interparticle charge transfer across the gap has been observed only for SAMs formed by higher conductance molecules [aromatic 1,4-benzenedithiolates (BDT) in this case]. For SAMs in the junction between the gold nanoparticle and the gold surface, the higher conductance biphenyl-4,4′-dithiol (BPDT) molecules induce a blue shift of the gap plasmon mode. In this case, the tunneling conditions correspond to the situation described with Equation (4.6). Tunneling current partially neutralizes the plasmon-induced screening charges at the junction interfaces, and thus reduces the attractive interaction between the nanoparticle and the substrate.[89] At this point, it is important to stress that using metal-oxide-metal type of junctions also allows to change the tunneling barrier depending on the band structure of the dielectric inside the junction. Indeed, the tunneling barrier is set by the energy of the oxide conduction band minimum in this case.[71,83,91] Last but not least, the examples above show that monitoring the emergence of the charge transfer plasmon and/or the blueshift of the gap plasmon provides an access to the tunneling characteristics of molecular layers. A sophistication of the connection between the conductance and the plasmonic response in SAMs could serve as a valuable tool to quantitatively measure molecular conductance at AC conditions.

4.5. Conclusions

In this chapter, we have described quantum effects in nanoplasmonics derived from the quantum nature of the many-body interactions between conduction electrons in metals. The presence of surfaces of metallic objects intrinsically brings up the effects of confinement, of dynamical screening of electromagnetic fields, of electron spill out, and of the atomistic arrangement at the interface between different materials. Plasmonic structures are key building blocks in nanophotonics where all these quantum effects can be tested. In particular, geometries characterized by narrow gaps between metallic surfaces (plasmonic gaps) emerge as canonical systems where optics and electron transport

can be connected with unprecedented levels of control and monitoring capability. As experimental architectures reach the subnanometric scale, accounting for quantum effects becomes progressively relevant for a correct interpretation of data obtained from optical spectroscopy. There are many situations in nanoplasmonics where classical and approximate non-local treatments correctly describe many of the spectral features of the far- and near-field response associated with plasmons; however, new atomic-scale realms in nanophotonics are firmly emerging, as robust technological platforms.

In this chapter, we have focused on quantum-size effects in metallic nanoparticles, and on the optoelectronics of plasmonic gaps. We have shown the differences between tackling the response of noble metals or free metals in the former, and we have established a description of the effect of a tunneling current in screening the gap plasmon, and eventually the formation of a charge transfer plasmon, for the latter. Our description targets the challenge of an accurate quantitative comparison between phenomena identified in cutting-edge experiments and state-of-the-art theoretical descriptions based on TDDFT techniques. The main trends in optical spectroscopy of metal nanoparticles can be unveiled in this way. However, much work remains to be done regarding more sophisticated aspects of optical spectroscopy and optoelectronics. A modification of the tunneling current, and thus of the optical response, in molecular layers has been shown here, but more active ways of controlling the tunneling barriers have called the attention of many researchers in molecular electronics[100–102] and in electrochemistry.[103,104] The possibility of controlling the optical response by application of an external bias to a metallic nanogap points out towards new possibilities in optoelectronic devices, and optical routing.[105] Another aspect of interest is that the tunneling process across a metallic junction is an intrinsic nonlinear process that could be exploited together with other sources of nonlinearity connected with structural and morphological properties of the particles and their environments. All these combined effects can be exploited in novel nanosystems that include plasmonic structures coupled at subnanometric scale. A substantial increase of

nonlinear signals and the possibility of its control can be expected in this situations.[30,84,106,107]

In addition to the bottom-up approaches to build subnanometric structures, based on wet chemistry methods, scanning probe microscopies also provide a very interesting platform to obtain control of surfaces and molecular species at the atomic-scale. A STM junction where a DC current produces plasmonic excitation and further decays into photons reveals interesting possibilities to modify molecular electroluminescence with unprecedented levels of control.[108,109] This can lead to the development of new sources of photons for quantum information. Molecular spectroscopy is another field where quantum atomistic effects could be key to understand the subnanometric resolution recently found in experiments which reveals intramolecular features.[110] Atomistic vertices, edges, and protrusions are able to localize optical fields beyond the standard plasmonic capabilities,[39] producing a new type of ultraconfined optical cavities which can boost plasmon-exciton coupling[111] and optomechanical coupling[112] to values and regimes not explored before.

We have limited the content of this chapter to metallic structures. It is worthwhile to mention that other materials with novel optical properties such as Van der Waals 2D materials,[113] are also offering the possibility to explore and reveal quantum effects derived from the many-body interaction between the electrons in the system. This is a topic of utmost relevance because of the important technological opportunities; however, we have shown here that standard metallic materials equally hold an incredible potential for optoelectronics and spectroscopy in the optical range of the spectrum.

References

1. García de Abajo, F. J. (2010). *Rev. Mod. Phys.* **82**, p. 209.
2. Kreibig, U. and Vollmer, M. (1995). *Optical Properties of Metal Clusters*, (Springer Series in Materials Science) 25.
3. Schuller, J. A., Barnard, E. S., Cai, W., Jun, Y. C., White, J. S. and Brongersma, M. L. (2010). *Nat. Mater.* **9**, p. 193.
4. Alvarez-Puebla, R., Liz-Marzán, L. M. and García de Abajo, F. J. (2010). *J. Phys. Chem. Lett.* **1**, p. 2428.

5. Stockman, M. I. (2011). *Optics Exp.* **19**, p. 22029.
6. Stewart, M. E., *et al.* (2008). *Chem. Rev.* **108**, p. 494.
7. Gonçalves, M. R. (2014). *J. Phys. D: Appl. Phys.* **47**, p. 213001.
8. Anker, J. N., Hall, W. P., Lyandres, O., Shah, N. C., Zhao, J. and Van Duyne, R. P. (2008). *Nat. Mater.* **7**, p. 442.
9. Kinkhabwala, A., Yu, Z., Fan, S., Avlasevich, Y., Müllen, K. and Moerner, W. E. (2009). *Nat. Photon.* **3**, p. 654.
10. Baffou, G. and Quidant, R. (2014). *Chem. Soc. Rev.* **43**, p. 3898.
11. Kauranen, M. and Zayats, A. V. (2012). *Nat. Photon.* **6**, p. 737.
12. Liu, N., Hentschel, M., Weiss, T., Alivisatos, A. P. and Giessen, H. (2011). *Science* **332**, p. 1407.
13. Knight, M. W., King, N. S., Liu, L., Everitt, H. O., Nordlander, P. and Halas, N. J. (2014). *ACS Nano* **8**, p. 834.
14. Halas, N. J., Lal, S., Chang, W.-S., Link, S. and Nordlander, P. (2011). *Chem. Rev.* **111**, p. 3913.
15. Savage, K. J., Hawkeye, M. M., Esteban, R., Borisov, A. G., Aizpurua, J. and Baumberg, J. J. (2012). *Nature* **491**, p. 574.
16. Scholl, J. A., García-Etxarri, A., Koh, A. L. P. and Dionne, J. A. (2013). *Nano Lett.* **13**, p. 564.
17. Tan, S. F., Wu, L., Yang, J. K. W., Bai, P., Bosman, M. and Nijhuis, C. A. (2014). *Science* **343**, p. 1496.
18. Hajisalem, G., Nezami, M. S. and Gordon, R. (2014). *Nano Lett.* **14**, p. 6651.
19. Zhu, W. and Crozier, K. B. (2014). *Nat. Commun.* **5**, p. 5228.
20. Apell, P. and Ljungbert, Å. (1982). *Solid State Commun* **44**, p. 1367.
21. Apell, P. and Penn, D. R. (1983). *Phys. Rev. Lett.* **50**, p. 1316.
22. Yannouleas, C. and Broglia, R. A. (1992). *Annal. Phys.* **217**, p. 105.
23. Monreal, C. R., Antosiewicz, T. J. and Apell, S. P. (2013). *New J. Phys.* **15**, p. 083044.
24. Tucker, J. R. and Feldman, M. J. (1985). *Rev. Mod. Phys.* **57**, p. 1055.
25. Platero, G. and Aguado, R. (2004). *Phys. Reports* **395**, p. 1.
26. Ciraci, C., Pendry, J. B. and Smith, D. R. (2013). *Chem. Phys. Chem.* **14**, p. 1109.
27. Toscano, G. *et al.* (2015). *Nat. Comm.* **6**, p. 7132.
28. Zuloaga, J., Prodan, E. and Nordlander, P. (2009). *Nano Lett.* **9**, p. 887.
29. Zuloaga, J., Prodan, E. and Nordlander, P. (2010). *ACS Nano* **4**, p. 5269.
30. Marinica, D. C., Kazansky, A. K., Nordlander, P., Aizpurua, J. and Borisov, A. G. (2012). *Nano Lett.* **12**, p. 1333.
31. Esteban, R., Borisov, A. G., Nordlander, P. and Aizpurua, J. (2012). *Nat. Commun.* **3**, p. 825.
32. Stella, L., Zhang, P., García-Vidal, F. J., Rubio, A. and García-González, P. (2013). *J. Phys. Chem. C* **117**, p. 8941.
33. Teperik, T. V., Nordlander, P., Aizpurua, J. and Borisov, A. G. (2013). *Phys. Rev. Lett.* **110**, p. 263901.

34. Teperik, T. V., Nordlander, P., Aizpurua, J. and Borisov, A. G. (2013). *Opt. Express* **21**, p. 27306.
35. Andersen, K., Jensen, K. L., Mortensen, N. A. and Thygesen, K. S. (2013). *Phys. Rev. B* **87**, p. 235433.
36. Kulkarni, V., Prodan, E. and Nordlander, P. (2013). *Nano Lett.* **13**, p. 5873.
37. Zhang, P., Feist, J., Rubio, A., García-González, P. and García-Vidal, F. J. (2014). *Phys. Rev. B* **90**, p. 161407(R).
38. Varas, A., García-González, P., García-Vidal, F. J. and Rubio, A. (2015). *J. Phys. Chem. Lett.* **6**, p. 1891.
39. Barbry, M., Koval, P., Marchesin, F., Esteban, R., Borisov, A. G., Aizpurua, J. and Sánchez-Portal, D. (2015). *Nano Lett.* **15**, p. 3410.
40. Feibelman, P. J. (1982). *Prog. Surf. Sci.* **12**, p. 287.
41. Liebsch, A. (1993). *Phys. Rev. B* **8**, p. 11317.
42. Liebsch, A. (1987). *Phys. Rev. B* **36**, p. 7378.
43. Haberland, H. (2013). *Nature* **494**, p. E1.
44. Meier, M. and Wokaun, A. (1983). *Opt. Lett.* **11**, p. 581.
45. Beck, D. E. (1987). *Phys. Rev. B* **35**, p. 7325.
46. Yannouleas, C., Vigezzi, E. and Broglia, R. A. (1993). *Phys. Rev. B* **47**, p. 9849.
47. Parks, J. H. and McDonald, S. A. (1989). *Phys. Rev. Lett.* **69**, p. 2301.
48. Borggreen, J. *et al.* (1993). *Phys. Rev. B* **48**, p. 17507.
49. Reiners, T., Ellert, C., Schmidt, M. and Haberland, H. (1995). *Phys. Rev. Lett.* **74**, p. 1558.
50. Charlé, K.-P., Schulze, W. and Winter, B. (1989). *Z. Phys. D -Atoms, Molecules and Clusters* **12**, p. 471.
51. Tiggesbäumker, J., Köller, L., Meiwes-Broer, K.-H. and Liebsch, A. (1993). *Phys. Rev. A* **48**, p. R1749.
52. Charlé, K.-P., König, L., Nepijko, S., Rabin, I. and Schulze, W. (1998). *Cryst. Res. Technol.* **33**, p. 1085.
53. Kasperovich, V. and Kresin, V. V. (1998). *Phil. Mag.* **78**, p. 385.
54. Nilius, N., Ernst, N. and Freund, H.-J. (2000). *Phys. Rev. Lett.* **84**, p. 3994.
55. Borisov, A. G., unpublished.
56. Lermé, J. (2011). *J. Phys. Chem. C* **115**, p. 14098.
57. Serra, L. and Rubio, A. (1997). *Phys. Rev. Lett.* **78**, p. 1428.
58. Jensen, L. L. and Jensen, L. (2009). *J. Phys. Chem. C* **113**, p. 15182.
59. Li, J.-H., Hayashi, M. and Guo, G.-Y. (2013). *Phys. Rev. B* **88**, p. 155437.
60. Iida, K., Noda, M., Ishimura, K. and Nobusada, K. (2014). *J. Phys. Chem. A* **118**, p. 11317.
61. Kuisma, M. *et al.* (2015). *Phys. Rev. B* **91**, p. 115431.
62. Romero, I., Aizpurua, J., Bryant, G. W. and García de Abajo, F. J. (2006). *Optics Exp.* **14**, p. 9988.
63. Hövel, H., Fritz, S., Hilger, A., Kreibig, U. and Vollmer, M. (1993). *Phys. Rev. B* **48**, p. 18178.

64. Klein-Wiele, J.-H., Simon, P. and Rubahn, H.-G. (1998). *Phys. Rev. Lett.* **80**, p. 45.
65. Baida, H. *et al.* (2009). *Nano Lett.* **9**, p. 3463.
66. Novo, C. *et al.* (2006). *Phys. Chem. Chem. Phys.* **8**, p. 3540.
67. Wokaun, A., Gordon, J. P. and Liao, P. F. (1982). *Phys. Rev. Lett.* **48**, p. 957.
68. Watanabe, K., Menzel, D., Nilius, N. and Freund, H.-J. (2006). *Chem. Rev.* **106**, p. 4301.
69. Brongersma, M. L., Halas, N. J. and Nordlander, P. (2015). *Nat. Nanotechnol.* **10**, p. 25.
70. Baffou, G. and Quidant, R. (2013). *La. Photon. Rev.* **7**, p. 171.
71. Clavero, C. (2014). *Nat. Photon.* **8**, p. 95.
72. Binnig, G. and Rohrer, H. (1986). *IBM J. R. Dev.* **30**, p. 355.
73. Kravtsov, V., Berweger, S., Atkin, J. M. and Raschke, M. B. (2014). *Nano Lett. 14*, p. 5270.
74. Mertens, J., *et al.* (2013). *Nano Lett.* **13**, p. 5033.
75. Cha, H., Yoon, J. H. and Yoon, S. (2014). *ACS Nano* **8**, p. 8554.
76. Bochterle, J., Neubrech, F., Nagao, T. and Pucci, A. (2012). *ACS Nano* **6**, p. 10917.
77. Ward, D. R., Hüser, F., Pauly, F., Cuevas, J. C. and Natelson, D. (2010). *Nat. Nanotechnol.* **5**, p. 732.
78. Stolz, A., Berthelo, J., Mennemanteuil, M.-M., Colas des Francs, G., Markey, L., Meunier, V. and Bouhelier, A. (2014). *Nano Lett.* **14**, p. 2330.
79. Ittah, N. and Selzer, Y. (2011). *Nano Lett.* **11**, p. 529.
80. Pedersen, M. H. and Büttiker, M. (1998). *Phys. Rev. B* **58**, p. 12993.
81. Grafström, S. (2002). *J. Appl. Phys.* **91**, p. 1717.
82. Dong, T., Ma, X. and Mittra, R. (2012). *Appl. Phys. Lett.* **101**, p. 233111.
83. Haus, J. W., de Ceglia, D., Vincenti, M. A. and Scalora, M. (2014). *JOSA B* **31**, p. 259.
84. Haus, J. W., de Ceglia, D., Vincenti, M. A. and Scalora, M. (2014). *JOSA B* **31**, p. A13.
85. Morton, S. M., Silverstein, D. W. and Jensen, L. (2011). *Chem. Rev.* **111**, p. 3962.
86. Fernández-Domínguez, A. I., Zhang, P., Luo, Y., Maier, S. A., García-Vidal, F. J. and Pendry, J. B. (2012). *Phys. Rev. B* **86**, p. 241110(R).
87. Kottmann, J. and Martin, O. J. F. (2001). *Optics Exp.* **8**, p. 655.
88. Lei, D. Y., Aubry, A., Luo, Y., Maier, S. A. and Pendry, J. B. (2011). *ACS Nano* **5**, p. 597.
89. Pérez-González, O., Zabala, N., Borisov, A. G., Halas, N. J., Nordlander, P., Aizpurua, J. (2010). *Nano Lett.* **10**, p. 3090.
90. Esteban, R., Zugarramurdi, A., Zhang, P., Nordlander, P., Garcia-Vidal, F. J., Borisov, A. G. and Aizpurua, J. (2015). *Faraday Discussions* **178**, p. 151.
91. Simmons, J. G. (1963). *J. Appl. Phys.* **34**, p. 1793.
92. Kulkarni, V. and Manjavacas, A. (2015). *ACS Photonics* **2**, p. 987.
93. Zapata, M., Camacho Beltrán, Á. S., Borisov, A. G. and Aizpurua, J. (2015). *Opt. Exp.* **23**, p. 8134.

94. Hohenester, U. and Draxl, C. (2016). *Phys. Rev. B* **94**, p. 165418.
95. Galperin, M. and Nitzan, A. (2012). *Phys. Chem. Chem. Phys.* **14**, p. 9421.
96. Song, P., Nordlander, P. and Gao, S. (2011). *J. Chem. Phys.* **134**, p. 074701.
97. Lutz, T. *et al.* (2013). *Nano Lett.* **13**, p. 2846.
98. Conklin, D. *et al.* (2013). *ACS Nano* **7**, p. 4479.
99. Benz, F. *et al.* (2015). *Nano Lett.* **15**, p. 669.
100. Große, C. *et al.* (2014). *Nano Lett.* **14**, p. 5693.
101. Emboras, A. *et al.* (2016). *Nano Lett.* **16**, p. 709.
102. Lerch, S. and Reinhard, B. (2016). *Adv. Mater.* **28**, p. 2030.
103. Byers, C. P. *et al.* (2015). *Sci. Adv.* **1**, p. e1500988.
104. Dong, S., Zhang, K., Yu, Z. and Fan, J. A. (2016). *ACS Nano* **10**, p. 6716.
105. Marinica, D. C. *et al.* (2015). *Sci. Adv.* **1**, p. e1501095.
106. Danckwerts, M., Novotny, L. (2007). *Phys. Rev. Lett.* **98**, p. 026104.
107. Cai, W., Vasudev, A. P. and Brongersma, M. L. (2011). *Science* **333**, p. 1720.
108. Chong, M. C. *et al.* (2016). *Phys. Rev. Lett.* **116**, p. 036802.
109. Imada, H., Miwa, K., Imai-Imada, M., Kawahara, S., Kimura, K. and Kim, Y. (2016). *Nature* **538**, p. 364.
110. Zhang, R. *et al.* (2013). *Nature* **498**, p. 82.
111. Chikkaraddy, R. *et al.* (2016). *Nature* **535**, p. 127.
112. Benz, F. *et al.* (2016). *Science* **354**, p. 726.
113. Basov, D. N., Fogler, M. M. and García de Abajo, F. J. (2016). *Science* **354**, p. 1992.

CHAPTER 5

Plasmonics and Transformation Optics

ANTONIO I. FERNÁNDEZ-DOMÍNGUEZ*

Universidad Autónoma de Madrid, Spain

YU LUO

Nanyang Technological University, Singapore

RONGKUO ZHAO

University of California, Berkeley, USA

SIR JOHN B. PENDRY

Imperial College London, UK

Transformation optics (TO) is a recently developed theoretical tool that makes possible an unprecedented control over the propagation and confinement of electromagnetic (EM) fields at both supra- and sub-wavelength scales. In the context of metamaterial science, TO establishes a direct link between a desired EM phenomenon and the material characteristics required for its occurrence. More recently, this powerful framework has been successfully applied to the study of surface-plasmon-assisted phenomena, such as light collection and concentration

*Corresponding author: a.fernandez-dominguez@uam.es

by nano-antennas, the emergence of non-local effects in the optical response of nanostructured metals, or the van der Waals interactions between particles separated by nanometric distances. In this chapter, the fundamentals of TO are reviewed, and recent advances on the TO design and description of plasmonic devices are presented with special focus on the deep insights provided by this largely analytic approach.

5.1. Introduction

Transformation optics[1,2] (TO) has played a key role in the fast development that metamaterial science and technology have experienced during the last decade.[3] This theoretical framework exploits the invariance of macroscopic Maxwell's equations under coordinate transformations to establish a direct link between an electromagnetic (EM) phenomenon, encoded in the transformation, and the material response required for its occurrence. Specifically, TO determines the way in which the EM constitutive relations, and therefore the permittivity and permeability tensors, must be tailored in space in order to obtain a desired optical effect. Thus, in recent years, advances in nanofabrication techniques have made possible the design of metamaterial devices implementing TO prescriptions and realizing technologically relevant functionalities[4,5] or striking phenomena such as invisibility cloaking.[6,7] Importantly, TO fully accounts for the vectorial and undulatory nature of EM fields,[8] which makes it exact not only at supra- but also at sub-wavelength scales.

TO theory demonstrates that, under a general spatial operation like the one in Fig. 5.1(a), EM fields are distorted in a way that is exactly equivalent to a transformation of the electric permittivity and magnetic permeability tensors components of the form[9]

$$\begin{aligned} \varepsilon'^{i'j'} &= [\det(\Lambda)]^{-1}\Lambda_i^{i'}\Lambda_j^{j'}\varepsilon^{ij} \\ \mu'^{i'j'} &= [\det(\Lambda)]^{-1}\Lambda_i^{i'}\Lambda_j^{j'}\mu^{ij}, \end{aligned} \tag{5.1}$$

where $\varepsilon^{ij}(\mu^{ij})$ and $\varepsilon'^{i'j'}$ $(\mu'^{i'j'})$ are the permittivity (permeability) elements in the original (r^i) and transformed $(r'^{i'})$ frames,

respectively, and $\Lambda_i^{i'} = \partial r'^{i'}/\partial r^i$ is the Jacobian matrix for the operation.

Equation (5.1) lie at the core of the TO methodology. From a metamaterial view, they can be understood as a link between material characteristics and the EM effect resulting from the spatial operation. However, from a purely computational electrodynamics perspective, Eq. (5.1) represent a recipe to exchange geometric and material characteristics of the EM system. In fact, this interpretation was the original motivation that led to the development of the TO framework, first thought as a strategy to ease the numerical solution of Maxwell's equations in acute geometries.[1] Figure 5.1(b) exemplifies how TO can be used to map a sharp conical tip into a blunt cylindrical one, for which Maxwell equations can be accurately discretized in a regular mesh. This advantage comes at the expense of more complex, non-uniform and anisotropic permittivity and permeability distributions.

Recently, this initial purpose of TO has been recovered under a different approach. Instead of using this theoretical tool to attack Maxwell's equations numerically, it has been exploited to shed analytical insight into plasmonic phenomena[10,11] taking place in metallo–dielectric systems at a deeply sub-wavelength scale.[12,13] In this nanometric regime, spatial derivatives in Maxwell's curl equations are much larger than temporal ones. Therefore, to a first-order approximation, the latter can be neglected, which translates into the decoupling of magnetic and electric fields. This is the so-called quasi-static regime,[11] in which electric fields can be described in terms of an electrostatic potential, $\mathbf{E}(\mathbf{r}) = -\nabla\Phi(\mathbf{r})$, satisfying Gauss' law

$$\nabla[\bar{\bar{\varepsilon}}(\mathbf{r})\nabla\Phi(\mathbf{r})] = 0, \tag{5.2}$$

where in the most general scenario, the permittivity is an inhomogeneous, anisotropic tensor. Note that in the case of a homogeneous scalar permittivity, Eq. (5.2) recovers Laplace's equation. Importantly, although the quasi-static approximation is only valid for systems of the order of tens of nanometers at visible frequencies,

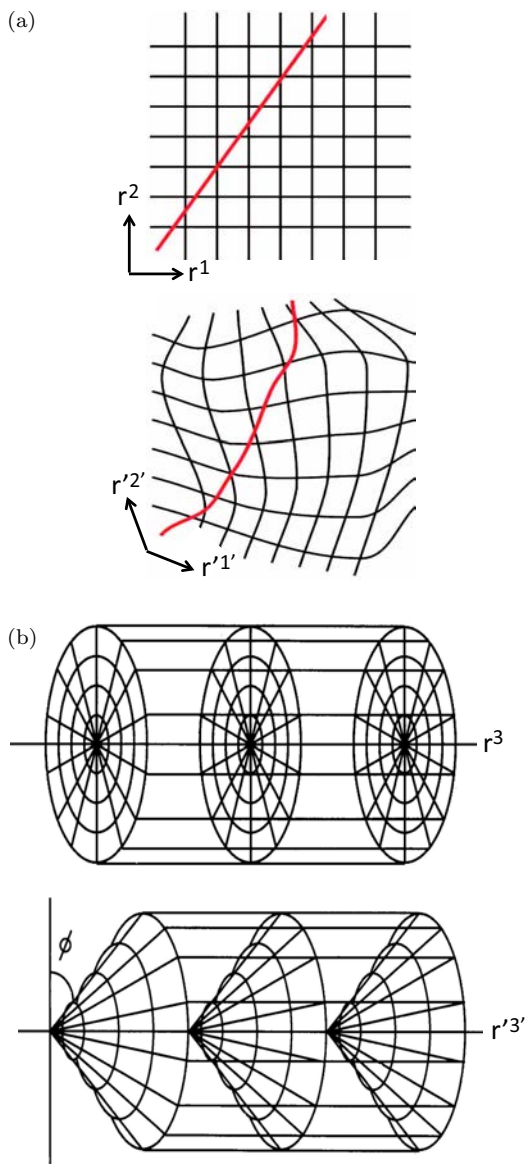


Fig. 5.1. (a) EM fields propagation in free space (sketched as a single field line) with a Cartesian grid in the background (top), and their distortion under an arbitrary geometric transformation (bottom).[2] (b) Diagram showing the conical distortion of a cylindrical mesh, parameterized through the angle ϕ. Note that for $\phi = 0$, the original grid is recovered.[1]

the validity of Eq. (5.2) can be pushed to dimensions up to a few hundreds of nanometers[14] by introducing radiation losses in a self-consistent fashion through the so-called radiative reaction concept[11] (see Section 5.2.4). For larger structure sizes, retardation effects become significant and a full electrodynamics treatment of optical effects is required.

In this chapter, we review a set of analytical and quasi-analytical, exact and approximate, TO solutions of Eq. (5.2) that have lately been exploited to investigate plasmonic phenomena such as the harvesting of light by 2D[15–22] and 3D[23–27] nano-antennas and rough surfaces and gratings,[28–30] the impact of non-local effects (originated from the spatially dispersive nature of the metal permittivity) in nano and sub-nanometric metallic gaps,[31–33] and the near-field van der Waals (vdW) interactions between particles separated by these tiny gap distances.[33–35]

5.2. Two-dimensional Conformal Mapping

In this section, we will show how conformal mapping theory[36,37] can be used to describe the optical properties of 2D metallic nanostructures. Such systems can be described by an isotropic scalar permittivity and present translational symmetry along a spatial direction. A conformal map is an analytic transformation of the form $z' = f(z)$, where $z^{(')} = x^{(')} + iy^{(')}$, that preserves local angles (everywhere except at geometric singularities). The holomorphic nature of the mapping translates into the fact that if a given function $\Phi(z)$ is a solution of Laplace's equation (for the unprimed variables), then $\Phi'(z') = \Phi(f^{-1}(z'))$ will be the solution of Laplace's as well (for the primed ones). This function can, therefore, represent a quasi-static potential in a 2D geometry, where the electric fields do not depend on the third spatial coordinate.

Moreover, as $z' = f(z)$ preserves local angles, the tangential component of the electric field, $E_{\parallel} = -\hat{n} \times \nabla\Phi$, and the normal component of the displacement field, $D_{\perp} = -\varepsilon(\hat{n} \cdot \nabla\Phi)$, which are continuous across any material interface, must be also conserved under the transformation. This implies that $\varepsilon'(z') = \varepsilon(f^{-1}(z))$, and

the magnitude of the permittivity is not affected under the mapping. This result can be also obtained from Eq. (5.1) in terms of the $(x^{(\prime)}, y^{(\prime)})$ spatial coordinates.

The convenient fashion in which quasi-static potentials and permittivity distributions transform under conformal operations, make them suitable candidates for the analytical study of plasmonic geometries. This approach consists in mapping an initial sub-wavelength device into an extended (much simpler) structure, whose analytical description is possible. Importantly, through a set of different mathematical transformations, various sub-wavelength geometries are mapped into the same extended one. This allows for unveiling structural symmetries and optical similarities among them which would otherwise remain hidden. This is illustrated in Fig. 5.2, which

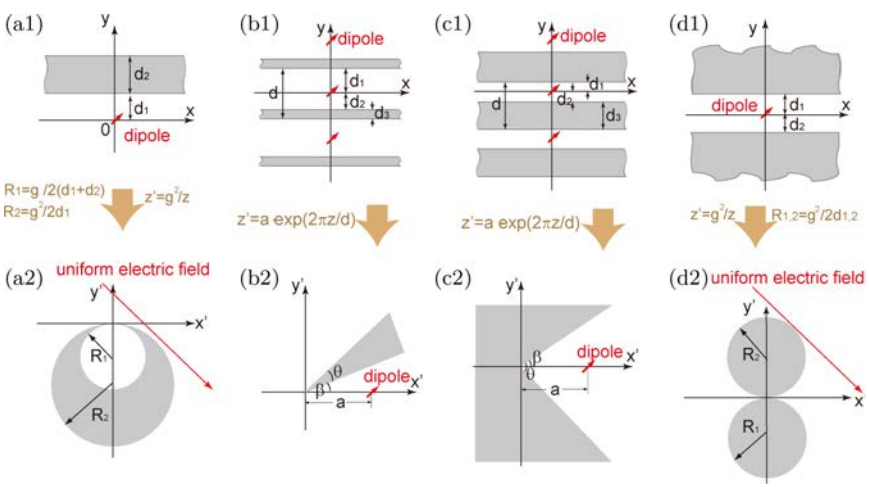


Fig. 5.2. Schematics of conformal transformations that map canonical extended metallic systems (top panels) to a whole category of singular structures (bottom panels).[13] (a1) A thin metal slab coupled to a 2D line dipole maps, under an inversion, into a crescent-shaped cylinder illuminated by a uniform electric field (a2). (b1)–(c1) Periodic metallic slabs excited by an array of line dipoles. An exponential transformation converts these two structures into a metallic wedge (b2) and a V-shaped metallic groove (c2) excited by a single line dipole. (d1) Two semi-infinite metal slabs separated by a thin dielectric film and excited by a 2D dipole source are inverted into two touching metallic nanowires illuminated by a uniform electric field (d2).

shows that a periodic metallo-dielectric system can be linked to a variety of metallic nanostructures, which have different geometric characteristics, but all belong to the same TO family.

In the following, we will describe how the TO strategy described above has been successfully applied to three different families of nanostructures: those presenting geometric singularities (such as touching points or infinitely sharp corners), those whose singular features have been rounded (such as edgeless tips or blunt claws), and those in which strong plasmon hybridization takes place (through vacuum gaps or circulating around metallic elements).[12,13,37]

5.2.1. *Singular geometries*

Over the last decades, much research attention has been focused on the theoretical description of the EM properties of metallic structures with sharp geometric features. This interest has been motivated by the ability of these devices to localize light into nanometric hotspots, where extremely large field enhancements are produced.[38,39] The numerical modelling of these systems[40–47] is very demanding in terms of computational resources, as it requires the accurate treatment of very different length scales. Note that the hotspot size can be 3 orders of magnitude smaller than the optical wavelength. Moreover, these heavy simulations neither provide deep physical insight into the mechanisms that govern the light localization phenomena nor allow a comprehensive configurational study of these systems due to their high time-consuming character. TO offers a way to circumvent these limitations.

As shown in Fig. 5.2, TO makes it possible revealing the optical properties of different sub-wavelength structures (bottom panels) having little symmetry and presenting sharp corners by solving Eq. (5.2) for a single transformed system (top panels). Importantly, the latter are planar and have translational symmetry along the x-direction, which allows expressing the solution to Eq. (5.2) in terms of a summation (or k-integration) over Bloch wave eigenstates.[13,15] Two different EM excitations are considered: a uniform electric field (the quasi-static analog of an incident plane wave) or a static

line dipole source (modelling a nano-emitter). Through inverse-like (exponential-like) transformations, the original plane wave (dipole) is mapped into a single (an array of) line dipole(s), whose projection over Bloch waves is well-established. This procedure maps Eq. (5.2) into a set of algebraic equations in the Bloch expansion coefficients which result from the EM continuity conditions at the boundaries of the structure.

Once the quasi-static potential is known in all space, analytical (or quasi-analytical) expressions not only for the electric field but also for other derived magnitudes can be obtained. We consider here the case of a pair of touching nanowires, which, as Figs. 5.2(d1) and (d2) show, can be mapped into a metal-vacuum-metal waveguide under an inversion of the form

$$z' = f(z) = \frac{g^2}{z}, \tag{5.3}$$

where g is an arbitrary length parameter characteristic of the mapping. This transformation allows obtaining the following expression for the absorption cross-section of a pair of touching nanowires of radius R illuminated by light polarized along the dimer axis[15,16]

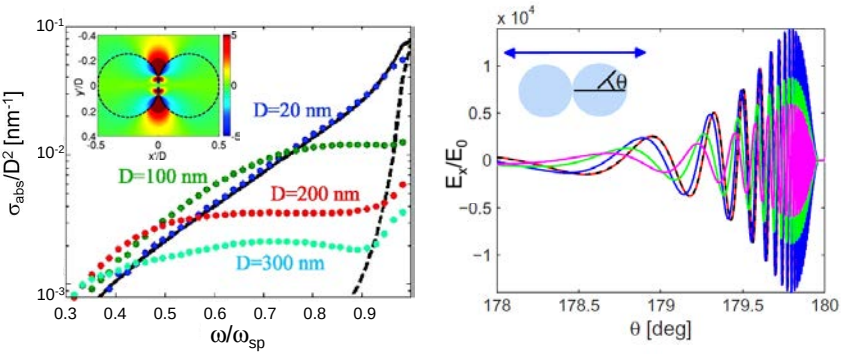
$$\sigma_{\text{abs}}(\omega) = \frac{\pi^2 D^2 \omega}{4c} \text{Re}\left\{\ln\left(\frac{\varepsilon(\omega)-1}{\varepsilon(\omega)+1}\right)\right\}, \tag{5.4}$$

where c is the speed of light and ω is the incident frequency. $D = 4R$ is the physical cross-section of the dimer and $\varepsilon(\omega)$ is its permittivity. This expression is formally approximate, as it neglects the contribution of lossy surface waves to the absorption spectrum of the dimer.[14]

Note the close connection between Eq. (5.4) and the quasi-static dispersion relation of the Surface Plasmon (SP) modes for a metal-vacuum-metal waveguide (the transformed counterpart of the touching dimer), which is given by[16,30]

$$k = \frac{1}{d}\ln\left(\frac{\varepsilon(\omega)-1}{\varepsilon(\omega)+1}\right), \tag{5.5}$$

where $d = d_1 + d_2$ is the slab width, and $\omega < \omega_{\text{sp}}$, the SP frequency for the metal satisfying $\text{Re}\{\varepsilon(\omega_{\text{sp}})\} = -1$. Figure 5.3(a) plots the



(a) (b)

Fig. 5.3. (a) Absorption cross-section normalized by the square of the physical cross-section as a function of frequency for touching nanowires.[15] The incident field is polarized along the dimer axis. The analytical absorption spectrum (continuous black line) is compared to numerical results (dots) for different structure dimensions (20, 100, 200 and 300 nm). The absorption spectrum of one individual cylinder[11] is also shown for comparison (dashed black line). The inset shows the axial component electric field for $\omega = 0.9\,\omega_{\rm sp}$. (b) Amplitude of the axial electric field at the boundary of one of the nanowires as a function of the angle theta evaluated at $\omega = 0.68\,\omega_{\rm sp}$.[16] The analytical electric field (continuous red line) is compared to numerical results for several D : 20 nm (dashed black curve), 100 nm (blue curve), 200 nm (green curve) and 300 nm (purple curve).

absorption cross-section predicted by Eq. (5.4) for a pair of touching Ag nanowires (permittivity taken from the Drude fitting to experimental data[48,49]) as the solid black line. The TO analytical result is compared against full numerical calculations, showing an excellent agreement for particle sizes within the range of validity of the quasi-static approximation ($R < 20$ nm). Figure 5.3(a) shows that touching dimers exhibit a broadband continuous spectra, presenting a cross-section several orders of magnitude larger than single nanowires (black dashed line) at frequencies well below $\omega_{\rm sp}$.

TO reveals the physical origin of the broadband response of touching nanowires. Note that this geometry is mapped into a structure (a metal-dielectric-metal waveguide) that supports SP modes within a wide spectral window $\omega < \omega_{\rm sp}$.[10] Thus, as spectral features are conserved under the transformation[12] (note that these are

fully given by the permittivity in the quasi-static regime), the subwavelength device inherits the broadband response of its extended counterpart. Moreover, TO demonstrates that this is a general characteristic of acute geometric features, in which SP waves are slowed down in their propagation toward the singularity. In the absence of losses, the SP group velocity will eventually vanish. This translates into the fact that the SP waves take infinite time to reach the singularity, in a similar way as the SP modes in the transformed extended structure travel for infinite time along the flat metal surfaces.[15] Importantly, this drastic reduction of group velocity leads to a strong accumulation of EM energy.

Once material losses are introduced in the metal permittivity, the concentration of SP fields in touching dimers is governed by the trade-off between the adiabatic focusing mechanism described above and absorption damping. These two effects can be observed in Fig. 5.3(b), which renders the TO result (black line) for the axial component of the electric field along the boundary of one of the nanowires at $\omega = 0.68\omega_{\rm sp}$.[16] As the SP wave approaches the touching point, its effective wavelength shrinks and the field amplitude increases. However, in the vicinity of the singularity, the small group velocity makes the SP mode extremely sensitive to absorption losses. As a result, the electric field amplitude decreases and vanishes just at the touching point. TO predicts field enhancements of the order of 10^3, which is in excellent agreement with numerical simulations for dimer sizes in the quasi-static regime. For larger D, radiation damping reduces the focusing capabilities of the device. The inset of Fig. 5.3(a) shows a complete 2D map of the field distribution in a nanometric sized dimer at $\omega = 0.9\omega_{\rm sp}$.

Importantly, although the TO treatment of light collection and concentration by SP resonances in 2D singular nanostructures has been discussed above for the particular case of a pair of touching nanowires,[15,16] the physical mechanisms behind this phenomenon are the same for a whole set of singular geometries, such as crescent-shaped nanowires,[17] open crescent structures[28] or overlapping nanowire dimers.[20] Remarkably, in the latter case, TO

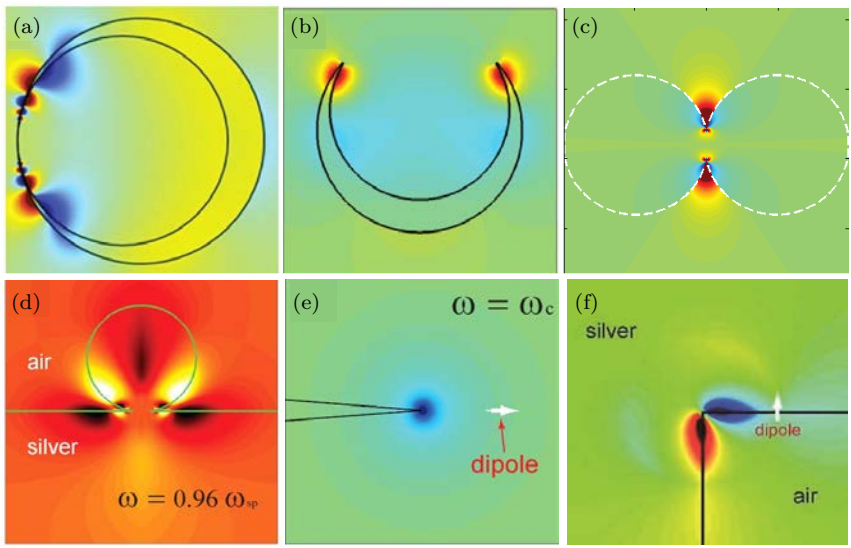


Fig. 5.4. TO Electric field distributions for (a) a crescent-shaped nanowire;[17] (b) an open-crescent structure,[21] (c) an overlapping nanowire dimer,[20] (d) a cylindrical bump carved on a flat metal surface[29] (e) a sharp metal wedge[28] (f) a metallic V-channel.[28]

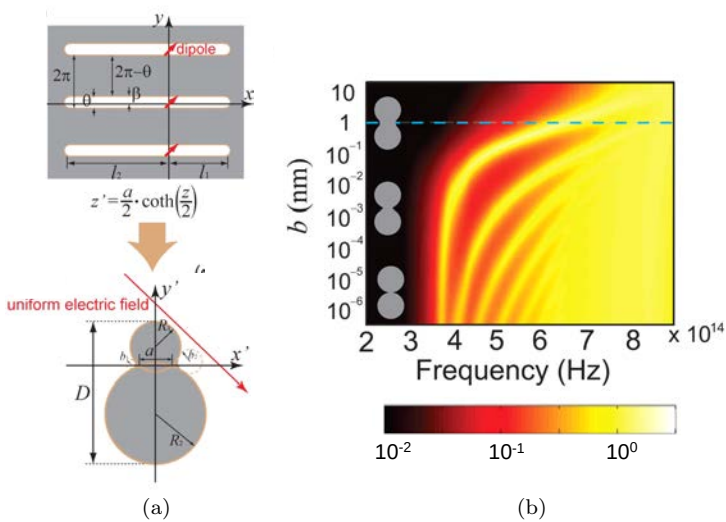
predicts diverging electric fields if the intersection angle is different from zero (as in the perfectly touching case), even with the inclusion of absorption losses in $\varepsilon(\omega)$. Figures. 5.4(a)–5.4(c) show the TO electric field maps associated to the SP modes excited under plane wave illumination in these sub-wavelength devices.

As anticipated in panels (b) and (c) of Fig. 5.2, TO has also been applied to describe the excitation and propagation of SP modes in singular extended geometries, such as concave and convex wedges[28] or cylindrical bumps and grooves carved on a flat surface.[29] In this context, TO allows elucidating the optimum way to decorate a flat metal surface to maximize its performance in EM-field-enhancement-based applications, such as surface-enhanced Raman spectroscopy[50] or SP optical sensing.[51] TO predictions for the electric field distributions in these systems at SP resonance are shown in Figs. 5.4(d)–5.4(f).

5.2.2. *Blunt structures*

As discussed in the previous subsection, TO allows mapping infinite plasmonic structures into finite singular ones (and *vice versa*), the latter inheriting the continuous and broadband SP spectrum of the former. Considering experimental applications, however, these singular structures will suffer from inevitable imperfections due to the nanofabrication process.[52–55] TO also enables us to quantitatively examine how the rounding at the sharp boundaries of the nanostructures will alter their optical response.[21,22]

Figure 5.5(a) shows an example of blunt device: overlapping nanowires connected through a rounded metallic neck. The optical response of this system, as well as other blunt geometries such as



(a) (b)

Fig. 5.5. (a) Mapping of a truncated periodic metallo–dielectric structure into a pair of overlapping nanowires connected through a rounded neck. The EM source transforms from an array of line dipoles into a uniform electric field. (b) Absorption cross-section of normalized to the physical size as a function of frequency and the bluntness dimension of the geometry corners. The overlapping distance is fixed as $a = 0.05D$, and the total structure dimension is set as $D = 100\,\text{nm}$ ($R_1 = R_2$). The blue dashed line indicates the optimal bluntness ($b = 1\,\text{nm}$) where the overlap of all the SP resonances results in a relatively continuous absorption spectrum.[22]

rounded crescents, can be described quasi-analytically through conformal mapping. Since these structures are free of singularities, SP waves do not take infinite time in their propagation along its surface. Instead, SP modes scatter with the truncating features. In the quasi-static regime, the scattered fields do not radiate, but excite reflected SP waves that propagate away from the blunt singularity. In the transformed frame, this translates into planar geometries which are no longer infinite, but finite in space, see Fig. 5.5(a). The superposition among the multiple reflected SP waves propagating across the transformed planar metal surfaces leads to plasmonic spectra that are no longer continuous, but quantized at discrete SP resonant frequencies. The Fabry–Perot-like resonance condition for the device in Fig. 5 is given by[13]

$$\left(\frac{\varepsilon(\omega)-1}{\varepsilon(\omega)+1}\right)\left[\exp\left(\frac{n\pi(2\pi-\theta)}{l_1+l_2}\right)-\exp\left(\frac{n\pi\theta}{l_1+l_2}\right)\right]^2 - \left[\exp\left(\frac{2\pi^2 n}{l_1+l_2}\right)-1\right]^2 = 0, \tag{5.6}$$

where n is an arbitrary integer which denotes the angular momentum of the SP mode, and l_1, l_2 and θ are geometric parameters measuring the length and height of the dielectric slabs in the transformed geometry.

The TO analysis of blunt structures indicates that the spectral position and bandwidth of each SP resonance mode can be controlled through the geometry of the rounded edges. Therefore, with appropriate design, broadband devices whose absorption properties are robust to edge rounding can be realized. Figure 5.5(b) shows that blunt overlapping Ag nanowires present an absorption spectrum that is nearly independent of the geometry of the connecting neck for small rounding radius.[22] For larger bluntness radii, b, the cross-section develops a discrete set of SP resonances. However, through the careful choice of rounding parameters, the broadband response characteristic of singular geometries can be preserved in these devices. Thus, for instance, the blue dashed line in Fig. 5.5(b)

shows a structural configuration for which σ_{abs} for a pair of blunt overlapping nanowires is continuous over a wide frequency window.

The TO treatment of structures with asymmetric bluntness (such as overlapping nanowires with different rounding radii at the left and right sides of the connecting neck) shows the manifestation of dark modes through symmetry breaking. Specifically, invisibility dips appear in the radiative spectra of these devices, which originate from the interference of these additional sub-radiant modes with neighboring dipolar modes. This effect gives rise to rapid changes in the scattering cross-section, similar to the phenomenology observed in nanoparticle configurations supporting plasmonic Fano resonances.[56]

5.2.3. *Plasmon hybridization*

There is another family of plasmonic structures which can also be treated within the TO framework: composite systems in which the localized SPs supported by individual nanoparticles interact, resulting in a new set of resonances supported by the system as a whole.[57–59] One of the most common approaches used to understand this interaction is the so-called plasmonic hybridization picture.[60,61] This elegant model establishes an analogy between the overlapping of atomic orbitals that gives rise to molecular states and the interaction between the localized SPs supported by different metallic elements in composite plasmonic nanostructures. This method still relies on numerical simulations to calculate the spectra of complex metallic devices. In contrast, TO makes it possible solving this problem quasi-analytically by, once again, transforming the interacting nanoparticles into planar layered structures, on which analytical solutions to Eq. (5.2) can be easily found without fitting parameters.[18,19]

Figure 5.6 depicts some examples of 2D nanoparticle systems where strong plasmonic hybridization takes place, which can be studied within the TO frame through conformal transformations involving hyperbolic functions.[13] These are gapped nanoparticles in different geometric configurations. Like blunt geometries, these structures present a discrete SP spectrum, composed of "bonding" and "anti-bonding" modes emerging from the coupling of the resonances

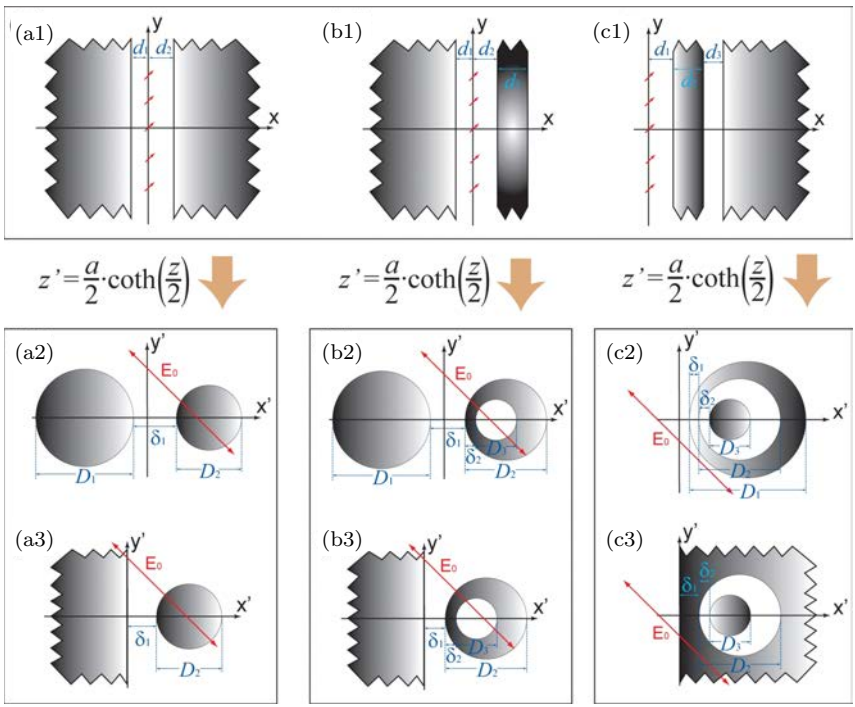


Fig. 5.6. A conformal transformation builds bridges between three planar metallo–dielectric structures (top panels) and three paradigmatic plasmonic nanoantenna geometries. The EM source in the initial coordinate frame is an array of line dipoles with a 2π pitch. Under the conformal mapping, this line dipole array is transformed into a uniform electric field.[13]

supported by each of their components. Under the TO picture, the origin of this spectral discretization can be found in their planar counterparts, see panels (a1)–(c1). These are periodic along the vertical axis; hence, the SP modes supported by these structures are discrete, characterized by the Bloch wave-vector $k_n = n$ (where n is an arbitrary integer). In the finite nanostructure frame, these linear momenta map into SP angular momenta waves, whose "quantization" originates from their circulation around the nanoparticles. The superposition principle gives each SP wave-vector a different resonance condition, which can be derived analytically in the planar frame.

We focus here on the structure in panels (a) of Fig. 5.6. Specifically, we consider pair of identical nanowires separated by a gap size δ. Within the quasi-static approximation, TO yields the following expression for the absorption cross-section for this geometry[18]

$$\sigma_{\text{abs}}(\omega) = 4\pi\frac{\omega}{c}\rho(\rho+1)D^2 \times \text{Im}\left\{\frac{\varepsilon(\omega)-1}{\varepsilon(\omega)-1}\sum_{n=1}\frac{n}{(\sqrt{\rho}+\sqrt{\rho+1})^{4n}-\frac{\varepsilon(\omega)-1}{\varepsilon(\omega)+1}}\right\}, \quad (5.7)$$

where $D = 4R$ is the physical size of the dimer and $\rho = \delta/D$ is the normalized gap size. This equation shows clearly that the spectrum for the system is composed a set of discrete absorption maxima given by the last term of the expression. This corresponds to the sum of contributions due to the different SP modes supported by the nanowire pair, denoted by their angular momentum n (number of effective wavelengths fitting in the nanowires' perimeter), and whose resonant condition is given by the denominators in Eq. (5.7).

Figure 5.7(a) renders the absorption spectrum for a pair of gapped Ag nanowires as a function of ρ. This spectral map demonstrates that, indeed, the cross-section for the system builds up from the summation of multiple SP resonances. Moreover, it can be observed that these SP modes redshift when the two metallic elements approach each other, in excellent agreement with the plasmonic hybridization phenomenology reported theoretically and experimentally for these systems.[45,46]

Figure 5.7(a) also clarifies the physical origin of the broadband continuous spectrum shown in Fig. 5.2(a) for touching dimers. As the gap between the nanowires diminishes, the number of circulating SP resonances increases, each redshifting without bound. It is the overlapping of all the absorption maxima associated with these SP modes that gives rise to the spectrum in Fig. 5.2(a). In section 5.4, we will describe how the introduction of non-local corrections to the metal permittivity modifies plasmon hybridization in sub-nanometric gap sizes.

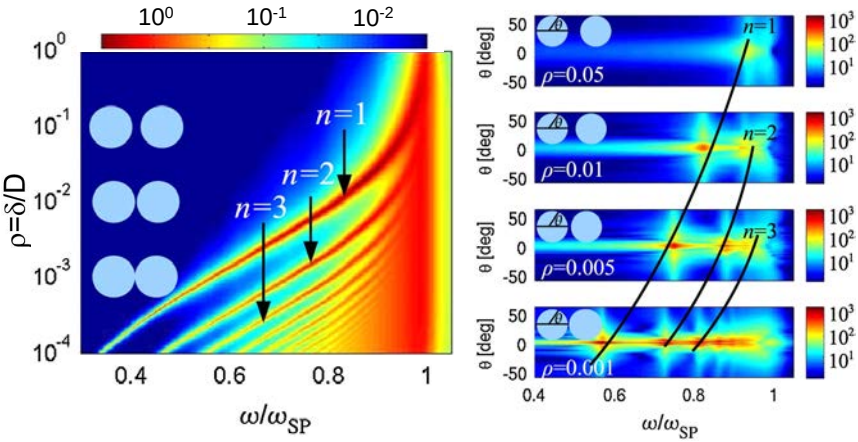


Fig. 5.7. Absorption cross-section normalized to the physical cross-section D as a function of frequency and normalized gap for a nanowire dimer with $D = 20\,\text{nm}$. Field enhancement at the surface of the nanowires versus polar angle θ and frequency for different gap sizes.[18]

Figure 5.7(b) presents a systematic investigation of the nanofocusing properties of separated nanowire dimers. The field enhancement, $|E'|/E_0$, along the nanowires' boundaries is plotted in log scale as a function of the angle θ for different gap distances. In accordance with Fig. 5.2(b), the electric field extends around the whole nanowire perimeter for a large gap. As δ is reduced, $|E'|/E_0$ is compressed around the point which eventually becomes the touching singularity in the $\rho \to 0$ limit.

In the usual plasmonic hybridization picture, the interaction among different SP modes takes place via the superposition of their evanescent tails penetrating into a sub-wavelength dielectric region. However, SP modes can also couple through metallic elements. The paradigmatic example of such coupling is a thin metal slab, where the SP modes at each metal–vacuum interface (satisfying the condition $\text{Re}\{\varepsilon(\omega)\} = -1$ in the quasi-static limit) hybridize. As a result, the SP modes of the slab acquire the spectral dispersion given by Eq. (5.5). Note that the dispersion relation is formally the same as for the complementary structure, a metal-vacuum-metal waveguide.

Panels (a) in Fig. 5.2 show that, under an inversion, such geometry transforms into a crescent-shaped nanowire, a singular device whose optical properties can be described in the TO frame.[15,17]

TO also allows mapping a simple metal slab structure into periodic gratings under more complex conformal transformations, like the one considered in Fig. 5.8(a),[30]

$$z' = \gamma \log\left(\frac{1}{e^z - iw_0} + iy_0\right). \tag{5.8}$$

Importantly, this transformation presents two free parameters, w_0 and γ, which means that the same original slab can be linked into a whole family of gratings with different modulations. As discussed above, conformal operations preserve spectral characteristics. This means that, within the quasi-static regime, the dispersion relation for all the members of one of these grating families must be the same and it must be given by Eq. (5.5). This prediction proves again the insightful power of TO, which reveal symmetries and similarities (or even the total equivalence) among systems that are apparently completely unrelated.

Figure 5.8(b) shows the dispersion relation of two metallic gratings belonging to the same TO family.[30] Solid lines plot Eq. (5.5) folded into the first Brillouin zone, and dots correspond to numerical simulations for deep (triangles) and shallow (circles) grooves. The good comparison between numerics and analytics demonstrates the validity of the physical picture emerging from TO. Note that the approach is not able to reproduce the numerical results for both grating geometries in two aspects: the overlapping of the first SP-band and the light line ($\omega = c\,k_y$, black dashed line) at small parallel wave-vectors, and the apparent gap which opens at the edge of the Brillouin zone for the first few SP-bands. The overlapping of the light line is due to the emergence of strong radiative effects at long wavelengths, which are beyond the quasi-static picture inherent to Eq. (5.5). On the other hand, the deviation at the band gap is due to neglect of field discontinuities across a branch point in the transformation, which vanish at the zone center but become increasingly significant towards the zone boundary.

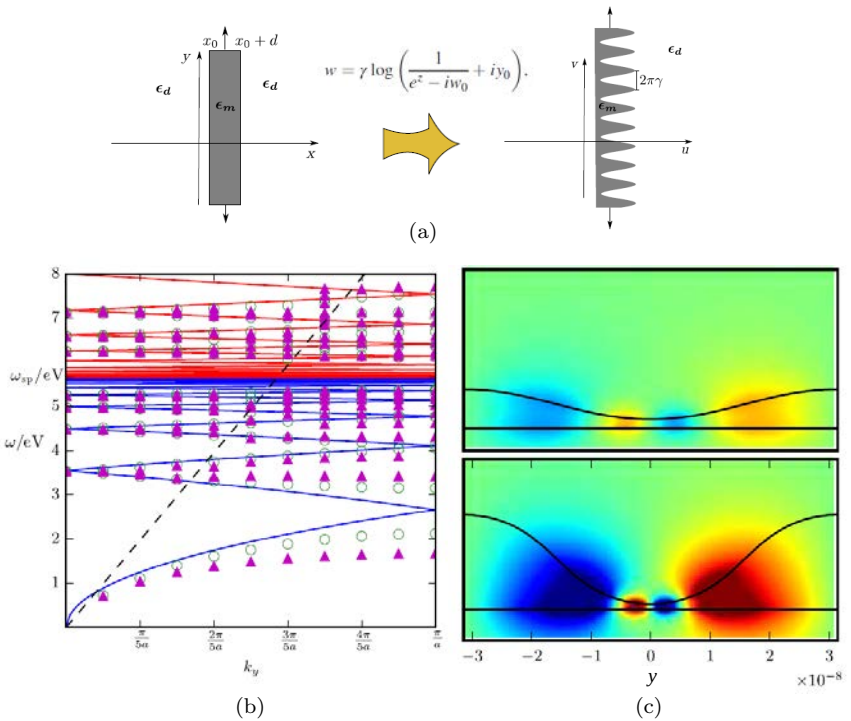


Fig. 5.8. (a) Set of conformal transformations that map a single-metallic slab into a whole family of periodic gratings. (b) Dispersion relation of two gratings of the same family in the first Brillouin zone. The magenta triangles and green open circles are numerical simulations for structures ($\gamma = 10^{-8}$) with $w_0 = 2.5$ and $w_0 = 1.5$, respectively ($x_0 = 1, d = 0.5$). (c) Mode profile for the SP modes supported by these structures at 4.5 eV.[30]

In Fig. 5.8(c), the mode profile for the SP modes at 4.5 eV is rendered for both grating structures, showing the strong EM focusing that takes place at the bottom of the grooves thanks to the evanescent SP coupling through the thin metal region.

5.2.4. *Radiative corrections*

Up to here, we have reviewed how conformal mapping makes it possible to describe quasi-analytically the optical properties of a wide range of 2D plasmonic nanostructures within the quasi-static approximation. This restricts the validity of the approach to system sizes of

very few tens of nanometers. In this subsection, we discuss how the insightful power of TO can be applied to larger device dimensions. This should be done rigorously by considering the dynamic coupling between electric and magnetic fields. However, whereas Eq. (5.1) preserve the permittivity tensor within the incidence plane, the permeability component along the direction of translational symmetry (the one relevant for transverse magnetic excitation) would acquire a spatial dependence.[62] As we will see in the following section, this complicates the TO treatment of the system significantly and prevents obtaining exact quasi-analytical expressions for the EM fields. However, an alternative route can be used to extend TO quasi-static predictions to larger structural sizes, up to a few hundreds of nanometers, through the introduction of radiative damping. Note that retardation effects, which can only be described within a full electrodynamics picture, play a key role in the optical response of systems comparable to the incident wavelength.

The radiation reaction concept[11] accounts for the fact that the EM radiation emitted by an oscillating charge distribution does not only dissipate energy away from the oscillator but also influences the charge motion itself. This back action force can be considered as originated by a self-induced electric field. For structures treated in the dipole limit, the inclusion in a self-consistent fashion of this new electric field component yields a corrected particle polarizability, $\alpha(\omega)$. Note that $\mu = \varepsilon_0 \alpha(\omega) \mathbf{E_0}$ is the dipole moment induced in the structure illuminated by the incident electric field $\mathbf{E_0}$. The expression reads[63]

$$\alpha_{\mathrm{corr}}(\omega) = \frac{\alpha_{\mathrm{QS}}(\omega)}{1 - i g_0(\omega) \alpha_{\mathrm{QS}}(\omega)}, \tag{5.9}$$

where $\alpha_{\mathrm{QS}}(\omega)$ is the polarizability obtained within the quasi-static approximation, $g_0(\omega) = \frac{1}{\varepsilon_0} \left(\frac{\omega}{c}\right)^2 \mathrm{Im}\left\{\mathrm{tr}\left\{\mathbf{G_0}(\mathbf{r}, \mathbf{r'} = \mathbf{r}, \omega)\right\}\right\}$, and $\mathbf{G_0}(\mathbf{r}, \mathbf{r'}, \omega)$ is the EM dyadic Green's function in free space. In the case of 2D geometries, $g_0^{2D}(\omega) = \omega^2/8c^2$.[14]

This refined induced polarizability allows introducing radiative corrections in all the EM quantities for the system. Thus, we can

write the radiative corrected absorption cross-section and near-field enhancement as

$$\sigma_{\mathrm{abs}}^{\mathrm{corr}}(\omega) = \frac{\sigma_{\mathrm{abs}}^{\mathrm{QS}}(\omega)}{\left|1-ig_0(\omega)\alpha_{\mathrm{QS}}(\omega)\right|},$$
$$\left|\frac{E_{\mathrm{corr}}(\omega)}{E_0}\right| = \frac{\left|E_{\mathrm{QS}}(\omega)/E_0\right|}{\left|1-ig_0(\omega)\alpha_{\mathrm{QS}}(\omega)\right|}. \tag{5.10}$$

TO offers a novel perspective on the radiative reaction scheme described above. It shows that the radiation losses experienced by a nanometric system can be directly mapped onto the power dissipated by a fictive absorbing particle located at the source position in its transformed extended counterpart. Figure 5.9(a) sketches the TO picture for radiation damping in the case of touching nanowires (bottom panel).[14] The self-consistent approach described above yields the dipolar moment of the absorbing particle (red) induced by the driving source (blue) in the metal-vacuum-metal geometry (top panel). The radiative reaction in the planar frame corresponds to the fields scattered by this fictive absorber which are back-emitted toward the metal slabs.

Figure 5.9(b) plots the absorption spectra for touching nanowires of different physical sizes: 40, 80, 120 and 160 nm. Analytical TO predictions obtained within the quasi-static approximation (black dashed line) and corrected through Eq. (5.10) (red line) are compared to numerical simulations (blue line). The cross-section for a single nanowire is rendered in green. These results show that the validity of TO calculations is pushed to larger nanostructure sizes through the inclusion of radiative damping.[14]

5.3. Full Three-dimensional Coordinate Transformations

In the previous section, we have shown that the TO description of 2D systems provides deep physical insight into the optical properties of plasmonic devices. However, these TO treatments assume translational invariance on not only the metal geometry (inherently 2D), but also the electric fields. This limits the validity of TO predictions

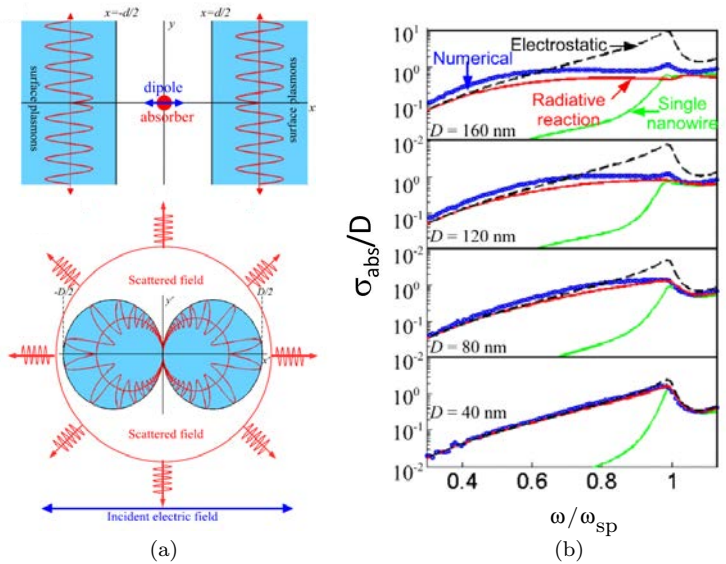


Fig. 5.9. (a) Conformal inversion that maps a metal-vacuum-metal waveguide into a pair of touching nanowires (and *vice versa*). In the top panel, the emitting dipole source (blue) is superimposed to a fictive absorbing particle (red). The former transforms into the uniform electric field illumination in the bottom panel, whereas the latter accounts for the radiative damping suffered by the touching dimer. (b) Normalized absorption cross section as a function of frequency for different physical sizes D. TO quasi-static (dashed black), TO radiation-reaction corrected quasi-static (red) and numerical (blue) calculations are plotted. The spectrum for a single nanowire is also shown for comparison.[14]

greatly, as they are restricted to the subset of EM modes with transverse magnetic polarization, and neglects polarization cross-coupling effects in the EM scattering at the structural boundaries along the third dimension.[13]

In this section, we discuss how the TO framework can be extended to 3D systems, allowing for a complete description of plasmonic phenomena in realistic devices. We briefly describe this 3D TO methodology, which exploits analogies with its 2D counterpart. Crucially, in contrast to 2D conformal maps, Eq. (5.1) imply that general 3D transformations act on both geometry and material properties. This complicates their mathematical treatment significantly

which, in most cases, translates into the fact that exact and tractable analytical expressions can no longer be obtained. We will present two different strategies to circumvent this limitation: through approximate analytical solutions or exact quasi-analytical treatments. Specifically, we will consider TO solutions for four different systems: singular touching nanospheres,[23,24] blunt spherical nanocrescents,[25] separated nanospheres,[26] and oblate and prolate spheroids.[27]

5.3.1. *Touching nanospheres*

A pair of touching nanospheres is the evident 3D counterpart of a pair of touching cylinders. Therefore, it seems natural that the TO description of this system should exploit the 3D version of the conformal inversion in Eq. (5.3), which has the form[23]

$$\mathbf{r}' = \frac{g^2}{r^2}\mathbf{r}. \tag{5.11}$$

Similarly to its 2D analog, this inversion maps two touching spheres into two semi-infinite planar slabs, as shown in Fig. 5.10(a).

Introducing Eq. (5.11) into Eq. (5.1), we find that the permittivity in the bottom panel of Fig. 5.10(a) acquires a spatial dependence of the form

$$\varepsilon'(\mathbf{r}') = \left(\frac{g}{r'}\right)^2 \varepsilon(\mathbf{r}(\mathbf{r}')), \tag{5.12}$$

where $\varepsilon(\mathbf{r})$ is a distribution of uniform permittivity regions. This means that the slabs in the transformed frame are not filled with (surrounded by) a conventional metal (dielectric), but with a modified material having a permittivity given by expression above. Equation (5.2) yields that the transformed quasi-static potential must be then a solution of

$$\nabla' \left(\left(\frac{g}{r'}\right) \nabla' \Phi'(\mathbf{r}') \right) = 0. \tag{5.13}$$

inside each of the different regions of the transformed geometry. Similarly to the 2D case, Eq. (5.11) maps the incident light (uniform

electric field) into a dipole-like source at the origin of the transformed frame. This source does not correspond to any conventional EM source, as it is also the solution of Eq. (5.13).

The general solution to Eq. (5.13) can be written as $\Phi'(\mathbf{r}') = r'\phi'(\mathbf{r}')$, where $\phi'(\mathbf{r}')$ satisfies Laplace's equation. This result sheds light on the physical interpretation of transformed potentials and provides us with an expansion basis for $\Phi'(\mathbf{r}')$ in terms of solutions of Laplace's equation in cylindrical coordinates. However, due to the presence of the r' pre-factor, the elements of the basis do not present the cylindrical symmetry of Fig. 5.10(a), and hence they

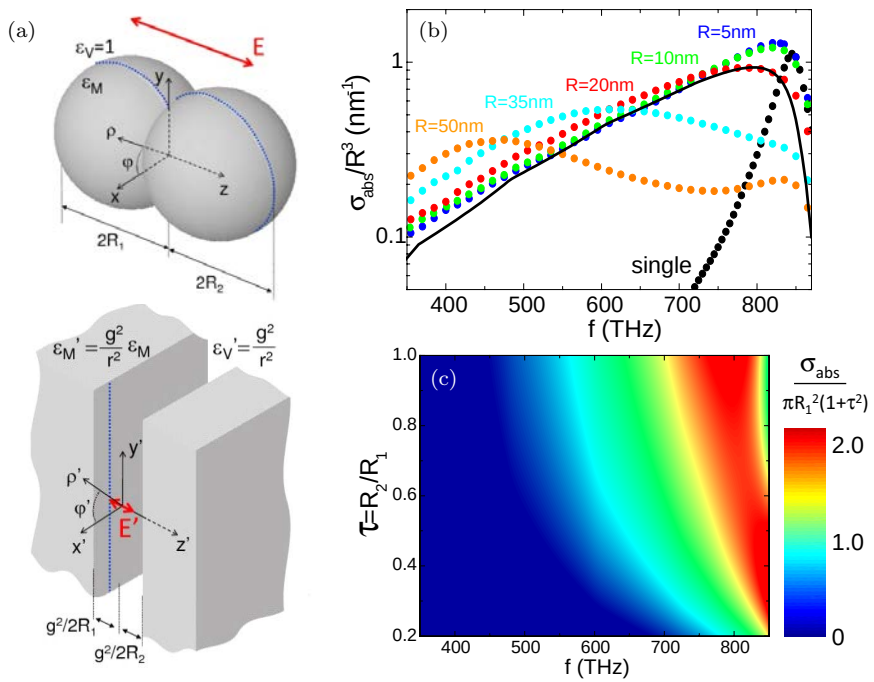


Fig. 5.10. (a) Top panel shows light impinging on a pair of metal nanospheres that touch at a single point. Bottom panel shows the transformed geometry, illumination, and permittivity obtained from Eq. (5.11). (b) Numerical absorption cross-section versus incident frequency for twin Ag[48] touching spheres of different radii. Black solid line plots the analytical result given by Eq. (5.14). (c) Normalized absorption cross-section as a function of frequency and the ratio between the two spheres forming the touching dimer.[23]

cannot satisfy EM fields continuity at every single point of the slab boundaries. An approximate analytical solution can be obtained in a self-consisted way by first ensuring continuity at the point that transforms into the spheres touching point (in whose vicinity we expect EM fields to accumulate, as it was reported in the 2D case, see Fig. 5.3), and then account for the fields mismatch in any other point of the boundaries as a surface charge correction to the transformed potential. This way, a single term in the solution basis can be used to obtain the following expression for the dimer's absorption cross-section[24]

$$\sigma_{\text{abs}}(\omega) = \frac{64\pi^2 R^3 \omega}{3c} \text{Re}\left\{ \ln\left(\frac{\varepsilon(\omega)-1}{\varepsilon(\omega)+1}\right)\left[\ln\left(\frac{\varepsilon(\omega)-1}{\varepsilon(\omega)+1}\right)-1\right]\right\}, \tag{5.14}$$

where R is the nanosphere's radius. Note the similarity between Eqs. (5.14) and (5.4), which anticipates that, like their 2D analog, the absorption properties of touching nanospheres can be described in terms of the SP modes supported by metal–dielectric–metal waveguides, see Eq. (5.5).

Figure 5.10(b) shows the comparison between numerical calculations (color dots) and the prediction from Eq. (5.14) (black solid line) for the absorption spectrum (normalized-to-volume cross-section) of two touching Ag[48] nanospheres under axial illumination. The spectra demonstrate that, similar to the 2D case, the TO result is valid up to ∼30 nm radii, where the quasi-static assumptions break down. This panel proves that the broadband response obtained for touching nanowires (see Fig. 5.3) is reproduced also for their 3D version. Note the narrowband character of the dashed black spectrum, which corresponds to a single sphere. Figure 5.10(c) renders TO results for the transition between a single sphere and a dimer of twin spheres through the variation of the relative size $\tau = R_2/R_1$ of the particles in the dimer.

Although not discussed here, the broadband absorption featured by touching spheres is accompanied by remarkable field enhancement capabilities at frequencies well below ω_{sp}.[23,24] A comprehensive TO analysis proves that touching spheres show an even better

nanofocusing performance than touching nanowires, yielding field enhancements of the order of 10^4 close to their geometric singularity. Importantly, this result can be generalized and extended to all 3D devices, which provide larger field localization and confinement than their 2D counterparts. This is a purely geometrical effect, originating from the fact that SPs are focused onto a point in 3D systems instead of a line, as it happens in 2D structures.[13]

5.3.2. *Blunt spherical nanocrescents*

Figure 5.2 shows that conformal inversions make the TO treatment of touching and crescent-shaped cylinders possible. Similarly, the geometric inversion given by Eq. (5.11) can be used to describe not only touching spheres, but also spherical crescents (nanoshells whose inner and outer boundaries are displaced). Fig. 5.11(a) renders a 3D nanocrescent device (left) and its mapping into a thin slab (right), whose spatially dependent permittivity is given by Eq. (5.12). If the original structure corresponds to a singular crescent, presenting a vertex where the thickness of the shell vanishes, the slab extends till infinity in the transformed frame. This corresponds to a vanishing bluntness radius, $R_b = 0$ in Fig. 5.11(a). On the contrary, if the singularity is removed by drilling a hole with $R_b > 0$, the transformed structure becomes a cylindrical plate of finite radius. Using the TO rationale described in section 5.2.2, we can expect that singular (blunt) spherical crescents present a continuous (discrete) spectral response. This is confirmed by Fig. 5.11(b), which renders the TO absorption cross-section of Au[48] nanocrescents with 10 nm outer and 8.5 nm inner radii. The structure presents a broadband spectrum for small bluntness radius, whereas a number of resonances emerge as R_b increases. Note that, whereas the optical properties of touching dimers depend strongly on the incident polarization, the response of nanocrescents is rather similar for the two polarizations in Fig. 5.11(a).[25]

The TO description of singular spherical crescents can be performed following an approximate, self-consistent procedure similar to the one leading to Eq. (5.14) for touching spheres. Alternatively,

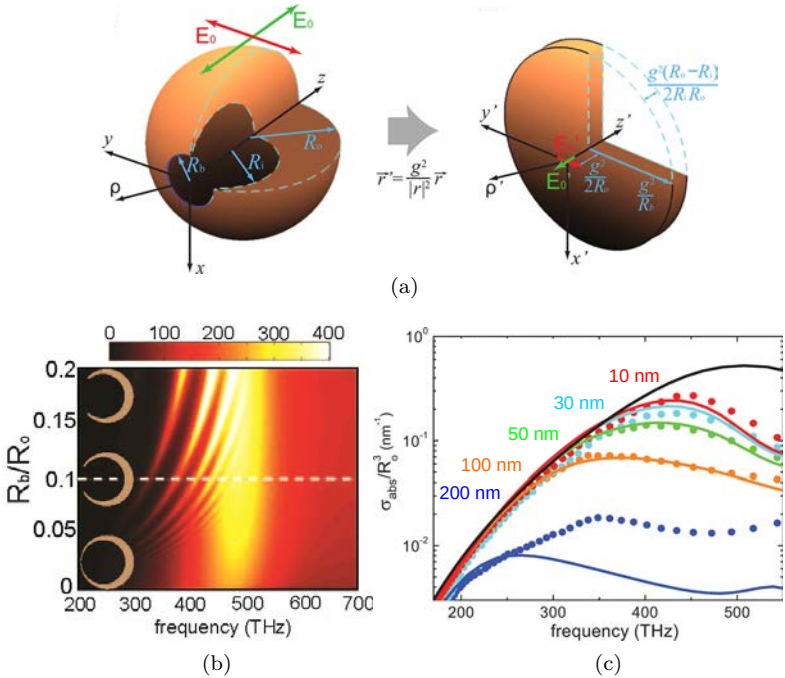


Fig. 5.11. (a) Left: Two different light polarizations incident on a gold spherical nanocrescent. Right: Geometry, illumination and permittivity obtained from the inversion of the crescent. (b) Absorption cross-section in nm^2 as a function of the incident frequency and the bluntness radius for a 10 nm outer, 8.5 nm inner radii crescent. (c) TO (solid lines) and numerical (dots) absorption spectrum for singular crescents with inner-to-outer-radius ratio or 0.9 and different outer radii.[25]

an exact solution to Eq. (5.13) can be built by expressing it as an expansion series on the basis $\Phi'_n(\mathbf{r}') = r'\phi'_n(\mathbf{r}')$ (where $\phi'_n(\mathbf{r}')$ are solution to Laplace's equation in cylindrical coordinates) and impose the appropriate boundary conditions to obtain the corresponding coefficients. Importantly, this approach, which is no longer analytical, is not restricted to singular crescents, but allows the treatment of blunt geometries as well. Note that this was the strategy followed to obtain the spectra in Fig. 5.11(b).

Figure 5.11(c) plots the absorption cross-section normalized to the structure volume for singular crescents with an inner-to-outer-radius ratio of 0.9 and outer radii ranging from 10 to 200 nm.[25]

Black solid line plots the purely quasi-static analytical TO solution obtained for the geometry, which is independent of the overall size of the structure. Color lines plot quasi-analytical solutions requiring the combination of ∼50 expansion terms. These include radiative corrections in the form of Eq. (5.10), where $g_0^{3D}(\omega) = \omega^3/6\pi c^3$. For comparison, numerical simulations for every outer radius value are shown in color dots. The agreement between quasi-analytical and numerical simulations is remarkable up to the crescent with 200 nm outer radius, for which retardation effects (not included in the TO solutions) are significant. The purely static analytical solution reproduces the low-frequency tail of the absorption spectrum, but fails to describe the cross-section at higher frequencies.

5.3.3. *Separated nanospheres*

A more general transformation than the one given by Eq. (5.11) results from the displacement of the inversion point away from the coordinate origin in both the original and transformed frames. Such transformation reads

$$\mathbf{r}' - \mathbf{R}_0' = \frac{g^2}{|\mathbf{r} - \mathbf{R}_0|^2} (\mathbf{r} - \mathbf{R}_0) . \tag{5.15}$$

Through an appropriate choice of the transformation parameters $\mathbf{R}_0$ and $\mathbf{R}_0'$, Eq. (5.15) makes it possible to map two separated spheres into a concentric annular geometry formed by an inner spherical core and a spherical void, as shown in Fig. 5.12.[26,34]

Similar to the previous 3D plasmonic designs, the permittivity in the transformed frame in Fig. 5.12 acquires a spatial dependence obtained from Eq. (5.12) by replacing $\mathbf{r}'$ by $\mathbf{r}' - \mathbf{R}_0'$. Therefore, although the concentric geometry presents spherical symmetry, the permittivity distribution does not. Consequently, the transformed potential can be expanded in terms of the wave-function basis $\Phi_{lm}'(\mathbf{r}') = |\mathbf{r}' - \mathbf{R}_0'|\, \phi_{lm}'(\mathbf{r}')$, where $\phi_{lm}'(\mathbf{r}')$ is solution of Laplace's Equation in spherical coordinates with polar and azimuthal indices l and m, respectively. Importantly, the radial pre-factor in $\Phi_{lm}'(\mathbf{r}')$ does not satisfy the spherical symmetry of the geometry, but it only couples basis elements with adjacent polar indices. As a result, the

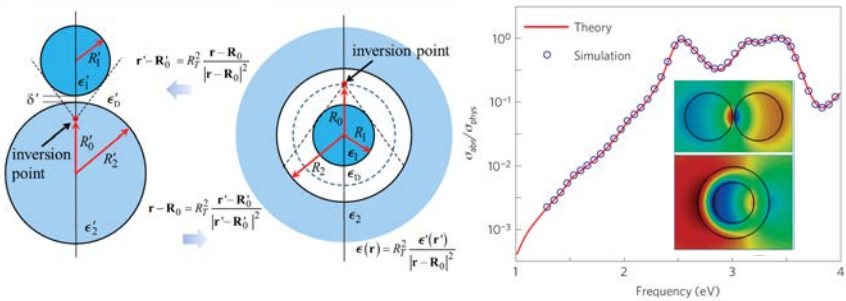


Fig. 5.12. Left: Eq. (5.15) maps a pair of separated spheres into a concentric spherical geometry, whose permittivity is spatially dependent. Right: TO absorption cross-section normalized to the physical size for a dimer of 5 nm radius Ag spheres separated by a 0.1 nm gap. The inset renders the quasi-static potential distribution in original and transformed frames for the lowest bright plasmonic mode supported by 5 nm Ag spheres separated by a 0.4 nm gap.[26,34]

EM continuity conditions for the system can be written in the form of a tri-diagonal matrix. This allows for the fast computation of the EM properties of a pair of separated nanospheres, a paradigmatic plasmonic structure.

The right panel of Fig. 5.12 shows the absorption cross-section for a pair of Ag[48] 5 nm radius nanospheres separated by a 0.1 nm gap.[26] The spectrum is normalized to the cross-sectional area of the dimer. The solid line corresponds to TO calculations, and dots show simulation results. The agreement is perfect. Importantly, the TO methodology is up to two orders of magnitude faster than conventional numerical techniques, even for this challenging geometry (note that the gap is 100 times smaller than the spheres diameter), for which 60 partial wave terms were included in the quasi-static potential expansion. A comprehensive analysis of the absorption spectrum dependence on the gap between the spheres, similar to Fig. 5.7(a), can be found in Ref. 63.

By removing the excitation source, the EM continuity conditions yield a homogeneous system of equations, whose non-vanishing solutions correspond to the plasmonic modes for the gapped dimer. Therefore, TO makes it possible to calculate the resonant frequency, mode profile and EM characteristics of the SPs supported by this

system. The inset of the left panel of Fig. 5.12 renders the quasi-static potential distribution for the lowest bright (dipole active) SPs sustained by a pair of 5 nm radius silver spheres separated by a 0.4 nm gap at 3.53 eV. The potential in the transformed frame is also shown. As discussed in Section 5.5, the high speed and accuracy of the TO approach allows for the investigation of the vdW interactions between the spheres.[33–35]

5.3.4. *Spheroids*

So far, we have shown that TO exploits the mapping of an original structure into another one presenting higher symmetry. These previously hidden symmetries make the quasi-analytical treatment of the optical properties of complex devices possible. The price to pay in this procedure for 3D geometries resides in the spatially dependent permittivity distribution resulting from the transformation. In this section, we will pay attention to the effect of this operation on the illumination. Specifically, we will show that the mapping of an asymmetric structure into a symmetric one can be carried out by coding the original lack of symmetry into the transformed source. Importantly, this can be done not only for both 3D structures, but also for 2D ones, where conformal transformations do not alter the permittivity.

The mapping in Fig. 5.13(a) can be interpreted both in 2D and 3D. In its 2D version, it maps an ellipse (primed coordinates) into a circular ring (unprimed coordinates) and corresponds to the inverse of the so-called Joukowski transformation[65]

$$z'(z) = \frac{c}{2}\left(z - \frac{1}{z}\right), \tag{5.16}$$

where c is an arbitrary length scale constant. Due to the holomorphic nature of this mapping, the permittivity of ellipse and ring are the same. On the contrary, the uniform electric field excitation in the primed frame, $\Phi'_S(z') = \mathrm{Re}\{E_0^* z'\}$ (where $E_0 = E_{0x} + iE_{0y}$) maps into an illumination of the form $\Phi_S(z) = \Phi_S(z'(z)) = \mathrm{Re}\{\frac{c' E_0^*}{2} z\} + \mathrm{Re}\{\frac{c' E_0^*}{2z}\}$. Note that the latter corresponds to the coherent superposition of a uniform electric field and a dipole source

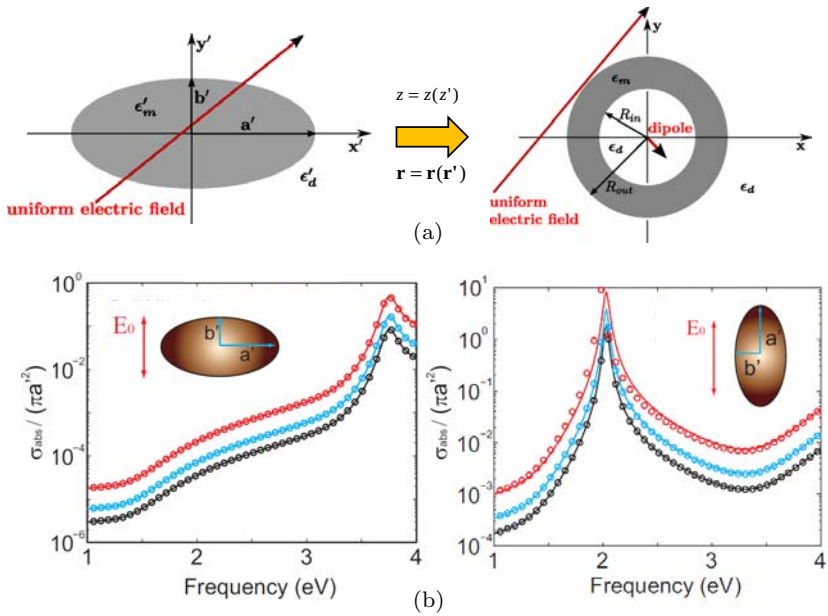


Fig. 5.13. (a) Sketch of the 2D (3D) mapping that transforms an ellipse (spheroid) into a circular ring (spherical shell). The transformed geometry is rotationally symmetric in xy-plane. (b) Normalized absorption spectra for silver oblate (left) and prolate (right) spheroids ($b' = 0.5a'$) for several a': 5 (black), 10 (cyan) and 30 (red) nm. Solid lines (dots) render TO (numerical) calculations.[27]

located at the coordinate origin. This shows that TO transfers the asymmetric character of the ellipse into a more complex EM excitation of the ring. Thus, it enables us to establish a link between the ellipse and the degenerate ring SP spectra, and interpret the plasmonic mode splitting in the former geometry in terms of their coupling to the external source.[27]

Figure 5.13(a) can be interpreted as the mapping of a spheroid into a spherical shell. This is possible thanks to the 3D version of the conformal transformation in Eq. (5.16), which reads

$$\rho' = \frac{c}{2}\left(r \mp \frac{1}{r}\right)\sin\theta, \quad z' = \frac{c}{2}\left(r \pm \frac{1}{r}\right)\cos\theta, \tag{5.17}$$

where primed (unprimed) spatial variables are expressed in cylindrical (spherical) coordinates, and the upper (lower) sign must be used to treat oblate (prolate) spheroids. Equation (5.17) results in tensorial permittivity and permeability distributions (depending on r and θ) in the ring frame. An analytical solution for Eq. (5.2) can be found for the complex $\bar{\varepsilon}(\mathbf{r})$ obtained under this transformation. It can be shown that the solution basis for the problem is given by Laplace's wave-functions in which the radial variable r is replaced by $r \pm 1/r$.[27]

Figure 5.13(b) plots the absorption cross-section for oblate (left) and prolate (right) Ag^{49} spheroids. The cross-sections are normalized by the effective area $\pi a'^2$ of the spheroids (where a' is the semi-major axis). The incident electric field is aligned along the vertical axis (see inset). Solid lines show TO predictions, and circles correspond to simulations. Different colors correspond to different a': 5 nm (black), 10 nm (cyan) and 30 nm (red). In all cases, $b' = 0.2a'$. The agreement between numerical and analytical spectra is excellent. Note that radiative corrections in the form of Eq. (5.10) were implemented self-consistently in the analytical cross-sections.

5.4. Non-local Effects

We have shown that TO is a very suitable framework for the theoretical investigation of light concentration phenomena in acute metal geometries, such as the touching point between nanoparticles or the apex of crescent-shaped tips. It provides design strategies for the optimization of these effects and sheds light into the sensitivity of plasmonic performance against geometric bluntness and imperfections. In this section, we discuss how the metal permittivity description can be improved beyond local electromagnetism in order to refine the TO recipes for plasmonic enhancement and localization.

Figure 5.3(b) illustrates the focusing performance of touching silver nanowires. It plots the electric field amplitude along the perimeter of the structure and shows that the drastic EM group velocity reduction and fast spatial SP oscillations take place within an azimuthal range smaller than 1°. For the geometry yielding the largest field enhancement, $D = 20$ nm, this angular distance

corresponds to 0.3 nm. This spatial extent is comparable to the Thomas–Fermi screening length in noble metals, $\delta_F \approx 0.1$ nm for silver.[66] Therefore, an accurate description of SP-assisted phenomena in this sub-nanometer regime requires the implementation of spatially dispersive (non-local) permittivities beyond the Drude free electron gas,[10] taking into account the effect of electron–electron interactions.[67–69]

The simplest treatment of the dielectric properties of metals that reflect the fermionic character of conduction electrons is the so-called hydrodynamical Drude model.[70–72] See Chapter 3 for a more comprehensive review on this topic. This introduces a pressure-like term in the Newtonian dynamics for the Fermi electron gas density, leading to a tensorial permittivity with transverse and longitudinal components of the form

$$\begin{aligned} \varepsilon_T(\omega) &= \varepsilon_\infty \left[1 - \frac{\omega_p^2}{\omega(\omega + i\gamma)}\right], \\ \varepsilon_L(k,\omega) &= \varepsilon_\infty \left[1 - \frac{\omega_p^2}{\omega(\omega + i\gamma) - \beta^2 k^2}\right]. \end{aligned} \tag{5.18}$$

Note that ε_∞, ω_p and γ are usual Drude constants obtained from the fitting to experimental data.[48,49] The parameter $\beta \approx 10^6$ m/s measures the degree of non-locality and is proportional to the Fermi velocity.[66] As discussed in Section 5.6, this parameter has also been determined experimentally for gold.[73] Equations (5.18) show that, within the hydrodynamic Drude model, the transverse permittivity retains its local form, whereas the longitudinal one becomes spatially dispersive and acquires an explicit dependence on the modulus of the EM wave-vector, $k = |\mathbf{k}|$.

The k-dependence of the longitudinal permittivity in Eq. (5.18) allows for the excitation of longitudinal plasmon (LP) modes at optical frequencies (below ω_p), whose dispersion relation, $\varepsilon_L(q_{LP}, \omega) = 0$, yields[72]

$$q_{\mathrm{LP}} = \frac{1}{\beta}\sqrt{\omega_p^2 - \omega(\omega + i\gamma)} \tag{5.19}$$

for the LP evanescent wave-vector. The LP decay into the metal translates into the thickening of the polarization charges induced at the metal surface.[72] In the local approximation, these charges accumulate abruptly at the structure boundaries, whereas in the non-local picture, the induced charge distribution acquires a finite width $\delta_{\rm LP} = 1/q_{\rm LP}$. Note that this width scales linearly with β, having $\delta_{\rm LP} \approx 0.1\,$nm for noble metals.

The TO approach introduced in Section 5.2.1 for the description of the optical properties of touching nanowires can be extended to investigate how the non-local blurriness of metal boundaries affects the focusing ability of this device. This requires the transformation of the permittivity tensor in Eq. (5.18) under the conformal inversion of Eq. (5.3). The mapping only modifies the longitudinal permittivity and yields a spatially dependent non-local parameter of the form

$$\beta'(z') = \left|\frac{{\rm d}z'}{{\rm d}z}\right| \beta = \frac{|z'|^2}{g^2}\beta. \tag{5.20}$$

Figure 5.14(a) shows the inversion of a pair of touching nanowires, similar to Fig. 5.2(d), but including the non-local surface charge redistribution according to Eq. (5.20). The original uniform charge thickness (along the nanowires boundaries) maps into a non-uniform width, $\delta'_{\rm LP}(z')$, that increases along the transformed flat metal surfaces away from the origin.

An analytical solution to Eq. (5.2) for the transformed system in Fig. 5.14(a) can be built through the so-called WKB or eikonal approximation, which assumes that $\delta'_{\rm LP}(z')$ varies in space much more slowly than the SPs oscillating fields.[31] The main panel of Fig. 5.14(b) shows the comparison between TO theory (solid lines) and numerical simulations (dots) for the absorption cross-section of 10 nm radius touching Ag nanowires and different degrees of non-locality (red color corresponds to realistic β). The agreement is remarkable. Note the significant impact of nonlocal effects on the absorption spectrum, leading to a set of maxima and minima, instead of the continuous broad response in the local approximation (black solid line). Importantly, the position of the absorption maxima blueshifts for larger β.

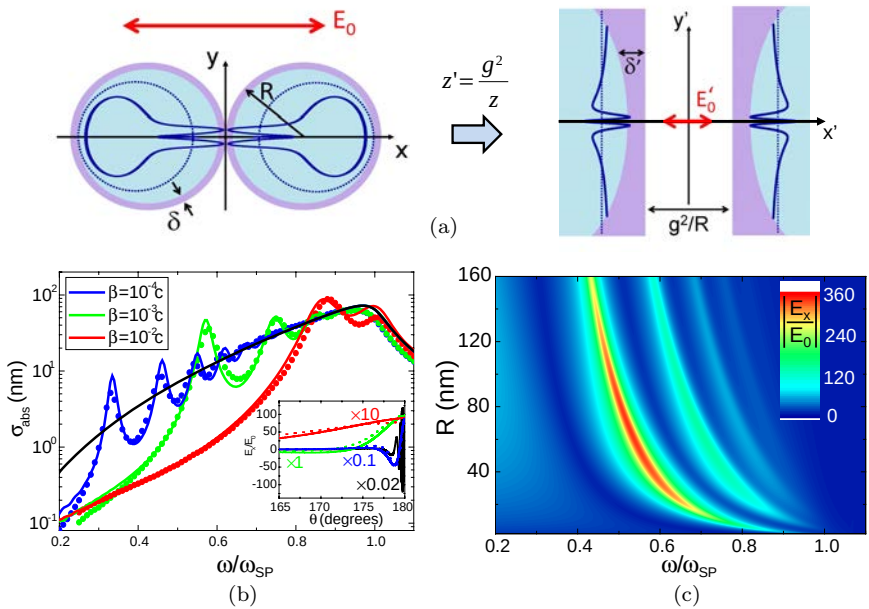


Fig. 5.14. Inversion mapping of two touching nanowires into a metal-vacuum-metal geometry. The non-local uniform surface charge thickening acquires a spatial dependence in the transformed frame. (b) Absorption cross-section of 10 nm radius Ag touching nanowires for different degrees of non-locality. The inset renders the electric field enhancement in the vicinity of the touching point for $\omega = 0.5\omega_{sp}$. (c) Electric field enhancement at the touching point versus frequency and nanowire radius.[31]

The inset of Fig. 5.14(b) renders the electric field amplitude in the vicinity of the touching point at $\omega = 0.5\,\omega_{\rm sp}$ for the various configurations in the main panel. It clarifies the optical implications of non-locality. Due to the blurriness of the metal boundaries, the touching point is no longer an EM singularity, the electric field amplitude does not vanish and SPs can circulate around the nanowire. This leads to the emergence of maxima and minima in the absorption cross-section. Moreover, plasmonic fields no longer oscillate in spatial ranges as small as 0.1 nm, but span into much larger azimuthal angles. This reduces the focusing efficiency and field enhancement ability of the structure (of the order of 10^2 for realistic Fermi velocity).[31]

Spatial dispersion in the metal permittivity diminishes the light collection and concentration ability of plasmonic devices, being the impact of non-locality larger for smaller structures. On the contrary, radiation losses diminish SP-assisted light focusing efficiency for large systems. The inclusion of both mechanisms into the TO frame makes it possible to determine the device dimensions maximizing its performance. Thus, Fig. 5.14(c) renders the field enhancement in touching dimers as a function of frequency and the nanowire radius, showing that the optimum device size lies between 30 and 80 nm.[31]

Non-local effects do not only alter the optical response of singular geometries, but also modify plasmon hybridization in nanoparticles separated by nanometric gaps. In a similar way as in the touching geometry, spatial dispersion in the metal permittivity can be introduced approximately in, for instance, the TO approach in Fig. 5.6(a). This allows describing the impact of non-locality in the optical properties of separated nanowires, such as the SP mode splitting shown in Fig. 5.7.[32]

A more general strategy for the inclusion of non-local effects in the TO frame can be done through a recently developed simplified modelling for non-locality.[74] This method consists in replacing the spatially dispersive metal by a compound system formed by a usual local metal and a thin fictitious dielectric cover. The thickness, Δd, and permittivity, $\varepsilon_{\mathrm{NL}}$, of this layer are designed so that the transmission and reflection coefficients for all incident wave-vectors and at all frequencies are the same as in the spatially dispersive metal. It can be demonstrated that for geometries with a radius of curvature larger than the LP decay length, δ_{LP} ($\sim$0.1 nm), these must fulfil

$$\frac{\varepsilon_{\mathrm{NL}}}{\Delta d} = \frac{\varepsilon(\omega)}{\varepsilon(\omega) - 1} q_{\mathrm{LP}}, \tag{5.21}$$

where $\varepsilon(\omega)$ is the local (spatially non-dispersive) metal permittivity and q_{LP} is given by Eq. (5.19). This model has been implemented in the TO description of separated nanowires,[74] showing an excellent agreement with numerical simulations and with previous TO non-local descriptions of the same geometry.[32]

Equation (5.21) provides the local equivalent to a non-local system, and maps a bulk spatially dispersive metal into a slightly more complicated (layered) local structure. Such geometry has been implemented in the inversion illustrated in Fig. 5.12, making possible the TO description of non-local effects in gapped nanospheres.[33] Note that the inclusion of the fictitious non-local dielectric layer translates into more complicated continuity equations for the electric field in the transformed frame. These are no longer tri-diagonal, but penta-diagonal expressed in matrix form. Figure 5.14 renders the absorption cross-section of a pair of gold spheres as a function of the incident frequency and the gap size, δ. Two different radii are considered: 5 nm (a) and 30 nm (b). The comparison between these contour plots and Fig. 5.7(a) clarifies the effect of spatial dispersion in the optical properties of separated dimers.

For large gaps, the TO theory reproduces the non-local blueshift experienced by the dipolar SPs of isolated cylinders and spheres.[67,71] More importantly, Figs. 5.15(a)–5.15(b) show that spatial dispersion truncates the continuous redshifting of the absorption maxima with decreasing δ obtained within the local approximation (see Fig. 5.7). This truncation also explains the emergence of maxima and minima in the spectra for touching nanowires (see Fig. 5.14). The LP decay length, δ_{LP}, introduces a new length scale in the system. For gap sizes below this distance, due to the non-local-induced metal boundary blurriness, the absorption cross-section saturates, and the optical properties of separated and touching dimers are exactly the same.

Figure 5.15 also shows that, as anticipated, the impact of spatial dispersion is larger for smaller particle size. For 5 nm radius spheres, the non-local truncation allows only for 2 SP resonances (a). On the contrary up to 4 distinct modes can be observed for 30 nm radius (b). This is even clearer in the lower panels of Fig. 5.15, which show the comparison between quasi-analytical calculations and local and non-local numerical simulations for a gap size $\delta = 0.2$ nm and for 5 nm (c) and 30 nm (d) radii.[33] In Section 5.6, the experimental verification of non-local effects in the scattering resonances of a gold nanosphere gold substrate configuration[73] will be discussed.

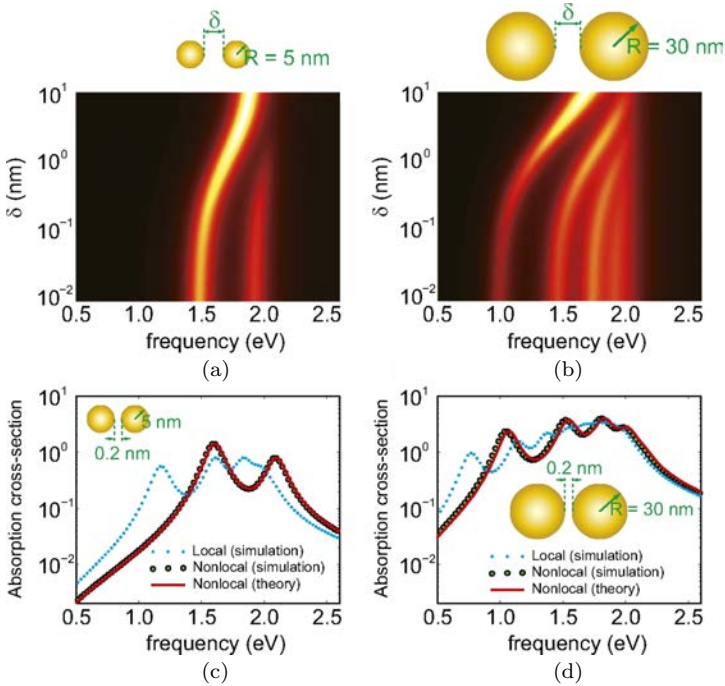


Fig. 5.15. Contour plots of absorption cross-section versus incident frequency and the separation between a pair of gold nanospheres with equal radii of 5 nm (a) and 30 nm (b). Comparison of TO quasi-analytical calculations with local and non-local numerical simulations for two closely separated ($\delta = 0.2$ nm) gold spheres with equal radii of 5 nm (c) and 30 nm (d).[33]

5.5. Van der Waals Interactions

vdW forces are caused by the EM interaction between quantum fluctuation-induced charges. These are ubiquitous but subtle forces which play an important role in a wide range of surface phenomena such as adhesion, friction and colloidal stability. In the limiting cases of small molecules and large macroscopic bodies, vdW forces can be calculated by asymptotic equations.[75] Moreover, despite the intricacy of the problem, there are various approaches available to compute vdW interactions between nanostructures whose size and separation are comparable. Traditional procedures, such as the Hamaker's method[76] or the proximity force approximation

(PFA),[77] compute vdW forces via summing over infinitesimal elements. However, due to the fact that vdW forces are inherently non-additive, these techniques are not always valid. Some recently developed methodologies[78] and numerical techniques[79] are, in principle, capable of dealing with sophisticate geometries. However, for very closely spaced plasmonic nanostructures with nanometer separations, the calculation of vdW forces is still challenging due to the high confinement of SP field in the gap as shown in Fig. 5.3(b).

Although vdW forces have a purely quantum origin, the crucial step in its calculation turns out to be a classical EM problem.[80,81] Intuitively, the vdW energy between two objects with a separation δ is calculated by summing over the zero-point energy of all the eigenmodes supported by the system.[83] This yields an expression of the form[35]

$$E(\delta) = \sum_n \frac{1}{2}\hbar\omega_n(\delta) - \sum_n \frac{1}{2}\hbar\omega_n(\delta \to \infty), \tag{5.22}$$

where ω_n is the eigenfrequency for eignenmode n. Then, the associated force can be calculated as $F(\delta) = -\mathrm{d}E/\mathrm{d}\delta$.

Equation (5.22) is extremely difficult to implement for plasmonic systems, as these support an infinite number of SP eigenmodes. Moreover, finding these resonances is a difficult task, since the corresponding eigenfrequencies lie within a 2D complex frequency plane due to the inherent lossy character of metals. In light of the fact that the roots of $\det(\mathbf{1} - \mathbf{R}_1\mathbf{R}_2) = 0$ give the resonant frequencies for all the SP eigenmodes (the $\mathbf{R}$'s are the reflection matrices for the interacting objects), and using the Cauchy's residue theorem, the summation of Eq. (5.22) can be converted into an integral weighted by the zero point energy,[82,83]

$$E(\delta) = \frac{\hbar}{2\pi}\int_0^\infty \mathrm{Im}\left\{\ln\left[\det(\mathbf{1} - \mathbf{R}_1\mathbf{R}_2)\right]\right\}\mathrm{d}\omega. \tag{5.23}$$

Using this scattering approach, Casimir's theory[84] can be extended to deal with systems whose reflection coefficients can be analytically

expressed (such as two dielectric planes and a molecule and a plane).

TO allows for the quasi-analytical calculation of the reflection coefficients in Eq. (5.23) for a range of complex geometries. For instance, the inversion transformation in Fig. 5.12 is exploited to obtain $\mathbf{R}_1$ and $\mathbf{R}_2$ for two spheres in close contact.[26] These reflection matrices are written in terms of the continuity equations for the electric field in the transformed frame. Importantly, in the limit of small separations, TO yields diagonal $\mathbf{R}$-matrices.[34] Plugging these in Eq. (5.23), vdW energies can be expressed as a simple summation. Moreover, as discussed in Section 5.3.3, TO provides an efficient strategy to identify and characterize all the SP resonances supported by the system, whose dispersion with the gap distance governs the vdW interactions between two spheres.

The left panel of Fig. 5.16 plots fully converged TO calculations of the absolute value of the vdW energy between two identical Ag[48] spheres and between a sphere and a plane as a function of the gap distance normalized to the sphere radius. The kinetic energy of a Brownian particle at room temperature, $k_BT/2 = 0.013\,\text{eV}$, is indicated by a black dashed line. $|E(\delta)|$ exceeds this value when $\delta/R < 0.86$ for the sphere–plane configuration, and $\delta/R < 0.42$ for the sphere–sphere geometry. In the small gap limit, the vdW energy for the former is roughly twice $|E(\delta)|$ for the latter. In this region, both vdW energies are inversely proportional to δ/R. In the large gap limit, the sphere-plane energy decays much more slowly than the sphere–sphere one, with characteristic decays $(\delta/R)^{-3}$ and $(\delta/R)^{-6}$, respectively. Importantly, the lack of geometric singularities in the transformed geometry of Fig. 5.12 translates into the fast convergence of TO-based vdW calculations. Thus, only 35 (100) $\Phi'_{lm}(\mathbf{r}')$ terms were required to reach 99.99% convergence for a gap-to-radius ratio of 10^{-2} (10^{-3}) in the sphere-plane geometry in the left panel of Fig. 5.16.

The right panel of Fig. 5.16 plots $E(\delta)$ obtained from Eq. (5.23) for a pair of 5 nm radius gold[48] spheres. The panel shows the comparison between TO predictions obtained from the local (dashed

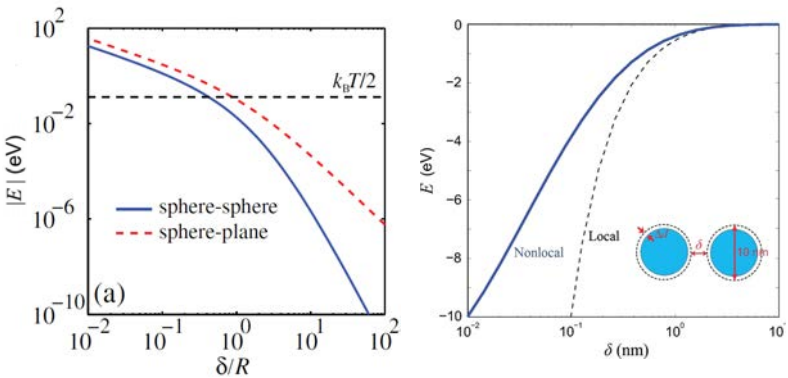


Fig. 5.16. Left: Absolute value of the vdW energy between two Ag spherical nanoparticles and between a sphere and a plane as a function of the separation (normalized to the radius) obtained using TO within the local approximation. The characteristic kinetic energy of Brownian particle at room temperature is indicated by the dashed line.[34] Right: vdW energy between two 5 nm radius Au nanospheres as a function of the separation. Solid (dashed) line plots TO nonlocal (local) predictions.[33]

line) and non-local (solid line) modelling of the metal permittivity in the calculation of the **R**-matrices for the system. Non-local effects are negligible for gap sizes larger than 1 nm, and both levels of description yield similar results. However, for δ comparable to $\delta_{\mathrm{LP}} \approx 0.1$ nm, spatial dispersion in the metal permittivity leads to a drastic reduction in the vdW energy between the nanospheres. The local $1/\delta$ dependence of $E(\delta)$ in the small gap limit is lost once non-local corrections are taken into account. Instead, TO calculations show that $E(\delta)$ saturates at the touching limit for gap sizes smaller than the LP decay length. This trend can be understood in terms of the non-local blurriness of metal boundaries discussed in Section 5.4.

5.6. Experimental Realizations

In the previous sections, we have shown that much theoretical effort has been invested to assess and analyze the role played by acute geometric features in TO designed plasmonic devices. Through the

inclusion of bluntness and non-local effects into the TO framework, more robust devices, not sensible to structural imperfections in the sub-nanometric scale, have been proposed. However, the experimental realization of TO designed nanostructures remains a challenge. On the one hand, their fabrication requires an enormous precision in order to define accurately their geometry.[85–88] On the other hand, their full optical characterization involves the probing of the local field enhancement with spatial resolution at the nanoscale.[89–91] Here, we discuss briefly three works that, going along these lines, found experimental evidence in agreement with TO predictions for the close encounter between plasmonic nanoparticles.

The left panel of Fig. 5.17(a) shows the dark-field scattering spectra measured for different incident polarizations (see insets) from a 155 nm diameter gold sphere on top of a gold film.[92] The spectra present three scattering maxima within the wavelength range under study. The right panel of Fig. 5.17(a) plots the scattering cross-sections obtained through local numerical simulations for a nanosphere separated by a 2 nm gap from the substrate. Note that both the position and shape of the measured scattering maxima are reproduced numerically. The insets render the electric field amplitude evaluated at the three resonances supported by the geometry, proving the strong field enhancement taking place at the gap. Through the analogy with the TO predictions for the 2D analogue for the system, a nanowire on top of an infinite substrate,[19] the nature of the SP modes behind these scattering peaks can be clarified. The long wavelength resonance, which is efficiently excited by 90^o illumination, is caused by a dipolar mode oriented normally to the substrate. The intermediate maximum, which is apparent for all incident polarizations, corresponds to a hybrid mode with bright dipolar components parallel (located at the nanosphere) and normal (at the gap) to the gold film. Finally, the origin of the short wavelength peak is a high-order (quadrupole) mode strongly confined at the gap of the geometry.

A strategy to ease the fabrication of TO designs consists in using semiconductor materials as the plasmonic platform.[93] The dielectric

properties of semiconductors, such as intrinsic InSb, can be described through a lossy Drude model with a plasma frequency lying at the THz range.[94] In contrast, the plasma frequency for noble metals is located at the ultraviolet range of the EM spectrum. Therefore, by replacing the plasmonic material, the operating wavelength of TO devices can be shifted from the sub-micron (metals) to the sub-millimeter (semiconductors) regime. This makes possible the realization of the plasmonic phenomena predicted within the TO framework in much larger structures.

The left panel of Fig. 5.17(b) shows the THz absorption cross-section measured from a square array of touching dimers of 10-micron radius InSb disks for different incident polarizations. Note that the InSb plasma frequency is 2.4 THz. The gap region at each dimer was patterned using electron beam lithography, giving rise to a V-shaped gap between the disks. See the right insets of Fig. 5.17(b) for the lateral and top SEM images of the experimental samples.

Although the experimental InSb disks have a finite thickness, the measured cross-sections in Fig. 5.17(b) demonstrate the broadening of the absorption spectrum of the disks as the incident polarization is rotated from normal to parallel to the dimer axis. This observation is in agreement with the broadband response obtained within the TO framework for touching nanowires, see Fig. 5.3(a), and nanospheres, see Fig. 5.10(b). The bottom insets of Fig. 5.17(b) render the charge distribution maps for the different SP resonances sustained by the experimental geometry. Note that, whereas the system supports only one resonance under normal illumination, two different modes emerge when the impinging fields are polarized along the dimer axis. In agreement with Fig. 5.7(a) in the region of separations $\sim$10 times smaller than the nanowire radius, it is the spectral overlapping between two distinct plasmonic modes which causes the broadening of the absorption cross-section.

Recently, much research attention has focused on the experimental exploration of the impact that spatial dispersion in the metal permittivity has on the optical properties of plasmonic nanostructures. Figure 5.18 shows an experimental configuration in which

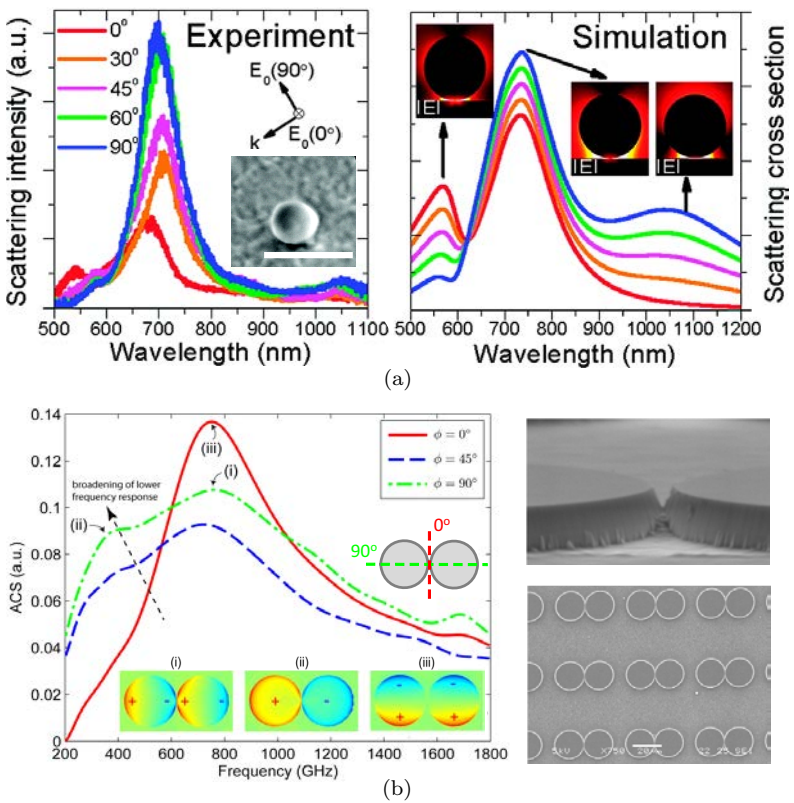


Fig. 5.17. (a) Measured (left) and calculated (right) dark-field scattering spectra for a 155 nm diameter gold sphere standing on top of a gold film. The various spectra correspond to different incident polarizations. The left inset shows a scanning electron microscopy (SEM) image of the experimental sample (white bar: 500 nm). The right insets render the field enhancement evaluated at the various scattering maxima.[92] (b) Left: measured THz absorption cross-section variation with the incident polarization angle for InSb touching disks of 10-micron radius. The labeled absorption peaks correspond to the charge distribution maps in the inset. Right: lateral and top SEM images of the experimental InSb samples.[93]

the deviation of empirical observations from local EM predictions has been quantified through far-field measurements.[73] The experimental samples consist of gold nanoparticles seating on top of a gold substrate. Importantly, the nanoparticle–substrate distance was controlled with sub-nanometric precision by chemical means. Spacer

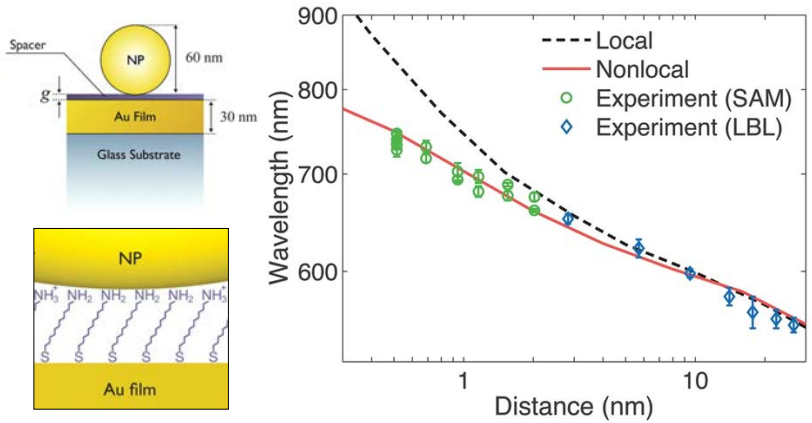


Fig. 5.18. Left: Geometry of the gold film gold nanoparticle experimental sample (top). Zoom in the gap region, showing the nanoparticle-film separation by an amine-terminated alkanethiol SAM. Right: Comparison of experimental measurements from SAM and layer-by-layer-type spacers with numerical local and non-local simulations.[73]

layers were grown on the flat gold surface using either layer-by-layer deposition of polyelectrolytes,[95] for separations that range from 2.8 to 26.6 nm, or by the formation of self-assembled monolayers (SAMs) of amine-terminated alkanethiols for even smaller separations, from 0.5 to 2.0 nm.

The optical properties of the nanoparticle–substrate system were characterized through single particle dark-field spectroscopy measurements. Figure 5.18 shows the spectral position of the lowest energy scattering resonance measured from the samples as a function of the gap distance (dots). For comparison, local EM predictions are plotted in black solid line. For gap sizes between 1 and 10 nm, experiments and theory show a remarkable agreement. On the contrary, local calculations yield a redshifting of SP resonances at subnanometric distances, which is not found in the experiments.

The red solid line in Fig. 5.18 renders numerical results obtained for the same geometric parameters as the local calculations, but refining the gold permittivity through a hydrodynamical Drude model. These reproduce the experimental trend not only at large but also at

small particle–substrate separations. Importantly, TO sheds physical insight into this deviation of experimental and theoretical non-local spectra from local predictions. Figures 5.15(a) and 5.15(b) show a similar result for nanosphere dimers. In Section 5.4, we associated this effect with the blurriness of metallic boundaries originated from the optical excitation of LP modes, which is only significant for gap sizes comparable to δ_{LP}. The numerical fitting to the experimental spectrum in Fig. 5.18 allowed for the empirical determination of the non-local parameter for gold, obtaining $\beta = 1.27 \times 10^6$ m/s.[73]

5.7. Conclusions

TO is an elegant, largely analytical tool for the design and analysis of plasmonic nanostructures. We have shown that it allows revealing the key features that optimize SP-assisted far-to-near-field EM energy conversion and plasmonic field enhancement and localization. 2D and 3D devices presenting geometric singularities and those in which strong plasmon hybridization phenomena take place were investigated in detail. Through the inclusion of geometric bluntness and non-local effects, this TO approach has been refined. Realistic TO models taking into account shape imperfections and material limitations mean a step forward of the framework toward practical applications. We have also shown that the range of applicability of this powerful method is beyond the description of the optical properties of metals. TO also allows for the quasi-analytical treatment of vdW interactions between objects separated by nanometric distances. Finally, we have reviewed recent experimental reports inspired by or in excellent agreement with TO predictions.

Acknowledgments

We gratefully acknowledge significant contribution and useful discussions with Alexandre Aubry, Dangyuan Lei, Yannick Sonnefraud, Matthias Kraft and Stefan A. Maier. We thank the following for support: the ESF Plasmon-Bionanosense Network, the EU Marie Curie IEF scheme, the Engineering and Physical Sciences Research Council

(EPSRC), the Gordon and Betty Moore Foundation, the Royal Commission for the Exhibition of 1851 and the Leverhulme Trust.

References

1. Ward, A. J. and Pendry, J. B. (1996). *J. Mod. Opt.* **45**, p. 773.
2. Pendry, J. B. Schurig, D. and Smith, D. R. (2006). *Science* **312**, p. 1780.
3. Chen, H. Y., Chan, C. T. and Sheng, P. (2010). *Nat. Mater.* **9**, p. 387.
4. Kundtz, N. B., Smith, D. R. and Pendry, J. B. (2011). *Proc. IEEE* **99**, p. 1622.
5. Kildishev, A. V. and Shalaev, V. M. (2011). *Physics-Uspekhi* **54**, p. 53.
6. Schurig, D., Mock, J. J., Justice, B. J., Cummer, S. A., Pendry, J. B., Starr, A. F. and Smith, D. R. (2006). *Science* **314**, p. 977.
7. Ergin, T. Stenger, N., Brenner, P., Pendry, J. B. and Wegener, M. (2010). *Science* **328**, p. 337.
8. Pendry, J. B. Luo, Y. and Zhao, R. (2015). *Science* **348**, p. 6234.
9. Schurig, D., Pendry, J. B. and Smith, D. R. (2006). *Opt. Express* **14**, p. 9794.
10. Maier, S. A. *Plasmonics: Fundamentals and Applications* (Springer-Verlag, New York 2007).
11. Novotny, L. and Hetch, B. *Principles of Nano-Optics* (Cambridge University Press, Cambridge 2006).
12. Pendry, J. B., Aubry, A., Smith, D. R. and Maier, S. A. (2012). *Science* **337**, p. 549.
13. Luo, Y., Zhao, R. Fernández-Domínguez, A. I., Maier S. A. and Pendry, J. B. (2013). *Sci. China Inf. Sci.* **56**, p. 12 0401.
14. Aubry, A., Lei, D. Y., Maier S. A. and Pendry, J. B. (2010). *Phys. Rev. B* **82**, p. 205109.
15. Aubry, A., Lei, D. Y., Fernández-Domínguez, A. I., Sonnefraud, Y., Maier S. A. and Pendry, J. B. (2010). *Nano Lett.* **10**, p. 2574.
16. Lei, D. Y., Aubry, A., Maier, S. A. and Pendry, J. B. (2010). *New J. Phys.* **12**, p. 093030.
17. Aubry, A., Lei, D. Y., Maier, S. A. and Pendry, J. B. (2010). *Phys. Rev. B* **82**, p. 125430.
18. Aubry, A., Lei, D. Y., Maier, S. A. and Pendry, J. B. (2010). *Phys. Rev. Lett.* **105**, p. 233901.
19. Aubry, A., Lei, D. Y., Maier, S. A. and Pendry, J. B. (2011). *ACS Nano* **5**, p. 3293.
20. Lei, D. Y., Aubry, A., Luo, Y., Maier, S. A. and Pendry, J. B. (2011). *ACS Nano* **5**, p. 597.
21. Luo, Y. Lei, D. Y., Maier, S. A. and Pendry, J. B. (2012). *Phys. Rev. Lett.* **108**, p. 023901.
22. Luo, Y., Lei, D. Y., Maier S. A. and Pendry, J. B. (2012). *ACS Nano* **6**, p. 6492.

23. Fernández-Domínguez, A. I., Maier S. A. and Pendry, J. B. (2010). *Phys. Rev. Lett.* **105**, p. 266807.
24. Fernández-Domínguez, A. I., Maier S. A. and Pendry, J. B. (2012). *Phys. Rev. B* **85**, p. 165148.
25. Fernández-Domínguez, A. I., Luo, Y. Wiener, A., Pendry, J. B. and Maier, S. A. (2012). *Nano Lett.* **12**, p. 5946.
26. Pendry, J. B., Fernández-Domínguez, A. I., Luo, Y. and Zhao, R. (2013). *Nat. Phys.* **9**, p. 518.
27. Kraft, M., Pendry, J. B., Maier, S. A. and Luo, Y. (2014). *Phys. Rev. B* **89**, p. 1098.
28. Luo, Y. Pendry, J. B. and Aubry, A. (2010). *Nano Lett.* **10**, p. 4186.
29. Luo, Y., Aubry A. and Pendry, J. B. (2011). *Phys. Rev. B* **83**, p. 155422.
30. Kraft, M., Luo, Y., Maier, S. A. and Pendry, J. B. (2015). *Phys. Rev. X* **5**, p. 2160.
31. Fernández-Domínguez, A. I., Wiener, A., García-Vidal, F. J., Maier, S. A. and Pendry, J. B. (2012). *Phys. Rev. Lett.* **108**, p. 106802.
32. Fernández-Domínguez, A. I., Zhang, P., Luo, Y., Maier, S. A., García-Vidal, F. J. and Pendry, J. B. (2012). *Phys. Rev. B* **86**, p. 241110.
33. Luo, Y., Zhao, R. and Pendry, J. B. (2014). *Proc. Natl. Acad. Sci. USA* **111**, p. 18422.
34. Zhao, R. Luo, Y., Fernández-Domínguez, A. I. and Pendry, J. B. (2013). *Phys. Rev. Lett.* **111**, p. 033602.
35. Zhao, R., Luo Y. and Pendry, J. B. (2015). *Sci. Bull.* DOI 10.1007/s11434-015-0958-x.
36. Schinzinger, R. and Laura, P. A. A. (2003). *Conformal Mapping Methods and Applications* (Dover, New York, 2003).
37. Zayats, A. V. and Maier, S. A. (2013). Active plasmonics and tunable plasmonic metamaterials (Wiley, New Jersey).
38. Mcphedran, R. C. and Perrins, W. T. (1981). *Appl. Phys.* **24**, p. 311.
39. Mcphedran, R. C. and Milton, G. W. (1987). *Proc. R. Soc. A-Math. Phys. Eng. Sci.* **411**, p. 313.
40. Garcia-Vidal, F. J. and Pendry, J. B. (1996). *Phys. Rev. Lett.* **77**, p. 1163.
41. Xu, H. X., Aizpurua, J., Kall, M. and Apell, P. (2000). *Phys. Rev. E* **62**, p. 4318.
42. Kottmann, J. P. and Martin, O. J. F. (2001). *Opt. Express* **8**, p. 655.
43. Talley, C. E. Jackson, J. B., Oubre, C., Grady, N. K., Hollars, C. W., Lane, S. M., Huser, T. R., Nordlander, P. and Halas, N. J. (2005). *Nano Lett.* **5**, p. 1569.
44. Sweatlock, L. A., Maier, S. A., Atwater, H. A., Penninkhof, J. J. and Polman, A. (2005). *Phys. Rev. B* **71**, p. 235408.
45. Romero, I., Aizpurua, J., Bryant, G. W. and de Abajo, F. J. G. (2006). *Opt. Express* **14**, p. 9988.
46. Lassiter, J. B., Aizpurua, J., Hernandez, L. I., Brandl, D. W., Romero, I., Lal, S., Hafner, J. H., Nordlander, P. and Halas, N. J. (2008). *Nano Lett.* **8**, p. 1212.

47. Atre, A. C., Garcia-Etxarri, A., Alaeian H. and Dionne, J. A. (2013). *Adv. Optical Mater.* **1**, p. 327.
48. Palik, E. D. *Handbook of Optical Constants of Solids* (Academic Press, New York, 1985).
49. Johnson, P. B. and Christy, R. (1972). *Phys. Rev. B* **6**, p. 4370.
50. Moskovits, M. (1985). *Rev. Mod. Phys.* **57**, p. 783.
51. Anker, J. N., Hall, W. P., Lyandres, O., Shah, N. C., Zhao, J. and Van Duyne, R. P. (2008). *Nat. Materials* **7**, p. 442.
52. Wu, L. Y., Ross, B. M. and Lee, L. P. (2009). *Nano Lett.* **9**, p. 1956.
53. Shumaker-Parry, J. S., Rochholz, H. and Kreiter, M. (2005). *Adv. Mater.* **17**, p. 2131.
54. Rochholz, H. Bocchio, N. and Kreiter, M. (2007). *New J. Phys.* **9**, p. 53.
55. Bukasov, R. and Shumaker-Parry, J. S. (2007). *Nano Lett.* **7**, p. 1113.
56. Lukyanchuk, B., Zheludev, N. I., Maier, S. A., Halas, N. J., Nordlander, P., Giessen, H. and Chong, C. T. (2010). *Nat. Materials* **9**, p. 707.
57. Halas, N. J., Lal, S., Chang, W. S., Link, S. and Nordlander, P. (2011). *Chem. Rev.* **111**, p. 3913.
58. Rycenga, M., Cobley, C. M., Zeng, J., Li, W. Y., Moran, C. H., Zhang, Q., Qin D., and Xia, Y. N. (2011). *Chem. Rev.* **111**, p. 3669.
59. Giannini, V., Fernandez-Dominguez, A. I., Sonnefraud, Y., Roschuk, T., Fernandez-Garcia R. and Maier, S. A. (2010). *Small* **6**, p. 2498.
60. Prodan, E., Radloff, C., Halas, N. J. and Nordlander, P. (2003). *Science* **302**, p. 419.
61. Nordlander, P. and Prodan, E. (2004). *Nano Lett.* **4**, p. 2209.
62. Pendry, J. B. (2003). *Opt. Express* **11**, p. 755.
63. Carminati, R., Greffet, J.-J., Henkel, C. and Vigoureux, J. M. (2006). *Opt. Commun.* **261**, p. 368.
64. Luo, Y. (2012). *Transformation Optics Applied to Plasmonics* (PhD Thesis, Imperial College London, London).
65. Schinzinger, R. and Laura, P. A. A. (2003). *Conformal Mapping, Methods and Applications* (Dover Publications, New York).
66. Kittel, C. (1996). *Introduction to Solid State Physics* (Wiley, New York).
67. Fuchs, R. and Claro, F. (1987). *Phys. Rev. B* **35**, p. 3722;
68. Rojas, R., Claro, F. and Fuchs, R. (1988). *Phys. Rev. B* **37**, p. 6799.
69. Ruppin, R. (2001). *Opt. Commun.* **190**, p. 205.
70. David, C. and García de Abajo, F. J. (2011). *J. Phys. Chem. C* **115**, p. 19470.
71. Raza, S., Toscano, G., Jauho, A.-P., Wubs, M. and Mortensen, N. A. (2011). *Phys. Rev. B* **84**, p. 121412(R);
72. Ciraci, C., Pendry, J. B. and Smith, D. R. (2013). *Chem. Phys. Chem.* **15**, p. 1109.
73. Ciraci, C., Hill, R. T., Mock, J. J., Urzhumov, Y., Fernández-Domínguez, A. I., Maier, S. A., Pendry, J. B., Chilkoti, A. and Smith, D. R. (2012). *Science* **337**, p. 1072.

74. Luo, Y., Fernández-Domínguez, A. I., Wiener, A., Maier, S. A. and Pendry, J. B. (2013). *Phys. Rev. Lett.* **111**, p. 093901.
75. Israelachvili, J. N. (1998). *Intermolecular and Surface Forces* (Academic, San Diego).
76. Hamaker, H. C. (1937). *Physica (Amsterdam)*, **4**, 1058.
77. Deriagin, B. V., Abrikosova, I. I. and Lifshitz, E. M. (1956). *Q. Rev. Chem. Soc.* **10**, p. 295.
78. Emig, T., Graham, N., Jaffel, R. L. and Karder, M. (2007). *Phys. Rev. Lett.* **99**, p. 170303.
79. Rodriguez, A. W., Capasso, F. and Johnson, S. G. (2011). *Nat. Photonics* **5**, p. 211.
80. Casimir, H. B. G. (1948). *Proc. K. Ned. Akad. Wet.* **51**, p. 793.
81. Dzyaloshinskii, I. E., Lifshitz, E. M. and Pitaevskii, L. P. (1961). *Sov. Phys. Usp.* **4**, p. 153.
82. Davies, B. (1972). *Chem. Phys. Lett.* **16**, p. 388.
83. Intravaia, F. and Behunin, R. (2012). *Phys. Rev. A.* **86**, p. 062517.
84. Landau, L. D., Pitaevskii, L. P. and Lifshitz, E. M. (1984). *Electrodynamics of Continuous Media* (Elsevier, Butterworth-Heinemann).
85. Atay, T., Song, J. H. and Nurmikko, A. V. (2004). *Nano Lett.* **4**, p. 1627.
86. Hill, R. T., Mock, J. J., Urzhumov, Y., Sebba, D. S., Oldenburg, S. J., Chen, S. Y., Lazarides, A. A., Chilkoti, A. and Smith, D. R. (2010). *Nano Lett.* **10**, p. 4150.
87. Brown, L. V., Sobhani, H. Lassiter, J. B., Nordlander, P. and Halas, N. J. (2010). *ACS Nano* **4**, p. 819.
88. Huang, F. M., Wilding, D., Speed, J. D., Russell, A. E., Bartlett, P. N. and Baumberg, J. J. (2011). *J. Nano Lett.* **11**, p. 1221.
89. Cang, H., Labno, A., Lu, C. G., Yin, X. B., Liu, M., Gladden, C., Liu, Y. M. and Zhang, X. (2011). *Nature* **469**, p. 385.
90. Koh, A. L., Fernández-Domínguez, A. I., McComb, D. W., Maier S. A. and Yang, J. K. W. (2011). *Nano Lett.* **11**, p. 1323.
91. Savage, K. J., Hawkeye, M. M., Esteban, R., Borisov, A. G., Aizpurua, J. and Baumberg, J. J. (2012). *Nature* **491**, p. 574.
92. Lei, D. Y., Fernández-Domínguez, A. I., Sonnefraud, Y., Appavoo, K., Haglund, R. F., Pendry, J. B. and Maier, S. A. (2012). *ACS Nano* **6**, p. 3537.
93. Hanham, S. M., Fernández-Domínguez, A. I., Teng, J. H., Ang, S. S., Lim, K. P., Yoon, S. F., Ngo, C. Y., Klein, N., Pendry, J. B. and Maier, S. A. (2012). *Adv. Mater.* **24**, p. 226.
94. Giannini, V., Berrier, A., Maier, S. A., Sanchez-Gil, J. A. and Rivas, J. G. (2010). *Opt. Express* **18**, p. 2797.
95. Mock, J. J., Hill, R. T., Tsai, Y.-J., Chilkoti, A. and Smith, D. R. (2012). *Nano Lett.* **12**, p. 1757.

CHAPTER 6

Ultrafast Spectroscopy and Nonlinear Control of Single Nanoparticles and Antennas

OTTO L. MUSKENS*

University of Southampton Highfield, UK

PETER R. WIECHA and ARNAUD ARBOUET†

CEMES-CNRS, Université de Toulouse, France

With a high surface to volume ratio and strongly varying properties on nanometer length scales, the dynamics of nanomaterials are often very different from those of bulk materials. Ultrafast single-particle spectroscopy provides a means of accessing these properties, but has to overcome challenges related to the diffraction limit and small cross-sections of single nanoobjects. Several far-field and near-field approaches have been introduced capable of achieving single-particle resolution in ultrafast spectroscopy. By designing plasmonic nanosystems as an antenna, one can exploit local-field enhancements and interactions

Corresponding authors: *o.muskens@soton.ac.uk
†arnaud.arbouet@cemes.fr

between the antenna and its environment to design ultrafast nonlinear response. Design considerations for hybrid systems combining plasmonic and nonlinear materials are discussed together with the progress made in this area.

6.1. Introduction

Plasmonics, the field dealing with the free-electron response in metal nanostructures, offers some key advantages in enhancing the way we can manipulate light on length scales much smaller than the optical wavelength. The combination of strong optical resonances and spatial field confinement results in a very high strength of local electromagnetic fields directly surrounding nanostructured metals. This large field enhancement drives many of the primary applications of metal nanoparticles in plasmon-enhanced sensing and spectroscopy.

The surface plasmon (SP) resonances supported by metallic nanoparticles strongly depend on the metal itself, the particle morphology (size, shape) and its dielectric environment. This extreme sensitivity limits the information obtainable from ensemble measurements and has motivated the development of ultrasensitive optical techniques capable of detecting individual nanoobjects and characterizing their linear and nonlinear optical properties,[1–3] Time-resolved single particle studies have naturally followed providing a unique insight in the ultrafast electronic and vibrational dynamics of individual plasmonic nanostructures. As will be discussed in this chapter, these techniques rely on the nonlinear interaction between one or several ultrashort laser pulses and the metallic nanoobject. They share several important features. First, they all benefit from the strong electric field enhancement associated with SP resonances that yield detectable signals from nanoscale volumes. Second, far-field nonlinear microscopies provide a spatial resolution superior to their linear counterparts due to their superlinear dependence on the incident intensity. Plasmon modes can, therefore, be mapped in submicronic particles by scanning ultrashort laser pulses and detecting the nonlinear emission at each point of the nanoobject. Finally, nonlinear techniques generally offer a greater versatility due to the many electric fields involved. From an experimental point of view,

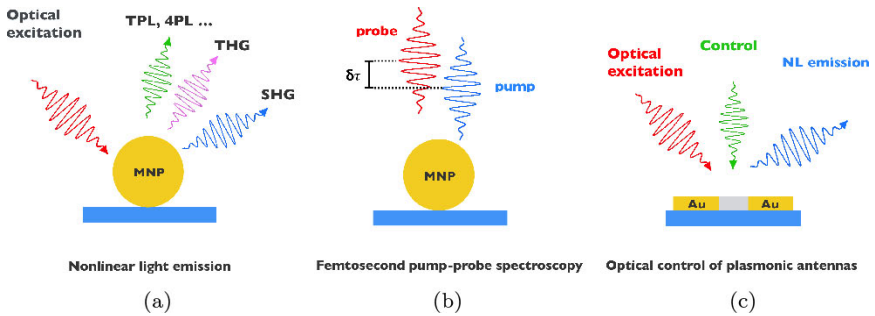


Fig. 6.1. (a) Examples of nonlinear light emission upon excitation of a metallic nanoobject by an ultrashort optical pulse: SHG, THG, Two-Photon induced Luminescence (TPL) and Four-Photon induced Photoluminescence (4PL). (b) Principle of time-resolved pump-probe spectroscopy. (c) Optical control of the nonlinear response of plasmonic antennas through the optically induced modification of their environment.

the availability of robust tabletop ultrafast laser systems, accurate nanopositioning systems and the improvement of detector technology have constantly pushed the limits of these demanding experiments.

Figure 6.1 shows several examples of ultrafast nonlinear optical interactions in individual nanoparticles. First, the nonlinear interaction of an ultrashort laser pulse with a metallic nanostructure can generate a radiation at a frequency different from in the incident waves. Second harmonic generation (SHG) and third harmonic generation (THG) are well-known examples of these processes. They respectively yield nonlinear emission at 2ω and 3ω (Fig. 6.1(a)). We will discuss them in the second part of this chapter, insisting on their differences with respect to the nonlinear interactions of propagating waves in bulk media. We will also show that a lot can be learnt about the optical resonances and ultrafast dynamics of a metallic nanoparticle from the characteristics of its nonlinear emission.

Time-resolved pump–probe spectroscopies are another important class of ultrafast nonlinear spectroscopies. Their sensitivity and versatility offer them a separate discussion in the third part of this chapter (Fig. 6.1(b)). In these experiments, a first ultrashort optical "pump" pulse is sent on a sample and the pump-induced modification of a selected optical property is analyzed by a second delayed

"probe" pulse. Whereas most pump–probe studies only involve optical pulses, sub-picosecond electron packets are now used to probe the optical near-fields of plasmonic nanostructures. This novel type of electron spectroscopy is based on recently developed ultrafast transmission electron microscopes. Finally, nonlinear interactions in plasmonic nanostructures can be used to optically control their ultrafast dynamics and optical properties (Fig. 6.1(c)).

The first part of this chapter introduces the ultrafast dynamics of metallic systems and presents the physical origin of their optical nonlinearities. The second part focuses on recent studies of the coherent or incoherent nonlinear light emission from individual plasmonic particles. Experiments addressing the electronic or vibrational dynamics of single metallic nanoparticles using time-resolved pump–probe spectroscopy are discussed in the third part. Finally, we review possible solutions for an optical control of the ultrafast dynamics in plasmonic nanostructures. We would like to emphasize that we do not consider what follows as an exhaustive review of the available literature. This is a subjective selection of examples mostly chosen for their ability to illustrate the richness of a dynamic research field. We apologize in advance for the inevitable omissions.

6.2. Ultrafast Dynamics and Optical Nonlinearities in Metallic Nanomaterials

We first present the main processes involved in the ultrafast dynamics of optically excited metallic nanomaterials and associated timescales. We then briefly introduce the two-temperature model (TTM) often used to describe the electronic dynamics on picosecond timescales. Finally, we discuss the physical origin of the optical nonlinearities in metals.

6.2.1. *Ultrafast dynamics in metallic nanostructures*

6.2.1.1. *Characteristic timescales*

An optical pulse incident on a metal nanostructure first excites the electrons of the metal. A minority of optically excited electrons then

redistributes the energy absorbed from the light beam to the rest of the electron gas by electron/electron interactions. Going back to a thermalized electronic distribution characterized by a Fermi–Dirac occupation function and an electronic temperature, T_e, involves a very large number of these elementary electronic interactions. Therefore, electronic thermalization occurs on a timescale of ∼500 fs (350 fs in bulk silver and 500 fs in bulk gold), much larger than the typical lifetime of a single electronic excitation.[4–6] Electronic thermalization only redistributes the excess energy among the electrons, but does not alter their total energy or temperature. The electron gas cools down via energy transfer to the metal lattice mediated by electron/phonon interactions. This electron/lattice thermalization occurs on a picosecond timescale in bulk and nanosized noble metals.[4,7–12] It is worth noting that the thermalization of the electron gas and the electron/lattice thermalization have an opposite dependence on the incident optical power. Whereas, the former is faster at higher fluences due to a larger fraction of optically excited electrons, the latter becomes slower due to the nonlinearity of the electronic heat capacity.

After a few picoseconds, the electron gas and the ionic lattice of the metallic nanoobject are thermalized but hotter than the environment (see Fig. 6.2). Complete thermalization is achieved by means of heat diffusion and radiation of sound waves due to the excitation of confined acoustic vibrations.[9,13–16] These latter processes occur on a 10 ps −1 ns timescale. The different processes and associated timescales are illustrated Fig. 6.2.

6.2.1.2. *The two-temperature model*

The relaxation of an optically excited metallic system involves energy transfer from the electrons to the lattice. After internal thermalization of the electron gas, the electron/lattice energy exchanges can be described using the TTM.[17] This model assumes that the lattice is kept in thermal equilibrium at a temperature T_L by anharmonic interactions. Meanwhile, it is assumed that the electron–electron interactions are fast enough to maintain the electron gas in thermal

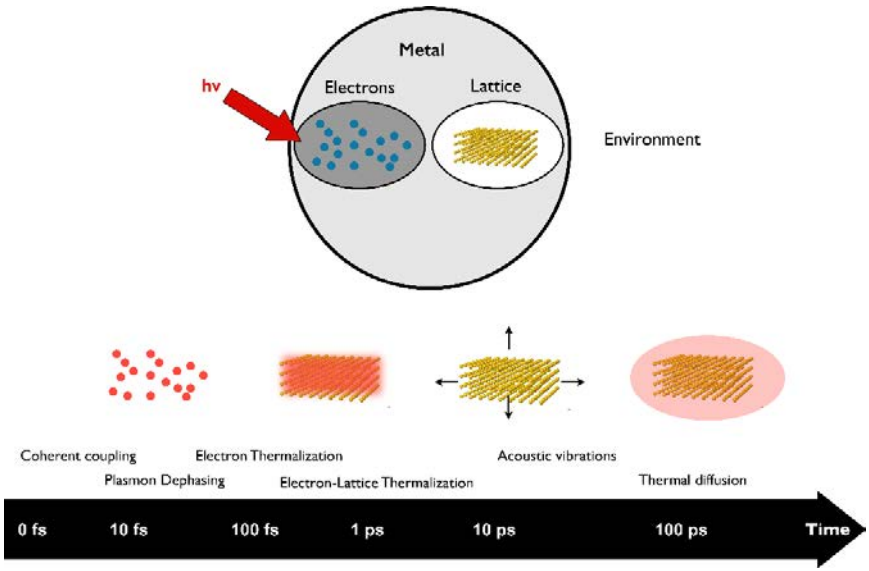


Fig. 6.2. Ultrafast dynamics in metallic systems: main processes and timescales.

equilibrium at a temperature $T_e > T_L$. The temporal evolution of the temperature of the electron gas and the lattice are given by two coupled differential equations:

$$\begin{aligned} C_e \frac{\mathrm{d}T_e}{\mathrm{d}t} &= -G(T_e - T_L) + p(t) \\ C_e \frac{\mathrm{d}T_L}{\mathrm{d}t} &= +G(T_e - T_L) \end{aligned} \tag{6.1}$$

in which G is the electron–phonon coupling constant and $p(t)$ represents the selective excitation of the electron gas by the pump pulse.

Figure 6.3(b) shows the ultrafast relaxation of the electronic distribution in gold mediated by electron–electron and electron–phonon interactions after the intraband absorption of a 20 fs infrared optical pulse (Fig. 6.3(a)). The electronic and lattice temperatures predicted by the TTM are presented in Fig. 6.3(c). Due to the small electronic heat capacity, subpicosecond optical excitation of a metal can induce a very high electronic temperature and a strongly out-of-equilibrium electronic distribution. As can be

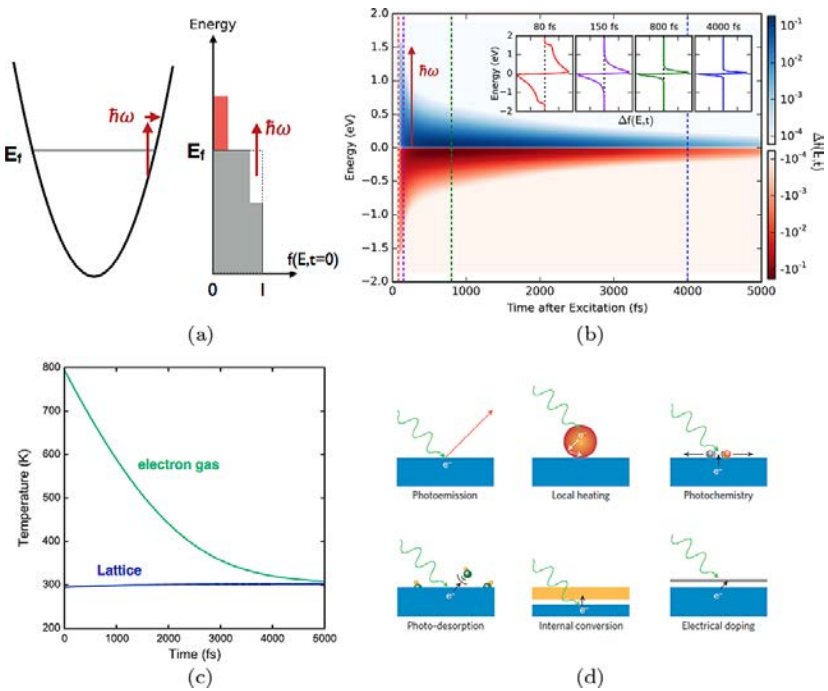


Fig. 6.3. Ultrafast dynamics of a free-electron gas: (a) Due to the Pauli exclusion principle, the absorption of an ultrashort optical pulse by a free-electron gas promotes electrons from states below the Fermi level E_f to states above the Fermi level. $f(E,t)$ denotes the electronic occupation number at energy E and time t. (b) Ultrafast electronic dynamics following excitation of gold by a femtosecond optical pulse (20 fs, $\hbar\omega = 1.55$ eV, $\Delta T = 500$ K). The relaxation dynamics is computed from a resolution of Boltzmann's equation taking into account electron–electron and electron–phonon interactions.[4,5] The colormap shows the difference between $f(E,t)$ and the initial unperturbed Fermi–Dirac distribution. At short delays after excitation, the electronic distribution is strongly athermal. Electronic interactions establish a new Fermi–Dirac distribution in about 500 fs. (c) Evolution of the temperatures of the electron gas and the lattice of gold predicted by the TTM for the same excitation conditions as in (b). Due to the small heat capacity of the electron gas, a large increase of the electronic temperature can be induced optically. (d) Possible applications of hot electrons.[18]

inferred from Fig. 6.3(b), excited electronic states lying far from the Fermi level are transiently populated on ultrashort timescales. These so-called "hot electrons" have an energy excess, which makes them useful for a variety of applications such as charge injection in

neighboring systems (e.g. quantum dots, molecules), photovoltaics, or photocatalysis (Fig. 6.3(d)).[18] This topic is currently being given a large interest in the plasmonics community for its potential applications.

Care must be taken that electronic and electron/lattice thermalizations occur on comparable characteristic timescales: electronic and lattice dynamics are strongly coupled in the early stages of relaxation. Therefore, the simulation of their dynamics on these timescales must take into account both processes. Whereas an extension of the TTM accounting for the athermal electron distribution can provide a satisfactory description, a more rigorous approach based on the computation of the electron distribution using Boltzmann's equation provides accurate results.[4,5]

6.2.2. *Optical nonlinearities in metallic nanostructures*

6.2.2.1. *Introduction to nonlinear optics*

An (EM) wave impinging on a material induces a polarization $\mathbf{P}(\mathbf{r},t)$ which can be written in the case of a lossless and dispersionless medium as a power series in the electric field strength (SI units)[19,20]:

$$\mathbf{P}(\mathbf{r},t) = \varepsilon_0(\chi^{(1)}\mathbf{E}(\mathbf{r},t) + \chi^{(2)}\mathbf{E}(\mathbf{r},t)\mathbf{E}(\mathbf{r},t) + \chi^{(3)}\mathbf{E}(\mathbf{r},t)\mathbf{E}(\mathbf{r},t)\mathbf{E}(\mathbf{r},t)) + \cdots \tag{6.2}$$

In the latter, $\chi^{(n)}$ are nonlinear susceptibility tensors of order $n+1$ which depend on the considered material and can be significantly enhanced by intrinsic resonances.

For sufficiently weak electric fields, only the first term is not negligible and we recover the well-known linear relation between polarization and local electric field:

$$\mathbf{P}(\mathbf{r},t) = \varepsilon_0\chi^{(1)}\mathbf{E}(\mathbf{r},t).$$

Reflection, scattering and absorption are all linear optics phenomena. In the case of a small nanoparticle, the optical wavelength is much larger than the object and its optical response is then entirely

captured in a single dipole moment (*quasi-static approximation*). For larger nanostructures, *retardation effects* yield more complex polarization distributions, but the linear optical response (e.g. scattering spectrum) can still be traced back to the spatial and spectral dependencies of the induced polarization $\mathbf{P}(\mathbf{r}, t)$. For instance, so-called *dark Surface Plasmon modes* arise from destructive interferences among induced dipoles that prevent their efficient coupling with far-field radiation.

When the electric field of the optical excitation becomes comparable in strength to the atomic electric field holding in equilibrium the electrons of the metal, the motion of the latter becomes sensitive to the anharmonicity of the lattice potential and the higher order terms of Eq. (6.2) must be taken into account. These additional terms yield spectral components at frequencies different from the incident optical excitation and are the sources of nonlinear emission. For instance, a monochromatic wave of angular frequency ω creates in a second-order nonlinear medium a polarization oscillating at twice the incident frequency and a static polarization. The former is associated with SHG and the latter with *optical rectification*. THG (radiation at 3ω) and the optical Kerr effect (light-induced modification of the refractive index) are examples of third-order nonlinear processes. In a quantum mechanical picture, SHG is a process in which two photons at the fundamental frequency are annihilated and a photon at twice the fundamental frequency is created.[19] SHG was actually the first nonlinear optical effect ever observed in the optical frequency range.[21]

In bulk materials, the nonlinear emission intensity is governed by the magnitude of the nonlinear susceptibility tensor components and phase-matching conditions. The latter follow from the requirement that the waves radiated by different parts of the material must add constructively to yield a detectable signal. In the case of nanoobjects, phase-matching is automatically fulfilled because of their subwavelength dimensions.

The electric fields in Eq. (6.2) are *local* fields and not the bare incident fields. This is extremely important as the resonant excitation

of SP resonances can lead to large electric field enhancements. The latter explain why, despite their nanoscale volumes, plasmonic nanoobjects can generate exploitable nonlinear signals. As will be illustrated in the following, this sensitivity of nonlinear optical processes to local electric fields is also extremely useful for the spatial characterization of plasmon modes in metallic nanoantennas.

The nonlinear susceptibility tensors depend on the symmetry of the material. For instance, it follows from inversion symmetry that centrosymmetric media must have a vanishing second-order susceptibility. This selection rule is valid in the *dipole approximation* which assumes a uniform electric field. However, surface plasmon resonances can confine the electric fields in subwavelength regions. The existence of strong electric field gradients, therefore, leads to a partial relaxation of the dipole approximation and non-negligible contributions from higher order terms such as electric quadrupole or magnetic dipole to the nonlinear emission. Furthermore, the presence of interfaces breaks inversion symmetry in centrosymmetric media and relaxes the selection rule valid for SHG in bulk materials. The proximity of other nanoobjects can also affect nonlinear emission.

In the third part of this chapter, we will show that the investigation of nonlinear emission from plasmonic nanostructures gives access to the spatial and spectral characteristics of their surface plasmon resonances and also allows characterizing their ultrafast dynamics.

6.2.2.2. *The anharmonic oscillator model*

A first insight in the nonlinear response of plasmonic nanostructures can be obtained from the anharmonic oscillator model. The latter is an extension of the conventional model which describes localized surface plasmon resonances as classical lorentz oscillators of charge $-e$ and density N excited by an electric field $E(t)$. A small perturbation is added to the equation of motion to account for the nonlinearity. In the scalar field approximation, the equation of motion for a second-order nonlinearity is:

$$\ddot{x} + \Gamma \dot{x} + \omega_0^2 x + a x^2 = -e\frac{E(t)}{m}. \tag{6.3}$$

In the latter, the parameters ω_0 and Γ describe the optical resonance frequency and linewidth, respectively, whereas a determines the amplitude of the nonlinearity.[19,20,22] We assume that the applied electric field oscillates at ω_1 and ω_2:

$$E(t) = 2E_1 \cos(\omega_1 t) + 2E_2 \cos(\omega_2 t).$$

The displacement $x(t)$ of the oscillator can be obtained from a perturbation expansion in the electric field strength. The lowest-order term of the displacement is the well-known solution of the harmonic oscillator model:

$$x^{(1)}(t) = x^{(1)}(\omega_1)e^{-i\omega_1 t} + x^{(1)}(\omega_2)e^{-i\omega_2 t} + c.c.$$

in which:

$$x^{(1)}(\omega_i) = \frac{-eE_i}{m}\frac{1}{(\omega_0^2 - \omega_i^2 - i\omega_i\Gamma)}.$$

The corresponding linear polarization and susceptibility follow:

$$P^{(1)}(\omega_i) = -Nex^{(1)}(\omega_i) = \varepsilon_0\chi^{(1)}(\omega_i)E_i$$

with:

$$\chi^{(1)}(\omega_i) = \frac{Ne^2}{m\varepsilon_0}\frac{1}{(\omega_0^2 - \omega_i^2 - i\omega_i\Gamma)}.$$

Inserting the expression of $x^{(1)}$ in Eq. (6.3), we note that the nonlinear term generates components of the second-order displacement $x^{(2)}$ oscillating at $2\omega_1$, $2\omega_2$, $\omega_1 + \omega_2$ and $\omega_1 - \omega_2$. The corresponding contributions to the second-order nonlinear polarization are for instance at $\omega_1 + \omega_2$:

$$P^{(2)}(\omega_1 + \omega_2) = -Nex^{(2)}(\omega_1 + \omega_2) = 2\varepsilon_0\chi^{(2)}(\omega_1 + \omega_2; \omega_1, \omega_2)E_1E_2$$

The second-order nonlinear susceptibility can be expressed as a function of the linear susceptibilities as:

$$\chi^{(2)}(\omega_1 + \omega_2; \omega_1, \omega_2) = \frac{\varepsilon_0^2 ma}{N^2e^3}\chi^{(1)}(\omega_1 + \omega_2)\chi^{(1)}(\omega_1)\chi^{(1)}(\omega_2). \quad (6.4)$$

This relation, known as *Miller's rule*, relates the nonlinear optical properties to their linear counterpart. It shows that optical

resonances revealed in the linear response of a nanoobject can also enhance its nonlinear emission. Miller's rule can be generalized to higher order nonlinearities. For instance, for THG, the nonlinear susceptibility is proportional to the product of four linear susceptibilities at the different involved frequencies. Whereas this simple model provides useful insight, it is important to keep in mind that a rigorous derivation of the nonlinear optical susceptibilities requires the density matrix formulation of quantum mechanics.

In the anharmonic model, the nonlinearity is introduced only phenomenologically. To gain an deeper understanding of (or about) the physical origin of optical nonlinearities in metallic systems, it is instructive to consider the hydrodynamic description of a free-electrons gas.

6.2.2.3. *Second-order nonlinearities from a free-electron gas*

Within the framework of the hydrodynamic model, the equation of motion for a gas of free-electrons excited by an EM wave is[19]:

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} + \gamma \mathbf{v} = -\frac{e}{m}[\mathbf{E} + \mathbf{v} \times \mathbf{B}] - \frac{1}{mn}\nabla p. \tag{6.5}$$

In the latter equation, the electron gas is characterized by its density n, pressure p and velocity field $\mathbf{v}$. γ is the scattering rate characterizing ohmic damping. $\mathbf{E}(\mathbf{r}, t)$ and $\mathbf{B}(\mathbf{r}, t)$ are the electric and magnetic fields of the EM wave. The continuity equation ensures charge conservation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0.$$

Combining these two equations with Maxwell's equations allows finding the polarization $\mathbf{P}(\mathbf{r}, t)$ inside the electron gas using a series expansion as a function of the electric field.[19] For instance, the second-order nonlinear polarization of a free-electron gas can be written as:

$$\mathbf{P}^{(2)}(2\omega) = \frac{e^3 n^{(0)}}{4m^2\omega^4}(\mathbf{E}^{(1)} \cdot \nabla)\mathbf{E}^{(1)} + \frac{e\varepsilon_0}{2m\omega^2}(\nabla \cdot \mathbf{E}^{(1)})\mathbf{E}^{(1)} - \frac{e^3 n^{(0)}}{4im^2\omega^3}(\mathbf{E}^{(1)} \times \mathbf{B}^{(1)}). \tag{6.6}$$

The previous expression shows that optical nonlinearities in a free-electron gas have two origins. The first is the *Lorentz force* term in the equation of motion (last term of Equation 6.6). The second lies in the *spatial variations of the electric field* $\mathbf{E}(r,t)$ (first and second term on the right-hand side). Electric field inhomogeneities are particularly important in the case of plasmonic nanostructures. Indeed, the nanoscale spatial confinement of optical fields enabled by SPs generates strong field gradients. The hydrodynamic model was used to interpret SHG experiments on metallic films.[23,24] Recently, a non-perturbative time-domain implementation of the hydrodynamic model has been developed to self-consistently describe nonlinear optical interactions in plasmonic nanostructures.[25]

6.3. Nonlinear Optical Interactions in the Near-field of Plasmonic Nanostructures

The nonlinear response of a metallic nanoobject strongly depends on the *local* electric field created by an EM solicitation. The spatial, spectral and temporal variations of nonlinear signals therefore carry a wealth of information about the optical resonances supported by the nanosystem. Two different kinds of nonlinear interactions will be discussed in this section. First, we consider nonlinear light emission from plasmonic nanostructures excited by an optical beam. Second, we describe recent electron spectroscopy experiments performed in ultrafast transmission electron microscopes in which relativistic electrons interact with the optical near-field of a metallic nanoparticle.

6.3.1. *Nonlinear light emission: Harmonic generation and multiphoton-induced luminescence*

In this first part, we address coherent nonlinear optical processes such as SHG and THG in plasmonic nanostructures. Then, we describe the mechanism and applications of multiphoton-induced luminescence in gold. We show that nonlinear optical signals can be exploited to yield useful information on the SP modes and ultrafast dynamics and review recent studies.

6.3.1.1. *Second-harmonic generation*

Plasmonic metals such as aluminium, copper, gold and silver have a face-centered cubic structure. We have seen in Section 6.2.2.1 that for such centrosymmetric media, the bulk second-order nonlinear susceptibility vanishes in the electric dipole approximation. As a consequence, only higher-order terms, such as electric quadrupole and magnetic dipole contribute to SHG in the bulk. Conversely, the presence of surfaces in plasmonic nanostructures breaks the inversion symmetry at the boundaries. Electric dipole terms are therefore allowed at the surface. The surface nonlinear sources can be described by a nonlinear dipole moment per unit area:

$$\mathbf{P}_{\mathrm{s}}^{(2)}(\mathrm{r}, 2\omega) = \chi_{\mathrm{s}}^{(2)} \vdots \mathbf{E}(\mathbf{r}, \omega)\mathbf{E}(\mathbf{r}, \omega). \tag{6.7}$$

Assuming that the metal surface is locally isotropic or equivalently has $C_{\infty\nu}$ symmetry, its susceptibility tensor $\chi^{(2)}$ reduces to three independent components : $\chi^{(2)}_{\mathrm{nnn}}$, $\chi^{(2)}_{\mathrm{nss}} = \chi^{(2)}_{\mathrm{ntt}}$, $\chi^{(2)}_{\mathrm{ssn}} = \chi^{(2)}_{\mathrm{sns}} = \chi^{(2)}_{\mathrm{ttn}} = \chi^{(2)}_{\mathrm{tnt}}$, n being the unit normal vector and s, t, two perpendicular tangent vectors. Due to the discontinuity of the normal component of the electric field at the surface, the normal component of the surface nonlinear polarization is expected to yield the largest contribution. Whereas the bulk nonlinear polarization arises from weaker higher order contributions, its contribution to the emitted SHG extends throughout the optically excited volume and could in principle represent a sizable fraction of the total emission. However, it can be shown in materials with a high permittivity, such as metals, that the bulk nonlinearity is an order of magnitude weaker than the surface contribution.[26] This predominance of surface effects is also supported by angle and polarization-resolved measurements[27] and two-beam SHG experiments on a gold film.[28] The latter study suggests that the nonlinear response is dominated by the normal surface component $\chi^{(2)}_{\mathrm{nnn}}$. The SHG from metallic nanoparticles in the Rayleigh limit has both dipolar and quadrupolar components.[29–32] The dipole component originates both from non-local bulk and local surface nonlinear sources, whereas the quadrupolar component only arises from surface sources. For increasing sizes, the contribution from an octupolar

plasmon arising from retardation effects can be detected through interferences with the dipolar contribution.[33,34]

The influence of surface plasmon resonances on SHG has motivated a rich literature. In agreement with the simple picture given by Miller's rule, an enhancement of SHG is observed whenever either the excitation or emission wavelength matches the frequency of a plasmon mode.[35–39] For instance, SHG from rough silver films is enhanced compared to smooth films due to the excitation of local resonances.[40,41] Correlated SHG and transmission electron microscope (TEM) experiments on individual metallic nanoparticles have demonstrated that the second harmonic (SH) emission from Ag nanorods and clusters (trimers and dimers) is greatly enhanced by one-photon excitation of the longitudinal plasmons.[35] SHG is enhanced in V-shaped gold nanostructures or aluminum nanoantennas resonant at the emission wavelength[42,39] (see Fig. 6.4(a)). In the case of sharply pointed structures such as nanotips, the electrostatic lightning-rod effect, due to the geometric singularity at the tip comes in addition to localized SP resonances. As a consequence, a strongly enhanced SHG is obtained when the nanotip is excited with an electric field parallel to the tip axis.[43]

However, in plasmonic nanostructures and metamaterials, the nonlinear response cannot always be inferred from the linear properties. In particular, the influence of the symmetry of the structure, which is not considered in the anharmonic model, can play a decisive role. The limits of the anharmonic oscillator model and Miller's rule have been recently illustrated in the case of SHG from nanostructures gradually evolving from a symmetric bar to an asymmetric U-shape. Whereas Miller's rule does not predict the optimum geometry, more sophisticated modelling based on nonlinear scattering theory provides a much better agreement with experiment.[44] Nanostructures supporting plasmon resonances at the same wavelength and belonging to the same symmetry group can yield very different SHG.

This effect is due to the cancellation of nonlinear dipoles created at different hot spots and depends on the particle geometry.[47] This silencing also explains why linear dimer antennas with small

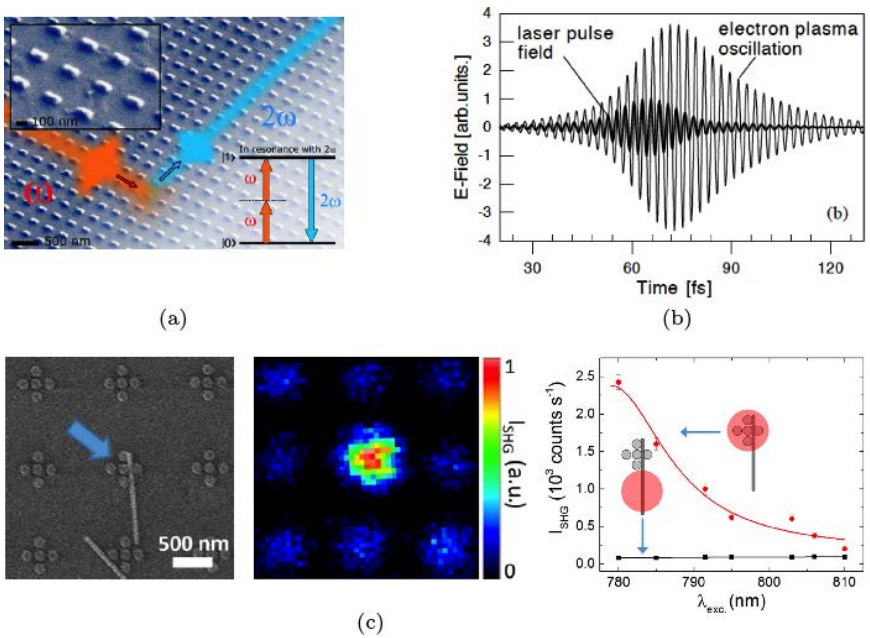


Fig. 6.4. (a) Enhancement of SHG by aluminum nanoantennas resonant at the harmonic frequency.[39] (b) Incident laser field and plasmon field deduced from third-order autocorrelation measurements on gold nanodisks.[45] (c) Left: scanning electron microscopy (SEM) image of a ZnO nanowire deposited on the hot-spot of a plasmonic oligomer. Center: map of the SHG intensity. Right: SHG intensity as a function of excitation wavelength for coupled and uncoupled hybrid ZnO NW and plasmon oligomer.[46]

gaps may not generate optimal SHG.[48] The combined influence of geometrical effects and local field distribution has been pointed out in the case of SHG from arrays of T-shaped nanostructures.[49] On a smaller scale, it has been shown that nanoscopic defects can have a drastic influence on the SHG.[50]

6.3.1.2. *Third-harmonic generation*

Contrary to SHG, THG is allowed in centrosymmetric media. Therefore, despite the much lower third-order susceptibilities, THG from the bulk of metallic nanostructures yields detectable signals. The first report of THG from individual gold plasmonic particles

was achieved in 2005.[51] In these experiments, gold nanospheres down to 40 nm in diameter were excited with 1 ps pulses at 1500 nm yielding a third harmonic resonant with the SP of the particles. The connection between the linear and nonlinear optical response has been used to quantitatively predict THG from gold rod-type and bowtie antennas.[52,53] In the latter case, despite the complex shape of the nanostructure, the measured spectra are in excellent agreement with predictions from the anharmonic oscillator model for a large range of different geometries. Deviations are only observed for sub-20 nm gaps and attributed to the antenna hot spot.[53] The anharmonic oscillator model also proved useful for the interpretation of polarization-resolved third harmonic spectroscopy experiments on arrays of dolmen-type nanostructures supporting Fano resonances.[54]

Contrary to dissipative processes such as multiphoton luminescence which will be discussed in the following section, parametric optical processes such as SHG and THG do not involve real intermediate electronic states. Therefore, they can be considered as instantaneous with respect to the timescales discussed at the beginning of this chapter and are often used for the characterization of the electron dynamics on subpicosecond timescales. For instance, by measuring the SHG from non-centrosymmetric nanoparticles fabricated by electron beam lithography, it is possible to give an estimate of the plasmon decay time (dephasing of the coherent electron plasma oscillation) in Au and Ag nanoparticles.[55,56] Similarly, the plasmon dephasing has been investigated in arrays of metal nanoparticles using third-order interferometric autocorrelation measurements[45] (see Fig. 6.4(b)).

6.3.1.3. *Enhancement of extrinsic nonlinearities using plasmonic nanostructures*

The electric field amplification by SPs can boost nonlinear optical emission from the metallic nanostructures themselves, but it is also interesting for nonlinear optical processes in molecules or quantum dots placed in their vicinity. For instance, a ZnO nanowire placed in the hot-spot of a gold pentamer oligomer exhibits a SH conversion

efficiency of $\sim 3 \times 10^{-5}\%$, among the highest values for a nanoscale object at optical frequencies. Due to the plasmonic enhancement of the optical near-field, the hybrid system yields an SH intensity ∼1700 times larger than that from the same nanowire excited outside the hot-spot[46] (see Fig. 6.4(c)). Similarly, dielectric ITO nanocrystals incorporated into the hot-spot of gold nanogap-antennas show an increase of the radiated third-harmonic intensity when compared to bare gold antennas.[57,58] The two studies disagree in their conclusions, possibly due to differences in antenna design and gap size used in experiments. Finite element simulations show that the enhancement of the third harmonic signal in Ref. 58 is related to changes in the linear optical properties of the plasmonic antenna resonances when the ITO nanocrystals are incorporated.

6.3.1.4. *Multiphoton-induced luminescence*

When exciting a gold nanostructure with ultrashort laser pulses, it is possible to get a spectrally broad visible luminescence with an intensity varying superlinearly with the optical excitation.[59] As will be discussed in more detail in the following, this so-called multiphoton-induced photoluminescence is fundamentally different from the parametric processes discussed so far such as SHG, or THG. Whereas the latter leave the quantum state of the emitter unchanged and only involve virtual energy levels, multiphoton-induced luminescence involves real intermediate electronic states and non-radiative decay steps. TPL is a third-order nonlinear process involving two sequential one-photon absorptions and therefore has a quadratic dependence on the light intensity. It shows great similarity with two-photon absorption from molecules. As sketched in Fig. 6.5(a), a first photon absorption excites an electron via an intraband transition within the *sp* conduction band. A second photon excites an electron from the *d*-bands to the conduction band. The hole created in the *d*-band can then decay radiatively and generate TPL.

The spectrum of TPL in gold has two distinctive features in the green (520 nm) and red (630 nm) spectral regions. This is due to the fact that electron-hole recombination occurs preferentially near

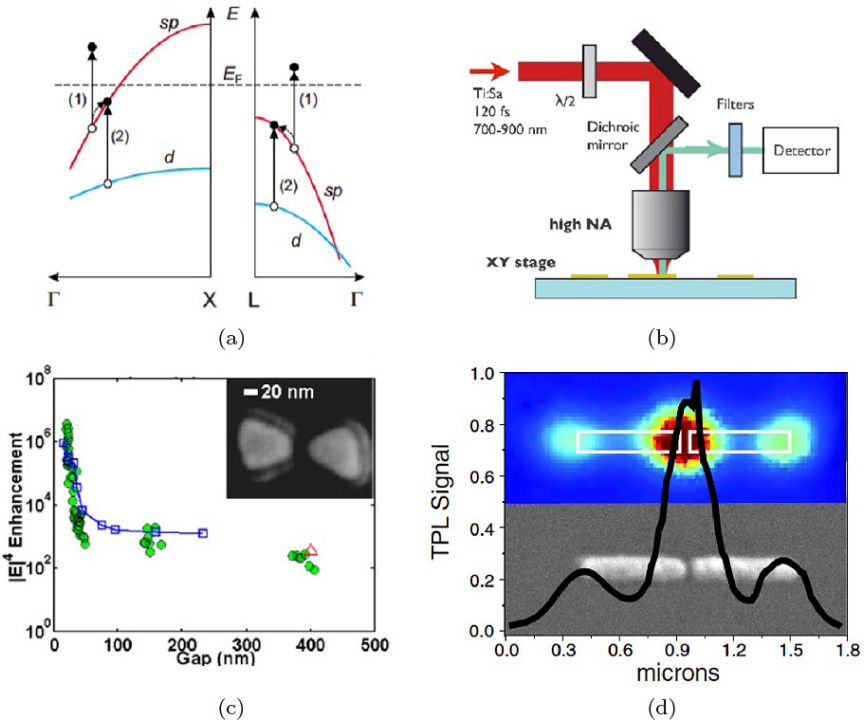


Fig. 6.5. (a) Mechanism of Two-Photon photoluminescence in gold.[61] (b) Typical experimental set-up for TPL imaging. (c) Determination of electric field enhancement in the gap of bow-tie nanoantennas from TPL measurements.[64] (d) TPL imaging of a gold dimer antenna showing the enhancement of the electric field in the gap.[65]

the L and X symmetry points of the Brillouin zone where the electronic density of states is larger. First proposed after polarization studies of the emitted light, the microscopic mechanism of TPL was confirmed by two pulse autocorrelation measurements.[60,61] These latter experiments investigated the influence of the duration of the excitation laser pulse on the TPL intensity and revealed the existence of a real intermediate electronic state created by the first photon absorption.[61] Further studies have shown that the relaxation of this intermediate electronic state is limiting the dynamics of both two-photon and four-photon photoluminescence in gold nanoantennas.[62]

Two-pulse autocorrelation measurements of the TPL intensity have shown that this relaxation occurs on a picosecond timescale, similar to the characteristic time of electron-lattice thermalization in gold (cf. part 2).[62]

Figure 6.5(b) shows a typical experimental set-up for TPL detection. The beam of a femtosecond infrared laser is focused using a high numerical aperture microscope objective on a sample placed on an XY nanopositioning stage. The TPL is separated from the incident beam using a dichroic mirror, filtered and sent on a photodetector (photomultiplier tube for analog detection or photon counter). An evidence of the influence of SPs in the TPL emission was given by comparing the spectra of linear scattering and two-photon photoluminescence from individual nanorods.[63] It was shown that the Two-Photon Luminescence displayed the same optical characteristics as the localized SPs supported by the particles, suggesting that the recombination of the *d*-hole involves the excitation of SPs and their subsequent radiative recombination.[63] As will be seen in the following, the TPL intensity is related to the fourth power of the local electric field amplitude in the nanoobject. This makes it extremely sensitive to the enhanced electric fields in plasmonic nanostructures. Even if bulk gold has a weak third-order nonlinear susceptibility, SPs can locally enhance the electric field and boost the TPL photon yield. It follows that Two-Photon Luminescence can be used to quantify the plasmon-induced *E*-field enhancement (Fig. 6.5(c))[64] and map the spatial distribution of the local electric field associated with plasmon modes (Fig. 6.5(d)).[65–71]

Investigations in plasmonic nanoantennas have initially focused on the properties of the lowest order hybridized mode (bonding mode) in dimer or bowtie antennas.[65] Confocal two-photon luminescence mapping was then used to experimentally demonstrate direct imaging and selective excitation of the "bonding" and "antibonding" plasmon mode in symmetric dipole nanoantennas (Fig. 6.6(a)).[67] Later, two-photon luminescence microscopy has been used to investigate higher order plasmon modes in gold nanorod antennas (Fig. 6.6(b))[71] and gold nanoplatelets (see Fig. 6.9

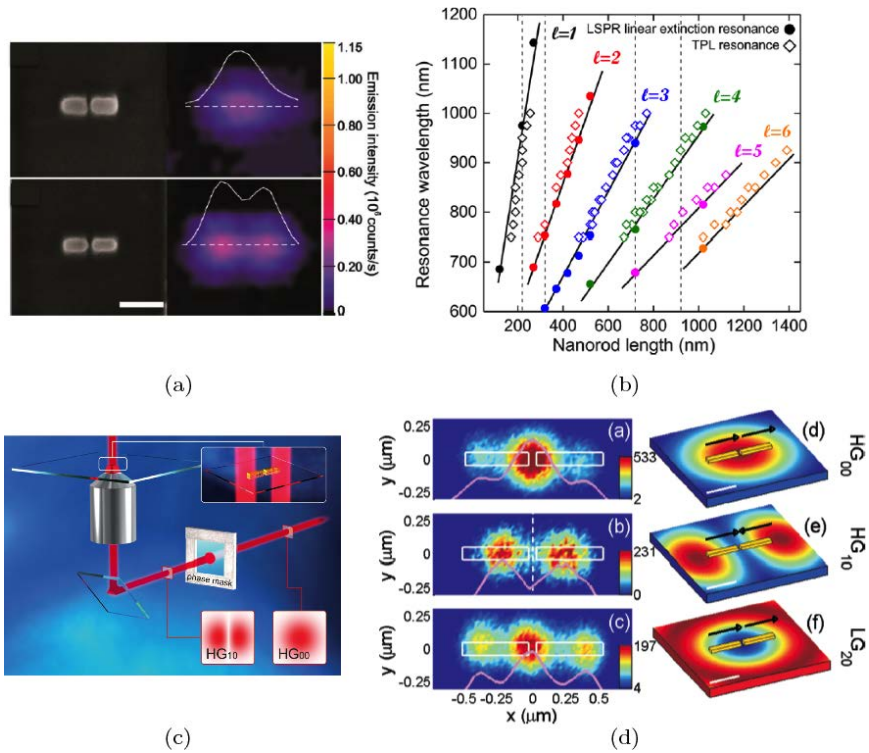


Fig. 6.6. (a) SEM images and corresponding TPL maps together with signal profiles (white dashed line) for two single crystalline nanoantennas having the same height (40 nm), length (383 nm) and gap size (20 nm) but different width. The TPL distribution changes from a single spot (top) to a two-lobed pattern (bottom) indicating the excitation of the antibonding mode.[67] (b) Resonance wavelength versus length of metallic nanorod antennas. A good agreement is observed between linear extinction (full circles) and nonlinear TPL resonances (open diamonds). Full black lines: linear fits to the extinction data.[71] (c) Experimental setup for optical control of near-field distribution: a Gaussian beam from a near-infrared laser is transformed into a higher-order beam using a phase mask before being focused onto a single gold optical antenna.[74] (d) Experimental TPL maps recorded on a single-gap antenna (located by the white rectangles) when driven by three different incident beams ($\lambda = 750$ nm): (a) HG_{00} (Gaussian), (b) HG_{10} and (c) LG_{20}. The color scale gives the TPL intensity in photon counts. In map (b), the vertical dashed line locates the π-shift position of the HG_{10} beam. (d)–(f) Associated computed intensity distribution of the incident field (scale bar = 500 nm). Each of the three beams was linearly polarized along the x-axis. The black arrows give the relative polarization orientation across the beam.[74]

and 6.10).[70,72] In the former case, the influence of SP resonances beyond the fundamental dipolar mode has been clearly evidenced by comparing the TPL emission intensity and extinction measurements for different rod lengths (Fig. 6.6(b)). The strong correlation between the TPL and linear spectroscopy results confirmed that the two-photon absorption cross-section is enhanced whenever the TPL excitation is resonant with a plasmon mode of the metallic nanorod. The dispersion of the *Fabry–Pérot* standing plasmon modes could even be deduced from the TPL measurements. Furthermore, a clear signature of the mode parity was found in the TPL images, the TPL maxima being located within the antenna for odd modes and outside for even modes. The interpretation of this effect required to carefully take into account retardation of the excitation electric field by computing the near-field for each location of a Gaussian light beam and integrating the near-field intensity in the volume of the nanostructure.[71,73]

TPL scanning microscopy has also been used to demonstrate the control of optical near fields using sub-wavelength spatial phase variations at the focus of high-order beams.[74] In these experiments, gold dimer antennas are illuminated with high-order beams such as Hermite–Gaussian (HG) and Laguerre–Gaussian (LG) beams obtained after insertion of a phase mask in the beam of a femtosecond laser (see Fig. 6.6(c)). The antenna is accurately positioned in the focus of the excitation beam, and the local electric field intensity at each point of the antenna is mapped by raster scanning the confocal detection volume using an automated steering mirror inserted in the detection path and measuring the TPL intensity. The TPL images of Fig. 6.6(d) demonstrate that the spatial polarization inhomogeneities of the incident beam can be used to turn on and off the "hot-spot" at the gap of the antennas by either driving in- or out-of-phase its constitutive bars. In particular, the π phase jump existing between the two intensity maxima of a HG_{10} beam allows to drive out of phase the two antenna arms and therefore to minimize the electric field in the gap region. Recently, multiphoton-induced luminescence from aluminum antennas has been reported.[75]

6.3.1.5. *Electrodynamical simulations of optical nonlinearities in plasmonic nanostructures*

Several methods can be used to calculate the nonlinear response of plasmonic nanostructures. Simple geometries such as nanospheres can be treated with analytical approaches. For instance, SHG from nanospheres has been modelled using nonlinear Mie theory.[29,76–78] However, particles having more complex morphologies or placed in complex environments (substrate, neighboring nanoobjects) require numerical techniques. A rich collection of methods such as discrete dipole approximation, finite element modelling, finite difference time domain, volume integral methods and boundary element method are currently available to predict the linear optical properties of metallic nanoobjects. For a comprehensive introduction to these methods and for references to the extensive literature on the subject, one may refer to Refs. 79–81. The numerical frameworks to compute the nonlinear optical properties are derived from these methods.[82–84]

6.3.1.5.1. The Green Dyadic Method

In the following, we present a frequency domain approach derived from the *Green Dyadic Method* to predict SHG and TPL from a plasmonic nanostructure as measured in nonlinear scanning microscopy. In these experiments, a laser beam is focused and raster scanned on the surface of the metallic particle and a nonlinear signal is collected at each point. The nanoobject is deposited on a substrate and the upper medium has a dielectric constant ε_{env}. Several properties can be measured, such as intensity, polarization or radiation pattern. We assume that the optical excitation is monochromatic and denote the incident electric field as $\mathbf{E}_0(\mathbf{r},\omega)$. Care is taken in the following to rigorously take into account the presence of the substrate and key experimental parameters such as the numerical aperture of the collecting optics.

The optical excitation $\mathbf{E}_0(\mathbf{r},\omega)$ induces in the metallic nanoobject a polarization $\mathbf{P}(\mathbf{r},\omega)$. Assuming a local response of the metal, this polarization is given by (CGS units):

$$\mathbf{P}(\mathbf{r},\omega) = \chi(\mathbf{r},\omega)\cdot\mathbf{E}(\mathbf{r},\omega), \tag{6.8}$$

where $\chi(\mathbf{r},\omega)$ is the linear susceptibility of the metal and $\mathbf{E}(\mathbf{r},\omega)$ is the total electric field at $\mathbf{r}$. The latter is solution of the Helmholtz equation:

$$\Delta \boldsymbol{E}(\mathbf{r}\cdot\omega) + k_0^2\epsilon_{\text{env}}\mathbf{E}(\mathbf{r},\omega) = -4\pi\left[k_0^2 + \frac{1}{\epsilon_{\text{env}}}\nabla\nabla\right]\cdot\mathbf{P}(\mathbf{r},\omega).$$

in which $k_0 = \omega/c$.

In the framework of the *Green Dyadic Method* or electrodynamical theory of field susceptibilities, the electric field created at $\mathbf{r}$ by an oscillating dipole $\mathbf{p}(\mathbf{r}',\omega)$ located at $\mathbf{r}'$ is given by:

$$\mathbf{E}(\mathbf{r},\omega) = \mathbf{S}(\mathbf{r},\mathbf{r}',\omega)\cdot\mathbf{p}(\mathbf{r}',\omega).$$

$\mathbf{S}(\mathbf{r},\mathbf{r}',\omega)$ is the *field susceptibility* of the environment. The total electric field $\mathbf{E}(\mathbf{r},\omega)$, solution of the Helmholtz equation, can then be written as the sum of the incident electric field $\mathbf{E}_0(\mathbf{r},\omega)$ and the electric field radiated by the polarization distribution created in the nanoobject:

$$\mathbf{E}(\mathbf{r},\omega) = \mathbf{E}_0(\mathbf{r},\omega) + \int_V \mathrm{d}\mathbf{r}'\mathbf{S}(\mathbf{r},\mathbf{r}',\omega)\cdot\mathbf{P}(\mathbf{r}',\omega). \tag{6.9}$$

Combining Eqs. (6.8) and (6.9) yields the following Lippmann–Schwinger equation:

$$\mathbf{E}(\mathbf{r},\omega) = \mathbf{E}_0(\mathbf{r},\omega) + \int_V \mathrm{d}\mathbf{r}'\mathbf{S}(\mathbf{r},\mathbf{r}',\omega)\cdot\chi(\mathbf{r}',\omega)\cdot\mathbf{E}(\mathbf{r}',\omega). \tag{6.10}$$

This self-consistent equation can be rewritten as a function of the incident electric field $\mathbf{E}_0(\mathbf{r},\omega)$:

$$\mathbf{E}(\mathbf{r},\omega) = \int_V \mathrm{d}\mathbf{r}'\mathbf{K}(\mathbf{r},\mathbf{r}',\omega)\cdot\mathbf{E}_0(\mathbf{r}',\omega). \tag{6.11}$$

In the latter, $\mathbf{K}(\mathbf{r},\mathbf{r}',\omega)$ is the so-called *generalized field propagator*, which can be obtained from the following relation:

$$\mathbf{K}(\mathbf{r},\mathbf{r}',\omega) = \mathbf{I}\boldsymbol{\delta}(\mathbf{r}-\mathbf{r}') + \chi(\mathbf{r},\omega)\cdot S(\mathbf{r},\mathbf{r}',\omega).$$

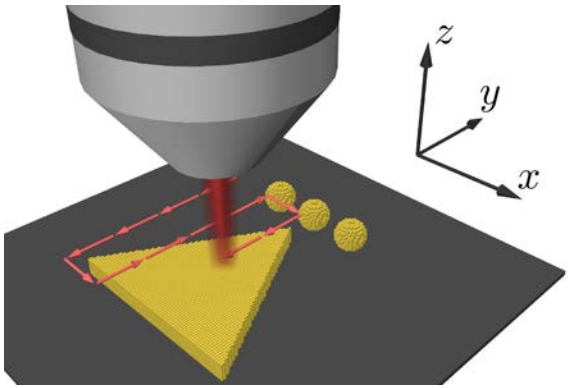


Fig. 6.7. An optical beam is focused and raster scanned on the surface of a metallic particle. At each location $\mathbf{R_0}$ of the excitation, a nonlinear signal (SHG, TPL...) is calculated. For the computation, the volume of the nanostructure is discretized.

The dyadic tensor $S(\mathbf{r}, \mathbf{r}', \omega)$ obeys Dyson's equation:

$$\mathcal{S}(\mathbf{r}, \mathbf{r}', \omega) = \mathbf{S}(\mathbf{r}, \mathbf{r}', \omega) + \int_{\mathcal{V}} \mathrm{d}\mathbf{r}'' \mathbf{S}(\mathbf{r}, \mathbf{r}'', \omega) \cdot \chi(\mathbf{r}'', \omega) \cdot \mathcal{S}(\mathbf{r}'', \mathbf{r}', \omega) \tag{6.12}$$

The numerical resolution of Dyson's equation is achieved after discretization of the volume of the nanostructure. As shown in Fig. 6.7, the investigated nanoobject is meshed in a lattice of a total of N points located at positions $\mathbf{r}_\mathrm{i}$. This procedure generates a system of $3N$ linear equations that can be self-consistently resolved by numerical inversion yielding the generalized field propagator and then the electric field $\mathbf{E}(\mathbf{r}_i, \omega)$ at each lattice location. We will not go further into the details of the numerical resolution of Dyson's equation. The interested reader is referred to Refs. 85, 86. It is important to keep in mind that the knowledge of $\mathbf{K}(\mathbf{r}, \mathbf{r}', \omega)$ enables the direct calculation of the *linear* optical response of a metallic nanosystem to arbitrary EM solicitations. Stated differently, the EM properties of the system (geometry, dielectric constants) are entirely captured in the generalized field propagator for a given frequency ω.

6.3.1.5.2. Second-harmonic generation

In the following, we show how this framework can be extended to describe SHG. The microscopic description involves three steps: (i) optical excitation of the nanoobject by the incident wave, (ii) creation of a nonlinear polarization and (iii) radiation of the harmonic wave by the nonlinear sources. In a first step, the total electric field in the nanostructure $\mathbf{E}(\mathbf{r},\omega)$ is computed from the generalized field propagator $\mathbf{K}(\mathbf{r},\mathbf{r}',\omega)$ (Fig. 6.8(b)). Then, the nonlinear sources are deduced from the fundamental electric field. Here, it is assumed that the normal component of the nonlinear surface tensor $\chi^{(2)}_{\mathrm{nnn}}$ yields the largest contribution to SHG. The same hypothesis was used in other studies of SHG from plasmonic nanostructures.[84,87] The contribution from other surface tensor components and even the bulk contribution could equally be computed following the same steps. We define complex surface nonlinear dipoles $\mathbf{p}_{\mathrm{NL}}(\mathbf{r},2\omega)$ associated to each discretization cell located on the surface of the antenna at position $\mathbf{r}$ (Fig. 6.8(c)):

$$\mathbf{p_{NL}}(\mathbf{r},2\omega) = \chi_{nnn}\mathbf{E_n}(\mathbf{r},\omega)^2\mathbf{n}(\mathbf{r}), \tag{6.13}$$

$\mathbf{E}_{\boldsymbol{n}}(\boldsymbol{r},\omega)$ is the complex component of the total electric field along the local surface normal $\mathbf{n}(\mathbf{r})$. The nonlinear sources described by Eq. (6.13) are complex, and therefore the phase information governing the far-field radiation is taken into account.

Finally, the energy radiated at the harmonic frequency by the nonlinear dipoles is computed from the far-field asymptotic limit of the field susceptibility of the environment $\mathbf{S}_\infty(\mathbf{r},\mathbf{r}',2\omega)$ at the SHG angular frequency 2ω[88]:

$$\boldsymbol{E}_{\mathbf{SHG}}(\mathbf{r}'\,2\omega) = \boldsymbol{S}_\infty(\mathbf{r},\mathbf{r}',2\omega)\cdot\mathbf{p_{NL}}(\mathbf{r},\,2\omega).$$

The contribution from the substrate is taken into account by including the corresponding field-susceptibility in the calculations:

$$\mathbf{S}_\infty(\mathbf{r},\mathbf{r}'2\omega) = \mathbf{S}_0(\mathbf{r},\mathbf{r}',2\omega) + \mathbf{S}_{\mathbf{surf},\infty}(\mathbf{r},\mathbf{r}',2\omega).$$

To compute the SHG intensity, the intensity scattered in the solid angle defined by the collecting optics is integrated (Fig. 6.8(d)). The

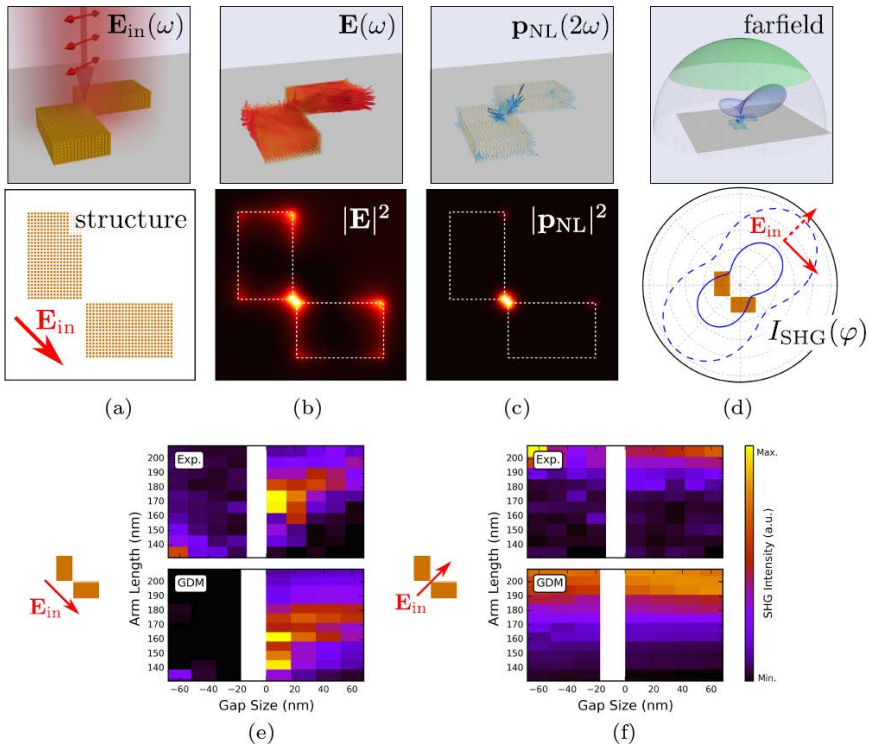


Fig. 6.8. Illustration of the numerical modelling of SHG from a gold L-shaped nanoantenna. (a) Discretization of the structure and modelling of the incident electric field. (b) Real parts of the fundamental electric field after excitation (top) and corresponding field intensity 10 nm below the structure (bottom). (c) Surface nonlinear dipoles (real parts, top) and corresponding squared modulus (bottom). (d) Radiation of the nonlinear dipoles to the far-field allows to calculate the radiation pattern (top) and extract other information like e.g. the polarization of SHG in the far-field (bottom), by integration over the collecting solid angle (green highlighted surface in (d), top). In (e) and (f) experimental results (top) are compared to simulated SHG intensities (bottom) for two different incident polarizations. Excitation at $\lambda = 900\,\mathrm{nm}$ on a glass substrate ($n = 1.5$).[90]

numerical aperture of a microscope objective can therefore easily be taken into account in the simulations. Also, polarization-resolved information can be obtained using this approach by considering only the radiated SHG electric field with the proper polarization.

We illustrate this method by computing SHG from L-shaped antennas. Due to their non-centrosymmetry, these structures have motivated several studies.[45,47,50,89] SHG intensities for a large parameter-range of L-shaped gold antennas from the connected regime denoted by negative gap sizes to capacitively coupled antenna arms (positive gaps) are shown in Figs. 6.8(e) and 6.8(f). The incident wavelength was $\lambda = 900\,\mathrm{nm}$ ($\lambda_{\mathrm{SHG}} = 450\,\mathrm{nm}$), and the structures were deposited on a glass substrate with a refractive index of $n = 1.5$. Comparing the experimental results from electron-beam-lithographically fabricated nanostructures (top colormaps) to simulations of the SHG intensities using the *Green Dyadic Method* (bottom colormaps), an excellent agreement is observed. Interestingly, the polarization of SHG radiation does not necessarily follow the polarization of the fundamental field, a behavior that has also been confirmed in experiments and can be explained by strong cancellation of nonlinear dipoles in the gap region due to the symmetry of the structure.[90]

We note, that the presence of the nanoobject at the harmonic radiation is neglected in the last step of the simulations (it is only taken into account in the computation of the total electric field at the fundamental frequency). This may become problematic in the case of near-field investigations and is rigorously taken into account in more sophisticated approaches.[82]

6.3.1.5.3. Two-photon luminescence from gold nanostructures

We now address scanning two-photon luminescence microscopy experiments. As explained in details earlier, TPL is a third-order incoherent nonlinear process arising from two sequential one-photon absorptions in the metal. Similar to two-photon absorption in molecular systems, the induced polarization can be described by a third-order nonlinear susceptibility $\chi^{(3)}$ [91,92]:

$$\mathbf{P}^{(3)}(\mathbf{r},\omega) = \chi^{(3)}(\omega,-\omega,\omega)\mathbf{E}(\mathbf{r},\omega)\mathbf{E}^{*}(\mathbf{r},\omega)\mathbf{E}(\mathbf{r},\omega). \tag{6.14}$$

In the latter equation, $\mathbf{E}(\mathbf{r},\omega)$ is the total electric field at the excitation frequency and $\mathbf{E}^{*}(\mathbf{r},\omega)$ its complex conjugate.

The energy transferred from the field to a volume element located at $\mathbf{r}$ per unit time can be calculated from the induced polarization $\mathbf{P}(\mathbf{r},t)$ and electric field $\mathbf{E}(\mathbf{r},t)$[93]:

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \left\langle \frac{\mathrm{d}\mathbf{P}(\mathbf{r},t)}{\mathrm{d}t} \cdot \mathbf{E}(\mathbf{r},t) \right\rangle$$

In the case of the linear response, $\mathbf{P}(\mathbf{r},t) = \chi(\mathbf{r},t)\mathbf{E}(\mathbf{r},t)$. For a time-harmonic electric field of frequency $\omega, \mathbf{E}(\mathbf{r},t) = \mathrm{Re}(\mathbf{E}(\mathbf{r},\omega)e^{-i\omega t})$. This leads to the well-known fact that the energy transfer between the optical wave and the metal is given by the imaginary part of the susceptibility:

$$\frac{\mathrm{d}E_{\mathrm{lin}}}{\mathrm{d}t} = \frac{\omega}{2}\mathrm{Im}(\chi(\mathbf{r},\omega))|\mathbf{E}(\mathbf{r},\,\omega)|^2.$$

Photoluminescence from gold nanoobjects is, therefore, proportional to the linear absorption and local electric field intensity. Similarly, in the case of TPL, the emission intensity is given by the nonlinear absorption rate:

$$I_{\mathrm{TPL}} \propto \frac{\mathrm{d}E_{\mathrm{nl}}}{\mathrm{d}t} = \frac{\omega}{2}\mathrm{Im}(\chi^{(3)}(\omega,-\omega,\omega))|\mathbf{E}(\mathbf{r},\,\omega)|^4. \tag{6.15}$$

In a scanning experiment, the laser beam is tightly focused by a high numerical aperture microscope objective and is scanned on the surface of the sample. The scanning procedure can be explicitly taken into account in the simulations by considering an excitation with a Gaussian beam centered at $\mathbf{R}_0$ denoted $\mathbf{E}_0(\mathbf{r},\mathbf{R}_0,\omega)$ and computing the total E-field $\mathbf{E}(\mathbf{r},\mathbf{R}_0,\omega)$ for each position of the excitation beam using the generalized propagator (Eq. 6.11). Finally, the two-photon luminescence collected from the entire nanostructure is obtained by summing the incoherent contributions from the different volume elements[73]:

$$I_{\mathrm{TPL}}(\mathbf{R}_0) \propto \int_V |\mathbf{E}(\mathbf{r},\mathbf{R}_0,\,\omega)|^4 \mathrm{d}V \tag{6.16}$$

The quadratic dependence of TPL on the intensity of the electric field explains its ability to map the plasmon modes with a spatial resolution typically better than 300 nm for excitation in the near-infrared. It is important to note that in the last expression, the

electric field is computed *for each position* of the excitation beam. In fact, a direct convolution of the optical near-field intensity distribution to account for the limited spatial resolution in TPL scanning microscopy can only be the first step in the interpretation of TPL images.[69,71,73]

Figure 6.9 shows the results of correlated scanning TPL microscopy and SEM experiments performed on gold particles.[72] The latter are 20-nm thick gold triangular, truncated triangular, and hexagonal nanoprisms of typical lateral size ranging from 0.5 to 1.0 μm prepared using a one-pot protocol based on the direct reduction of Au precursors by polyvinylpyrrolidone (PVP) (see Figs. 6.9(a)–(c)). The TPL maps of Figs. 6.9(d)–(i) clearly show a well-resolved patterning of the TPL emission strongly dependent on the shape of the nanoobject and polarization of the incident optical wave. Figures 6.9(j)–(o) demonstrate that the numerical framework discussed above reproduces the observed dependencies.

One important question is the connection between the observed TPL patterns and the SP modes of the nanostructures, or stated differently the impact of the limited spatial resolution on the collected signal. Electron energy loss spectroscopy (EELS) experiments have indeed shown that the plasmon modes supported by gold platelets of micronic size in the 1–2 eV range are higher order plasmon modes with a finer spatial structure.[94]

In fact, it is possible to demonstrate that the experimental TPL maps are the convolution of the surface plasmon local density of states (SP-LDOS) in the metallic particle with the spatial profile of the excitation beam.[70] First, let us remember that the *photonic local density of states* (ph-LDOS) $\rho_{\mathrm{ph}}(\mathbf{r},\omega)$ is defined as the number of EM modes of energy $\hbar\omega$ per unit volume at position $\mathbf{r}$, $\rho_{\mathrm{ph}}(\mathbf{r},\omega)$ can be related to the field susceptibility tensor as[95]:

$$\rho_{\mathrm{ph}} = \frac{1}{2\pi^2\omega}\mathrm{Im}\{\mathrm{Tr}\mathbf{S}(\mathbf{r},\mathbf{r},\omega)\}. \tag{6.17}$$

Similarly, for planar plasmonic structures such as the gold platelets above, it is possible to define inside the metallic nanoobject an in-plane SP-LDOS $\rho_{\mathrm{sp},\parallel}(\mathbf{r},\omega)$. $\rho_{\mathrm{sp},\parallel}(\mathbf{r},\omega)$ is the number of SP modes

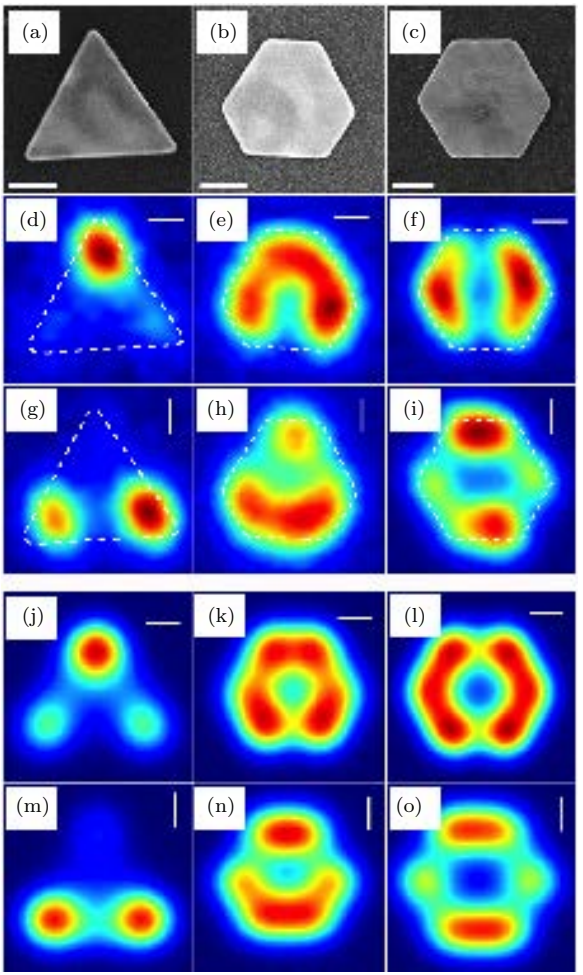


Fig. 6.9. (a)–(c): SEM images of triangular, truncated triangular and hexagonal gold nanoplatelets. (d)–(i): Corresponding TPL maps acquired with $\lambda = 700\,\text{nm}$ and linear polarization of 0° (d)–(f) and 90° (g)–(i), as indicated by the white bars in the upper right corners. (j)–(o): corresponding simulated maps. Scale bars are 200 nm.[72]

of energy $\hbar\omega$ at position $\mathbf{r}$ inside the metal. For a circularly polarized excitation, the number of plasmon excitations of energy $\hbar\omega$ created at $\mathbf{r}$ is then proportional to $\rho_{sp,\parallel}(\mathbf{r},\omega)$ and to the number of incident photons of the same energy. Equivalently, it is possible to relate the

local electric field $\mathbf{E}(\mathbf{r}, \mathbf{R}_0, \omega)$ to the excitation field $\mathbf{E}_0(\mathbf{r}, \mathbf{R}_0, \omega)$ by:

$$|\mathbf{E}(\mathbf{r}, \mathbf{R}_0, \omega)|^2 \propto |\mathbf{E}_0(\mathbf{r}, \mathbf{R}_0, \omega)|^2 \rho_{\mathrm{sp},\parallel}(\mathbf{r}, \omega).$$

A complete derivation of this relation is provided in Ref. 70. It is now possible to write differently the TPL intensity:

$$I_{\mathrm{TPL}}(\mathbf{R}_0) \propto \int_V |\mathbf{E}_0(\mathbf{r}, \mathbf{R}_0, \omega)|^4 \rho^2_{\mathrm{sp},\parallel}(\mathbf{r}, \omega) dV. \tag{6.18}$$

This last equation provides a new interpretation of the TPL maps as the convolution between the squared SP-LDOS with the Gaussian profile of the light beam. This is illustrated in Figs. 6.10(b)–(e) in which the beam waist is varied between the experimental value of 300 nm and 50 nm. The former gives an excellent agreement with the experiments (Fig. 6.10(a)) while the last value is equivalent to a Dirac delta-function and therefore provides a map of the SP-LDOS squared (Fig. 6.10(e)). To go further, we have computed the photonic local density of states (LDOS) from Eq. 6.17 at different heights above the gold nanoparticle. The results shown in Figs. 6.10(f)–(h) demonstrate that the photonic LDOS becomes increasingly similar to the SP-LDOS when we get closer to the metallic particle. This reveals that the contribution of the SP modes of the particle to the photonic LDOS is restricted to distances shorter than the decay length of the evanescent optical fields. In the case of linearly polarized excitation along (OX), the same relation holds but with a projected SP-LDOS $\rho_{\mathrm{sp},x}(\mathbf{r}, \omega)$ instead of the total SP-LDOS $\rho_{\mathrm{sp},\parallel}(\mathbf{r}, \omega)$. Different linear polarizations, therefore, probe different projections of the SP-LDOS.[70]

6.3.2. *Nonlinear optical interaction of fast electrons with plasmonic nanostructures*

The nonlinear optical processes discussed so far arise from the interaction between light fields and electrons confined in the metallic nanoobjects. Their extreme sensitivity to the local electric field strength explains their enhancement by SPs and makes them very useful to map the spatial distribution of these modes. However,

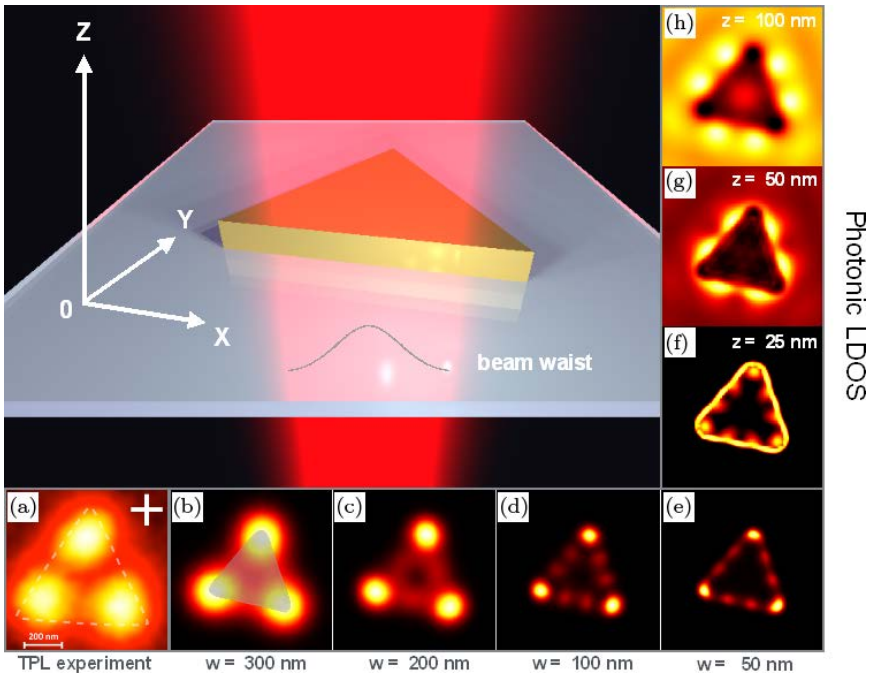


Fig. 6.10. (a) Sum of the TPL signal obtained on a triangular gold nanoplatelet for a polarization along either X or Y. (b) TPL map computed using the numerical framework discussed in the text for a beam waist $w = 300$ nm. (c)–(e) Same for $w = 200$, 100 and 50 nm, respectively. (f)–(h) are the ph-LDOS computed from Eq. 6.17 in a plane located at $z = 25$, 50 and 100 nm, respectively.

far-field nonlinear scanning microscopies suffer from a limited spatial resolution. We now discuss a different kind of nonlinear optical interaction which involves ultrashort packets of relativistic electrons in TEMs. As will be shown, the analysis of the interaction between free electrons and the optical near-fields confined around plasmonic nanoobjects holds promise for E-field mapping with deeply subwavelength spatial resolution and coherent manipulation of ultrashort electron packets.

6.3.2.1. *Interaction of free electrons with light*

Energy and momentum conservation laws prohibit linear interaction between free electrons and free-space light.[96–99] However, it is

possible in some cases to satisfy these conservation laws and yield detectable effects. Thus, Kapitza and Dirac predicted in 1933 that the elastic interaction between a beam of free electrons and a stationary light wave could cause electron diffraction.[100] The absence of sufficiently intense light sources has long impeded the experimental verification of this effect, but it has since been detected unambiguously.[101,102] An electron passing near a nanostructured surface may also interact inelastically with confined light fields. The energy losses are the best-known manifestation of this coupling but it is also possible that electrons absorb one or more photons and can thus be accelerated.[96,98,99] The inverse Smith–Purcell[103] or inverse Cherenkov effects[104] are examples of such interactions.

6.3.2.2. *Electron energy gain spectroscopy*

In 2009, Barwick *et al.* have demonstrated the immense possibilities of electron energy gain spectroscopy (EEGS) (also termed *PINEM* for *Photon Induced Near-field Microscopy*).[97] This technique relies on the analysis of the energy spectrum of fast electrons incident on an optically excited nanoobject.

It is based on a pump–probe scheme in which a first femtosecond optical pulse focused on the nanoobject generates an electric field confined around the nanoobject and a second probe pulse consisting of a subpicosecond packet of relativistic electrons is synchronized with the pump pulse (Figs. 6.11(a)–(b)). The electronic probe pulse is generated by a delayed optical pulse focused on a photocathode inside the TEM. The electric fields confined around the optically excited nanoobject provide the required momentum to satisfy conservation laws and permit interaction between the free electrons and the light field. This interaction results in energy exchanges between the free electrons and the light field in multiples of the incident photon energy $\hbar\omega$ as shown in Figs. 6.11(c)–(d).

The energy gain probability directly depends on the amplitude of the electric field along the electron trajectory.[96–99,105,106] Therefore, by analyzing the energy of the transmitted electrons, it is possible to map the optical near-field with nanometer spatial resolution. For

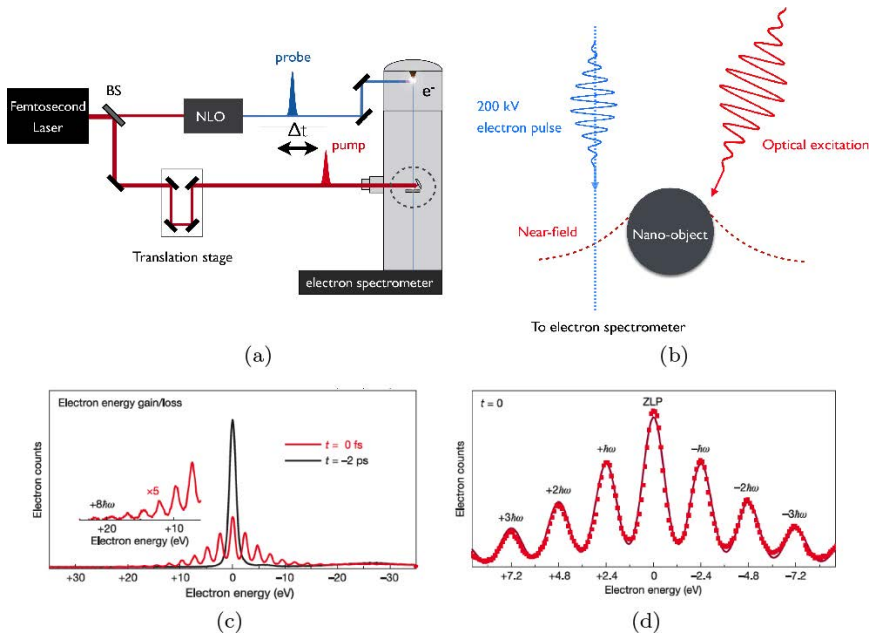


Fig. 6.11. (a) Schematic of an ultrafast transmission electron microscope. (b) Close-up of the interaction between an ultrashort electron pulse and an optically excited nanoobject. (c) Comparison between the electron energy spectrum before (black) and during (red) optical excitation of a bundle of carbon nanotubes. (d) Magnified version of (c) showing absorption and emission of multiples of the photon energy $\hbar\omega = 2.4\,\mathrm{eV}$.[97]

instance, the probability of absorbing one photon can be derived from first-order perturbation theory as[96]:

$$P_{\mathrm{EEGS}}(E) = \left(\frac{2\pi e}{\hbar\omega}\right)^2 \left| \int_{-\infty}^{+\infty} \mathrm{d}z E_z(\mathbf{r}_\mathrm{e}(t), E) e^{-i\omega z/v} \right|^2$$

Figure 6.12(a) shows a map of the PINEM signal on a dimer of silver nanoparticles which illustrates the nanometer spatial resolution of the technique.[107] Figure 6.12(b) shows the gain probability computed from first-order perturbation theory using an extension of the *Green Dyadic Method* discussed above.[108] Two important differences between EEGS/PINEM and EELS or cathodoluminescence are worth being emphasized. First, the observed signal can be amplified

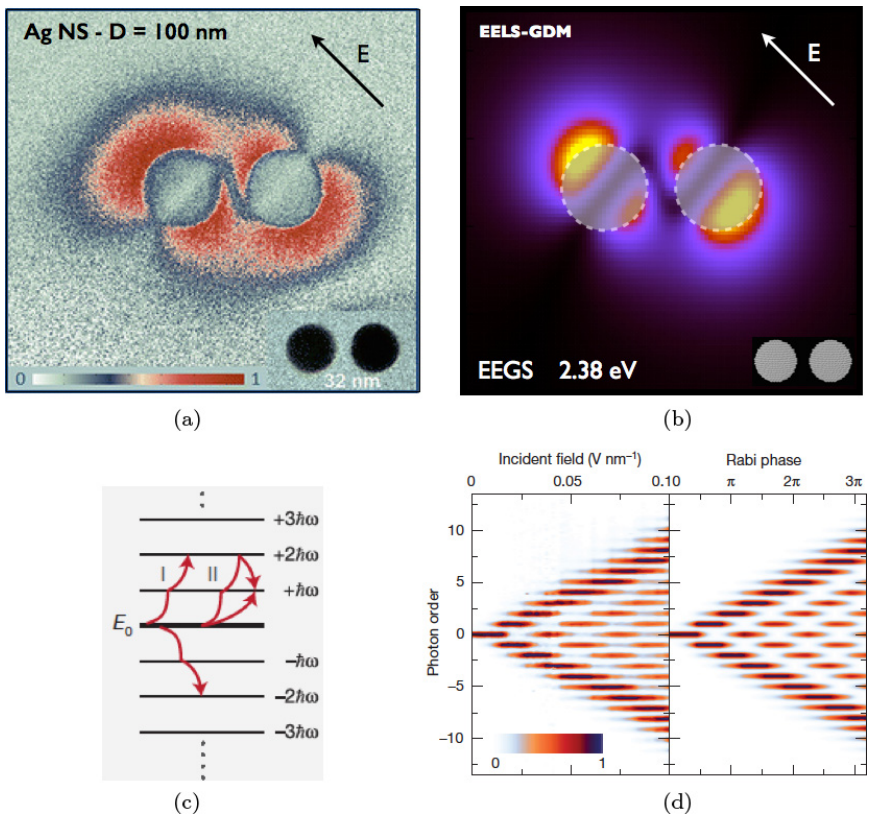


Fig. 6.12. (a) Map of the electron energy gain signal on a dimer of silver nanospheres.[107] (b) Numerical simulation of (a) based on a first-order perturbation theory and computation of the optical near-field using EELS-GDM.[108] (c) Energy diagram showing the different possible interactions and interference between quantum mechanical paths.[106] (d) Evolution of the electron-energy spectrum for different incident electric field strength (left) and corresponding numerical simulations of the quantum mechanical evolution of the energy components of the ultrashort electron packet.[106]

by increasing the power sent on the sample within the limits set by the damage threshold. Second, whereas EELS is sensitive to all EM modes supported by a plasmonic nanoparticle, EEGS will only probe the plasmon modes effectively excited by the pump pulse. Furthermore, the PINEM signal being a cross-correlation between the electron and optical pulse, its temporal width could in principle

give access to the lifetime of SPs. However, these dephasing times are of the order of tens of femtoseconds and it is, therefore, difficult to imagine accessing their dynamics using this technique. In the case of high-power optical excitation, higher-order interactions develop between the optical field and the free electrons, yielding a more complex dynamics characterized by multiple photon sidebands. The occupation probability $P_{EEGS}(N\hbar\omega)$ of the Nth photon sideband is then given by a Bessel function of the first kind of order N.[98,99] It shows an oscillatory behavior with excitation intensity which can be interpreted as multilevel Rabi oscillations. These oscillations arise from interferences between different quantum mechanical paths contributing to the transition probability (Figs. 6.12(c) and 6.12(d)).[106] These experiments are extremely promising steps towards the quantum coherent manipulation of free-electron states.

6.4. Femtosecond Time-resolved Spectroscopy of Individual Plasmonic Nanoparticles and Antennas

Time-resolved spectroscopy is a valuable tool for obtaining information on the dynamics and processes involving electrons, lattice vibrations, and other excitations of materials. As discussed in section 6.2, most electronic and vibrational processes take place on time scales ranging from a few femtoseconds to hundreds of picoseconds. Most of the time-resolved investigations performed on individual nanoparticles are therefore based on femtosecond pump–probe spectroscopy. Femtosecond pump–probe spectroscopy is a special case of four-wave mixing or transient grating techniques[109] which involve a third-order nonlinear interaction between the electric field of a delayed probe pulse and a stronger pump pulse in the metallic particle.

6.4.1. *Ultrafast electronic dynamics of individual small metal nanoparticles*

Compared to spectroscopy of bulk systems, challenges appear in the detection of small single nanoparticles because of the fact that

the optical cross-section of nanosystems is typically much less than the geometrical cross-section of diffraction-limited far-field radiation. The nonlinear perturbations are generally only a fraction of the linear cross-section, and therefore require very sensitive detection techniques based on lock-in amplification. As an alternative to far-field spectroscopy, near-field optical techniques can be used to overcome the diffraction limit and achieve more efficient coupling and high-resolution readout.[110] However, the added complexity and presence of tip–sample interactions associated with near-field techniques in many cases is unfavorable.

Femtosecond studies of single metal nanoparticles using far-field transmission have been demonstrated on metal nanoparticles with optical cross-sections of several thousand nanometer squared[2]. Figure 6.13(a) shows the experimental arrangement used in Ref. 12, where a single silver nanoparticle of 30 nm diameter was positioned in the tight focus of a 0.9 N.A. microscope objective. The mismatch of the illumination area S and particle cross-section σ_{ext} in this case results in a fraction of incident light absorbed of order 10^{-3}. In addition, photoexcitation of hot electrons results in a nonlinear modulation of the extinction $\Delta\sigma_{\text{ext}}/\sigma_{\text{ext}}$ of several percent. The pump-induced transmission change $\Delta T/T$ can be written as the product of the cross-sectional mismatch and the pump-induced perturbation according to

$$\Delta T/T = -(\Delta\sigma_{\text{ext}}/\sigma_{\text{ext}})(\sigma_{\text{ext}}/S).$$

The resulting normalized change in transmission $\Delta T/T$ in the range $10^{-6} - 10^{-4}$ can be recovered using sensitive lock-in detection (Fig. 6.13(b)). The time-resolved data show a fast rise in the nonlinear absorption, caused by excitation of hot electrons, followed by a picosecond decay due to hot carrier relaxation, following the TTM of section 6.2.1.2.

Varying the power of the excitation laser pulses shows the dependence of the electron–phonon relaxation time on absorbed energy (Fig. 6.13(c)), through the temperature dependence of the

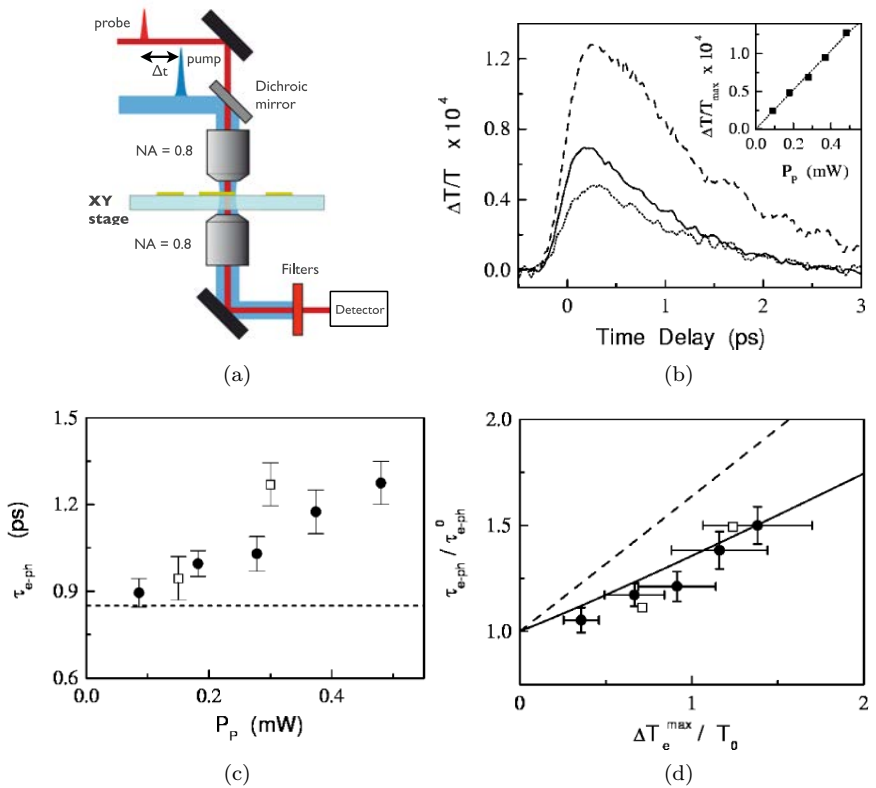


Fig. 6.13. (a) Experimental setup for ultrafast single-nanoparticle spectroscopy. (b) Time-resolved differential transmission $\Delta T/T$ for single 30 nm silver nanoparticle for different pump powers 180 μW (dotted line), 280 μW (solid line) and 480 μW (dashed line). (c) Electron–phonon relaxation time $\tau_{\mathrm{e-ph}}$ against pump power for 30 nm (dots) and 20 nm (squares) silver particle. (d) same as (c) but against calculated maximum electron temperature rise $\Delta T_e/T_e$. Lines: analytical model for electron excess energy (solid) and electron temperature (dashed).[12]

electronic heat capacity,

$$C(T_c) = c_0 T_e$$

with $c_0 = 65\,\mathrm{J/(m^3K^2)}$ for silver and T_e the electronic heat capacity. Figure 6.13(d) shows the same increase against maximum temperature rise estimated using the absolute cross-section of the single nanoparticle. A universal scaling is found, which is independent

of particle size and only depends on the amount of temperature increase, in agreement with the TTM (solid line). Thus, the quantitative *in situ* studies of both the linear and nonlinear response of a single nanoparticle allows to calibrate the TTM for the strong excitation regime.

In addition to revealing fundamental physical processes, the quantitative spectroscopy of single nanoparticles allows extracting the ultrafast changes in extinction cross-section.[111] This information allows quantitative modelling of the SP resonance perturbation using the real and imaginary permittivity changes, under the assumption that only the permittivity of the nanoparticle is modified and its local environment remains constant. An extensive analysis of electronic excitation and cooling dynamics was reported for single gold plasmonic nanoantennas fabricated using electron-beam lithography.[112]

6.4.2. *Coherent vibrational modes of nanoparticles*

The near-instantaneous change in electron temperature following femtosecond pulsed laser excitation sets up an impulsive strain in the nanoparticle.[113] This impulsive strain launches a coherent wavepacket of vibrational eigenstates of the nanoparticle. Ensemble measurements generally suffer from dephasing of the signals caused by the inhomogeneous distribution of resonance frequencies. Single-particle spectroscopy allows measurements of the vibrational modes of nanoparticles without the effects of inhomogeneous distributions, providing direct access to the homogeneous damping rates and relative contributions of the vibrational eigenmodes.[2,3,15] It was found that the damping of vibrational modes in nanoparticles strongly depends on its coupling with the surrounding medium.[114] Figure 6.14 shows an experiment where individual 80 nm gold nanospheres were captured in an optical trap suspended in water. The time response of each particle was measured to obtain both its resonance frequency and homogeneous damping rate. The effect of the surrounding water was observed by comparing with gold nanoparticles deposited on a substrate, where the coupling with the underlying material is weak.

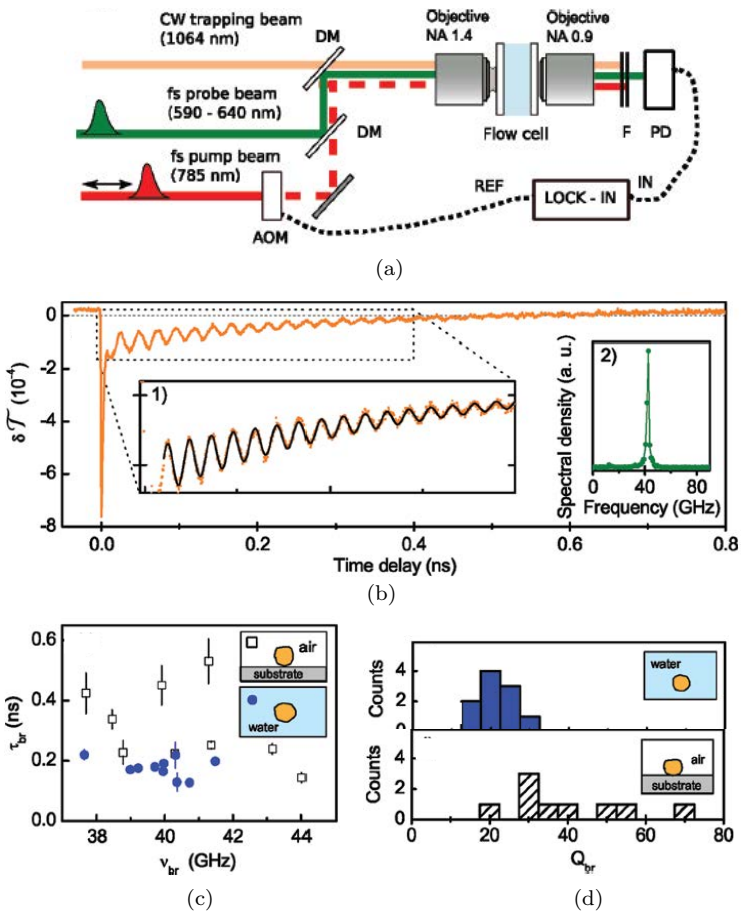


Fig. 6.14. (a) Experimental arrangement for vibrational spectroscopy of single nanoparticles trapped in an optical trap. (b) Example of time-resolved transmission of the trapped nanoparticle, revealing a 40 GHz vibrational eigenmode. (c) Distribution of eigenmode frequencies and damping rates for nanoparticles in water (blue circles) and in air (squares). (d) Histograms of the quality factors obtained from (c).[114]

6.4.3. *Single plasmonic nanoantenna ultrafast nonlinear devices*

Next to the use of single-particle time-resolved spectroscopy of plasmonic nanoparticles in studying the physical processes of electrons

and phonons in confined geometries, the high sensitivity of plasmonic modes to small perturbations is of interest to achieve new types of devices for nanoscale sensing and nonlinear control of light. As the size of nanoparticles is scaled up, the extinction cross-section approaches the geometrical area of the optical wavelength and order-unity coupling efficiency can be achieved. Combined to bulk materials, the design of optical nonlinearity at the nanoscale can exploit combinations of effects and materials.

6.4.3.1. *Acousto-plasmonics and antenna vibrational sensors*

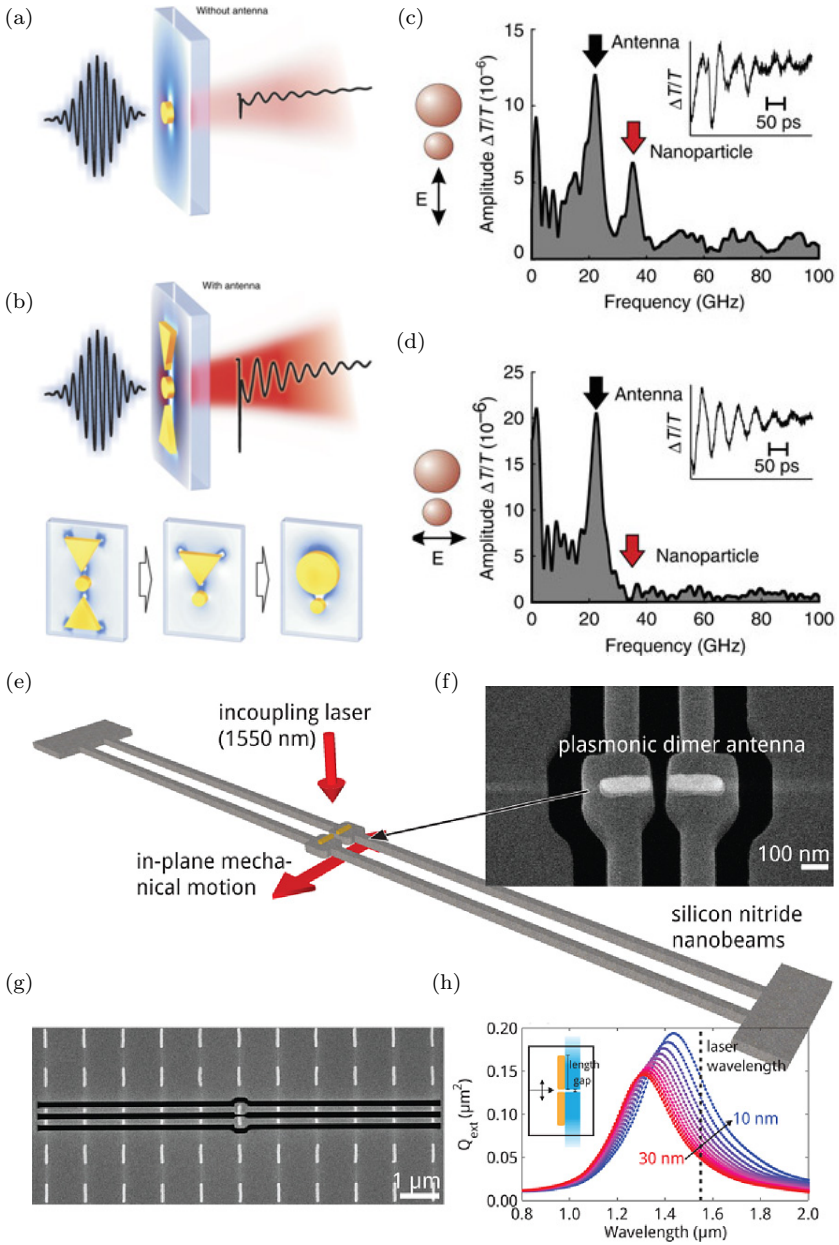
Similar techniques as for single small colloidal nanoparticles can be applied to much bigger plasmonic antenna structures defined using electron-beam lithography.[115–118] The effect of vibrational response here strongly depends on the precise structuring of the nanoantenna.[119] Vibrational modes could be designed for crossed antennas with different arm lengths in two polarizations, allowing selective coupling of light to either of the vibrational modes and coherent control.[117] In Figs. 6.15(a)–(d), the interaction between closely spaced elements in a nanoantenna arrangement was used to enhance the vibrational response of a small nanoparticle, by coupling its dipole moment to a much larger antenna structure.[118] Here, the plasmonic sensitivity for the local near-field environment is used to enhance acousto-optical effects. The field of acousto-plasmonics aims to exploit such combined effects to achieve a significant increase in the optical response. Modulation of plasmonic modes using vibrational coupling connects to the field of optomechanics, where external mechanical modes are used to drive the system. Typically, the systems under study have much lower resonance frequencies than those of single nanoparticles. In recent work,[120] a single plasmonic nanoantenna on a silicon nitride mechanical resonator was used to achieve plasmo-mechanical coupling at a mechanical resonance frequency of around 9 MHz (Figs. 6.15(e)–(h)). Such nanomechanical resonators are receiving great interest as they can potentially be cooled down to the quantum mechanical ground state. Furthermore, the compact design and plasmonic readout makes acousto-plasmonic

systems an interesting novel platform for nanomechanical (thermal) force sensing.

6.4.3.2. *Toward order-unity modulation of plasmonic nanoantennas*

Mechanical forces can potentially be an efficient route to controlling plasmonic response; however, their response is fundamentally limited by the vibrational frequencies of nanomechanical resonators to the MHz–GHz regime. Optical driving of nonlinear phenomena in solids can be based on a range of effects including Kerr-nonlinearity, photorefractive effects, free-carrier, thermo-optic response and phase-change. In bulk form, many of these phenomena are relatively small and require long interaction lengths and/or high field strengths to accumulate sufficient phase- or amplitude response. At the nanoscale, it may be possible to use strong local field confinement in space and/or optical resonance (i.e. enhancement in time) to achieve locally much higher effects than in bulk materials using far-field optics.[89]

Antenna-sensitized response A first strategy for enhancing optical modulation at the nanoscale makes use of the antenna as sensitizer for both the excitation and detection of nonlinear optical response of an external medium. An example of this approach is shown in Fig. 6.16(a), where an antenna is placed on top of an otherwise transparent substrate consisting of a transparent conducting oxide (TCO) on glass.[121,122] Optical excitation of the antenna results in a strong optical absorption and photoexcitation of hot carriers. Due to the high DC-electrical conductivity and ohmic contact between gold and the TCO substrate, electron-mediated heat conductivity results in a very fast spreading of heat into the TCO film in an area of several 100 nm around the antenna. Unlike conventional glass, TCO shows a strong free-carrier nonlinearity, which has been shown to result in order unity changes in permittivity when modulating the free-carrier density. The free-carrier nonlinearity is mediated by a spill out of hot carriers in the localized hot area, governed by the Seebeck effect, resulting in a local reduction of the electron density.


(a)
Without antenna
(b)
With antenna
(c)
Antenna
Nanoparticle
ΔT/T
50 ps
E
Amplitude ΔT/T (10⁻⁶)
Frequency (GHz)
(d)
Antenna
Nanoparticle
ΔT/T
50 ps
E
Amplitude ΔT/T (10⁻⁶)
Frequency (GHz)
(e)
incoupling laser
(1550 nm)
in-plane mecha-
nical motion
silicon nitride
nanobeams
(f)
plasmonic dimer antenna
100 nm
(g)
1 μm
(h)
length
gap
laser
wavelength
10 nm
30 nm
Q_ext (μm²)
Wavelength (μm)

←

Fig. 6.15. (*Figure on facing page*) Detection of the vibrational modes of a small nanoparticle (a) can be enhanced by placing the particle in the near field of a plasmonic nanoantenna (b). Modulation of the nanoparticle response is translated into a larger modulation of the antenna response through near-field interaction, resulting in a polarization-dependent enhancement of the vibrational signal (c,d).[118] (e) Single plasmonic dimer antenna on a Si_3N_4 nanomechanical resonator, with SEM images (f, g) and effect of modulation of nanobeams on optical resonance frequency.[120]

This local perturbation of the substrate is subsequently read out by the shifting of the plasmonic antenna using single-antenna quantitative spectroscopy (Figs. 6.16(b) and (c).[121] By carefully optimizing the antenna design to include an asymmetric gap, different types of nonlinear response were investigated such as the Fano-resonant antenna shown in Fig. 6.16(d).[123]

Next to relatively slow, picosecond nonlinear modulation, much faster optical response could be achieved using femtosecond optical pumping of antenna–ITO hybrids.[124] It was demonstrated that the enhanced local fields around the antenna ends produce a local, ultrafast perturbation of the ITO, which could subsequently be read out by another, perpendicularly oriented antenna positioned within several tens of nanometers distance from the excited antenna. This configuration allows to separate the ultrafast nonlinearity of the substrate from the intrinsic effects associated with photoexcitation of the metal nanoparticle.

Metallic nonlinearity Next to the use of an external nonlinear material, the strong optical nonlinearity of the metal itself produces a large optical modulation. The combination of free-electron response and interband transitions with the SP resonance in small metallic nanosystems provides very large optical Kerr-nonlinearities.[125] Quantitative experiments on small single nanoparticles showed that nonlinear modulation can amount up to several percent of the cross-section.[12,111] The free-carrier nonlinearity in the near-infrared range, far away from any interband transitions, is governed by temperature-dependent changes in the electron–electron scattering rate. Observed changes for single gold antennas on glass amounted

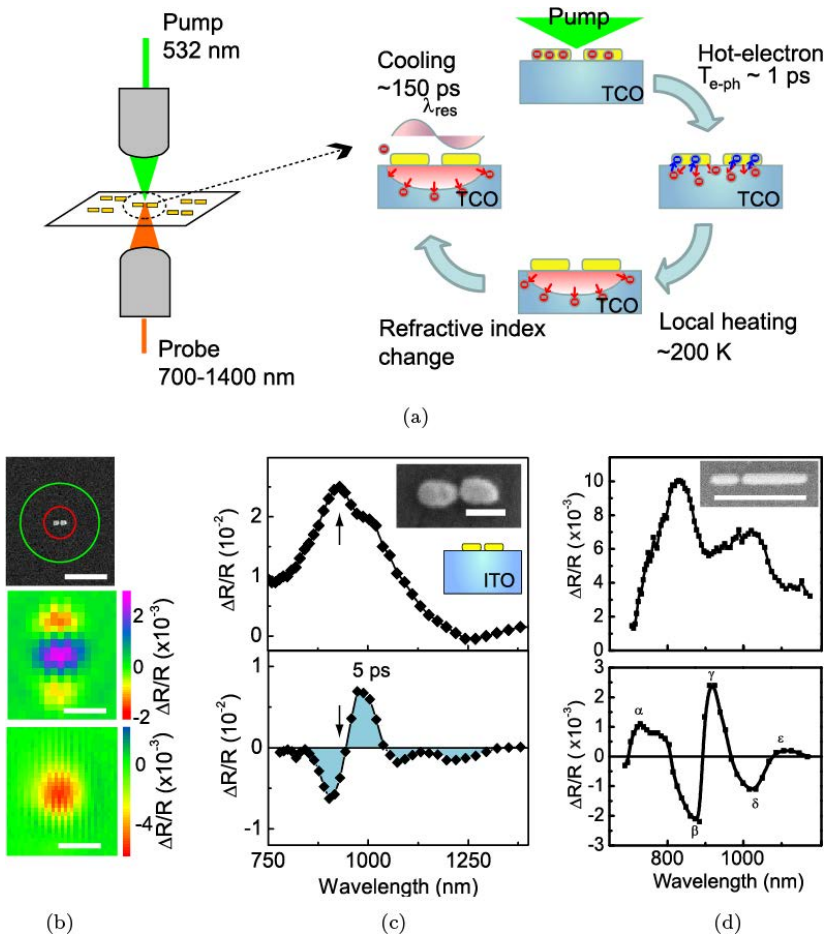


Fig. 6.16. (a) Experimental arrangement for two-color, picosecond optical spectroscopy of individual antennas with cartoon showing the effect of antenna-sensitized optical response (see description in text).[122] (b) SEM image of single antenna (top) with spatial map of optical extinction (middle) and picosecond nonlinear response (bottom) measured in reflection. (c) Optical extinction spectrum (top) and picosecond nonlinear modulation (bottom) for selected antenna–ITO hybrid.[121] (d) Same for asymmetric dimer antenna showing Fano-resonance dip (labelled as γ).[123]

to over 20% modulation of the linear cross-section due to transient bleaching.[126] Next to relatively slow free-carrier nonlinearity, two-photon absorption of energy below the interband transitions offers a fast, sub 100 fs third-order nonlinearity which can be greatly

enhanced by structuring of metal nanostructures into antennas or metamaterials.[127] Next to the response of the thermalized electron gas, recent work has claimed another, fast nonlinearity caused by hot carrier effects.[128] This contribution was observed to occur only for very narrow (<5 nm) gaps between a nanoantenna and a conducting substrate, and had a near-instantaneous response time. The intrinsic metallic nonlinearity can be combined with particular metamaterial response, such as non-locality,[129] epsilon near-zero response[130] and hyperbolic dispersion,[131] to achieve further enhancement of the nonlinear properties.

Coupling plasmonic and photonic resonators Plasmonic antennas provide strong mode confinement below the diffraction limit over a relatively broad resonance with quality factors typically around 10 or below. Photonic resonators offer high-quality resonances, but are limited in mode confinement by the diffraction limit. Coupling of photonic and plasmonic resonators offers an interesting playground for achieving hybrid modes which retain part of both characteristics and offer distinct nonlinearities. Positioning of vertically stacked, parallel nanowire nanoantennas in the standing waves of a microcavity results in symmetric and anti-symmetric combinations where the anti-symmetric mode couples primarily to magnetic resonances of the antenna.[132] Coupling of plasmonic systems to a dielectric waveguide results in sharp Fano resonances in the optical extinction associated with strong optical nonlinearity.[133] Hybridization of microsphere resonators with plasmonic nanoparticles has been developed as a sensitive platform for optical sensing exploiting the confinement of whispering gallery modes on the metal nanoparticles.[134] The effect of photonic resonances on the ultrafast nonlinear modulation of plasmonic nanoparticles was explored by positioning small colloidal nanoparticles in the center of a photonic microcavity.[135] The hybridization with the photonic cavity mode resulted in a narrowing of the SP resonance, which could be interpreted as an increased lifetime of the hybrid photon-plasmon state, and a concomitant enhancement of the optical modulation amplitude. While the dissipative plasmon mode reduces the quality of photonic resonances,

an overall improvement of the nonlinear response can be achieved by optimizing both the plasmonic and photonic systems and their coupling.

Photoconductive loading of antennas Originally developed in the microwave and THz domain, the use of photoconductive materials or varactors in controlling the optical resonances of antennas has been well established.[136] In the optical domain, photoconductive effects require very high carrier densities exceeding $10^{21}\,\mathrm{cm}^{-3}$. Such carrier densities are achievable using pulsed laser excitation of electron-hole plasma in semiconductors. Theoretical studies have explored the effects of photoconductive materials placed in the feed gap of plasmonic antennas.[137,138] It was found that conductive loading of a plasmonic dimer antenna results in a short-circuiting of the antenna, giving rise to large changes in the optical resonance spectrum. In this process, the fundamental mode is changed from two half-wave resonances over the individual antenna arms to a single, half-wave resonance of the full antenna length. The new, long-wavelength mode corresponds to the formation of a long-wavelength charge-transfer plasmon, where the increased conductivity enables a displacement current between the two antenna elements.

A related, yet very distinct nonlinearity is formed by currents caused by quantum tunneling between closely spaced (sub-nm) metal particles.[139–141] Here, the quantum conductivity is essentially nonlinear in the local electric field, leading to pronounced multiphoton absorption and nonlinear susceptibilities exceeding the intrinsic classical coefficients of the constituting metal. Experimental observation of these effects will require exquisite control over the nanoscale gaps, for example using single monolayer materials such as graphene.[142]

Gate tuning Electrical tuning of plasmonic devices works well up to mid-infrared wavelengths for carrier densities up to $10^{19} - 10^{20}\,\mathrm{cm}^{-3}$. Next to the use of semiconductors for the tuning of infrared plasmonics through the free-carrier density, gate tuning of free carriers can be done very easily using two-dimensional materials like graphene

where the optical permittivity depends strongly on the applied bias voltage.[143] Toward the near-infrared, gate tuning requires the use of highly doped materials like ITO and high-k dielectrics like HfO_2 to achieve measurable effects. Several percent modulation were achieved on a silicon-based MOS device platform incorporating plasmonic antennas at telecommunication wavelengths.[144]

Semiconductor dielectric antennas Next to the use of free-carrier effects in semiconductors in controlling the plasmon resonance of metallic antennas, there is an increasing interest in using the resonances of high-index dielectric nanostructures for optical antenna applications. The optical scattering properties of dielectric antennas could be controlled by exciting an ultrafast electron-hole plasma, achieving over 20% modulation of the forward to backward scattering ratio through the dynamical switching of the magnetoelectric interference of modes known as the Huygens source regime.[145] Self-modulation of magnetic Mie resonances in a-Si:H disks was achieved with a depth of 60% using sub-100 fs laser pulses, exploiting the strong spatiotemporal confinement of light in the Mie resonances of the dielectric structure.[146]

Phase change materials Very large optical modulation can be achieved by exploiting structural phase transitions in materials systems such as chalcogenides, or transition metal oxides. Direct optical writing of phase change materials like GST offers a route to non-volatile optically addressable antennas and metamaterials.[147,148] Some of the challenges are related to the relatively high temperatures required to set and reset the crystalline to amorphous phase transition, which is associated with a macroscopic structural rearrangement of atoms. Another opportunity is provided by purely electronic phase transitions found in tetravalent vanadium oxide, VO_2. Here, the phase change occurs at a much lower temperature of 68°C and is associated with a local rearrangement within the unit cell from a monoclinic to rutile structure. The phase change has been used to obtain fast, reversible switching of plasmonic antennas[149,150] and

variable level optical memory functionality by continuously tuning the temperature.[151]

6.5. Conclusions

We have discussed experiments and modelling of single plasmonic nanoparticles and antennas using a range of ultrafast techniques. Metal nanoparticles show pronounced nonlinear optical responses including harmonic generation and two-photon photoluminescence. Femtosecond pulsed excitation of electrons results in dynamics on time scales ranging from several femtoseconds to hundreds of picoseconds, involving plasmon relaxation, electron–electron, electron–phonon scattering and coherent vibrational modes. All these processes give rise to changes in the optical properties, which can be read out and used to analyze the system.

Apart from using techniques to study the nanoparticles, another line of research aims to use plasmonic nanostructures as antennas for increasing light–matter interaction. We have discussed several directions of how a nanoantenna can be used to enhance the nonlinear response of their environment, sensitize and read out external nonlinear media, and act as a nonlinear material by itself. In order to achieve order unity efficiency in optical modulation, a number of strategies can be exploited which may involve combinations of novel materials, nanoscale control in fabrication and quantum phenomena. There is a significant scope for further developing this field, which could result in a next generation of nanoscale photonic devices for controlling and redirecting optical radiation in technological applications.

Acknowledgements

The authors acknowledge fruitful discussions with Christian Girard, Vincent Paillard, Leo-Jay Black and Javier Aizpurua. Otto Muskens acknowledges financially support by EPSRC through grants EP/J016918/1 and EP/M009122/1.

References

1. Zijlstra, P. and Orrit, M. (2011). *Rep. Progr. Phys.* **74**, p. 106401.
2. Crut, A., Maioli, P., del Fatti, N. and Vallee, F. (2014). *Chem. Soc. Rev.* **43**, p. 3921.
3. Major, T. A., Lo, S. S., Yu, K. and Hartland, G. V. (2014). *J. Phys. Chem. Lett.* **5** p. 866.
4. Sun, C. K., Vallée, F., Acioli, L. H., Ippen, E. P. and Fujimoto, J. G. (1994). *Phys. Rev. B* **50**, p. 15337.
5. Del Fatti, N., Voisin, C., Achermann, M., Tzortzakis, S., Christofilos, D. and Vallée, F. (2000). *Phys. Rev. B* **61**, p. 16956.
6. Voisin, C., Christofilos, D., Del Fatti, N., Vallée, F., Prével, B., Cottancin, E., Lermé, J., Pellarin, M. and Broyer, M. (2000). *Phys. Rev. Lett.* **85**, p. 2200.
7. Groeneveld, R. H. M., Sprik, R. and Lagendijk, A. (1990). *Phys. Rev. Lett.* **64**, p. 784.
8. Link, S. and El-Sayed, M. A. (1999). *J. Phys. Chem. B* **103**, pp. 8410–8426.
9. Hodak, J., Martini, I. and Hartland, G. (1998). *Chem. Phys. Lett.* **284**, p. 135.
10. Hodak, J., Henglein, A. and Hartland, G. (2000). *J. Phys. Chem. B* **104**, p. 9954.
11. Arbouet, A., Voisin, C., Christofilos, D., Langot, P., Del Fatti, N., Vallée, F., Lermé, J., Celep, G., Cottancin, E., Gaudry, M., Pellarin, M., Broyer, M., Maillard, M., Pileni, M. P. and Treguer, M. (2003). *Phys. Rev. Lett.* **90**, p. 177401.
12. Muskens, O. L., Del Fatti, N. and Vallée, F. (2006). *Nano Lett.* **6**, p. 552.
13. Del Fatti, N., Voisin, C., Chevy, F., Vallée, F. and Flytzanis, C. (1999). *J. Chem. Phys.* **110**, p. 11484.
14. Hartland, G. V. (2006). *Ann. Rev. Phys. Chem.* **57**, p. 403.
15. van Dijk, M. A., Lippitz, M. and Orrit, M. (2005). *Phys. Rev. Lett.* **95**, p. 267406.
16. Juvé, V., Crut, A., Maioli, P., Pellarin, M., Broyer, M., Del Fatti, N. and Vallée, F. (2010). *Nano Lett.* **10**, p. 1853.
17. Kaganov, M. I., Lifshitz, I. M. and Tanatarov, L. V. (1957). *Sov. Phys. JETP* **4**, p. 173.
18. Brongersma, M. L., Halas, N. J. and Nordlander, P. (2015). *Nat. Nanotech.* **10**, p. 2534.
19. Shen, Y. R. (1984). *The Principles of Nonlinear Optics*, Wiley, J. New York.
20. Boyd, R. W. (2008). *Nonlinear Optics*, 3rd Edition, (Academic Press, USA).
21. Franken, P. A., Hill, A. E., Peters, C. W. and Weinreich, G. (1961). *Phys. Rev. Lett.* **7**, p. 118.
22. Yariv, A. (1967). *Quantum Electronics*, (Wiley).
23. Bloembergen, N., Chang, R. K., Jha, S. S. and Lee, C. H. (1968). *Phys. Rev.* **174**, p. 813.

24. Sipe, J. E., So, V. C. Y. and Fukui, M. G. I. (1980). *Phys. Rev. B* **21**, p. 4389.
25. Ginzburg, P., Krasavin, A. V., Wurtz, G. A. and Zayats, A. V. (2015). *ACS Photon.* **2**, p. 8.
26. Guyot-Sionnest, P., Chen, W. and Shen, Y. (1986). *Phys. Rev. B* **33**, p. 8254.
27. Krause, D., Teplin, C. W. and Rogers, C. T. (2004). *J. Appl. Phys.* **96**, p. 3626.
28. Wang, F., Rodriguez, F., Albers, W., Ahorinta, R., Sipe, J. and Kauranen, M. (2009). *Phys. Rev. B* **80**, p. 233402.
29. Dadap, J. I., Shan, J. and Heinz, T. F. (2004). *J. Opt. Soc. Am. B* **21**, p. 1328.
30. Russier-Antoine, I., Benichou, E., Bachelier, G., Jonin, C. and Brevet, P. F. (2007). *J. Phys. Chem. C* **111**, p. 9044.
31. Butet, J., Duboisset, J., Bachelier, G., Russier-Antoine, I., Benichou, E. and Jonin, C. (2010). *Nano Lett.* **10**, p. 1717.
32. Capretti, A., Pecora, E. F., Forestiere, C., Dal Negro, L. and Miano, G. (2014). *Phys. Rev. B* **89**, p. 125414.
33. Bachelier, G., Butet, J., Russier-Antoine, I., Jonin, C., Benichou, E. and Brevet, P. F. (2010). *Phys. Rev. B* **82**, p. 235403.
34. Butet, J., Bachelier, G., Russier-Antoine, I., Jonin, C., Benichou, E. and Brevet, P.-F. (2010). *Phys. Rev. Lett.* **105**, p. 077401.
35. Jin, R., Jureller, J. E., Kim, H. Y. and Scherer, N. F. (2005). *J. Am. Chem. Soc.* **127**, p. 12482.
36. Chen, K., Durak, C., Heflin, J. R. and Robinson, H. D. (2007). *Nano Lett.* **7**, p. 254.
37. Slablab, A., Le Xuan, L., Zielinski, M., de Wilde, Y., Jacques, V., Chauvat, D. and Roch, J.-F. (2012). *Opt. Exp.* **20**, p. 220.
38. Thyagarajan, K., Butet, J. and Martin, O. J. F. (2013). *Nano Lett.* **13**, p. 1847.
39. Metzger, B., Gui, L., Fuchs, J., Floess, D., Hentschel, M. and Giessen, H. (2015). *Nano Lett.* **15**, p. 3917.
40. Chen, C. K., de Castro, A. R. B. and Shen, Y. R. (1981). *Phys. Rev. Lett.* **46**, p. 145.
41. Chen, C. K., Heinz, T. F., Ricard, D. and Shen, Y. R. (1983). *Phys. Rev. B* **27**, p. 1965.
42. Linnenbank, H. and Linden, S. (2015). *Optica* **2**, p. 698.
43. Bouhelier, A., Beversluis, M., Hartschuh, A. and Novotny, L. (2003). *Phys. Rev. Lett.* **90**, p. 013903.
44. O'Brien, K., Suchowski, H., Rho, J., Salandrino, A., Kante, B., Yin, X. and Zhang, X. (2015). *Nat. Mater.* **14**, p. 379.
45. Lamprecht, B., Krenn, J. R., Leitner, A. and Aussenegg, F. R. (1999). *Phys. Rev. Lett.* **83**, p. 4421.

46. Grinblat, G., Rahmani, M., Cortés, E., Caldarola, M., Comedi, D., Maier, S. A. and Bragas, A. V. (2014). *Nano Lett.* **14**, p. 6660.
47. Czaplicki, R., Mäkitalo, J., Siikanen, R., Husu, H., Lehtolahti, J., Kuittinen, M. and Kauranen, M. (2015). *Nano Lett.* **15**, p. 530.
48. Berthelot, J., Bachelier, G., Song, M., Rai, P., Colas des Francs, G., Dereux, A. and Bouhelier, A. (2012). *Opt. Express* **20**, p. 10498.
49. Husu, H., Canfield, B. K., Laukkanen, J., Bai, B., Kuittinen, M., Turunen, J. and Kauranen, M. (2008). *Metamaterials* **2**, p. 155.
50. Czaplicki, R., Zdanowicz, M., Koskinen, K., Laukkanen, J., Kuittinen, M. and Kauranen, M. (2011). *Opt. Express* **19**, p. 26866.
51. Lippitz, M., van Dijk, M. A. and Orrit, M. (2005). *Nano Lett.* **5**, p. 799.
52. Metzger, B., Hentschel, M., Lippitz, M. and Giessen, H. (2012). *Opt. Lett.* **37**, p. 4741.
53. Hentschel, M., Utikal, T., Giessen, H. and Lippitz, M. (2012). *Nano Lett.* **12**, p. 3778.
54. Metzger, B., Schumacher, T., Hentschel, M., Lippitz, M. and Giessen, H. (2014). *ACS Photonics* **1**, p. 471.
55. Lamprecht, B., Leitner, A. and Aussenegg, F. (1999). SHG studies of plasmon dephasing in nanoparticles, *Appl. Phys. B: Lasers and Opt.* **68**, pp. 419–423.
56. Anderson, A., Deryckx, K. S., Xu, X. G., Steinmeyer, G. and Raschke, M. B. (2010). *Nano Lett.* **10**, p. 2519.
57. Aouani, H., Rahmani, M., Navarro-Cia, M. and Maier, S. A. (2014). *Nat. Nanotech.* **9**, pp. 290–294.
58. Metzger, B., Hentschel, M., Schumacher, T., Lippitz, M., Ye, X., Murray, C. B., Knabe, B., Buse, K. and Giessen, H. (2014). *Nano Lett.* **14**, p. 2867.
59. Beversluis, M. R., Bouhelier, A. and Novotny, L. (2003). *Phys. Rev. B* **68**, p. 115433.
60. Imura, K., Nagahara, T. and Okamoto, H. (2005). *J. Phys. Chem. B* **109**, p. 13214.
61. Biagioni, P., Celebrano, M., Savoini, M., Grancini, G., Brida, D., Mátéfi-Tempfli, S., Mátéfi-Tempfli, M., Duò, L., Hecht, B., Cerullo, G. and Finazzi, M. (2009). *Phys. Rev. B* **80**, p. 045411.
62. Biagioni, P., Brida, D., Huang, J. S., Kern, J., Duo, L., Hecht, B., Finazzi, M. and Cerullo, G. (2012). *Nano Lett.* **12**, p. 2941.
63. Bouhelier, A., Bachelot, R., Lerondel, G., Kostcheev, S., Royer, P. and Wiederrecht, G. P. (2005). *Phys. Rev. Lett.* **95**, p. 267405.
64. Schuck, P. J., Fromm, D. P., Sundaramurthy, A., Kino, G. S. and Moerner, W. E. (2005). *Phys. Rev. Lett.* **94**, p. 017402.
65. Ghenuche, P., Cherukulappurath, S., Taminiau, T. H., van Hulst, N. F. and Quidant, R. (2008). *Phys. Rev. Lett.* **101**, p. 116805.
66. Beermann, J., Novikov, S. M., Söndergaard, T., Boltasseva, A. and Bozhevolnyi, S. I. (2008). *Opt. Exp.* **16**, p. 17302.

67. Huang, J.-S., Kern, J., Geisler, P., Weinmann, P., Kamp, M., Forchel, A., Biagioni, P. and Hecht, B. (2010). *Nano Lett.* **10**, p. 2105.
68. Sanchot, A., Baffou, G., Marty R., Arbouet, A., Quidant, R., Girard, C. and Dujardin, E. (2012). *ACS Nano* **6**, p. 3434.
69. Jägeler-Hoheisel, T., Cordeiro, J., Lecarme, O., Cuche, A., Girard, C., Dujardin, E., Peyrade, D. and Arbouet, A. (2013). *J. Phys. Chem. C* **117**, p. 23126.
70. Viarbitskaya, S., Teulle, A., Marty, R., Sharma, J., Girard, C., Arbouet, A. and Dujardin, E. (2013). *Nat. Mater.* **12**, p. 426.
71. Verellen, N., Denkova, D., Clercq, B. D., Silhanek, A. V., Ameloot, M., van Dorpe, P. and Moshchalkov, V. V. (2015). *ACS Photonics* **2**, p. 410.
72. Viarbitskaya, S., Teulle, A., Cuche, A., Sharma, J., Girard, C., Dujardin, E. and Arbouet, A. (2013). *Appl. Phys. Lett.* **103**, p. 131112.
73. Teulle, A., Marty, R., Viarbitskaya, S., Arbouet, A., Dujardin, E., Girard, C. and Colas des Francs, G. (2012). *J. Opt. Soc. Am. B* **29**, p. 2431.
74. Volpe, G., Cherukulappurath, S., Juanola Parramon, R., Molina-Terriza, G. and Quidant, R. (2009). *Nano Lett.* **9**, p. 3608.
75. Schwab, P. M., Moosmann, C., Wissert, M. D., Schmidt, E. W.-G., Ilin, K. S., Siegel, M., Lemmer, U. and Eisler, H.-J. (2013). *Nano Lett.* **13**, p. 1535.
76. Pavlyukh, Y. and Hübner, W. (2004). *Phys. Rev. B* **70**, p. 245434.
77. de Beer, A. G. F. and Roke, S. (2009). *Phys. Rev. B* **79**, p. 155420.
78. Butet, J., Russier-Antoine, I., Jonin, C., Lascoux, N., Benichou, E. and Brevet, P. F. (2012). *Nano Lett.* **12**, p. 1697.
79. Novotny, L. and Hecht, B. (2006). *Principles of Nano-optics*, (Cambridge University Press, Cambridge).
80. Girard, C. and Dereux, A. (1996). *Rep. Progr. Phys.* **59**, p. 657.
81. Girard, C. (2005). *Rep. Prog. Phys.* **68**, p. 1883.
82. Bozhevolnyi, S. I. and Lozovski, V. Z. (2000). *Phys. Rev. B* **61**, p. 11139.
83. Mäkitalo, J., Suuriniemi, S. and Kauranen, M. (2011). *Opt. Exp.* **19**, p. 23386.
84. Butet, J., Thyagarajan, K. and Martin, O. J. F. (2013). *Nano Lett.* **13**, p. 1787.
85. Martin, O. J. F., Girard, C. and Dereux, A. (1995). *Phys. Rev. Lett.* **74**, p. 526.
86. Girard, C., Dujardin, E., Baffou, G. and Quidant, R. (2008). *New J. Phys.* **10**, p. 105016.
87. Butet, J. and Martin, O. J. F. (2014). *ACS Nano* **8**, p. 4931.
88. Novotny, L. (1997). *J. Opt. Soc. Am. A* **14**, p. 105.
89. Kauranen, M. and Zayats, A. V. (2012). *Nat. Photon.* **6**, p. 737.
90. Black, L.-J., Wiecha, P. R., Wang, Y., de Groot, C. H., Paillard, V., Girard, C., Muskens, O. L. and Arbouet, A. (2015). *ACS Photon.* **2**, p. 1592.

91. Bhawalkar, J. D., He, G. S. and Prasad, P. N. (1996). *Rep. Prog. Phys.* **59**, p. 1041.
92. Harutyunyan, H., Volpe, G., Novotny, L., Agio, M., Alu, A., (Eds.), (2013). *Optical Antennas*, (Cambridge University Press, Cambridge).
93. Jackson, J. (1998). *Classical Electrodynamics*, Third Edition, (Wiley, USA).
94. Gu, L., Sigle, W., Koch, C. T., Ögüt, B., van Aken, P. A., Talebi, N., Vogelgesang, R., Mu, J., Wen, X. and Mao, J. (2011). *Phys. Rev. B* **83**, p. 195433.
95. Girard, C., David, T., Chicanne, C., Mary, A., Colas des Francs, G., Bourillot, E., Weeber, J. C. and Dereux, A. (2004). *Europhys. Lett.* **68**, p. 797.
96. Garcia de Abajo, F. J. and Kociak, M. (2008). *New J. Phys.* **10**, p. 073035.
97. Barwick, B., Flannigan, D. J. and Zewail, A. H. (2009). *Nature* **462**, p. 902.
98. Garcia de Abajo, F. J. (2010). *Rev. Mod. Phys.* **82**, p. 209.
99. Park, S. T., Lin, M., Zewail, A. H. (2010). *New J. Phys.* **12**, p. 123028.
100. Kapitza, P. L. and Dirac, P. A. M. (1933). *Math. Proc. Cambridge Phil. Soc.* **29**, p. 297.
101. Bucksbaum, P. H., Schumacher, D. W. and Bashkansky, M. (1988). *Phys. Rev. Lett.* **61**, p. 1182.
102. Freimund, D. L., Aflatooni, K. and Batelaan, H. (2001). *Nature* **413**, p. 142.
103. Mizuno, K., Pae, J., Nozokido, T. and Furuya, K. (1987). *Nature* **328**, p. 45.
104. Kimura, W. D., Kim, G. H., Romea, R. D., Steinhauer, L. C., Pogorelsky, I. V., Kusche, K. P., Fernow, R., Wang, X. and Liu, Y. (1995). *Phys. Rev. Lett.* **74**, p. 546.
105. Piazza, L., Lummen, T., Quiñonez, E., Murooka, Y., Reed, B., Barwick, B. and Carbone, F. (2015). *Nat. Comm.* **6**, p. 6407.
106. Feist, A., Echternkamp, K. E., Schauss, J., Yalunin, S. V., Schafer, S. and Ropers, C. (2015). *Nature* **521**, p. 200.
107. Yurtsever, A. and Zewail, A. H. (2012). *Nano Lett.* **12**, p. 5027.
108. Arbouet, A., Mlayah, A., Girard, C. and Colas des Francs, G. (2014). *New J. Phys.* **16**, p. 113012.
109. Mukamel, S. (1995). *Principles of Nonlinear Optical Spectroscopy*, (Oxford University Press, UK).
110. Okamoto, H. and Imura, K. (2009). *Progr. Surf. Sci.* **84**, pp. 199–229.
111. Baida, H., Mongin, D., Christofilos, D., Bachelier, G., Crut, A., Maioli, P., del Fatti, N., Vallée, F. (2011). *Phys. Rev. Lett.* **107**, p. 057402.
112. Zavelani-Rossi, M., Polli, D., Kochtcheev, S., Baudrion, A.-L., Béal, J., Kumar, V., Molotokaite, E., Marangoni, M., Longhi, S., Cerullo, G., Adam, P.-M. and Della Valle, G. (2015). *ACS Photonics* **2**, p. 521.
113. Perner, M., Gresillon, S., März, J., von Plessen, G., Feldmann, J., Porstendorfer, J., Berg, K.-J. and Berg, G. (2000). *Phys. Rev. Lett.* **85**, p. 792.
114. Ruijgrok, P. V., Zijlstra, P., Tchebotareva, A. L. and Orrit, M. (2012). *Nano Lett.* **12**, p. 1063.
115. Burgin, J., Langot, P., Del Fatti, N., Vallée, F., Huang, W. and El-Sayed, M. A. (2008). *J. Phys. Chem. C* **112**, p. 11231.

116. Marty, R., Arbouet, A., Girard, C., Mlayah, A., Paillard, V., Lin, V. K., Teo, S. L. and Tripathy, S. (2011). *Nano Lett.* **11**, p. 3301.
117. O'Brien, K., Lanzillotti-Kimura, N. D., Rho, J., Suchowski, H., Yin, X. and Zhang, X. (2014). *Nature Comm.* **5**, p. 4042.
118. Schumacher, T., Kratzer, K., Molnar, D., Hentschel, M., Giessen, H. and Lippitz, M. (2011). *Nat. Comm.* **2**, p. 333.
119. Large, N., Saviot, L., Margueritat, J., Gonzalo, J., Afonso, C. N., Arbouet, A., Langot, P., Mlayah, A. and Aizpurua, J. (2009). *Nano Lett.* **9**(11), p. 3732.
120. Thijssen, R., Kippenberg, T. J., Polman, A. and Verhagen, E. (2015). *Nano Lett.* **15**, p. 3971.
121. Abb, M., Albella, J., Aizpurua P. J. and Muskens, O. L. (2011). *Nano Lett.* **11**, p. 2457.
122. Abb, M., Sepulveda, B., Chong, H. M. H. and Muskens, O. L. (2012). *J. Opt.* **14**, p. 114007.
123. Abb, M., Wang, Y., Albella, P., de Groot, C. H., Aizpurua, J. and Muskens, O. L. (2012). *ACS Nano* **6**, p. 6462.
124. Abb, M., Wang, Y., de Groot, C. H. and Muskens, O. L. (2014). *Nat. Comm.* **5**, p. 4869.
125. Hache, F., Ricard, D., Flytzanis, C. and Kreibig, U. (1988). *Appl. Phys. A: Solids Surf.* **47**, p. 347.
126. Wang, Y., Abb, M., Boden, S. A., Aizpurua, J., de Groot, C. H. and Muskens, O. L. (2013). *Nano Lett.* **13**, p. 5647.
127. Ren, M., Plum, E., Xu, J. and Zheludev, N. (2012). *Nat. Comm.* **3**, p. 833.
128. Harutyunyan, H., Martinson, A. B. F., Rosenmann, D., Khosravi Khorashad, L., Besteiro, L. V., Govorov, A. O. and Wiederrecht, G. P. (2015). *Nat. Nanotech.* **10**, p. 770.
129. Wurtz, G. A., Pollard, R., Hendren, W., Wiederrecht, G. P., Gosztola, D. J., Podolskiy, V. A. and Zayats, A. V. (2011). *Nat. Nanotech.* **6**, p. 107.
130. Kinsey, N., DeVault, C., Kim, J., Ferrera, M., Shalaev, V. M. and Boltasseva, A. (2015). *Optica* **2**, p. 616.
131. Neira, A. D., Wurtz, G. A., Ginzburg, P. and Zayats, A. V. (2014). *Opt. Expr.* **22**, p. 10987.
132. Ameling, R. and Giessen, H. (2010). *Nano Lett.* **10**, p. 4394.
133. Utikal, T., Zentgraf, T., Paul, T., Rockstuhl, C., Lederer, F., Lippitz, M. and Giessen, H. (2011). *Phys. Rev. Lett.* **106**, p. 133901.
134. Ahn, W., Boriskina, S. V., Hong, Y. and Reinhard, B. M. (2012). *ACS Nano* **6**, p. 951.
135. Wang, X., Morea, R., Gonzalo, J. and Palpant, B. (2015). *Nano Lett.* **15**, p. 2633.
136. Chen, H.-T., Padilla, W. J., Zide, J. M. O., Gossard, A. C., Taylor, A. J. and Averitt, R. D. (2006). *Nature* **444**, p. 597.
137. Large, N., Abb, M., Aizpurua, J. and Muskens, O. L. (2010). *Nano Lett.* **10**, p. 1741.

138. Maksymov, I. S., Miroshnichenko, A. E. and Kivshar, Y. S. (2012). *Opt. Expr.* **20**, p. 8929.
139. Haus, J. W., de Ceglia, D., Vincenti, M. A. and Scalora, M. (2014). *J. Opt. Soc. Am* **31**, p. A13.
140. Marinica, D. C., Kazansky, A. K., Nordlander, P., Aizpurua, J. and Borisov, A. G. (2012). *Nano Lett.* **12**, p. 1333.
141. Esteban, R., Zugarramurdi, A., Zhang, P., Nordlander, P., Garcia-Vidal, F. J., Borisov, A. G. and Aizpurua, J. (2015). *Faraday Discuss.* **178**, p. 151.
142. Mertens, J., Eiden, A. L., Sigle, D. O., Huang, F., Baumberg, J. J., Lombardo, A., Sun, A., Sundaram, R. S., Milana, S., Ferrari, A. C., Colli, A., Tserkezis, C. and Aizpurua, J. (2013). *Nano Lett.* **13**, p. 5033.
143. Yao, Y., Kats, M. A., Genevet, P., Yu, N., Song, Y., Kong, J. and Capasso, F. (2013). *Nano Lett.* **13**, p. 1257.
144. Olivieri, A., Chen, C., Hassan, S., Lisicka-Skrzek, E., Tait, R. N. and Berini, P. (2015). *Nano Lett.* **15**, p. 2304.
145. Makarov, S., Kudryashov, S., Mukhin, I., Mozharov, A., Milichko, V., Krasnok, A. and Belov, P. (2015). *Nano Lett.* **15**, p. 6187.
146. Shcherbakov, M. R., Vabishchevich, P. P., Shorokhov, A. S., Chong, K. E., Choi, D.-Y., Staude, I., Miroshnichenko, A. E., Neshev, D. N., Fedyanin, A. A. and Kivshar, Y. S. (2015). *Nano Lett.* **15**, p. 6985.
147. Michel, A.-K. U., Zalden, P., Chigrin, D. N., Wuttig, M., Lindenberg, A. M. and Taubner, T. (2014). *ACS Photon.* **1**, p. 833.
148. Gholipour, B., Zhang, J., MacDonald, K. F., Hewak, D. W. and Zheludev, N. I. (2013). An all-Optical, *Adv. Mater.* **25**, p. 3050.
149. Appavoo, K., Wang, B., Brady, N. F., Seo, M., Nag, J., Prasankumar, R. P., Hilton, D. J., Pantelides, S. T. and Haglund, R. F. (2014). *Nano Lett.* **14**, p. 1127.
150. Appavoo, K. and Haglund, R. F. (2014). *Sc. Rep.* **4**, p. 6771.
151. Lei, D. Y., Appavoo, K., Ligmajer, F., Sonnefraud, Y., Haglund, R. F. and Maier, S. A. (2015). *ACS Photon.* **2**, p. 1306.

CHAPTER 7

Ultrafast Plasmonics

JUE-MIN YI, PETRA GROß, and CHRISTOPH LIENAU*

Carl von Ossietzky Universität Oldenburg, Germany

7.1. Introduction

It has been known since the pioneering work of Ritchie that metal/dielectric interfaces support coherent collective charge density oscillations, and surface plasmons (SPs).[1,2] These surface charge oscillations couple strongly with light and the resulting coupled modes between charges and fields are termed surface plasmon polaritons (SPPs). As discussed in many of the chapters of this handbook, SPPs have two characteristic properties which make them highly interesting for applications.[3–6] They can propagate freely, at almost the speed of light, along planar interfaces. These SPP waves are then evanescently confined in the direction perpendicular to the interface and penetrate the dielectric on a scale given by their wavelength whereas the skin depth gives their penetration depth into the metal. As such, they can be used to guide and

*Corresponding author: christoph.lienau@uni-oldenburg.de

transport light over mesoscopic distances in the form of surface-bound SPP modes. At the same time, these SPP modes can be localized into exceedingly small nanometer-sized volumes, well below the diffraction limit, by tailoring the geometry of the metal. This offers the ability to confine and manipulate light fields at nanometer length scales, to locally strongly enhance optical fields and to create fields with interesting and non-trivial polarization states of the local electromagnetic field. The unique properties of SPPs open the door to optical spectroscopy with nanometer-resolution,[7–10] ultra-sensitive chemical sensing,[11–13] plasmon-enhanced photochemistry[14] or photovoltaics,[15] or plasmonic applications in, for instance, bio-medicine[16] or optical communication.[17] Hence, plasmonics has been a highly active research area in optics during the past decade.

All those intriguing optical properties of plasmonic materials come at a price. SPP fields inevitably penetrate the supporting metallic structure and hence are quickly transformed into heat due to their coupling with electrons and nuclei in the metal. Typical time scales for these couplings range from a few femtoseconds to a few hundreds of femtoseconds at the most, depending on the type of metal and, importantly, its size and shape.[18] Hence, these absorptive "Ohmic losses" put severe constraints on the distance over which SPP plasmons can transport energy. In addition, evanescent SPP modes may rapidly lose energy by being scattered into radiating far-field modes. While such radiative losses are generally negligible for small metal particles with a size of few nanometers,[19] they become important for larger particles and in particular for disordered or corrugated surfaces.[20] Together, absorptive and radiative losses result in short, femtosecond lifetimes of SPP modes. As such, it is evident that ultrafast, time-domain methods can provide important insight into the dynamics of SPP fields. Yet, the short femtosecond lifetimes make it difficult to trace the dynamics of plasmon excitations, in particular, their propagation over mesoscopic distances and their coupling with active quantum emitters, directly in the time domain.

Progress in the development of ultrafast laser sources and the use of advanced coherent spectroscopy techniques made it

possible to perform such time-resolved studies during the last years. This has led, among many other advances, to spatiotemporal imaging of SPP propagation,[21–23] coherent, spatiotemporal control of plasmonic fields in nanostructures,[24,25] time-resolved studies of optical nonlinearities of single plasmonic antennas,[26–31] the observation of coherent Rabi oscillations in strongly coupled exciton-plasmon systems,[32,33] the coherent control of tunneling currents in nanoantennas[34] or the discovery of strong-field photoemission from plasmonic antennas[35] and metallic nanotips.[36–43] In particular, ultrafast photoemission from metallic tips may find intriguing applications in next-generation ultrafast electron microscopes.[44]

It is beyond the scope of this chapter to provide a comprehensive review of this emerging field of ultrafast nanooptics.[45] Instead, we will focus on giving an overview of the most important experimental tools that have so far been employed to study the femtosecond dynamics of plasmonic excitations in metallic nanostructures in the time domain. For each of the methods, we will briefly discuss the underlying experimental concepts and present selected applications to plasmonics. The application examples are mostly from recent work of our own group and reference to related work of other groups is necessarily incomplete. They are mainly chosen to illustrate what type of new information can be obtained from the specific technique and to identify the respective pros and cons of the different methods. With this, we aim at providing the reader with an overview of what can be learned when applying the rapidly advancing toolbox in ultrafast nanooptics to plasmonic systems. We intend to show that such time-domain studies can provide detailed microscopic insight into the fundamental physical processes that are underlying the optical properties of plasmonic nanostructures and hope that this shall stimulate further research in this emerging field.

This chapter is structured as follows: In Section 7.2, Fourier transform spectroscopy (FTS) and spectral interferometry (SI) are introduced as time-domain methods to study the linear optical response of plasmonic systems. In Section 7.3, photoelectron

emission microscopy (PEEM) and interferometric frequency-resolved autocorrelation (IFRAC) spectroscopy are discussed as time-domain nonlinear optical techniques to study spatiotemporal SPP propagation and the nonlinear optical response of single plasmonic nanoantennas. In Section 7.4, coherent ultrafast pump-probe spectroscopy is applied to study strong coupling phenomena between excitons and SPPs in metal/J-aggregate hybrid systems. This chapter concludes with a brief outlook on future challenges.

7.2. Linear Optical Response of Plasmonic Systems

In this section, FTS is introduced to study the optical extinction spectrum of a single metallic nanoparticle. We will then show that SI provides a measurement of the time-domain response function of the plasmonic system and will apply the technique to study light transmission through plasmonic crystals and coherent SPP propagation.

7.2.1. *Fourier transform spectroscopy: Extinction spectrum of a single gold nanorod*

FTS[46] is arguably the most straightforward time-domain method to obtain spectroscopic information in a broad spectral range. Pulses from a broadband light source are sent into a Michelson interferometer, generating a phase-locked pair of optical pulses with time-dependent electric fields $E_0(t) + E_0(t+\tau)$.[a] Here, τ denotes the time delay between the two pulses that is introduced by the interferometer. Both pulses interact with the sample and create time-varying polarizations in the materials. These polarizations then induce the reemission of a pair of secondary pulses $E_1(t) + E_1(t+\tau)$. The total field $E = E_0 + E_1$ that is transmitted through (or reflected from) the sample is the sum of the incident and reemitted fields. It is related to the incident fields by convolution with a time-domain response

[a]For simplicity, scalar fields are assumed. An extension to vectorial fields is straightforward and can be found in standard textbooks.[80,122]

function $r(t)$

$$E(t) = \int_{-\infty}^{\infty} r(t-t')E_0(t')\mathrm{d}t'. \tag{7.1}$$

This response function fully characterizes the optical response of the sample. For a spatially homogeneous and isotropic sample studied in transmission geometry $r(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{t}(\omega)e^{-i\omega t}\mathrm{d}\omega$ is the Fourier transform of the complex transmission coefficient $\tilde{t}(\omega)$.

In FTS, a photodetector then measures the power of the transmitted field created by the two pulses for a series of different time delays τ. The intensity of these pulses, averaged over the integration time of the photodetector, is

$$I(\tau) = \frac{1}{2}\varepsilon_0 c\langle|E(t)+E(t+\tau)|^2\rangle = 2I' + 2C(\tau). \tag{7.2}$$

Here, $\langle ...\rangle$ denotes the time-average and I' is the intensity of just one of the pulses $E(t)$ of the pair. The interesting quantity is the field autocorrelation function

$$C(\tau) = \frac{1}{2}\varepsilon_0 c\mathrm{Re}\langle E(t)E^*(t+\tau)\rangle. \tag{7.3}$$

The Wiener–Khinchin theorem dictates that its Fourier transform $I(\omega) = \varepsilon_0 c|\tilde{E}(\omega)|^2 = \int_{-\infty}^{\infty} C(\tau)e^{i\omega t}\mathrm{d}t$ provides a measure of the power spectrum of $E(t)$. Since the amplitude spectrum of the transmitted field $\tilde{E}(\omega) = \tilde{t}(\omega)\tilde{E}_0(\omega)$ is related to that of the incident field by the complex transmission coefficient of the sample, this is identical to the usual transmission spectrum of a sample measured with a spectrometer,

$$I(\omega) = |\tilde{t}(\omega)|^2|\tilde{E}_0(\omega)|^2 = T(\omega)I_0(\omega). \tag{7.4}$$

Proper normalization then gives the sample transmission $T(\omega)$. It is seen that FTS does not provide the complete time-domain response function $r(t)$ and thus cannot give access to the dynamics of the transmitted field. Equation (7.4) also shows that the result of this measurement does not depend on the spectral phase of the incident pulses and thus, linear FTS spectra can be recorded with both coherent and incoherent light.

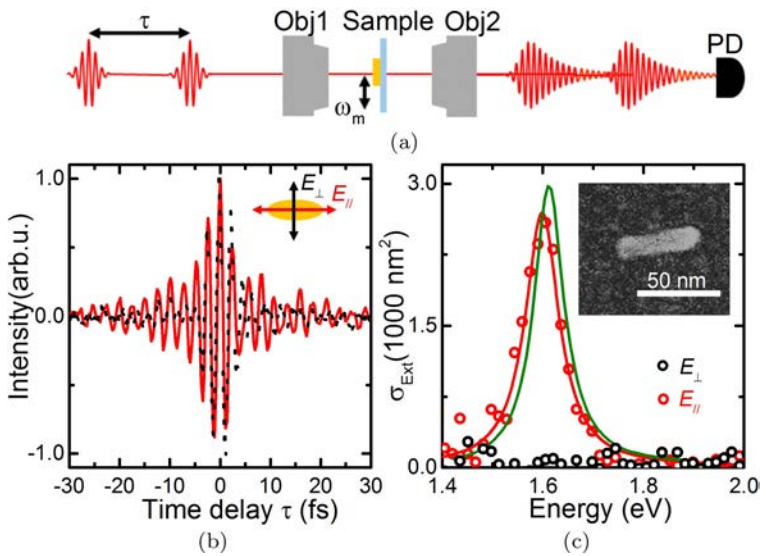


Fig. 7.1. (a) Spatial modulation FTS (SM-FTS) of single gold nanorods. (b) Linear field autocorrelation $C(\tau)$ of the scattering signal from an elliptical nanorod with 10×40 nm dimensions as a function of time delay τ for incident field polarizations parallel ($\parallel$, red curve) and perpendicular ($\perp$, black dots) to the nanorod long axis. (c) Deduced extinction spectra of the nanorods for $\parallel$-(red circles) and $\perp$-polarized (black circles) incident light. Red solid line: Fit to a Lorentzian line shape function. Green line: Finite element simulation of the p-polarized extinction spectrum. Inset: SEM image of a nanorod.

In nanooptics, FTS has quite recently found significant applications in the mid- and far-infrared spectral region for spectroscopic nanoimaging of, e.g. polymer vibrational modes,[47] individual protein complexes,[48] phonon–polariton modes in dielectric materials[49] or thermal near fields.[50] So far, however, FTS has hardly been used for nanospectroscopy in the visible or near-ultraviolet spectral region. This is likely due to the much higher demands on the stability of the Michelson interferometer used to create the time delay τ.

The first application of FTS to the measurement of extinction spectra of single gold nanorods is shown in Fig. 7.1. We study chemically synthesized, elliptical nanorods with dimensions of approximately 40×10 nm, dispersed at low density on a glass substrate. Their longitudinal SPP resonance is centered around 1.6 eV

(800 nm) and strongly red shifted from the much weaker transversal SPP resonance.[51] Commonly, dark-field spectroscopy is employed to measure the scattering cross-section of plasmonic nanoparticles. Since the scattering cross-section of such small rods is quite low ($\sim$60 nm^2), it is challenging to detect them by this method. This can be achieved, however, by spatial modulation (SM) spectroscopy,[52,53] which is a powerful frequency-domain technique providing a quantitative measurement of the extinction cross-section. For our elliptical nanorods, the extinction cross-section is almost the same as the absorption cross-sections and almost two orders of magnitude larger than the scattering cross-sections. This makes SM spectroscopy the method of choice for probing their optical spectra. It has indeed been successfully applied to measuring single nanoparticle scattering cross-sections, but it typically requires a continuous tuning of a narrowband excitation laser. Here, we show that such measurements can be performed without any tuning of the laser by using a broadband excitation source and combining SM and FTS.

For broadband SM-FTS of single nanorods, linearly polarized light pulses from a supercontinuum source are sent through a Michelson interferometer. The resulting time-delayed pulse pair, with a spectrum covering a range from 600 nm to 1000 nm, is focused to a 1-μm spot size using an all-reflective objective.[43] The focused pulses induce collective, dipolar oscillations of the electron gas in the gold rod at frequencies near the longitudinal SP resonance. The electric field that is emitted from the nanorod is collected with a second all-reflective objective, together with the transmitted fraction of the incident pulses. The total power of the transmitted field is measured with a photodiode. To suppress the strong background signal I' in Eq. (7.2), the sample position is periodically modulated at a frequency of a few kHz. The modulation amplitude is set to $\sim$300 nm and the resulting periodic modulation of the recorded photodiode signal is demodulated using lock-in detection. Under these experimental conditions, the recording of a single FTS spectrum at fixed average relative position between sample and focus takes a few to a few tens of seconds. This is sufficiently fast to take spatially resolved

maps of the SM-FT spectra. The data presented in Fig. 7.1 are taken from one of such maps at a position that maximizes the interaction with a specifically selected nanorod.

The resulting field autocorrelation function $C(\tau)$ is shown in Fig. 7.1(b). When choosing an incident polarization perpendicular to the nanorod axis (black dots), the correlation function is short lived, with a duration given by the inverse bandwidth of the laser. For parallel polarization (red line), finite oscillations of the correlation function are detected for delay times of up to 30 fs. A Fourier transform of $C(\tau)$ indeed reveals a pronounced, spectrally narrow peak at the SPP resonance ($\sim$1.6 eV) for parallel polarization. This peak vanishes completely when rotating the incident polarization by 90°. The resulting spectrum is well reproduced by a Lorentzian line shape function with a quality factor of about 15, matching the results of finite element simulations for this type of nanorod (green line in Fig. 7.1(c)).

A more detailed analysis of the data shows that the method provides a quantitative measurement of the extinction cross-section of the nanorod. For the specific example of the rod shown in Fig. 7.1(c), a value of 2700 nm^2 is deduced, matching simulated values to within 10–20%. The Lorentzian lineshape of the resonance is a strong indication that the data show the response of a single nanorod. Thus, SM FTS can provide quantitative time-domain measurements of the local optical extinction spectra of single metallic nanoparticles in a broad spectral range. Since it measures the amplitude of the spectrum, but not its phase, it does not give direct access to the field dynamics.

7.2.2. *Spectral interferometry: Probing plasmon propagation in the time-domain*

SI is a powerful linear optical technique to measure both amplitude and phase of the transmission spectrum and thus reconstruct the dynamics of the transmitted electric field in the time domain.[54–56] For these, a broadband light pulse is split into a sample and a reference pulse in a Mach–Zehnder interferometer (Fig. 7.2(a)). The sample pulse is transmitted through the sample and the transmitted

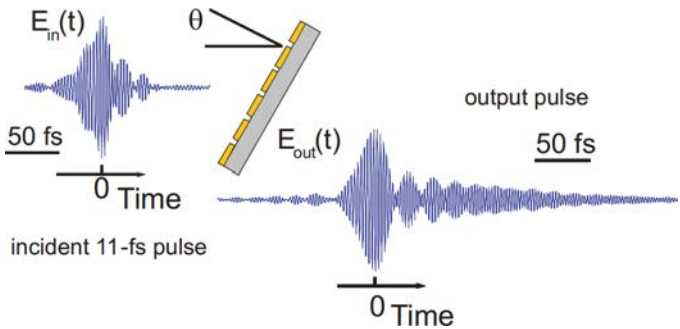


Fig. 7.2. (a) Light transmission through plasmonic nanoslit gratings studied by spectral interferometry. A 650-nm period grating of 50-nm wide slits in a thin gold film is illuminated with 11-fs pulses from a Titanium:sapphire laser with known time profile of their electric field $E_{\rm in}(t) = E_0(t)$. The time structure of the field $E_{\rm out}(t) = E(t)$ is reconstructed from the spectral interferogram between transmitted and incident pulses. The output pulse shows an initial, ultrafast burst reflecting light that is directly transmitted through the slits, with negligible coupling to SPPs. This is followed by a long-lived emission from different, interfering SPP resonances of the grating, persisting for more than 200 fs.[63]

field $E(t)$ is overlapped with the time-shifted reference pulse at the exit of the interferometer. Both beams are sent into a spectrometer which measures the intensity of the sum of the amplitude spectra of these pulses

$$I_{SI}(\omega) \propto |\tilde{E}_0(\omega)e^{-i\omega\tau} + \tau(\omega)\tilde{E}_0(\omega)|^2. \tag{7.5}$$

The phase factor $e^{-i\omega\tau}$ introduced by the time-shift results in pronounced fringes of the SI spectrum. These fringes carry the desired information on the spectral phase $\varphi(\omega)$ of the transmission coefficient $\tau(\omega) = |\tau(\omega)|e^{i\varphi(\omega)}$. To illustrate this, we rewrite $\tilde{E}_0(\omega) = |E_0(\omega)|e^{i\varphi_0(\omega)}$ and obtain

$$\begin{aligned} I_{\rm SI}(\omega) &\propto |\tilde{E}_0(\omega)|^2(1 + T(\omega) + |\tau(\omega)|(e^{-i\omega\tau}e^{i\varphi} + e^{i\omega\tau}e^{-i\varphi})) \\ &\propto I_0(\omega)(1 + T(\omega) + 2|\tau(\omega)|\cos(\omega\tau - \varphi(\omega)) \end{aligned} \tag{7.6}$$

The interference term leads to a high-frequency, quasi-periodic modulation of the recorded spectrum. A straightforward measurement of the frequency spacing $\Delta\omega(\omega) = \omega_2 - \omega_1$ between adjacent

zero crossings of the cosine modulation gives $\frac{\partial\varphi(\omega)}{\partial\omega} = \tau - \frac{\pi}{\Delta\omega(\omega)}$ and, after integration, the spectral phase $\varphi(\omega)$. Alternatively, a Fourier transform of the normalized interferogram I_{SI}/I_0 gives two side peaks centered at $\pm\tau$. The peak at τ contains the time-shifted response function $r(t-\tau)$. Careful peak isolation, time-shift by $-\tau$ and back transformation[57] gives $\tau(\omega)$.

SI is a versatile and flexible technique for phase-resolved measurements of optical spectra and for reconstructing time-domain response functions. Being a linear spectroscopic technique, signal levels are high and high-resolution spectra may be recorded within a few milliseconds or possibly even faster. Data can be analyzed on the fly and amplitude- and phase-resolved spectra may be visualized at video rate. As seen from Equation (7.6), SI is insensitive to the spectral phase of the reference pulse and hence may be used with both coherent and incoherent light sources. The high signal levels make it comparatively easy to obtain information in a broad spectral range and combine the technique with microscopy techniques. In plasmonics, SI far-field microscopy has been used in several groups to study or control plasmon propagation along wires[58] or nanoantennas[59] or to measure the group velocity of SPPs.[60] So far, combinations of SI with near-field microscopy are limited.[61,62]

Figure 7.2 illustrates the first application of SI to studying the dynamics of plasmonic excitations in nanostructures. In these experiments, we have transmitted 11-fs pulses from a Titanium:sapphire laser through a periodic array of nanometric slits in a 150-nm thin gold film. The time-dependent electric field of the incident pulses has been fully characterized prior to the measurements. These one-dimensional slits support a propagation mode so that a fraction of the light can directly pass to the backside of the grating. In addition, scattering of the incident light at the grating can launch SPP excitations at either the front or back side of the grating. These two types of excitations are clearly separated by measuring the time structure of the electric field of the transmitted pulses (blue curve in Fig. 7.2). The reconstructed electric field contains two distinct contributions: first, an ultrafast burst of light with a time structure which is very

similar to that of the incident pulse. This burst reflects a spectrally broadband, non-resonant contribution to the transmitted signal due to direct light transmission through the slit.[64,65] Second, the measurement reveals a long-lived transmission signal persisting for several hundreds of femtoseconds and showing an initial beating pattern at early times. This resonant contribution to the transmission stems from the excitation of standing wave SPP modes at the slit grating. Depending on incidence angle, these SPP modes can be very long-lived, with lifetimes that can even exceed those of SPPs at a planar gold/air interface. The origin of these long lifetimes can be uncovered by measuring the spatial mode profiles of the SPP fields. Long-lived SPP modes have field profiles that show little intensity at the slits, strongly suppressing radiative damping of the SPP fields.[63] In the spectral domain, the interference between the resonant SPP reemission and the off-resonant direct transmission gives rise to highly asymmetric, Fano-like lineshapes.[63,65] These Fano-type resonances are characteristic for this type of plasmonic gratings.[66] as well as for many other plasmonic nanostructures.[67] In the time domain, they are clearly revealed as a biexponential decay of the associated reemitted electric field.

A Fourier transformation of the time-domain field gives amplitude and phase of the transmission coefficient in a broad spectral range and, when recorded for a sufficiently large range of incidence angles, a precise measurement of the complete bandstructure of this type of plasmonic crystal.[63,68]

As a second example, we discuss the use of SI for probing the propagation of SPP pulses in plasmonic devices. We consider a slit-grating structure[23,60] depicted in Fig. 7.3(a,b). Short, broadband light pulses are transmitted through a 150-nm wide slit in a 200-nm thick epitaxially grown gold film. This launches SPP pulses at the back side of the film. These SPP waves partially propagate towards a 150-nm wide and 60-nm deep groove where they are scattered into the far field. The collected field is collimated and overlapped with a time-shifted replica of the incident pulse (shown in blue in Fig. 7.3(a)). The resulting spectral interferograms are analyzed and

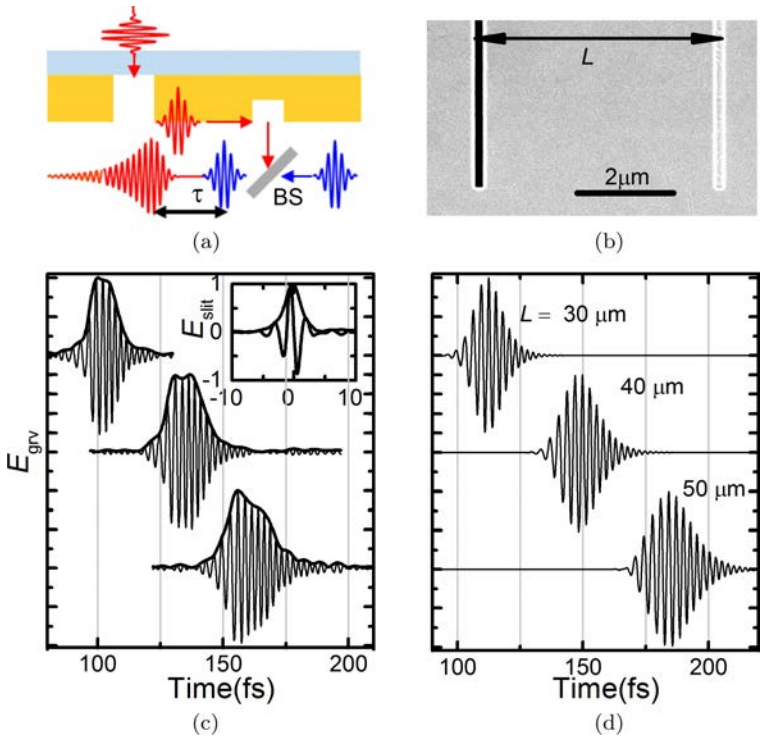


Fig. 7.3. (a) Sketch of excitation, propagation and recording of SPP pulses in a slit-groove structure. (b) Scanning electron microscope (SEM) image of a slit–groove structure in a 200-nm thick epitaxially grown gold film. (c) Reconstructed time structure of the propagating SPP field at grooves with a distance of $L = 30$, 40, 50 μm from the launching slit, respectively. Inset: Field directly transmitted through the slit. (d) Calculated electric fields at the grooves with varying slit–groove distance L by FDTD simulations.[23]

the time structures of the time-domain response functions $r(t)$ are reconstructed for the light that is transmitted through the slit or scattered from the groove.

This response function corresponds to the shortest possible pulse that can be generated at the different sample positions. The measurements are performed for several different distances L between slit and groove, varying between 30 μm and 50 μm, to visualize the effect of the SPP propagation of the pulse profile. As incident light,

pulses from a supercontinuum light source are used. Since the slit supports a propagating mode, essentially the entire bandwidth of the incident light is transmitted, resulting in an extremely short response function measured at the slit exit that supports a single-cycle SPP pulse. The effect of the propagation between slit and groove on the time-structure of such a short SPP pulse is clearly visualized in Fig. 7.3(c). The measurements not only reveal the increasing time delay of the pulses with increasing distance but they also show the reshaping and stretching of the pulse. A detailed data analysis shows that the deduced average group velocity ($\sim$0.9 c_0, c_0: speed of light in vacuum) agrees well with the tabulated values for the refractive index of gold and the predicted dispersion relation of a gold/air interface. Also, the measured values for the frequency-dependent group velocity dispersion $(\sim 0.5\ \mathrm{fs}^2\mu\mathrm{m})^{-1}$ agree reasonably well with those deduced from the tabulated refractive index. This shows that such measurements of plasmon propagation at the field level can give important information, with attosecond precision, on how this propagation affects the shape of the pulse. This may be helpful for probing the effects of optical nonlinearities on plasmon propagation and to develop and implement novel classes of ultrafast plasmonic switches.

7.3. Ultrafast Nonlinear Plasmonics

The linear spectroscopic techniques discussed in the previous chapter can provide time-domain measurements of the linear response of a plasmonic structure. As such, they give access to the electric field dynamics and — if combined with microscopy methods — to the local electric field dynamics in such structures. Yet, being linear optical techniques, they often suffer from interference from background or stray fields which may be difficult to eliminate even when using advanced modulation techniques.

One of the characteristic features of plasmonic samples is the strong local enhancement of the optical fields at the surface of the nanostructure. This can locally enhance optical nonlinearities by

orders of magnitude. This unique field localization feature and the short lifetimes of the localized plasmonic modes make it attractive to employ time-domain nonlinear spectroscopy techniques for characterizing plasmon dynamics. Such methods have been used by a number of research groups to characterize optical nonlinearities in single nanostructures[12,26–28,30,31,6–72] or to probe plasmon propagation and localization in a variety of plasmonic systems.[21,22,24,25,73,75] Here, we will discuss two of these methods, interferometric frequency-resolved nonlinear correlation spectroscopy and PEEM in more detail and give selected examples for their application.

7.3.1. *Interferometric frequency-resolved autocorrelation: Nonlinearities of single nanostructures*

IFRAC, or interferometric frequency-resolved optical gating, is a powerful method to retrieve the time structure of the electric field of an unknown ultrashort laser pulse.[75–77] The IFRAC measurements have also been helpful in determining the carrier-envelope offset (CEO) frequency of phase-stabilized pulses.[78]

Experimentally, IFRAC measurements are performed again by making use of a collinearly propagating, phase-locked pair of time-delayed pulses that is generated in a Michelson interferometer. This pulse pair is focused onto a sample and induces a nonlinear optical polarization.[79] The electric field that is reemitted by this nonlinear material polarization is spectrally dispersed in a monochromator and the intensity of the field is recorded with a charge-coupled device (CCD) camera as a function of the time delay τ between the two pulses.

To be specific, let us now assume that the pulses are inducing second harmonic (SH) radiation in an off-resonant and isotropic material. Then, the material is characterized by a second-order optical nonlinear susceptibility $\chi^{(2)}$ which does not vary strongly with laser wavelength. In this limit, the emitted SH field is simply proportional to the square of the total incident laser field and the

recorded interferometric autocorrelation signal can be written as

$$I_{\mathrm{IF}}(\omega,\tau) \propto |\int_{-\infty}^{\infty} (E(t)+E(t+\tau))^2 \exp(-i\omega t)\mathrm{d}t|^2 \tag{7.7}$$

It can be seen in Fig. 7.4, that this is quite a complex interferogram, showing many different types of fringes and carrying very detailed information. On the horizontal axis, it displays high-contrast temporal fringes along the delay axis. The vertical axis shows how these fringes vary with the detection wavelength $\lambda = 2\pi c/\omega$. Most importantly, as shown by Stibenz *et al.*,[76] this interferogram can be used to reconstruct the time dynamics of the electric field $E(t)$ of one of the incident pulses. To see this, we write $E(t) = A(t)\exp(i\omega_0 t)$ as the product of a complex amplitude $A(t)$ and a carrier wave at angular frequency ω_0 and introduce the difference frequency $\Delta\omega = \omega - 2\omega_0$. Then, Eq. (7.7) can be transformed into

$$\begin{aligned} I_{\mathrm{IF}}(\omega,\tau) = {} & 2|E_{\mathrm{SH}}(\Delta\omega)|^2 + 4|E_{\mathrm{FROG}}(\Delta\omega,\tau)|^2 \\ & + 8\cos\left[\left(\omega_0+\frac{\Delta\omega}{2}\right)\tau\right]\mathrm{Re} \\ & \times\left[E_{\mathrm{FROG}}(\Delta\omega,\tau)E^*_{\mathrm{SH}}(\Delta\omega)\exp\left(i\frac{\Delta\omega}{2}\tau\right)\right] \\ & + 2\cos[(2\omega_0+\Delta\omega)\tau]|E_{\mathrm{SH}}(\Delta\omega)|^2 \end{aligned} \tag{7.8}$$

Here,

$$E_{\mathrm{SH}}(\Delta\omega) = \int_{-\infty}^{\infty} A(t)^2 \exp(-i\Delta\omega t)\mathrm{d}t \tag{7.9}$$

is the amplitude spectrum of SH field of a single pulse, and

$$E_{\mathrm{FROG}}(\Delta\omega,\tau) = \int_{-\infty}^{\infty} A(t)A(t-\tau)\exp(-i\Delta\omega t)\mathrm{d}t \tag{7.10}$$

is the so-called FROG (frequency-resolved optical gating) field of the two pulses.[80,81] This FROG field is of particular interest for us since the spectrogram $I_{\mathrm{FROG}}(\Delta\omega,\tau) = |E_{\mathrm{FROG}}(\Delta\omega,\tau)|^2$ can be used to reconstruct amplitude and phase of the incident pulse. "FROG" is currently the most commonly used technique for the

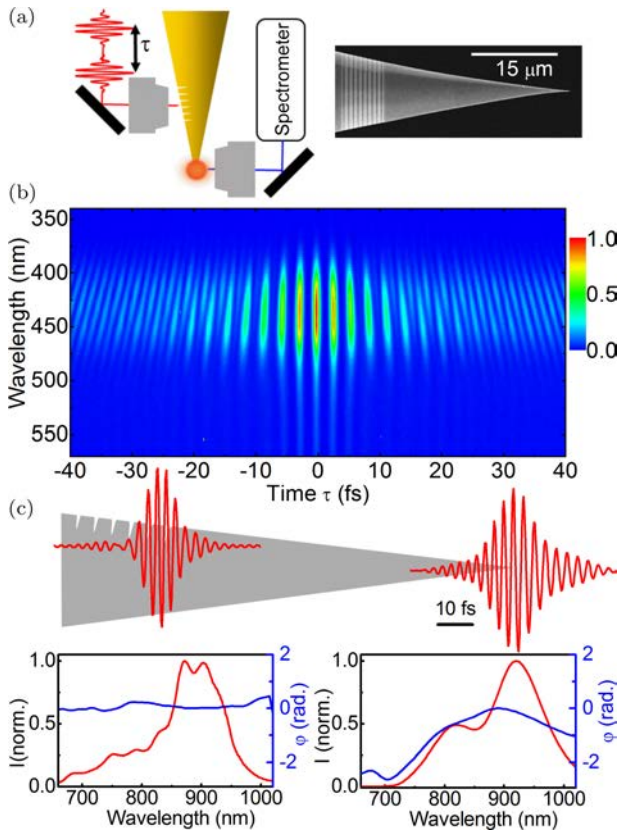


Fig. 7.4. (a) Schematic of the IFRAC microscope for time-resolved characterization of adiabatically nanofocused plasmonic light spots at the apex of a conical gold taper. A scanning electron microscope image of a taper with a grating coupler is shown. (b) Experimental IFRAC trace of the second harmonic emission from the plasmonic light spot at the taper apex. (c) Time profile of the electric fields (red curves) of the incident laser and the nanofocused field at the taper apex. The intensity spectra and spectral phases of both fields, retrieved from the IFRAC measurements, are shown as red and blue lines, respectively. In contrast to the flat phase of the incident pulse, the nanofocused pulse shows a small second-order dispersion of 25 fs^2.

analysis of ultrashort optical pulses and a variety of phase-retrieval codes[82,83] are publicly or commercially available to obtain $E(t)$ from $I_{\mathrm{FROG}}(\Delta\omega, \tau)$.

The FROG signal can now be isolated from the IFRAC interferogram by calculating a Fourier transform of I_{IF} along the delay

time axis. This Fourier transform shows well-separated peaks around 0 (the DC component), $\pm\omega_0$ (the fundamental sidebands) and $\pm 2\omega_0$ (the second-order sidebands). The individual peaks in the Fourier transform can be isolated by, e.g. super-Gaussian filtering and zero padding. Fourier back transformation of the DC peak gives the sum of the first two terms in Eq. (7.8) and, after subtraction of the delay-independent background $2|E_{\mathrm{SH}}|^2$, the desired spectrogram $I_{\mathrm{FROG}}(\Delta\omega, \tau)$. Amplitude and phase of the incident field $E(t)$ can then be reconstructed using standard phase-retrieval software.

It is important to note that also the first-order sideband can be used for phase retrieval.[77] and that the second-order sidepeaks provide a phase reference which can be helpful for extracting additional information from the first-order peak. Also, we note that the quantitative analysis of such IFRAC data is of course not limited to material systems that are showing a quasi-instantaneous nonlinear optical response. In such cases, a quantitative analysis of the data requires a more advanced modeling of the transient nonlinear optical response of the system. This may, for example, be achieved by density-matrix-based simulations of optical Bloch equations. Such a more advanced approach has successfully been applied, for example, to analyze carrier-wave Rabi oscillations in semiconductors,[84,85] to distinguish between SH emission and two-photon induced luminescence[86] or to study light localization in zinc oxide nanoneedle arrays.[87]

Here, we use IFRAC to study the pulse duration of ultrashort, nanolocalized light spots that are created at the apex of a sharp, needle-like gold taper by plasmonic nanofocusing.[7,31] Plasmonic nanofocusing is an emerging, powerful technique to create isolated and ultrashort nanometer-sized optical[27,86,88] and electron pulses[89–91] by using plasmonic superlenses as focusing element. Its applications are rapidly increasing in different areas of nanooptics.[29,92–97]

Plasmonic nanofocusing combines two of the outstanding properties of metallic nanostructures, their ability to coherently transport energy in the form of propagating SPP waves over mesoscopic distances and to localize this energy into nanometric spots by coupling SPPs with localized surface plasmon (LSP)

resonances. In particular, sharp conical gold tapers, such as those shown in Fig. 7.4(a), are of interest for plasmonic nanofocusing. Theoretical work by Babadjanyan *et al.*[98,99] and Stockman[100] showed that the lowest order eigenmode of such a conical metallic taper remains a bound, evanescent SPP mode even for vanishingly small taper radii. The analysis of the eigenmodes of such tapers indicate that the effective refractive index of this mode increases strongly with decreasing tip radius and diverges in the limit of vanishing radius. Consequently, the local SPP wavelength decreases to zero when approaching the taper apex. For the ideal limit of an infinitely sharp tip and a sufficiently slow (adiabatic) spatial variation of the taper radius, both phase and group velocities decrease to zero, and hence the SPP wave packet is predicted to slow down and come to a complete halt as it approaches the tip apex.[100] The adiabatic decrease in group velocity is related to a divergence of the electric field strength near the tip apex and to an extreme concentration of the electromagnetic energy stored in the SPP wave packet into a vanishingly small spot size.[100] Effectively, the spatial extent of the SPP wave packet reduces as it approaches the apex and is transformed adiabatically from a propagating into a localized mode. Even if the field singularity at the tip apex is removed by limiting the minimum tip radius to 2 nm,[101,102] a strong enhancement of the field intensity by more than three orders of magnitude remains. This field localization is expected for adiabatic tapers whose change in taper diameter is small on a scale of the wavelength.

This raises the question whether this adiabatic decrease of the SPP group velocity can be seen experimentally and to what extent this affects SPP propagation along such conical tapers.[31,103,104] To analyse these dynamics, we used the experimental setup schematically depicted in Fig. 7.4(a). Laser pulses at a center wavelength of about 800 nm and with a duration of only 6 fs are derived from a Titanium:sapphire oscillator operating at a repetition rate of 82 MHz. The dispersion of those pulses is controlled by a pair of chirped mirrors and their spectral amplitude and phase is precisely measured,

as shown in Fig. 7.4(c). This pulse is sent through a Michelson interferometer and the resulting time-delayed pulse pair is focused onto the shaft of a conical gold taper with an opening angle of about 30°. An all-reflective objective is used for focusing in order to avoid a chirping of the incident pulse.[105] The gold taper is prepared from a single-crystalline gold wire by chemical etching and has a particularly smooth surface. It is equipped with a slit grating with a subwavelength period positioned at a distance of about 30 μm from the taper apex. This grating is used to launch SPPs at the taper surface which then propagate as evanescent waves towards the taper apex. Away from the apex, the intensity of these pulses is so low that all nonlinear optical effects can safely be neglected. Close to the apex, however, the spot size of the SPP fields gradually decreases. This results in a local enhancement of the field intensity that is so significant that quite intense SH radiation is generated in the apex region. The emitted SH is collected with a high numerical aperture objective with a focal spot size of less than 1 μm, imaged onto a monochromator and detected with a CCD camera. Typically, we collect in these experiments about 10^6–10^7 SH photons per second, even though the apex radius is quite small — less than 10 nm as confirmed independently by high-resolution near-field scattering optical microscopy of single gold nanoparticles.[86]

Such signal levels are well sufficient to record noise-free IFRAC traces, as illustrated in Fig. 7.4(b). Compared to the IFRAC traces of the incident pulses,[43,86] the interference fringes in the signals from the tip apex extend to slightly longer time delays, pointing to a slight temporal stretching of the pulse at the apex of the taper. The electric field profile of the pulse at the taper apex that is deduced from the IFRAC measurements is shown in Fig. 7.4(c), together with that of the 6-fs incident pulses. A slight increase of the pulse duration and a minor chirp of the pulse is seen. Quantitative values of the intensity spectra and spectral phases of both fields are depicted in the insets in Fig. 7.4(c) as blue and red lines, respectively. It is seen that the intensity spectrum of the pulses reaching the apex is very similar

to that of the incident pulses, apart from a cutoff at short wavelengths. This cutoff is due to the strong increase in SPP absorption in the gold film in the blue part of the pulse spectrum. Such a broad intensity spectrum at the apex is achieved by rather tight focusing onto the grating coupler which largely increases the bandwidth of the incident light that is coupled with SPPs on the taper shaft. The bandwidth of the spectrum is more than 250 nm wide, which is sufficient to create 3-cycle pulses. In contrast to the flat spectral phase of the incident pulses, the retrieved spectral phase of the pulses at the apex shows a small, yet clear second-order dispersion of 25 fs^2. This stretches the pulse duration at the apex to more than 10 fs. Only a part of this pulse stretching is due to the measured chirp of the pulse. Our results suggest that the measured value of the second-order dispersion would result in a stretching of a chirp-free 6-fs pulse to slightly less than 8 fs. The additional increase in pulse duration at the apex is due to the finite spectral bandwidth of the grating coupler.

It seems tempting to assign this chirp to the predicted slowdown of the SPP group velocity in the apex region. A comparison of the experimental data with fully three-dimensional finite difference time-domain (FDTD) simulations of the SPP wavepacket dynamics indicates, however, that the actual spatiotemporal SPP propagation dynamics is more complex. Since the SPP pulses are only launched in a small area of a few square-microns on the grating coupler, the resulting SPP wavepacket will not only propagate towards the apex but also spread along the circumference of the taper (a movie of the SPP propagation dynamics is shown in the Supporting Information of Ref. [69]). This results in a spatial and temporal spread of the SPP wavepacket and may lead to local pulse durations that are even larger than those at the apex. Hence, the chirp that is seen in Fig. 7.4(c) mostly results from SPP propagation along the taper shaft. Importantly, the FDTD simulations support the main conclusion that plasmonic nanofocusing leads to the formation of high-intensity few-cycle SPP pulses at the very apex of the taper, with spot sizes solely limited by the taper apex diameter.

Similar studies of the SPP pulse propagation have also been reported in Ref. [104]. Here, group delays of about 9 fs have been observed in the apex region of chemically etched polycrystalline gold tapers and have been assigned to a nanofocusing-induced slowdown of the group velocity to about $0.2\,c_0$. Recent electron energy-loss spectroscopy (EELS) experiments[106–108] suggest that such SPP localization phenomena may strongly depend on the geometry of these tapers, in particular their opening angle and surface roughness. The exact shape of the taper in a subwavelength region around the apex may have a profound influence and result in complex reflection and localization phenomena that deserve further experimental investigations.[97,109,111]

In summary, the presented results show that IFRAC measurements are a powerful method to accurately measure the time structure of the local electric field at the surface of a plasmonic nanostructure. In contrast to spectral interferometry, discussed in Section 7.2.2, it does not measure the linear response function of a certain plasmonic system, but allows reconstructing the time-dependent SPP field by means of fast and efficient retrieval algorithms. This pulse characterization at the field level can only be achieved by nonlinear optical techniques.[111] It is important that such a retrieval cannot be unambiguously achieved on the basis of more conventional interferometric autocorrelation measurements, even if the intensity spectrum of the SPP pulse is known.[82] As such IFRAC measurements or other variants of nonlinear, frequency-resolved optical grating techniques are of paramount importance for providing full temporal information about SPP fields. Their sensitivity is so high that even the SPP field dynamics of single nanoparticles can be analyzed.[26,27,31]

Being far-field optical techniques, their spatial resolution is inherently limited by the wavelength of light. Photoelectron emission microscopy (PEEM), discussed in the next section, can overcome this limitation and can greatly increase the spatial resolution of time-domain SPP field characterization techniques, in principle even to less than 10 nm.

7.3.2. *Ultrafast photoemission electron microscopy: Controlling optical fields on the nanoscale*

Photoemission electron microscopy (PEEM) is a well-established technique in surface science in which the spatial distribution and, possibly, kinetic energy of photoelectrons that are emitted from a surface are imaged with high resolution.[112,113] It is a surface-sensitive technique since the photoelectrons are emitted from a shallow layer near the surface. Photoelectron microscopes operate under ultrahigh vacuum conditions and require a high bias voltage between the sample and the lens system that images the photoelectrons onto a two-dimensional detector screen. Under favorable conditions, the spatial resolution of such PEEM images can reach 10 nm or potentially even less. Thus, it can exceed that of far-field optical techniques discussed in Section 7.3.1. by approximately one order of magnitude.

In plasmonics, PEEM has first been used by Kubo *et al.* to study photoemission from a structured silver film.[21] In their experiment, they used a phase-locked pair of ultrashort 10-fs pulses at a wavelength of 400 nm to emit photoelectrons from a silver line grating. Since the photon energy of the incident pulses of 3.1 eV is much lower than the work function of silver ($\sim$4.2 eV), photoemission is a nonlinear, two-photon process. The photoemission yield scales with the fourth power of the incident field, as in SH emission discussed in Section 7.3.1. Due to this nonlinearity of the photoemission yield, the images show photoemission from local "hot spots" of the plasmon field at the surface (Fig. 7.5(a)). Hot spot formation is a consequence of SPP localization resulting from multiple scattering of SPPs in the rough and disordered line grating. It is clearly revealed by the local enhancement of the electron yield at certain positions along the line grating. Kubo *et al.* have now recorded such images as a function of the time-delay between the two pulses that were used for photoexcitation. A typical interferogram is shown in Fig. 7.5(b). When comparing interferograms recorded at different hot spots, slight variations in the fringe spacing were observed at delay times of more than 20 fs. These variations were taken as a signature that localized SPP modes with varying resonance frequencies were associated to

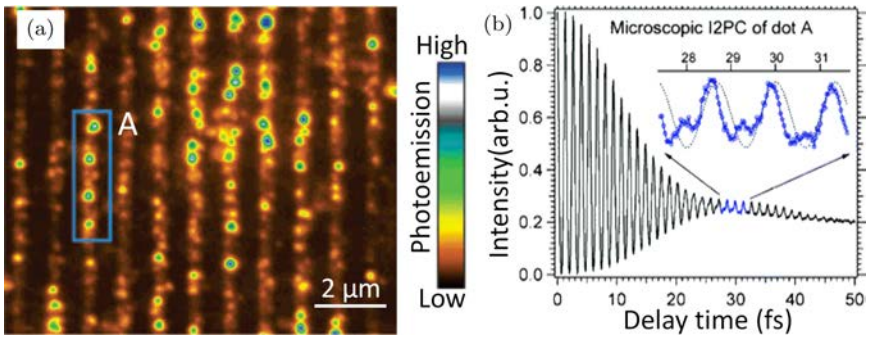


Fig. 7.5. (a) The PEEM micrograph of a silver grating illuminated with 400-nm femtosecond laser, linearly polarized perpendicular to the grating lines. The surface roughness gives rise to excitation of the localized SP modes seen as hot spots. (b) Microscopic interferometric two-pulse correlation scan from the individual hot spot A in the blue rectangle of (a). Figure reproduced with permission from Ref. [21].

the different hot spots. The lifetime of these localized SPP modes is sufficiently long to result in lasting fringes in the interferogram. It was difficult to extract more quantitative information about the dynamics of the local hot spots from those measurements.

In later experiments from the same group,[22] a phase-locked pair of similar 10-fs pulses at 400-nm wavelength was used to illuminate a line defect in a silver film. The PEEM images of the local photoelectron emission yield are recorded as a function of the time delay between the pulses and are presented in the form of a movie in the Supporting Information of Ref. [22]. The movie shows a spatially modulated emission pattern that propagates away from the slit as the time delay increases. Light scattering at the slit launches propagating SPP fields at the surface of the silver film that propagate away from the slit at a group velocity of about 0.6 c_0. At 400-nm wavelength, the SPP propagation length is rather short, $\sim$3 μm, corresponding to an SPP lifetime of 12 fs. This limits the experimentally observed interference patterns to a narrow region around the slit. Much longer propagation lengths are observed when using longer wavelength pulses for excitation.[114] The analysis of the results in Ref. [22] shows that the interference between the excitation field,

incident under a small angle, and the launched SPP field needs to be properly taken into account to match the experimental observations and extract quantitative information about the plasmon propagation. Generally, such PEEM measurements provide very detailed insight into the spatiotemporal propagation of SPP waves in plasmonic structures. Ultrafast PEEM can reveal the excitation and propagation of SPP waves even in complex and highly spatially inhomogeneous structures and has been employed, e.g. to probe the dynamics and lifetimes of SPP excitations in a number of different structures.[72,73,115,116].

An entirely new class of experiments becomes possible when combining PEEM with tailor-made ultra-fast optical pulses such as those shown in Fig. 7.6(a). Powerful pulse shaping tools have been developed over the last years to sensitively control the amplitude and phase of each of the frequency components of an ultrashort optical pulse.[117] By making use of advanced spatial light modulation techniques, Fourier-domain pulse shaping provides exquisite control over the temporal structure of the electric field of ultrashort optical pulses. Moreover, optical pulses with complex, time-varying polarization states can be synthesized by controlling the polarization state of each of the spectral components of the pulse.[118] This polarization pulse shaping thus does control the temporal shape not only of the amplitude, but also of the vectorial properties of an ultrashort optical pulse. An illustration of a complex time-varying electric vector field that can result from such a polarization pulse shaping is presented in Fig. 7.6(a).

This offers the exciting opportunity to coherently control the spatiotemporal dynamics of SPP fields in plasmonic nanostructures. Let us assume that such a polarization-tailored optical pulse interacts with an extended plasmonic nanostructure and that the spectrum of the pulse overlaps with several plasmonic eigenmodes of the nanostructure. This interaction then leads to a complex SPP field pattern at the surface. The spatiotemporal properties of this field depend, in general, sensitively on the polarization properties of the input pulse. Polarization pulse shaping thus offers an intriguing perspective to

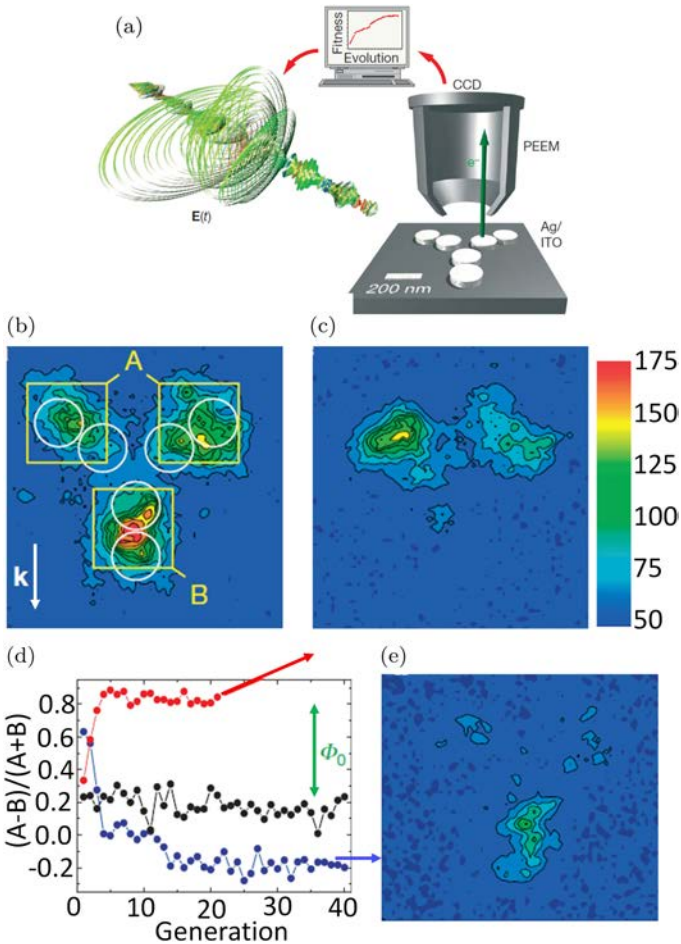


Fig. 7.6. (a) Experimental scheme of a PEEM combined with an ultrafast pulse shaper. The polarization shaper for ultrashort laser pulses controls the temporal evolution of the vectorial electric field $E(t)$ on a femtosecond scale. These pulses illuminate plasmonic nanostructures in an ultrahigh-vacuum chamber with a PEEM. A CCD camera records the photoemission image and provides a feedback signal for an evolutionary learning algorithm. The experimental photoelectron distribution of the star-shaped nanostructure is displayed under p-polarized excitation with unshaped pulses (b), after adaptive A/B maximization (c), and after A/B minimization (e), respectively. (d) Adaptive optimization of the A/B photoemission ratio leads to increased (red) and decreased (blue) contrast of electron yields from the upper and lower regions as compared to unshaped laser pulses recorded as a reference (black). Figure reproduced with permission from Ref. [24].

spatially and temporally shape the SPP field at the surface of the nanostructure.[119] This concept has been demonstrated in a seminal paper by Aeschlimann *et al.*[24] Pulses from a Titanium:sapphire laser are sent through a polarization pulse shaper. They are then focused onto a plasmonic nanostructure that consists of three dimers of circular silver disks with 180 nm diameter (Fig. 7.6(b)). A PEEM is used to image the emission of photoelectrons from the nanodisks. The number of photoelectrons that is emitted from either the upper two dimers (region A) or the lower dimer (region B) is recorded. The signal provides a feedback signal for an evolutionary learning algorithm. Its output controls the settings of the polarization pulse shaper. By means of a genetic feedback algorithm, the polarization pulse is modified until either emission from region A or B is maximized. The PEEM image that results for optimization of region A is shown in Fig. 7.6(c). Photoemission stems almost exclusively from the upper two dimers. In contrast (Figs. 7.6(d,e)), the photoemission is mostly dominated by the lower dimer after the learning algorithm was told to optimize emission from region B. Such a polarization pulse shaping thus offers an exciting perspective for controlling local optical fields on the nanoscale. It forms the basis for a spatiotemporal control of plasmonic nanofocusing.[120,121] or — more generally — for a coherent control of the transmission of light through multiply scattering media.[122,123]

Very interesting perspectives also emerge when combining PEEM with two-color excitation. In a theoretical paper, Stockman *et al.* suggested the development of a new type of ultrafast photoelectron microscope that allows for probing the dynamics of SPPs on a plasmonic structure on nanometer length and sub-femtosecond time scales.[124] The idea is to use a strong ultrashort near-infrared driving pulse that resonantly excites the structure, and a high-energy attosecond, extreme ultraviolet (XUV) pulse to photoemit electrons from the excited structure. The XUV pulse duration is shorter than one oscillation cycle of the near-infrared plasmonic field. Photoemission occurs mainly when both the driving pulse and the XUV pulse are present. Photoelectrons emitted by the XUV pulse in the presence

of local SPP field induced by the strong NIR driving pulse will, as a result, be accelerated, and the resulting change in the kinetic energy of the photoelectrons is directly related to the vector potential of the local SPP field at the moment of photoemission. By simultaneously measuring the spatial distribution and kinetic energy of the photoemitted electrons, the spatiotemporal dynamics of local near-infrared SPP fields can be directly probed on their natural ultrashort length and time scales. Experimental tests of these ideas, which are technically rather demanding and require advanced laser technology, are currently being pursued in different laboratories.[125–129].

7.4. Ultrafast Pump-probe Spectroscopy: Exciton–Plasmon Rabi Oscillations in Hybrid Metal-molecular Nanostructures

The measurement techniques that we have discussed so far give access to the dynamics of local electric fields at or near the surface of plasmonic nanostructures. We have seen that these coherent fields are usually very short-lived. Depending on the geometry, shape and composition of the plasmonic material or nanostructure, they persist for a few to a few tens of femtoseconds, at most slightly more than 100 fs. Often, it is of interest to not only know the dynamics of these ultrafast fields but also know whether the optical excitation of a plasmonic material results in a transfer of energy (or charges/spins) to a neighboring material. Measurements of the dynamics and yield of these processes can give important insight into the underlying physical processes. The experimental method of choice for probing such transfer processes is ultrafast pump-probe spectroscopy.[130,131] In pump-probe spectroscopy, a sample is illuminated with (a sequence of) ultrafast pump-pulses. The effect of this excitation on the transmission of a time-delayed probe pulse is measured by recording a differential transmission spectrum

$$\frac{\Delta T(\lambda,\tau)}{T(\lambda)} = \frac{T_{\text{on}}(\lambda,\tau) - T_{\text{off}}(\lambda)}{T_{\text{off}}(\lambda)}. \tag{7.11}$$

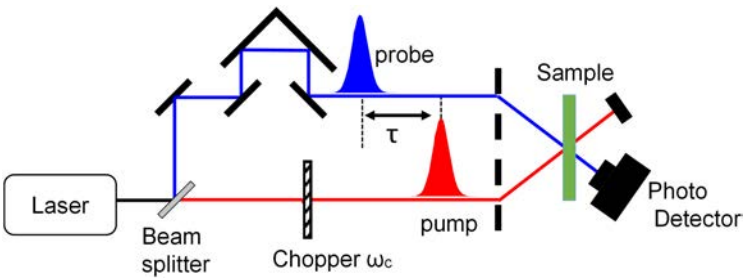


Fig. 7.7. Schematic illustration of a typical pump-probe setup. Ultrashort pulses are split into a pair of pump and probe pulses with variable time delay τ. The pump laser pulses are periodically switched on and off by a chopper, and differential transmission spectra $\Delta T/T$ are recorded.

Here, $T_{\mathrm{on}}(\lambda, \tau)$ and $T_{\mathrm{off}}(\lambda, \tau)$ denote the transmission of the probe laser through the sample in the presence and absence of the pump laser, respectively. Such pump-probe spectra are typically measured by spectrally dispersing the transmitted probe laser spectrum in a monochromator and recording spectra with a photodiode detector or CCD camera with electronics that is specially designed for fast read-out and low noise.[132] The pump laser is periodically blocked by a fast chopper (Fig. 7.7). Such pump-probe measurements can also be performed in a reflection geometry.

Importantly, the time resolution in such measurements is only limited by the duration of the pump and probe pulses. Using, e.g. non-collinear optical parametric amplifiers as spectrally tunable laser sources,[133] the time resolution can reach 10 fs or even less. This pulse duration is shorter than the period of even the fastest vibrational modes in molecules or solids. As such, the "impulsive" excitation of a material system with such short optical pulses generally induces a variety of high-frequency coherent electronic, nuclear and spin oscillations. The dynamics of those oscillations can conveniently be traced in the time domain by interrogating the optical response of the system with a time-delayed probe pulse. This can also sensitively probe incoherent relaxation pathways between different quantum states of the system.

More formally, such a pump-probe experiment measures a higher-order nonlinear material polarization $P^{(n)}$ of the system under

investigation. In the simplest case (in the limit of sufficiently weak pump- and probe pulses), the third-order polarization $P^{(3)}$ is measured. This nonlinear polarization can be written as the convolution of a third-order nonlinear response function $S^{(3)}(t_3, t_2, t_1)$ with the three interacting electric fields $E(t_i)$, $i = 1, 2, 3$. Here, $E(t_i)$ denotes the sum of all incident fields at moment t_i and

$$P^{(3)}(t) = \int_0^\infty dt_3 \int_0^\infty dt_2 \int_0^\infty dt_1 E(t - t_3)E(t - t_3 - t_2) \cdot E(t - t_3 - t_2 - t_1) \cdot S^{(3)}(t_3, t_2, t_1) \tag{7.12}$$

Hence, pump-probe spectroscopy probes specific types of nonlinear response functions, in contrast to the linear response functions that are measured with the techniques discussed in Sections 7.2–7.4. The theoretical description of these response functions is highly developed and can be found in several excellent textbooks.[131,134] These response functions are commonly used to calculate the nonlinear polarization and to simulate pump-probe spectra.

Alternatively, the nonlinear material polarization can be obtained by directly evaluating the time-evolution of the density matrix ρ of the system by solving the Liouville–von Neumann equation

$$\frac{\partial \rho}{\partial t} = -\frac{i}{\hbar}[H, \rho] + \left.\frac{\partial \rho}{\partial t}\right|_{\text{rel}}. \tag{7.13}$$

Here, H denotes the full Hamiltonian of the system including the light-matter coupling and the final term in Eq. (7.13) includes possible dephasing and relaxation phenomena. The macroscopic material polarization is then given by the expectation value of the dipole operator $\boldsymbol{\mu}$ as

$$\mathbf{P}(t) = \text{Tr}(\boldsymbol{\mu}\rho(t)), \tag{7.14}$$

where Tr denotes the trace of a matrix. The nth-order polarization $\mathbf{P}^{(n)}(t) = \text{Tr}(\boldsymbol{\mu}\rho^{(n)}(t))$ can be conveniently obtained by a perturbative analysis of the dynamics of the density matrix. For this, the density matrix is expanded in order of powers n of the incident electric field, $\rho^{(n)}(t)$.

To illustrate the vast amount of information that can be obtained from such pump–probe studies, we focus on a specific example, the coherent coupling of excitons in molecular aggregates to surface plasmon polariton fields in a plasmonic nanoresonator.[32,33,135] We will show how this coupling results in a periodic transfer of energy between plasmonic and excitonic system and how this coherent energy transfer can directly be seen in time-domain pump–probe experiments.

As an excitonic model system, we study the cyanine dye 2,2′-dimethyl-8-phenyl-5,6,5′,6′-dibenzothiacarbocyanine chloride. This dye was dissolved in a solution containing polyvinyl alcohol, water and methanol. The solution was spin-coated onto an optically thick gold film to form a 50-nm thick dye film. When deposited onto a planar, unstructured gold surface, this film shows very pronounced light absorption at the excitonic resonance of the *J*-aggregated dye film (∼700 nm, Fig. 7.8(b,c)). *J*-aggregation of the highly concentrated dye forms exciton states with excitonic wavefunctions that are delocalized over several monomer units of the dye. This exciton delocalization results in a pronounced red shift of the *J*-aggregate exciton with respect to that of the monomer. Also, the *J*-aggregated excitonic spectrum shows a very strong light absorption (absorption coefficient $\sim 2 \cdot 10^5$ cm^{-1}) and a particularly narrow linewidth. This strong coupling to light makes it an interesting candidate for exciton-plasmon coupling.[136] As can be seen in Fig. 7.8(b), the excitonic resonance of the dye film is essentially independent of the angle of incidence under which light reflection from the film is probed. This shows that the localization length of *J*-aggregate exciton is much smaller than one optical wavelength and that its effective mass is too large to induce a sizeable angle-dependence of the spectrum.

To form a plasmonic nanoresonator, a periodic array of nanometer-sized slits is milled into an optical gold film. The array period is 430 nm. The width (∼45 nm) and depth (∼30 nm) of the array was chosen to minimize radiative damping[20] of SPP excitations of the film and to increase the SPP lifetime. As is evident from the angle-dependent linear reflectivity spectra shown in Fig. 7.8(e), these

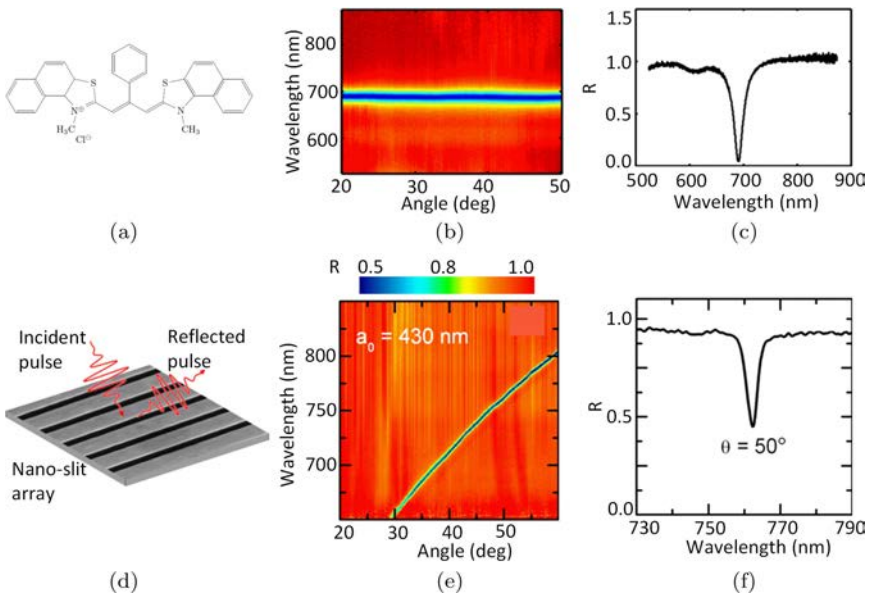


Fig. 7.8. Angle-resolved, p-polarized linear reflectivity spectra of J-aggregate dye molecules on a plasmonic nano-slit array. (a) Molecular structure of the dye. (b) Spectral reflectivity of a 50 nm-thick film of J-aggregate molecules in a polymer matrix coated onto a planar gold surface as a function of incident angle between 20° and 50°. (c) A typical single reflectivity spectrum at $\theta = 40°$. (d) Schematic of the grating structure excited by an ultrashort pulse. (e) Spectral reflectivity measurements of a grating with a period of 430 nm. (f) A typical reflectivity spectrum of the grating, taken at $\theta = 50°$.

nanoslit arrays show a very pronounced absorption dip at the SPP resonance. The linewidth of the SPP resonance is extremely narrow (∼5 nm), giving evidence for a long lifetime of the SPP mode. The steep angular dispersion of the SPP mode results in pronounced shifts of the SPP resonance with incidence angle.

These reflectivity spectra change considerably when depositing the J-aggregated dye film onto the nanoslit array (Fig. 7.9). Now, one still sees a remaining absorption at the resonance energy of the bare J-aggregate exciton. The dispersion of the strongly angle-dependent resonances, however, is drastically different from that of the SPP resonance at the bare slit grating. Near a crossing angle of ∼30 degrees,

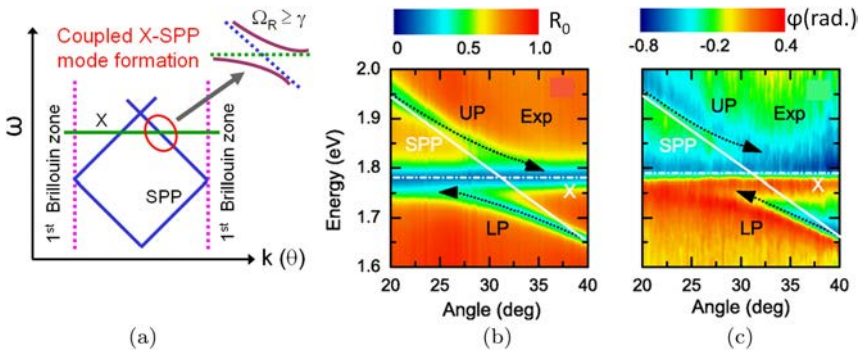


Fig. 7.9. (a) Schematic of the formation of coupled exciton–plasmon-polariton modes in a hybrid structure consisting of a gold nano-slit array and a *J*-aggregated dye film. Near the crossing of the dispersion diagram of an uncoupled dye film and a plasmonic periodic grating, exciton–plasmon coupling leads to the formation of coupled polariton modes. In case of strong coupling, the two branches split into distinctly observable upper and lower polaritons. (b) and (c) Angular-resolved reflectivity amplitude and spectral phase of the hybrid structure with a 50-nm-thick *J*-aggregated dye film. The black (white) lines in (b) and (c) mark the coupled (uncoupled) mode dispersion curves. (Adapted with permission from ACS Nano, Article ASAP, DOI: 10.1021/nn405981k. Copyright (2013) American Chemical Society.)

one sees a clear bending of the high-energy resonance (above the *J*-aggregate exciton) and the low-energy resonance. At the crossing angle, both modes are clearly energy separated by a splitting energy $\hbar\Omega_{\mathrm{NMS}}$, the normal mode splitting. In the present sample, the splitting energy is about 110 meV.

This splitting is a clear signature of a strong coupling between excitons and surface plasmon polaritons.[136–141] The local SPP field in the vicinity of the nanoslits couples to the optical transition dipole moments of those *J*-aggregate excitons that have spatial and orientational overlap with the SPP field. This results in a transfer of energy from the SPP field to the excitonic system: the SPP field induces coherent dipole oscillations in the excitonic layer. *Vice versa*, the optical near fields created by the excitonic dipole polarization induce SPP oscillations in the plasmonic structure and thus transfer energy back to the plasmonic system. In the limit of strong coupling, i.e.

if the normal mode splitting exceeds the losses of both the exciton and SPP mode, one expects coherent population oscillations, Rabi oscillations, between the excitonic and the SPP mode. The period of these Rabi oscillations should be $2\pi/\Omega_{\mathrm{NMS}}$, which, in the present sample is $\sim$30 fs.

The strength of this exciton–SPP coupling, the Rabi coupling $\Omega_R = \Omega_{\mathrm{NMS}}/2$, scales with the scalar product of the excitonic transition dipole moment and the local electric field times the square root of the number of J-aggregate excitons that are coupled to single SPP modes. In case that the energies of the uncoupled exciton and SPP modes are brought into resonance and that the Rabi coupling strength exceeds the losses of both the excitonic and the SPP system, this coupling results in the formation of new upper (UP) and lower (LP) polaritons. These polariton modes are delocalized, hybrid modes of excitons and SPPs and are the new normal modes of the coupled system. At the crossing between the exciton and SPP dispersion relations, the UP–LP splitting is approximately twice the Rabi energy $\hbar\Omega_R$. It is important that the existence and the strength of this polariton splitting do not depend on the presence of an external laser field. It may be understood as coupling of excitonic dipole moments to vacuum fluctuations of the SPP field. In the present experiments, a large number of excitons are coupled to the SPP field, and hence the optical signatures of this coupling are similar to those of two coupled classical oscillators. In this limit, the Rabi splitting is therefore often called normal mode splitting.[142] In the exciting limit of the strong coupling between a single exciton and a single SPP mode,[143–145] the vacuum Rabi splitting is expected to depend sensitively on the number of SPPs that are present in the SPP mode. The normal mode splitting between excitons and SPP fields occurs only in the regions of the sample with a sufficiently high-field amplitude of the SPP mode. In our grating structure, this is fulfilled inside or in the vicinity of the nanoslits. In between two slits, the field amplitude is much lower so that excitons in these regions will couple only weakly to the SPP field. Hence, in these regions, no normal

mode splitting appears and the excitons absorb light at the energy of the bare J-aggregate exciton resonance.

We now perform a pump–probe experiment to study the periodic energy transfer between SPPs and excitons in the time domain. For this, we excite the sample with ultrashort light pulses with duration of less than 15 fs. This pulse duration is much shorter than the expected Rabi oscillation period and, thus, short enough to temporally resolve the periodic energy transfer between excitons and SPPs. The center wavelength of these pulses is tuned to the uncoupled exciton resonance and their spectrum is sufficiently broad to cover both the upper and lower polariton resonances. The pulses are weakly focused onto the sample at an incidence angle of 39°. This angle is chosen in order to spectrally shift the LP away from the bare J-aggregate exciton. Under these excitation conditions, both bare J-aggregate exciton and UP mode overlap spectrally. Hence, the ultrafast pump pulse creates a coherent superposition of UP, LP and bare exciton states. To monitor the time dynamics of the coherent polaritons wavepackets, we probe the pump-induced change of the reflectivity of a second, somewhat weaker and time-delayed, probe pulse. This pulse is a replica of the pump laser and is also incident under the same angle as the pump laser. The probe beam that is reflected from the sample is spectrally dispersed in a monochromator, and the pump-induced change in probe reflectivity $\Delta R/R$ is monitored in a broad energy range, covering both polarition resonances (Fig. 7.10(a)). In the energy range around 1.8 eV, the differential reflectivity is mostly dominated by that of the bare J-aggregate excitons. The $\Delta R/R$ spectrum shows a dispersive lineshape, well known for such J-aggregated system, which reflects, at lower energies, the pump-induced bleaching of the J-aggregate exciton resonance. In the higher energy wing of the spectrum, one sees a pump-induced decrease in reflectivity since the probe light is absorbed by pump-induced exciton to biexciton transitions.[146] The time dynamics of the pump–probe spectra show, at negative time delays (when the probe laser arrives before the pump laser), the so-called perturbed free-induction decay of the excitonic polarization induced by the probe laser.[147,148]. This transient

persists during the dephasing time of the coherent exciton population of about 50 fs. At positive delay times, the decay of the pump–probe signal is given by the decay of the incoherent exciton population that is induced by the pump laser. In the present sample, this lifetime is rather short, about 100–200 fs.[33,140]

Very different and interesting dynamics are observed when probing the differential reflectivity near the lower-energy LP resonance at around 1.65 eV. Again, one finds a distinct dispersive lineshape of the spectra (Fig. 7.10(b)) reflecting exciton bleaching at lower energies and biexciton absorption at higher energies. When following the time dynamics of the differential reflectivity, one finds pronounced temporal oscillations of the pump–probe signal, both at positive and negative delay times. The oscillations have a period of about 26 fs, which matches quite well the expectation for the Rabi oscillation period. They persist for about 50 fs at positive and negative time delays, a time scale that is shorter than the population decay time of the excitons and reflects the dephasing time of the polariton resonances. Interestingly, the largest change in differential reflectivity occurs at around time zero, i.e. when pump and probe lasers interact simultaneously with the sample. This signal at around time zero does not stem from the coherent interference of pump and probe lasers on the sample (an effect that is often called a "coherent artifact") since care has been taken in suppressing such interference effects experimentally.[32] When simultaneously changing the incidence angle of both lasers, the time-domain oscillations sill persist (Fig. 7.10(c)) but their oscillation period is slightly changed. Measured oscillation periods for different incidence angles are shown in Fig. 7.10(f) as open triangles. The observed values match quite well with those anticipated from the dispersion relation of the coupled polariton modes (solid line in Fig. 7.10(f)).

How can we understand these time-domain oscillations and what can be learned from them about the physics of exciton–plasmon interactions? For this, it is important that the scattering cross-section for SPP excitation by the pump laser is much larger than the exciton absorption cross-section. Hence, the impulsive excitation of the

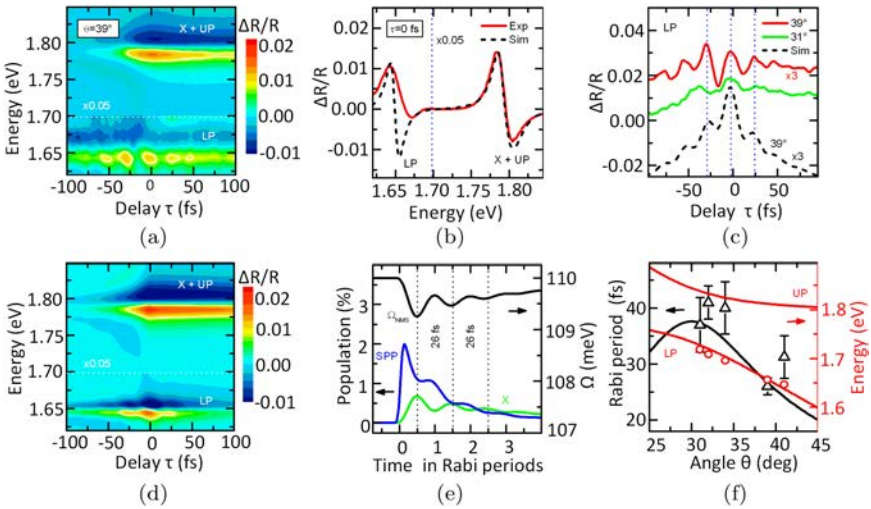


Fig. 7.10. Coherent dynamics of the hybrid exciton/SPP system. (a) Measured differential reflectivity map for a hybrid structure (J-aggregated dye on a grating with period 430 nm), recorded at an angle of incidence of 39 deg. Clear temporal Rabi oscillations are seen near the LP resonance at 1.64 eV. (b) Comparison of measured (solid line) and simulated (dashed line) differential reflectivity spectra at zero delay. (c) Time evolution of the differential reflectivity signal near the LP resonance, measured at $\theta = 39$ deg and $\theta = 31$ deg. Pronounced sub-40fs Rabi oscillations are noticeable. The shorter oscillation period for the larger angle of incidence reflects the increased detuning of the hybrid exciton/SPP system. The dashed line (vertically displaced by -0.025) represents simulated dynamics for an incidence angle of 39 deg. (d) Simulated differential reflectivity map. (e) Pump-induced SPP and exciton population dynamics at $\theta = 39$ deg. (f) Comparison of calculated (solid lines) and observed (open symbols) oscillation periods and LP resonance energies as functions of θ. The simulated dispersion relations are included. (Reprinted from Nature Photonics 7, 128132 (2013).)

sample by the pump laser mainly launches SPP excitations in the nanoslit, while initially the probability to create excitons is rather low. This is visualized by the blue and green curves in Fig. 7.10(e), showing the time evolution of the SPP population and the exciton population, respectively. These curves are simulated on the basis of a quantum-mechanical density matrix model for the dynamics of the hybrid exciton/SPP system.[32] This shows that the nanoslit grating essentially acts as an antenna that transforms the incident light into

SPPs, whereas the direct coupling of the incident far-field light to the J-aggregate excitons is so weak that it can basically be neglected. One sees that after impulsive SPP excitation, the exciton population is created within half a Rabi period. Concomitantly, the SPP population decays rapidly. This is followed by periodic, out-of-phase oscillations of both populations with the Rabi period. These oscillations, which reflect the periodic exchange of the energy between the coupled SPP and exciton modes, are then damped on a timescale given by the lifetime of the coupled polariton modes.

This nicely explains the appearance of Rabi oscillations in hybrid exciton–plasmon system. Yet, it cannot explain the time-dependent differential reflectivity that is seen in Fig. 7.10(a). For this, it is important to understand that the optical nonlinearities of excitons and SPPs are fundamentally different. J-aggregate excitons have very pronounced optical nonlinearities, and their optical excitation results in a bleaching of the exciton absorption and in the appearance of a transient biexciton absorption that is slightly blue-shifted with respect to the fundamental exciton resonance. These strong excitonic nonlinearities can account for the differential reflectivity near 1.80 eV, where the contribution from the UP resonance to the nonlinearity is weak. In contrast, surface polariton excitations in metallic nanostructures can essentially be understood as photonic modes with very weak optical nonlinearities. To first order, their optical properties are well-described by that of a linear harmonic oscillator. In this (somewhat simplified) picture, the optical excitation of an SPP mode should not lead to a differential reflectivity signal. Based on this very simple model for the optical nonlinearities of the hybrid system (bleaching of the excitonic mode but vanishing nonlinearity of the SPP mode), one would expect the following differential reflectivity dynamics. At time zero, SPP modes are excited (Fig. 7.10(e)). This excitation does not instantaneously induce nonlinearities, and hence the ΔR at time zero should be weak. After half a Rabi cycle, energy is transferred from the SPP mode to the excitonic system. This now results in a strong ΔR at this time delay, which decays again when the energy is transferred back to the SPP mode. Similar dynamics are

expected at negative time delays since both pump and probe laser can launch SPP modes. Density matrix simulations support these expectations.[32]

Yet, this is in striking contrast with the experimental data which in fact show a maximum in ΔR at time zero and minima in ΔR at half the Rabi period. This apparent discrepancy can be resolved by taking into account a second, and much less understood, optical nonlinearity of the coupled exciton–SPP system.[140,149,150] We have discussed that, in the present system, the SPP mode is coupled to a large number of exciton states and that the normal mode splitting scales with the square root of this number.[136] The optical excitation of a fraction n_X of these exciton states by pump-induced transitions from the exciton ground to the excited state will transiently alter this normal mode splitting. Those excitons that have been optically excited will no longer be able to absorb light from the SPP field. Instead, their coupling to the SPP field will either induce their transition to higher lying biexciton states or will give rise to the stimulated emission of SPP fields.[151] This will, therefore, transiently reduce the normal mode splitting $\Omega_{\mathrm{NMS}}(t) = \Omega_{\mathrm{NMS}}^{(0)}\sqrt{1-2n_x(t)}$ by a factor that depends on the fraction of optically created excitons. Here, $\Omega_{\mathrm{NMS}}^{(0)}$ is the normal mode splitting in the absence of an optical excitation of the system.

It turns out that this insight is the key to understanding the unexpected ΔR dynamics in Fig. 7.10(a). When taking such a transient reduction of the normal mode splitting into account in our density matrix simulations, we can nicely and almost quantitatively reproduce the experimental results (Fig. 7.10(d)). It appears that the effective reduction in normal mode splitting — and not the excitonic bleaching — is the dominant optical nonlinearity of the hybrid system. The simulations suggest that the change in normal splitting that is induced by the pump and seen by the probe laser is largest when the probe is time-delayed by an integer multiple of the Rabi period. Hence, the nonlinearity is maximum at time zero, satisfactorily explaining the experimental observation.

This example shows that nonlinear pump–probe spectroscopy, in particular when being performed with ultrashort optical pulses with a duration that is much shorter than the femtosecond lifetime of short-lived plasmonic modes, provides exquisitely detailed insight into the elementary energy transfer processes in plasmonic nanostructures that are coupled to other types of quantum emitters. It not only reveals ultrafast Rabi oscillations as a sign of a periodic energy exchange between plasmons and (molecular) excitons but also gives direct access to the physical mechanisms that govern the optical nonlinearities of such hybrid systems. In the present example, we have learned that the dominant optical nonlinearity is not the bleaching of the excitonic absorption or a pump-induced excited state absorption, as might have been expected on the basis of the optical nonlinearities of the uncoupled systems. Instead, it appears that the photo-induced generation of excitons results in a transient reduction of the normal mode splitting in the system. Exciton creation reduces the number of final states that can absorb the electromagnetic energy that is stored in surface plasmon polaritons. This mechanism provides a new ultrafast and highly efficient switch for evanescent SPP fields. Interestingly, the efficiency of this switch increases the smaller the dimensions of the hybrid exciton/plasmon nanoresonator get. In the fundamental limit of strong coupling between a single exciton and single SPP mode,[152] already the absorption of a single photon by the exciton would suffice to fundamentally alter the optical properties of the coupled exciton/plasmon nanoresonator. This opens up the exciting possibility to realize ultrafast, all-optical and highly efficient plasmonic nanoswitches, transistors or gates that might even consume less power than current electronic transistors, yet operate on ultrafast, femtosecond times scale that are only accessible by optical means. Quantum optics dictates that the optical nonlinearities in this ultimate limit of quantum plasmonics should be governed by the Jaynes–Cummings model[153,154] describing the coupling between a single photonic mode and a single "two-level system". It appears extraordinarily interesting to explore this new regime of light-matter

interaction. Ultrafast pump–probe spectroscopy should provide the necessary tools to explore those avenues.

Much more generally, advanced ultrafast pump–probe spectroscopy provides all the necessary tools for studying the dynamics of ultrafast charge, spin and energy transfer processes in various types of nanosystems. This is important, for instance, for understanding the effects of hot electron excitations on the catalytic activity of nanoparticles[155] or for unraveling the disputed role of coherent vibronic couplings on light-to-current conversion in natural or artificial light harvesting systems.[156–159] This holds in particular when using advanced ultrafast spectroscopic tools such as 2D[160,161] or 3D[162] electronic spectroscopy. Yet, even when using such advanced spectroscopic techniques, an outstanding scientific challenge remains: Ideally, we would love to probe the coherent dynamics of such energy-, charge- and spin transfer processes not only on femtosecond time scales but at the level of a single nanostructure. Only true single nanoparticle studies will enable us to unravel the effects of structural nanoscale inhomogeneities on the dynamics of optical excitations in hybrid and/or plasmonic nanosystems. Such ultrafast single nanoparticle studies are experimentally challenging and have so far only been performed in a limited number of favorable cases.[29,72,163–167] As such, the combination of ultrafast coherent pump–probe spectroscopy and single nanoparticle spectroscopy remains an outstanding scientific challenge.

7.5. Summary and Conclusions

In summary, this chapter provides an overview about different time-resolved spectroscopic techniques to study the dynamics of optical excitations in plasmonic nanosystems. Since plasmonic excitations are inherently short-lived, such experiments should be performed with femtosecond time resolution to provide direct insight into those dynamics. Starting from well-established FTS and SI, we have shown how time-domain spectroscopy can probe the linear optical response of a single nanoparticle. It is explained how SI

provides a direct visualization of the effects of SPP propagation on the temporal shape of an ultrafast SPP pulse. Ultrafast nonlinear optical spectroscopic techniques such as IFRAC and photoemission electron microscopy are introduced. We have discussed how these methods — if combined with sufficiently short, pulsed laser excitation — give direct access to the time structure of the local electric field at the surface of a plasmonic nanostructure. Exciting perspectives emerge when combining such nonlinear spectroscopic methods with advanced pulse shaping techniques as this allows for coherently controlling SPP fields in nanoscale dimensions and on ultrashort time scales. Finally, the last section of this chapter discusses a specific example of how ultrafast pump-probe spectroscopy can be used to track the periodic exchange of energy between a plasmonic nanoresonator and a thin film of molecular excitons. This provides fundamentally new insight into the exciting regime of strong coupling between plasmons and excitons and indicates strategies of how such hybrid nanostructures can be used to design new classes of ultrafast all-optical nanophotonic switches. To make full use of these interesting new hybrid devices, it seems important to explore the dynamics of strong coupling in hybrid nanostructures that couple a few — or possibly even only one — quantum emitter with a plasmonic cavity mode. This poses quite significant experimental challenges but opens up the exciting perspective to develop new types of nanophotonic structures that truly merge the best of the photonics and electronics world. We conclude by mentioning that yet another important and highly exciting research field has emerged from the combination of ultrafast optics and plasmonics over the last few years: strong-field plasmonics. It employs the strong electric field of an ultrashort light pulse to coherently control the photoemission of electrons from a plasmonic nanostructure and their motion in the plasmonic near-field. A recent overview over the field may be found, e.g. in a collection of research articles in Ref. [129]. Strong-field plasmonics may be a key enabler for solid-state attosecond physics and opens the door to new classes of ultrafast electron microscopes.

Acknowledgments

We thank all present and former members of the Ultrafast Nano-Optics group for their contributions to this research. We acknowledge the financial support by the Deutsche Forschungsgemeinschaft (SPP1391, SPP1839, SPP1840, GRK1885), by the EU (Cronos), GIF (Grant no. 1256), and by GRL. J.-M. Y. acknowledges a personal fellowship by the Hanse-Wissenschaftskolleg.

References

1. Ritchie, R. H. (1957). Plasma losses by fast electrons in thin films, *Phys. Rev.* **106**, pp. 874–881.
2. Ritchie, R. H. and Eldridge, H. B. (1962). Optical emission from irradiated foils, 1, *Phys. Rev.* **126**, pp. 1935–1947.
3. Atwater, H. A. (2007). The promise of plasmonics, *Scientific American* **296**, pp. 56–62.
4. Ozbay E. (2006). Plasmonics: Merging photonics and electronics at nanoscale dimensions, *Science* **311**, pp. 189–193.
5. Schuller, J. A., Barnard, E. S., Cai, W. S., Jun, Y. C., White, J. S., and Brongersma, M. L. (2010). Plasmonics for extreme light concentration and manipulation, *Nat. Mater.* **9**, pp. 193–204.
6. Stockman, M. I. (2011). Nanoplasmonics: Past, present, and glimpse into future, *Opt. Express* **19**, pp. 22029–22106.
7. Groß, P., Esmann, M., Becker, S. F., Vogelsang, J., Talebi, N., and Lienau, C. (2016). Plasmonic nanofocusing — grey holes for light, *Adv. Phys.: X* **1**, pp. 297–330.
8. Hillenbrand, R., Taubner, T., and Keilmann F. (2002). Phonon-enhanced light-matter interaction at the nanometre scale, *Nature* **418**, pp. 159–162.
9. Zhang, R., *et al.* (2013). Chemical mapping of a single molecule by plasmon-enhanced Raman scattering, *Nature* **498**, pp. 82–86.
10. Zhang, Y., *et al.* (2016). Visualizing coherent intermolecular dipole–dipole coupling in real space, *Nature* **531**, pp. 623–627.
11. Hartschuh, A., Sanchez, E. J., Xie, X. S., and Novotny L. (2003). High-resolution near-field Raman microscopy of single-walled carbon nanotubes, *Phys. Rev. Lett.* **90**, p. 0955503.
12. Benz, F., *et al.* (2016). Single-molecule optomechanics in "picocavities", *Science* **354**, pp. 726–729.
13. Stöckle, R. M., Suh, Y. D., Deckert, V., and Zenobi R. (2000). Nanoscale chemical analysis by tip-enhanced Raman spectroscopy, *Chem. Phys. Lett.* **318**, pp. 131–136.

14. Brongersma, M. L., Halas, N. J., and Nordlander, P. (2015). Plasmon-induced hot carrier science and technology, *Nat. Nanotechnol* **10**, pp. 25–34.
15. Atwater, H. A. and Polman, A. (2010). Plasmonics for improved photovoltaic devices, *Nat. Mater.* **9**, pp. 205–213.
16. Hirsch, L. R., *et al.* (2003). Nanoshell-mediated near-infrared thermal therapy of tumors under magnetic resonance guidance, *Proc. Natl. Acad. Sci. USA* **100**, pp. 13549–13554.
17. Engheta, N. (2007). Circuits with light at nanoscales: Optical nanocircuits inspired by metamaterials, *Science* **317**, pp. 1698–1702.
18. Khurgin, J. B. (2015). How to deal with the loss in plasmonics and metamaterials, *Nat. Nanotechnol.* **10**, pp. 2–6.
19. Boyer, D., Tamarat, P., Maali, A., Lounis, B., and Orrit M. (2002). Photothermal imaging of nanometer-sized metal particles among scatterers, *Science* **297**, pp. 1160–1163.
20. Kim, D. S., *et al.* (2003). Microscopic origin of surface-plasmon radiation in plasmonic band-gap nanostructures, *Phys. Rev. Lett.* **91**, p. 143901.
21. Kubo, A., Onda, K., Petek, H., Sun, Z. J., Jung, Y. S., and Kim, H. K. (2005). Femtosecond imaging of surface plasmon dynamics in a nanostructured silver film, *Nano. Lett.* **5**, pp. 1123–1127.
22. Kubo, A., Pontius, N., and Petek, H. (2007). Femtosecond microscopy of surface plasmon polariton wave packet evolution at the silver/vacuum interface, *Nano. Lett.* **7**, pp. 470–475.
23. Yi, J.-M., *et al.* (2017). Probing coherent surface plasmon polariton propagation using ultrabroadband spectral interferometry, *ACS Photonics* **4**, pp. 347–354.
24. Aeschlimann, M., *et al.* (2007). Adaptive subwavelength control of nano-optical fields, *Nature* **446**, pp. 301–304.
25. Aeschlimann, M., *et al.* (2010). Spatiotemporal control of nanooptical excitations, *Proc. Natl. Acad. Sci. USA* **107**, pp. 5329–5333.
26. Hanke, T., Krauss, G., Trautlein, D., Wild, B., Bratschitsch, R., and Leitenstorfer, A. (2009). Efficient nonlinear light emission of single gold optical antennas driven by few-cycle near-infrared pulses, *Phys. Rev. Lett.* **103**, p. 257404.
27. Anderson, A., Deryckx, K. S., Xu, X. J. G., Steinmeyer, G., and Raschke, M. B. (2010). Few-femtosecond plasmon dephasing of a single metallic nanostructure from optical response function reconstruction by interferometric frequency resolved optical gating, *Nano. Lett.* **10**, pp. 2519–2524.
28. Kollmann, H., *et al.* (2014). Toward plasmonics with nanometer precision: nonlinear optics of helium-ion milled gold nanoantennas, *Nano. Lett.* **14**, pp. 4778–4784.
29. Kravtsov, V., Ulbricht, R., Atkin, J., and Raschke, M. B. (2016). Plasmonic nanofocused four-wave mixing for femtosecond near-field imaging, *Nat. Nanotechnol.* **11**, pp. 459–464.

30. Biagioni, P., *et al.* (2012). Dynamics of four-photon photoluminescence in gold nanoantennas, *Nano. Lett.* **12**, pp. 2941–2947.
31. Schmidt, S., *et al.* (2012). Adiabatic nanofocusing on ultrasmooth single-crystalline gold tapers creates a 10-nm-sized light source with few-cycle time resolution, *ACS Nano* **6**, pp. 6040–6048.
32. Vasa, P., *et al.* (2013). Real-time observation of ultrafast Rabi oscillations between excitons and plasmons in metal nanostructures with J-aggregates, *Nat. Photonics* **7**, pp. 128–132.
33. Vasa, P., *et al.* (2015). Optical stark effects in j-aggregate-metal hybrid nanostructures exhibiting a strong exciton-surface-plasmon-polariton interaction, *Phys. Rev. Lett.* **114**, p. 036802.
34. Rybka, T., Ludwig, M., Schmalz, M. F., Knittel, V., and Brida, D., Leitenstorfer A. (2016). Sub-cycle optical phase control of nanotunnelling in the single-electron regime, *Nat. Photonics* **10**, pp. 667–670.
35. Dombi, P., *et al.* (2013). Ultrafast strong-field photoemission from plasmonic nanoparticles, *Nano Lett.* **13**, pp. 674–678.
36. Hommelhoff, P., Sortais, Y., Aghajani-Talesh, A., and Kasevich, M. A. (2006). Field emission tip as a nanometer source of free electron femtosecond pulses, *Phys. Rev. Lett.* **96**, p. 077401.
37. Ropers, C., Solli, D. R., Schulz, C. P., Lienau, C., and Elsaesser, T. (2007). Localized multiphoton emission of femtosecond electron pulses from metal nanotips, *Phys. Rev. Lett.* **98**, p. 043907.
38. Herink, G., Solli, D. R., Gulde, M., and Ropers, C. (2012). Field-driven photoemission from nanostructures quenches the quiver motion, *Nature* **483**, pp. 190–193.
39. Schenk, M., Krüger, M., and Hommelhoff, P. (2010). Strong-field above-threshold photoemission from sharp metal tips, *Phys. Rev. Lett.* **105**, p. 257601.
40. Krüger, M., Schenk, M., Förster, M., and Hommelhoff, P. (2012). Attosecond physics in photoemission from a metal nanotip, *J. Phys. B At. Mol. Opt.* **45**, p. 074006.
41. Bormann, R., Gulde, M., Weismann, A., Yalunin, S. V., and Ropers, C. (2010). Tip-enhanced strong-field photoemission, *Phys. Rev. Lett.* **105**, p. 147601.
42. Park, D. J., Piglosiewicz, B., Schmidt, S., Kollmann, H., Mascheck, M., and Lienau, C. (2012). Strong field acceleration and steering of ultrafast electron pulses from a sharp metallic nanotip, *Phys. Rev. Lett.* **109**, p. 244803.
43. Piglosiewicz, B., *et al.* (2014). Carrier-envelope phase effects on the strong-field photoemission of electrons from metallic nanostructures, *Nat. Photonics* **8**, pp. 37–42.
44. Feist, A., Echternkamp, K. E., Schauss, J., Yalunin, S. V., Schäfer, S., and Ropers, C. (2015). Quantum coherent optical phase modulation in an ultrafast transmission electron microscope, *Nature* **521**, pp. 200–203.

45. Vasa, P., Ropers, C., Pomraenke, R., and Lienau, C. (2009). Ultra-fast nano-optics, *Laser Photon. Rev.* **3**, pp. 483–507.
46. Bell, R. J. (1972). *Introductory Fourier Transform Spectroscopy* (Academic Press, USA).
47. Huth, F., Govyadinov, A., Amarie, S., Nuansing, W., Keilmann, F., and Hilenbrand, R. (2012). Nano-FTIR absorption spectroscopy of molecular fingerprints at 20 nm spatial resolution, *Nano. Lett.* **12**, pp. 3973–3978.
48. Amenabar, I., *et al.* (2013). Structural analysis and mapping of individual protein complexes by infrared nanospectroscopy, *Nat. Commun.* **4**, p. 3890.
49. Bechtel, H. A., Muller, E. A., Olmon, R. L., Martin, M. C., and Raschke, M. B. (2014). Ultrabroadband infrared nanospectroscopic imaging, *Proc. Natl. Acad. Sci. USA* **111**, pp. 7191–7196.
50. Jones, A. C. and Raschke, M. B. (2012). Thermal infrared near-field spectroscopy, *Nano. Lett.* **12**, pp. 1475–1481.
51. Sonnichsen, C., *et al.* (2002). Drastic reduction of plasmon damping in gold nanorods, *Phys. Rev. Lett.* **88**, p. 077402.
52. Muskens, O. L., *et al.* (2008). Quantitative absorption spectroscopy of a single gold nanorod, *J. Phys. Chem. C* **112**, pp. 8917–8921.
53. Crut, A., Maioli, P., Del Fatti, N., and Vallee, F. (2014). Optical absorption and scattering spectroscopies of single nano-objects, *Chem. Soc. Rev.* **43**, pp. 3921–3956.
54. Piasecki, J., Colombeau, B., Vampouille, M., Froehly, C., and Arnaud, J. A. (1980). New method for measuring impulse response of optical fibers, *Appl. Opt.* **19**, pp. 3749–3755.
55. Reynaud, F., Salin, F., and Barthelemy, A. (1989). Measurement of phase-shifts introduced by nonlinear optical phenomena on subpicosecond pulses, *Opt. Lett.* **14**, pp. 275–277.
56. Lepetit, L., Cheriaux, G., and Joffre, M. (1995). Linear techniques of phase measurement by femtosecond spectral interferometry for applications in spectroscopy, *J. Opt. Soc. Am. B.* **12**, pp. 2467–2474.
57. Dorrer, C., Belabas, N., Likforman, J. P., and Joffre, M. (2000). Spectral resolution and sampling issues in Fourier-transform spectral interferometry, *J. Opt. Soc. Am. B.* **17**, pp. 1795–1802.
58. Rewitz, C., *et al.* (2012). Ultrafast plasmon propagation in nanowires characterized by far-field spectral interferometry, *Nano Lett.* **12**, pp. 45–49.
59. Rewitz, C., *et al.* (2014). Coherent control of plasmon propagation in a nanocircuit, *Phys. Rev. Applied* **1**, p. 014007.
60. Temnov, V. V., Woggon, U., Dintinger, J., Devaux, E., and Ebbesen, T. W. (2007). Surface plasmon interferometry: measuring group velocity of surface plasmons, *Opt. Lett.* **32**, pp. 1235–1237.
61. Tragardh, J. and Gersen, H. (2013). Combining near-field scanning optical microscopy with spectral interferometry for local characterization of the optical electric field in photonic structures, *Opt. Express* **21**, pp. 16629–16638.

62. Bowlan, P., Fuchs, U., Trebino, R., and Zeitner, U. D. (2008). Measuring the spatiotemporal electric field of tightly focused ultrashort pulses with sub-micron spatial resolution, *Opt. Express* **16**, pp. 13663–13675.
63. Ropers, C., *et al.* (2005). Femtosecond light transmission and subradiant damping in plasmonic crystals, *Phys. Rev. Lett.* **94**, p. 113901.
64. Sarrazin, M., Vigneron, J. P., and Vigoureux, J. M. (2003). Role of wood anomalies in optical properties of thin metallic films with a bidimensional array of subwavelength holes, *Phys. Rev. B* **67**, p. 085415.
65. Genet, C., van Exter, M. P., and Woerdman, J. P. (2003). Fano-type interpretation of red shifts and red tails in hole array transmission spectra, *Opt. Commun.* **225**, pp. 331–336.
66. Ebbesen, T. W., Lezec, H. J., Ghaemi, H. F., Thio, T., and Wolff, P. A. (1998). Extraordinary optical transmission through sub-wavelength hole arrays, *Nature* **391**, pp. 667–669.
67. Luk'yanchuk, B., *et al.* (2010). The Fano resonance in plasmonic nanostructures and metamaterials, *Nat. Mater.* **9**, pp. 707–715.
68. Ropers, C., *et al.* (2006). Ultrafast dynamics of surface plasmon polaritons in plasmonic metamaterials, *Appl. Phys. B-Las. Opt.* **84**, pp. 183–189.
69. Lamprecht, B., Krenn, J. R., Leitner, A., and Aussenegg, F. R. (1999). Resonant and off-resonant light-driven plasmons in metal nanoparticles studied by femtosecond-resolution third-harmonic generation, *Phys. Rev. Lett.* **83**, 4421–4424.
70. Schumacher, T., Kratzer, K., Molnar, D., Hentschel, M., Giessen, H., and Lippitz, M. (2011). Nanoantenna-enhanced ultrafast nonlinear spectroscopy of a single gold nanoparticle, *Nat. Commun.* **2**, p. 333.
71. Metzger, B., Hentschel, M., and Giessen, H. (2016). Ultrafast nonlinear plasmonic spectroscopy: from dipole nanoantennas to complex hybrid plasmonic structures, *ACS Photonics* **3**, pp. 1336–1350.
72. Sun, Q., Yu, H., Ueno, K., Kubo, A., Matsuo, Y., and Misawa, H. (2016). Dissecting the few-femtosecond dephasing time of dipole and quadrupole modes in gold nanoparticles using polarized photoemission electron microscopy, *ACS Nano* **10**, pp. 3835–3842.
73. Dabrowski, M., Dai, Y. N., Argondizzo, A., Zou, Q., Cui, X. F., and Petek, H. (2016). Multiphoton photoemission microscopy of high-order plasmonic resonances at the Ag/Vacuum and Ag/Si interfaces of epitaxial silver nanowires, *ACS Photonics* **3**, pp. 1704–1713.
74. Razinskas, G., *et al.* (2016). Normal-incidence PEEM imaging of propagating modes in a plasmonic nanocircuit, *Nano Lett.* **16**, pp. 6832–6837.
75. Amat-Roldán, I., Cormack, I., Loza-Alvarez, P., Gualda, E., and Artigas, D. (2004). Ultrashort pulse characterisation with SHG collinear-FROG, *Opt. Express* **12**, pp. 1169–1178.
76. Stibenz, G. and Steinmeyer, G. (2005). Interferometric frequency-resolved optical gating, *Opt. Express* **13**, pp. 2617–2626.

77. Stibenz, G., *et al.* (2006). Advanced methods for the characterization of few-cycle light pulses: a comparison, *Appl. Phys. B* **83**, pp. 511–519.
78. Mücke, O. D., Tritschler, T., Wegener, M., Morgner, U., and Kärtner, F. X. (2002). Determining the carrier-envelope offset frequency of 5-fs pulses with extreme nonlinear optics in ZnO, *Opt. Lett.* **27**, pp. 2127–2129.
79. Boyd, R. W. (2003). *Nonlinear Optics* (Academic Press, USA).
80. Trebino, R. and Kane, D. J. (1993). Using phase retrieval to measure the intensity and phase of ultrashort pulses — frequency-resolved optical gating, *J. Opt. Soc. Am. A-Opt. Image Sci. Vis.* **10**, pp. 1101–1111.
81. Kane, D. J. and Trebino, R. (1993). Characterization of arbitrary femtosecond pulses using frequency-resolved optical gating, *IEEE J. Quantum Elect.* **29**, pp. 571–579.
82. Trebino, R. (2000). *Frequency-Resovled Optical Gating: The Measurement of Ultrashort Laser Pulses*. (Kluwer Academic, The Netherlands).
83. Stibenz, G. and Steinmeyer, G. (2006). Structures of interferometric frequency-resolved optical gating, *IEEE J. Sel. Top Quantum Elect.* **12**, pp. 286–296.
84. Mücke, O. D., Tritschler, T., Wegener, M., Morgner, U., and Kärtner, F. X. (2001). Signatures of carrier-wave Rabi flopping in GaAs, *Phys. Rev. Lett.* **87**, p. 057401.
85. Tritschler, T., Mücke, O. D., Morgner, U., Kärtner, F. X., and Wegener, M. (2003). Camouflage third-harmonic generation in the nonperturbative few-cycle regime, *Phys. Stat. Sol. B-Bas. Res.* **238**, pp. 561–567.
86. Schmidt, S., *et al.* (2010). Distinguishing between ultrafast optical harmonic generation and multi-photon-induced luminescence from ZnO thin films by frequency-resolved interferometric autocorrelation microscopy, *Opt. Express* **18**, pp. 25016–25028.
87. Mascheck, M., *et al.* (2012). Observing the localization of light in space and time by ultrafast second-harmonic microscopy, *Nat. Photonics* **6**, pp. 293–298.
88. Ropers, C., Neacsu, C. C., Elsaesser, T., Albrecht, M., Raschke, M. B., and Lienau, C. (2007). Grating-coupling of surface plasmons onto metallic tips: A nanoconfined light source, *Nano Lett.* **7**, pp. 2784–2788.
89. Vogelsang, J., *et al.* (2015). Ultrafast electron emission from a sharp metal nanotaper driven by adiabatic nanofocusing of surface plasmons, *Nano Lett.* **15**, pp. 4685–4691.
90. Schröder, B., Sivis, M., Bormann, R., Schäfer, S., and Ropers, C. (2015). An ultrafast nanotip electron gun triggered by grating-coupled surface plasmons, *Appl. Phys. Lett.* **107**, p. 231005.
91. Müller, M., Kravtsov, V., Paarmann, A., Raschke, M. B., and Ernstorfer, R. (2016). Nanofocused plasmon-driven sub-10 fs electron point source, *ACS Photonics* **3**, pp. 611–619.
92. De Angelis, F., *et al.* (2011). Breaking the diffusion limit with super-hydrophobic delivery of molecules to plasmonic nanofocusing SERS structures, *Nat. Photonics* **5**, pp. 682–687.

93. Sadiq, D., Shirdel, J., Lee, J. S., Selishcheva, E., Park, N., and Lienau, C. (2011). Adiabatic nanofocusing scattering-type optical nanoscopy of individual gold nanoparticles, *Nano Lett.* **11**, pp. 1609–1613.
94. Neacsu, C. C., Berweger, S., Olmon, R. L., Saraf, L. V., Ropers, C., and Raschke, M. B. (2010). Near-field localization in plasmonic superfocusing: a nanoemitter on a tip, *Nano. Lett.* **10**, pp. 592–596.
95. Umakoshi, T., Saito, Y., and Verma, P. (2016). Highly efficient plasmonic tip design for plasmon nanofocusing in near-field optical microscopy, *Nanoscale* **8**, pp. 5634–5640.
96. Lindquist, N. C., Nagpal, P., Lesuffleur, A., Norris, D. J., and Oh, S. H. (2010). Three-dimensional plasmonic nanofocusing, *Nano Lett.* **10**, pp. 1369–1373.
97. Sondergaard, T., *et al.* (2012). Plasmonic black gold by adiabatic nanofocusing and absorption of light in ultra-sharp convex grooves, *Nat. Commun.* **87**, p. 969.
98. Babadjanyan, A. J., Margaryan, N. L., and Nerkararyan, K. V. (2000). Superfocusing of surface polaritons in the conical structure, *J. Appl. Phys.* **87**, pp. 3785–3788.
99. Janunts, N. A., Baghdasaryan, K. S., Nerkararyan, K. V., and Hecht, B. (2005). Excitation and superfocusing of surface plasmon polaritons on a silver-coated optical fiber tip, *Opt. Commun.* **253**, pp. 118–124.
100. Stockman, M. I. (2004). Nanofocusing of optical energy in tapered plasmonic waveguides, *Phys. Rev. Lett.* **93**, p. 137404.
101. Stockman, M. I., Bergman, D. J., Anceau, C., Brasselet, S., and Zyss, J. (2004). Enhanced second-harmonic generation by metal surfaces with nanoscale roughness: Nanoscale dephasing, depolarization, and correlations, *Phys. Rev. Lett.* **92**, p. 057402.
102. Durach, M., Rusina, A., and Stockman, M. I. (2007). Toward full spatiotemporal control on the nanoscale, *Nano Lett.* **7**, pp. 3145–3149.
103. Berweger, S., Atkin, J. M., Xu, X. J. G., Olmon, R. L., and Raschke, M. B. (2011). Femtosecond nanofocusing with full optical waveform control, *Nano Lett.* **11**, pp. 4309–4313.
104. Kravtsov, V., Atkin, J. M., and Raschke, M. B. (2013). Group delay and dispersion in adiabatic plasmonic nanofocusing, *Opt. Lett.* **38**, pp. 1322–1324.
105. Piglosiewicz, B., *et al.* (2011). Ultrasmall bullets of light-focusing few-cycle light pulses to the diffraction limit, *Opt. Express* **19**, pp. 14451–14463.
106. Talebi, N., *et al.* (2015). Excitation of mesoscopic plasmonic tapers by relativistic electrons: Phase matching versus eigenmode resonances, *ACS Nano* **9**, pp. 7641–7648.
107. Schröder, B., *et al.* (2015). Real-space imaging of nanotip plasmons using electron energy loss spectroscopy, *Phys. Rev. B* **92**, p. 085411.
108. Guo, S. R., *et al.* (2016). Reflection and phase matching in plasmonic gold tapers, *Nano Lett.* **16**, pp. 6137–6144.

109. Jang, M. S. and Atwater, H. (2011). Plasmonic rainbow trapping structures for light localization and spectrum splitting, *Phys. Rev. Lett.* **107**, p. 207401.
110. Tsakmakidis, K. L., Boardman, A. D., and Hess, O. (2007). 'Trapped rainbow' storage of light in metamaterials, *Nature* **450**, pp. 397–401.
111. Walmsley, I. A. and Wong, V. (1996). Characterization of the electric field of ultrashort optical pulses, *J. Opt. Soc. Am. B* **13**, pp. 2453–2463.
112. Bauer, E. (1994). Low energy electron microscopy, *Rep. Prog. Phys.* **57**, pp. 895–938.
113. Schmidt, O., *et al.* (2002). Time-resolved two photon photoemission electron microscopy, *Appl. Phys. B-Lasers and Opt.* **74**, pp. 223–227.
114. Lemke, C., *et al.* (2012). Mapping surface plasmon polariton propagation via counter-propagating light pulses, *Opt. Express* **20**, pp. 12877–12884.
115. Chelaru, L. I. and Heringdorf, F. (2007). In situ monitoring of surface plasmons in single-crystalline Ag-nanowires, *Surf. Sci.* **601**, pp. 4541–4545.
116. Marsell, E., *et al.* (2015). Nanoscale imaging of local few-femtosecond near-field dynamics within a single plasmonic nanoantenna, *Nano Lett.* **15**, pp. 6601–6608.
117. Weiner, A. M. (2000). Femtosecond pulse shaping using spatial light modulators, *Rev. Sci. Instrum.* **71**, pp. 1929–1960.
118. Brixner, T. and Gerber, G. (2001). Femtosecond polarization pulse shaping, *Opt. Lett.* **26**, pp. 557–559.
119. Brixner, T., de Abajo, F. J. G., Schneider, J., and Pfeiffer, W. (2005). Nanoscopic ultrafast space-time-resolved spectroscopy, *Phys. Rev. Lett.* **95**, p. 093901.
120. Gjonaj, B., Aulbach, J., Johnson, P. M., Mosk, A. P., Kuipers, L., and Lagendijk, A. (2011). Active spatial control of plasmonic fields, *Nat. Photonics* **5**, pp. 360–363.
121. Ertsgaard, C. T., McKoskey, R. M., Rich, I. S., and Lindquist, N. C. (2014). Dynamic placement of plasmonic hotspots for super-resolution surface-enhanced Raman scattering, *ACS Nano* **8**, pp. 10941–10946.
122. Mosk, A. P., Lagendijk, A., Lerosey, G., and Fink, M. (2012). Controlling waves in space and time for imaging and focusing in complex media, *Nat. Photonics* **6**, pp. 283–292.
123. Katz, O., Small, E., Bromberg, Y., and Silberberg, Y. (2011). Focusing and compression of ultrashort pulses through scattering media, *Nat. Photonics* **5**, pp. 372–377.
124. Stockman, M. I., Kling, M. F., Kleineberg, U., and Krausz, F. (2007). Attosecond nanoplasmonic-field microscope, *Nat. Photonics* **1**, pp. 539–544.
125. Lin, J. Q., *et al.* (2009). Time of flight-photoemission electron microscope for ultrahigh spatiotemporal probing of nanoplasmonic optical fields, *J. Phys-Condens. Mat.* **21**, p. 314005.
126. Chew, S. H., *et al.* (2012). Time-of-flight-photoelectron emission microscopy on plasmonic structures using attosecond extreme ultraviolet pulses, *Appl. Phys. Lett.* **100**, p. 051904.

127. Forg, B., *et al.* (2016). Attosecond nanoscale near-field sampling, *Nat. Commun.* **7**, p. 11717.
128. Mikkelsen, A., *et al.* (2009). Photoemission electron microscopy using extreme ultraviolet attosecond pulse trains, *Rev. Sci. Instrum.* **80**, p. 123703.
129. Hommelhoff, P. and Kling, M. F. (eds.) (2015). *Attosecond Nanophysics: From Basic Science to Applications* (Wiley-VCH, Germany).
130. Weiner, A. M. (2009). *Ultrafast Optics* (Wiley, New Jersey).
131. Mukamel, S. (1999). *Principles of Nonlinear Optical Spectroscopy* (Oxford University Press, Oxford).
132. Polli, D., Luer, L., and Cerullo, G. (2007). High-time-resolution pump-probe system with broadband detection for the study of time-domain vibrational dynamics, *Rev. Sci. Instrum* **78**, p. 103108.
133. Cerullo, G., Baltuska, A., Mucke, O. D., and Vozzi, C. (2011). Few-optical-cycle light pulses with passive carrier-envelope phase stabilization, *Laser Photon Rev.* **5**, pp. 323–351.
134. Hamm, P. and Zanni, M. (2011). *Concepts and Methods of 2D Infrared Spectroscopy.* (Cambridge University Press, Cambridge).
135. Wang, W., *et al.* (2014). Interplay between strong coupling and radiative damping of excitons and surface plasmon polaritons in hybrid nanostructures, *Acs. Nano* **8**, pp. 1056–1064.
136. Dintinger, J., Klein, S., Bustos, F., Barnes, W. L., and Ebbesen, T. W. (2005). Strong coupling between surface plasmon-polaritons and organic molecules in subwavelength hole arrays, *Phys. Rev. B* **71**, p. 035424.
137. Bellessa, J., Bonnand, C., Plenet, J. C., and Mugnier, J. (2004). Strong coupling between surface plasmons and excitons in an organic semiconductor, *Phys. Rev. Lett.* **93**, p. 036404.
138. Fofang, N. T., Grady, N. K., Fan, Z. Y., Govorov, A. O., and Halas, N. J. (2011). Plexciton dynamics: exciton-plasmon coupling in a j-aggregate-au nanoshell complex provides a mechanism for nonllinearity, *Nano. Lett.* **11**, pp. 1556–1560.
139. Schwartz, T., Hutchison, J. A., Genet, C., and Ebbesen, T. W. (2011). Reversible switching of ultrastrong light-molecule coupling, *Phys. Rev. Lett.* **106**, p. 196405.
140. Vasa, P., *et al.* (2010). Ultrafast manipulation of strong coupling in metal-molecular aggregate hybrid nanostructures, *ACS Nano* **4**, pp. 7559–7565.
141. Törmä, P. and Barnes, W. L. (2015). Strong coupling between surface plasmon polaritons and emitters: a review, *Rep. Prog. Phys.* **78**, p. 013901.
142. Weisbuch, C., Nishioka, M., Ishikawa, A., and Arakawa, Y. (1992). Observation of the couplex exciton-photon mode splitting in a semicondcutor quantum microcavity, *Phys. Rev. Lett.* **69**, pp. 3314–3317.
143. Thompson, R. J., Rempe, G., and Kimble, H. J. (1992). Observation of normal-mode splitting for an atom in an optical cavity, *Phys. Rev. Lett.* **68**, pp. 1132–1135.

144. Yoshie T., *et al.* (2004). Vacuum Rabi splitting with a single quantum dot in a photonic crystal nanocavity, *Nature* **432**, pp. 200–203.
145. Reithmaier, J. P., *et al.* (2004). Strong coupling in a single quantum dot-semiconductor microcavity system, *Nature* **432**, pp. 197–200.
146. Fidder, H., Knoester, J., and Wiersma, D. A. (1993). Observation of the one-exciton to 2-exciton transition in a J-aggregate, *J. Chem. Phys.* **98**, pp. 6564–6566.
147. Hamm, P. (1995). Coherent effects in femtosecond infrared spectroscopy, *Chem. Phys.* **200**, pp. 415–429.
148. Guenther, T., *et al.* (2002). Coherent nonlinear optical response of single quantum dots studied by ultrafast near-field spectroscopy, *Phys. Rev. Lett.* **89**, p. 057401.
149. Dintinger, J., Robel, I., Kamat, P. V., Genet, C., Ebbesen, and T. W. (2006). Terahertz all-optical molecule-plasmon modulation, *Adv. Mater.* **18**, pp. 1645–1648.
150. Dintinger, J., Klein, S., and Ebbesen, T. W. (2006). Molecule-surface plasmon interactions in hole arrays: Enhanced absorption, refractive index changes, and all-optical switching, *Adv. Mater.* **18**, pp. 1267–1270.
151. Bergman, D. J. and Stockman, M. I. (2003). Surface plasmon amplification by stimulated emission of radiation: Quantum generation of coherent surface plasmons in nanosystems, *Phys. Rev. Lett.* **90**, p. 027402.
152. Chikkaraddy, R., *et al.* (2016). Single-molecule strong coupling at room temperature in plasmonic nanocavities, *Nature* **535**, pp. 127–130.
153. Jaynes, E. T. and Cummings, F. W. (1963). Comparison of quantum and semiclassical radiation theories with application to beam maser, *P IEEE* **51**, pp. 89–109.
154. Shore, B. W. and Knight, P. L. (1993). The Jaynes-Cummings model, *J. Mod. Opt.* **40**, pp. 1195–1238.
155. Wu, K., Chen, J., McBride, J. R., and Lian, T. (2015). Efficient hot-electron transfer by a plasmon-induced interfacial charge-transfer transition, *Science* **349**, pp. 632–635.
156. Engel, G. S., *et al.* (2007). Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems, *Nature* **446**, pp. 782–786.
157. Collini, E. and Scholes, G. D. (2009). Coherent intrachain energy migration in a conjugated polymer at room temperature, *Science* **323**, pp. 369–373.
158. Rozzi, C. A., *et al.* (2013). Quantum coherence controls the charge separation in a prototypical artificial light-harvesting system, *Nat. Commun.* **4**, p. 1602.
159. Falke, S. M., *et al.* (2014). Coherent ultrafast charge transfer in an organic photovoltaic blend, *Science* **344**, pp. 1001–1005.
160. Jonas, D. M. (2003). Two-dimensional femtosecond spectroscopy, *Ann. Rev. Phys. Chem.* **54**, pp. 425–463.
161. De Sio, A., *et al.* (2016). Tracking the coherent generation of polaron pairs in conjugated polymers, *Nat. Commun.* **7**, p. 13742.

162. Tollerud, J. O., Cundiff, S. T., and Davis, J. A. (2016). Revealing and characterizing dark excitons through coherent multidimensional spectroscopy, *Phys. Rev. Lett.* **117**, p. 097401.
163. Unold, T., Mueller, K., Lienau, C., Elsaesser, T., and Wieck, A. D. (2005). Optical control of excitons in a pair of quantum dots coupled by the dipole-dipole interaction, *Phys. Rev. Lett.* **94**, p. 137404.
164. Brinks, D., *et al.* (2010). Visualizing and controlling vibrational wave packets of single molecules, *Nature* **465**, pp. 905–U905.
165. Hildner, R., Brinks, D., Nieder, J. B., Cogdell, R. J., and van Hulst, N. F. (2013). Quantum coherent energy transfer over varying pathways in single light-harvesting complexes, *Science* **340**, pp. 1448–1451.
166. Haedler, A. T., *et al.* (2015). Long-range energy transport in single supramolecular nanofibres at room temperature, *Nature* **523**, pp. 196–199.
167. Sun, Q., Ueno, K., Yu, H., Kubo, A., Matsuo, Y., and Misawa, H. (2013). Direct imaging of the near field and dynamics of surface plasmon resonance on gold nanostructures using photoemission electron microscopy, *Light-Sci. Appl.* **2**, p. e118.

CHAPTER 8

Graphene Plasmonics

ALEXEY YU. NIKITIN

IKERBASQUE, Basque Foundation for Science,
Spain and CIC NanoGUNE, Spain
a.nikitin@nanogune.eu

Graphene plasmons (GPs) — electromagnetic fields coupled to Dirac electrons in graphene sheet — have recently attracted a great attention in different scientific fields, particularly in condensed matter physics and photonics. In this chapter, we aim to review the most relevant results in graphene plasmonics reported during the last few years. We consider important properties of GPs, such as their dispersion, confinement, losses, etc. We discuss various mechanisms of efficient coupling between free-space waves and GPs. In more detail, we theoretically illustrate an example of GPs excitation by a diffraction grating.

8.1. Introduction

Graphene plasmons (GPs) present confined electromagnetic fields coupled to charge carriers (dirac electrons or holes) oscillations in atomically thick carbon layer — graphene sheet. Apart from propagating along a purely two-dimensions (2D) material, GPs are significantly different form conventional surface-plasmon polaritons (SPPs) existing on metal–dielectric interfaces in several key aspects. First, while SPPs on metals correspond to the visible and near infra–red (IR) spectral regions, the frequency (energy) range of GPs is

$3 \div 2000\,\mathrm{cm}^{-1}$ $(0.4 \div 250\,\mathrm{meV})$ covering the mid- and far-IR and terahertz (THz) bands. Second, GPs have extremely short wavelengths λ_p (down to 100 times shorter that the wavelength of light in free space, $\lambda_p \ll \lambda_0$) and therefore large confinement. Third, and most important, since the GPs present the oscillations of dirac particles with the liner dispersion, the wavelength of GPs allows for an efficient electrostatic tuning via changing the Fermi level E_F by external gate voltage. Both the tunability of GPs (enabling an active control over the GPs' wavelength) and their extraordinary confinement afford the relevance of GPs for various vital applications in photonics and optoelectronics: photodetection,[1,2] molecular sensing (especially in the mid-IR fingerprint spectral region),[3–8] signal processing for THz communications,[9] etc.

The dispersion of GPs in a free-standing graphene presents a typical for 2D electron gases square-root dependency of frequency, ω, upon the wavevector,[10] k, as $\omega \propto \sqrt{k}$, being well separated from the light line $\omega = kc$. Free-standing graphene, however, is barely used and is usually grown (or transferred) on substrates. In the mid-IR range, substrates typically have phonons which can strongly couple to GPs. The phonons can thus significantly modify the dispersion of GPs around certain frequencies (anti-crossing of phonon and plasmon dispersion curves) and lead to the formation of graphene plasmon–phonon polaritons.[11–13]

Many theoretical and experimental studies of GPs have been reported during a few recent years,[1,14–20] revealing exciting and fundamental physical phenomena, which GPs involve.[a] Some of these (both linear and nonlinear) phenomena are similar to those known in metal plasmonics, e.g. plasmon-enhanced light absorption,[21,22] subwavelength waveguiding of light,[9,23,24] high harmonic generation,[25,26] etc. At the same time, strong confinement of GPs leads to unexpected and contra-intuitive new physical effects, e.g. enormous spontaneous emission rates $\sim 10^6$ (favoring strong

[a]Note that graphene plasmonics was positioned at the 8th place in the top 10 research fronts in Physics according to Thomson Reuters' Research Fronts 2014 report.

light/mater interaction),[27] anomalously high GP reflection by nanometric discontinuities,[28–30] giant Faraday rotation,[31,32] etc.

Deeply subwavelength character of GPs in many cases allows for an elegant analytical treatment of the GP-related electromagnetic problems by means of quasi-static approximation.[15,21,33] However, extremely short wavelength of GPs implies a huge momentum mismatch with free-space waves, and therefore the experimental observation of GPs faces substantial difficulties. The momentum mismatch can be overcome by the near-field microscopy. Particularly, scattering-type near-field microscopy (s-SNOM) is currently the only available tool that allows for imaging and characterizing graphene plasmons in real space.[34–36]

8.2. Conductivity of Graphene and Dispersion Relation of Graphene Plasmons

The GPs can be described by the classical electromagnetic theory based on Maxwell's equations, assuming that the graphene layer presents an infinitesimally thin conducting sheet with the 2D complex conductivity σ. The validity of this approximation is justified by significantly larger wavelengths of GPs compared to the atomic distances in the graphene monolayer. Within the random-phase approximation (accounting for a weak screened Coulomb interaction), the dynamic optical response of graphene can be obtained from Kubo formula yielding $\sigma = \sigma_{\text{inter}} + \sigma_{\text{intra}}$, with the terms σ_{inter} and σ_{intra} being the contributions from interband and intraband electron–photon scattering processes, respectively[37–39]:

$$\sigma_{\text{intra}} = \frac{2ie^2t}{\hbar\pi(\Omega + i\gamma)} \ln\left[2\cos h\left(\frac{1}{2t}\right)\right], \tag{8.1}$$

$$\sigma_{\text{inter}} = \frac{e^2}{4\hbar}\left[\frac{1}{2} + \frac{1}{\pi}\arctan\left(\frac{\Omega - 2}{2t}\right) - \frac{i}{2\pi}\ln\frac{(\Omega + 2)^2}{(\Omega - 2)^2 + (2t)^2}\right].$$

In these expressions, Ω, γ and t present frequency, decay rate and temperature, respectively, normalized to the Fermi energy E_F: $\Omega = \hbar\omega/|E_F|$, $\gamma = \hbar/\tau|E_F|$, $t = T/|E_F|$ (T is given in units of

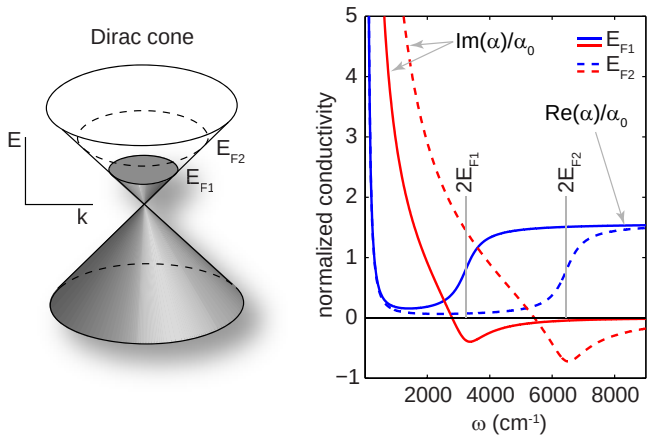
energy). Note that the mobility of the charge carriers μ_e is related to the relaxation time τ via $\tau = \mu_e |E_F|/ev_F^2$, where v_F is the Fermi velocity and e is the elementary charge. Recall that the real part of the complex conductivity, σ, characterizes the energy loss in the system, while the imaginary part of σ characterizes the phase shift between the electric field and the induced current ("inertness" of the charge carriers to respond to changing electric field). In the following, in order to facilitate readability of equations, we will use the normalized conductivity α expressed by

$$\alpha = \frac{2\pi\sigma}{c}. \tag{8.2}$$

Figure 8.1(b) shows both real (red curves) and imaginary (blue curves) part of α as a function of frequency for two Fermi energies: $E_{F1} = 0.2\,\text{eV}$ and $E_{F2} = 0.4\,\text{eV}$. According to the model given by Eq. (8.1), the intraband part of σ dominates in the frequency region $\omega < 2E_F$, and both $\text{Re}(\alpha)$ and $\text{Im}(\alpha)$ are positive and increase with the frequency decrease. In contrast, the interband contributions are dominant for $\omega > 2E_F$ resulting in negative $\text{Im}(\alpha)$ and the saturation of $\text{Re}(\alpha)$ to the value $\pi\alpha_0/2$, with $\alpha_0 = e^2/\hbar c \simeq 1/137$ being the fine structure constant.[b] For ω close to the "bounding" frequency, $\omega = 2E_F$ (separating the spectral regions with dominating either σ_{inter} or σ_{intra} terms in Eq. (8.1)), $\text{Im}(\alpha)$ changes the sign and $\text{Re}(\alpha)$ shows a step-like behavior (presenting an abrupt jump for zero temperature, $T = 0$), see Fig. 8.1(b).

The properties of the plasmons in graphene are mostly defined by the conductivity σ. Within the electromagnetic theory based on Maxwell's equations, the conductivity appears in the boundary conditions for the electromagnetic fields. Namely, σ provides the jump of the parallel components of the magnetic fields on the graphene sheet, related with the presence of the surface current

[b]The fine-structure constant, α_0, can be interpreted as a ratio of the velocity of the electron in the first circular orbit (the innermost electron) of the Bohr model of the atom, $v_e = c/137 = 2.19 \cdot 10^6\,\text{m/s}$ to the speed of light in vacuum, i.e. $\alpha_0 = v_e/c$.



(a) (b)

Fig. 8.1. Dirac cone and the conductivity of graphene. (a) Schematics for the dispersion of electrons in graphene near the Dirac point (in energy-momentum space). The electron filling is shown by the gray color for $E_F = E_{F1}$. (b) Real part (blue curves) and imaginary part (red curves) of the normalized graphene optical conductivity α (shown in units of the fine structure constant $\alpha_0 \simeq 1/137$) as a function of frequency ω, according to Eqs. 8.1 and 8.2. The solid curves correspond to $E_{F1} = 0.2\,\text{eV}$, while the dashed ones to $E_{F2} = 0.4\,\text{eV}$, respectively. The vertical solid lines mark the double Fermi energies. The mobility of the charge carriers μ_e and the temperature T are taken to be $10{,}000\,\text{cm}^2/(\text{Vs})$ and $300\,\text{K}$, respectively.

$\boldsymbol{j} = \sigma \boldsymbol{E}_t$ (subscript "t" means the component parallel to the graphene sheet). Assuming that the graphene sheet is placed at $z = 0$, the boundary conditions for the magnetic and electric fields read:

$$\begin{aligned} \boldsymbol{e}_z \times [\boldsymbol{H}_1 - \boldsymbol{H}_2] &= 2\alpha \cdot \boldsymbol{e}_z \times [\boldsymbol{e}_z \times \boldsymbol{E}_1], \\ \boldsymbol{e}_z \times [\boldsymbol{E}_1 - \boldsymbol{E}_2] &= 0, \end{aligned} \tag{8.3}$$

where with subscripts 1 and 2, we indicate the fields in the half-space "1" (positive values of the coordinate z) and "2" (negative values of the coordinate z) separated by the graphene sheet, respectively.

The GP present the wave decaying perpendicularly to the graphene sheet. Assuming that the GP is propagating along the x-axis, the magnetic fields of GP in the graphene sheet placed in

free space (vacuum) can be written in the following form

$$\begin{aligned} \boldsymbol{H}_1(x,z) &= \boldsymbol{e}_y H_0 e^{ik_p x + ik_{pz} z}, \quad z > 0, \\ \boldsymbol{H}_2(x,z) &= -\boldsymbol{e}_y H_0 e^{ik_p x - ik_{pz} z}, \quad z < 0, \end{aligned} \tag{8.4}$$

where $k_{pz} = \sqrt{k_0^2 - k_p^2}$ is the z-component of the GP wavevector and $k_0 = \omega/c = 2\pi/\lambda_0$ is the wave-vector in the free space. The monochromatic time dependence, $e^{-i\omega t}$, is assumed in Eq. (8.4) and everywhere in the following. The electric fields can be found from Eq. (8.4) using Maxwell's equations $\boldsymbol{E} = (i/k_0)\nabla \times \boldsymbol{H}$:

$$\begin{aligned} \boldsymbol{E}_1(x,z) &= \frac{1}{k_0} H_0 (\boldsymbol{e}_x k_{pz} - \boldsymbol{e}_z k_p) e^{ik_p x + ik_{pz} z}, \quad z > 0, \\ \boldsymbol{E}_2(x,z) &= \frac{1}{k_0} H_0 (\boldsymbol{e}_x k_{pz} + \boldsymbol{e}_z k_p) e^{ik_p x - ik_{pz} z}, \quad z < 0. \end{aligned} \tag{8.5}$$

From Eqs. (8.4) and (8.5), it follows that the magnetic field and z-component of the electric field are anti-symmetric with respect to the graphene sheet, while the x-component of the electric field is symmetric due to continuity given by the boundary conditions. Connecting the fields in the upper and lower half-spaces, "1" and "2", from Eqs. (8.4) and (8.5), through the boundary conditions given by Eq. (8.3), we can derive the dispersion relation for the GPs:[40,41]

$$k_p(\omega) = \frac{\omega}{c} \sqrt{1 - \frac{1}{\alpha^2}}, \tag{8.6}$$

(which can also be alternatively written as $k_{pz} = -1/\alpha$). In frequency regions, where $|\alpha| \ll 1$, the dispersion relation simplifies to

$$\frac{k_p}{k_0} = q_p \simeq \frac{i}{\alpha}. \tag{8.7}$$

In order to guarantee the exponential decay of the GP electromagnetic fields away from the graphene sheet (taking place when $\mathrm{Im}(k_{pz}) \geq 0$), the imaginary part of α must be positive. This limits the frequency range of GP to approximately $\omega < 2E_F$. On the other hand, the real part of α is always positive, so that the positiveness

of $\mathrm{Im}(k_p)$ is guaranteed (the energy in the system is not amplified). Due to small α, the transversal and longitudinal components of the GP wavevector are virtually equal in magnitude (and phase-shifted by $\pi/2$), i.e. $k_{pz} \simeq ik_p$, so that k_p characterizes both the wavelength of GPs, $\lambda_p = 2\pi/\mathrm{Re}(k_p)$ and their vertical confinement, $\delta_p = 1/\mathrm{Im}(k_{pz}) \simeq 1/\mathrm{Re}(k_p)$. Both λ_p and δ_p can reach ultrasmall values (down to two orders of magnitude) compared to free space wavelength, λ_0.

In the low-frequency region, where the intraband electron–photon scattering processes dominate, the relation between the frequency, Fermi level and GP wavevector can be explicitly found by substituting σ_{intra} from Eq. (8.1) into Eq. (8.7) (taking into account the definition of α in Eq. (8.2)). Neglecting the temperature ($t = 0$), we have

$$k_p \simeq \frac{\hbar}{2e^2}.\frac{\hbar\omega(\omega + i/\tau)}{|E_F|}. \tag{8.8}$$

We see, thus, that the real part of the GP momentum decreases with E_F and is quadratic in ω, i.e. $\mathrm{Re}(k_p) \propto \omega^2/|E_F|$. This dependency is illustrated by the dispersion curves of GPs, shown in the left panel of Fig. 8.2(b), at two different Fermi levels, typical for the graphene used in experiments.[34–36] Each of the dispersion curves is strongly separated from the light cone, $k = \omega/c$ (black solid line in the left panel of Fig. 8.2(b)), revealing extremely large momentums of GPs. The large momentums (short wavelengths) of GPs, as well as their high confinement to the graphene sheet can be also appreciated in Fig. 8.2(c) where the GP field snapshots are shown together with the half-wavelength in free space. The strong dependency of the GP wavelengths upon the Fermi level reveals a high tunability of GPs (for example, by means of electrostatic gating) being the key in many optoelectronic applications requiring an active control over the electromagnetic fields in the nanoscale.

As can be seen in Fig. 8.2(c), the intensity of blue and red fringes diminishes from left to right. This indicates that the amplitude of GP decreases from the excitation point due to the finite propagation length. Note that the success of many potential applications of

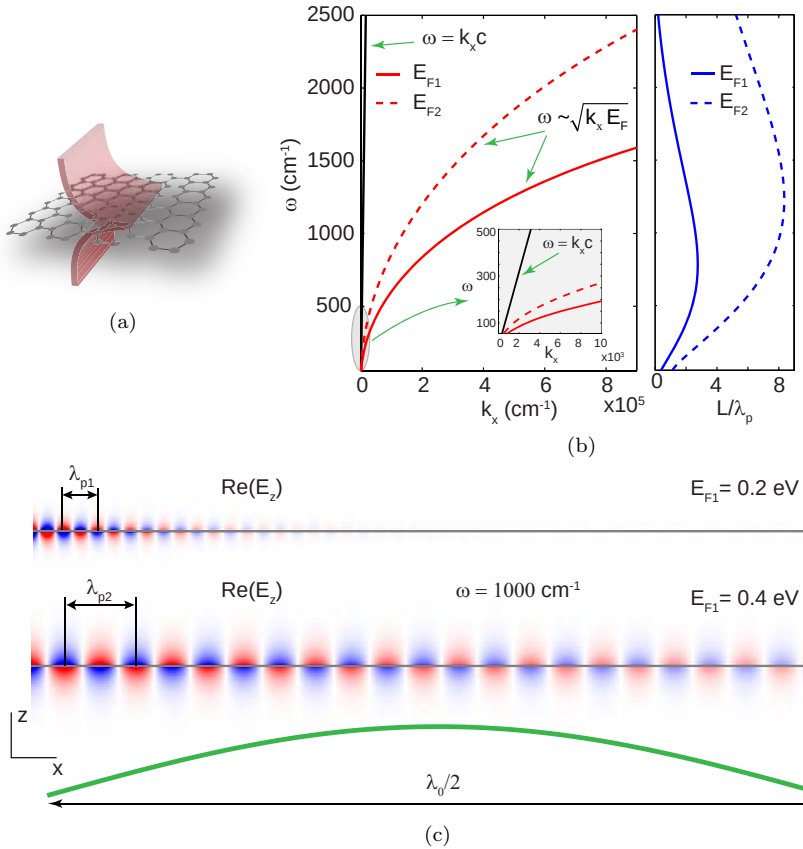


Fig. 8.2. The dispersion and electric fields of graphene plasmons. (a) Schematics for a propagating plasmon in graphene. (b) The dispersion curves (left panel) and the normalized propagation lengths (right panel) for GPs calculated for two Fermi levels: $E_{F1} = 0.2\,\text{eV}$ and $E_{F2} = 0.4\,\text{eV}$. The light cone is shown in the left panel by the black solid line (the inset shows the zoom-in of the low-frequency — far IR and THz — region to better illustrate the separation of the GP dispersion from the light cone). (c) The instant field snapshots for the GP fields at the fixed frequency $\omega = 1000\,\text{cm}^{-1}$ ($\lambda_0 = 10\,\mu\text{m}$) for E_{F1} and E_{F2}, corresponding to the GP wavelengths $\lambda_{p1} = 213\,\text{nm}$ and $\lambda_{p2} = 460\,\text{nm}$, respectively. For comparison, the half-period of the free-space wave oscillation (with the wavelength λ_0) is shown by the green curve. The rest of the parameters for σ are the same as in Fig. 8.1.

GPs (e.g. photonic circuits for communication technologies, sensing, light absorbers, etc.) is essentially dependent upon GP propagation lengths (or decay times). The GP amplitude propagation length, $L_p = 1/\mathrm{Im}(k_p)$, defined as the length at which the amplitude of the fields (see Eqs. (8.4) and (8.5)) decays e times along the propagation direction, can be explicitly found from Eq. (8.8):

$$L_p \simeq \frac{2e^2|E_F|}{\hbar}\frac{\tau}{\hbar\omega} = v_e\tau\frac{2|E_F|}{\hbar\omega}, \quad \frac{L_p}{\lambda_p} \simeq \frac{\tau\omega}{2\pi}, \tag{8.9}$$

where recall that $v_e = c/137$. While the small absolute values of L_p are related to the small wavelength of the GPs, the relative values, L_p/λ_p (the "figures of merit" for GPs) can potentially reach large values, which would correspond to long propagation lengths in units of GP wavelength. According to Eq. (8.9), the figures of merit are a liner function of frequency. Nevertheless, the right panel of Fig. 8.2(b) shows strongly non-monotonic behavior of L_p/λ_p with ω. This disagreement can be explained by a limited applicability of Eq. (8.9) (low frequencies, where the interband part of the conductivity dominates): it is not valid in the mid-IR range, where the intraband part of the conductivity in Eq. (8.1) starts playing an important role.

Note that the mobility of today's standard graphene samples (particularly, chemical vapor deposition (CVD) graphene transferred onto oxides) are rather low, being of order of $1000\,\mathrm{cm}^2/(\mathrm{V}\cdot\mathrm{s})$. Low mobility results in GP figures of merit compared or smaller than 1, thus limiting the performance of the potential GP-based devices. In the simulations shown in Fig. 8.2, the mobility is assumed to be $10{,}000\,\mathrm{cm}^2/(\mathrm{V}\cdot\mathrm{s})$, resulting in figures of merit 2.6 at E_{F1} and 8.1 at E_{F2}, respectively. The samples with the record mobilities, up to approximately $200{,}000\,\mathrm{cm}^2/(\mathrm{V}\cdot\mathrm{s})$, are graphene encapsulated between BN layers.[42,43] However, at present, the fabrication of the encapsulated graphene samples is a sophisticated process, staying far away from the mass production.

Apart from the importance of the propagation lengths in graphene plasmonics, some applications (related to signals propagation) can also require large amplitude decay time, τ_a, defined as

the time during which the amplitude of the GP pulse reduces e times. For the pulses with small frequency spreads, the lifetime can be calculated as $\tau_a = L_a/v_{\rm gr}$, where $v_{\rm gr}$ is the group velocity given by $v_{\rm gr} = {\rm d}\omega/{\rm d}k_x$ (the pulse is supposed to propagate along the x-axis). Both $v_{\rm gr}$ and τ_a can be explicitly found from Eq. (8.8) and Eq. (8.9):

$$v_{\rm gr} \simeq \frac{e^2}{\hbar}\frac{|E_F|}{\hbar\omega} = v_e\frac{|E_F|}{\hbar\omega}, \quad \tau_a \simeq 2\tau. \tag{8.10}$$

According to Eq. (8.10), the amplitude decay time is not dependent upon the frequency and is fully defined by the scattering time of the charge carriers, τ: the frequency dispersion of L_p is compensated by the dispersion of the group velocity, $v_{\rm gr}$. Namely, with the frequency increase, the GPs propagate smaller distance but with lower group velocity, so that the amplitude decay time remains constant.

As a last note in this section, we would like to mention that the plasmon dispersion given by Eq. (8.7) can be easily generalized to the GPs in graphene placed between two media with arbitrary dielectric permittivities, ε_1 and ε_2:

$$q_p = \frac{\varepsilon_1 + \varepsilon_2}{2} \cdot \frac{i}{\alpha}. \tag{8.11}$$

Analogously to Eq. (8.7), the validity of Eq. (8.11) is limited by the condition $|\alpha| \ll 1$. The exact solution for the momentum of GPs in a non-symmetric environment cannot be explicitly written in a simple form. The equation for the dispersion relation of GPs in a non-symmetric environment will be shown in Section 8.5 and can also be found, for instance, in [Refs. 40, 41].

8.3. Approximation of the Graphene Sheet by a Layer of Finite Thickness in the Electromagnetic Simulations

In the calculations of the electromagnetic fields in graphene structures, the graphene is taken into account via the boundary conditions

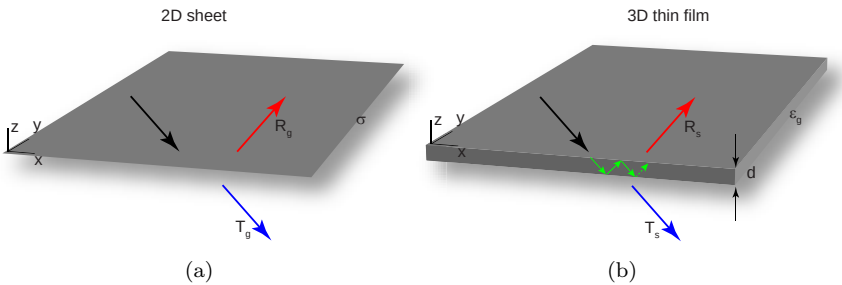


Fig. 8.3. Approximation of the graphene sheet by a finite-thickness film. (a) Schematics for the reflection of a plane wave from a graphene sheet with the optical conductivity σ. (b) Schematics for the reflection of a plane wave from a layer of thickness d with the dielectric permittivity ε_g.

given by Eqs. (8.3). On the one hand, in many cases, it simplifies the analytical treatment of the electromagnetic problems, since one does not need to calculate the fields inside the graphene layer. But on the other hand, some complex problems (involving, for instance, sophisticated shapes of either graphene patches or those of adjusted metallic antennas, etc.) can only be solved by 3D numeric simulations. Generally (and every time more frequently), the commercial software for full-wave simulations is used. However, in existing commercial solvers based on finite-element methods, setting up the boundary conditions on an infinitesimally thin conducting layer is not straightforward (or not even implemented yet).

Let us consider how to approximate a conducting sheet with the conductivity σ by a thin layer of thickness d with an effective dielectric permittivity ε_g. The correspondence between σ, ε_g and d can be established by comparing the Fresnel coefficients for a conducting 2D sheet with those for a thin 3D film. For simplicity, let us consider only the case of the normal incidence and the vacuum environment of both the conducting sheet and film. The final results for the relation between σ and ε_g will be valid for graphene sheet in any environment (for multilayer structures), arbitrary polarization and oblique incidence (we skip the proof here). The Fresnel reflection, r_F, and transmission, t_F, coefficients for graphene read simply as

(see Section 8.5):

$$r_F = \frac{-\alpha}{1+\alpha}, \quad t_F = \frac{1}{1+\alpha}. \tag{8.12}$$

We would like to remark that Fresnel coefficients can be used to easily obtain the absorption in graphene, A, as $A = 1 - |r_F|^2 - |t_F|^2$. Assuming that $|\alpha| \ll 1$, the absorption, according to Eq. (8.12), reads: $A \simeq \mathrm{Re}(\alpha)$. In the frequency region, where the interband conductivity dominates, particularly, in the visible region, A simplifies to $A = \pi\alpha_0$, yielding approximately 0.023.

The Fresnel reflection, r_{Fd}, and transmission, t_{Fd}, coefficients for a slab with the thickness d and dielectric permittivity ε_g, surrounded by air, have the following form:

$$r_{Fd} = r\frac{1 - e^{2i\varphi}}{1 - r^2 e^{2i\varphi}}, \quad t_{Fd} = \frac{\sqrt{\varepsilon_g}t^2 e^{i\varphi}}{1 - r^2 e^{2i\varphi}}, \tag{8.13}$$

where $\varphi = \mathrm{d}k_0\sqrt{\varepsilon_g}$ is the optical path length inside the slab. $r = \frac{1-\sqrt{\varepsilon_g}}{1+\sqrt{\varepsilon_g}}$ and $t = \frac{2}{1+\sqrt{\varepsilon_g}}$ are the reflection and transmission coefficients, respectively, corresponding to the interface between air and a semi-infinite bulk medium with the dielectric permittivity, ε_g. In order to establish the correspondence between Eqs. (8.13) and (8.12), we will perform some simplifications.

First of all, let us remember that in the macroscopic electromagnetic theory, the polarization vector inside a bulk medium (and therefore, the dielectric permittivity ε_g) is proportional to the number of polarized molecules, or in case of conductors, free charges per volume,[44] i.e. $\varepsilon_g \propto n_{3D}$, with $n_{3D} = N/\Delta V$. Let us assume that in a thin slab, which will approximate the graphene layer, N remains unchanged independently upon the thickness of the slab, d, so that the number of charges per unit area, $n_{2D} = N/\Delta S$, is constant, where ΔS is the area parallel to the slab surfaces. Then for n_{3D} in the slab, we can write $n_{3D} = N/(\Delta S \cdot d) = n_{2D}/d$. From the last equation, we conclude that the effective dielectric permittivity should be inversely proportional to d, and thus be a large number, e.g. $\varepsilon_g \propto 1/d \gg 1$. The large value of ε_g can greatly simplify Eq. (8.13).

Another simplification that can be done in Eq. (8.13) is related to small optical path lengths inside the slab. Indeed, taking into account $\varepsilon_g \propto 1/d$, from the definition of φ (given below Eq. (8.13)), we see that $\phi \sim \sqrt{d}/\lambda_0$. Then, approximating the exponentials by their Tailor series up to the first-order terms and retaining further in the numerators and denominators of Eq. (8.13) only the first non-vanishing terms (assuming d to be a small parameter), we arrive at

$$r_{Fd} = \frac{i\varphi\sqrt{\varepsilon_g}/2}{1 - i\varphi\sqrt{\varepsilon_g}/2}, \quad t_{Fd} = \frac{1}{1 - i\varphi\sqrt{\varepsilon_g}/2}. \tag{8.14}$$

Comparing r_{Fd}, t_{Fd} in Eq. (8.14) with r_F, t_F in Eq. (8.12), we obtain the explicit expression for ε_g:

$$\varepsilon_g = \frac{i\alpha\lambda_0}{\pi d}, \tag{8.15}$$

where α is given by Eq. (8.2).

Note that the result obtained in this section can be also used for the "inverse" problem: approximation of a thin slab by an effective infinitesimally thin conductive layer. In this case, Eq. (8.15) can be used to relate the dielectric permittivity and thickness of a thin slab with the effective conductivity. Such approximation was used in the theory of surface polaritons in macroscopic transition layers,[41] which established effective boundary conditions for the electromagnetic fields (analogs to Eq. (8.3)) in thin polar slabs.

8.4. Optical Excitation of Graphene Plasmons

As we have seen in Section 8.2, GPs have momenta much larger than those corresponding to free-space waves. In other words, at a fixed frequency, ω, the phase velocity of GPs, $v_{ph} = \omega/k_x$, is significantly smaller than that of a wave propagating in free space. Therefore, due to a strong phase mismatch, a direct coupling between free-space waves and GPs is not possible. In order to provide an interaction between plane waves and GPs, one has to compensate the momentum mismatch, $\Delta k = k_x - k_p$, by breaking the spatial symmetry.

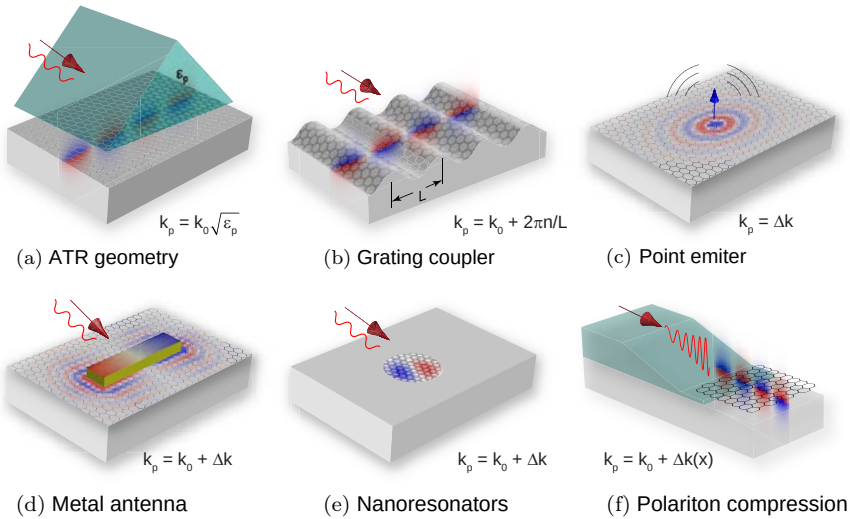


Fig. 8.4. Various methods for linear optical excitation of GPs. (a) Attenuated total reflection with a prism (Otto geometry[45]). (b) Coupling to GPs via a periodic grating. (c) Excitation of GPs by a localized (point) source. (d) Coupling to GPs from a resonant metal antenna. (e) Coupling to GP Fabry–Perot modes in graphene resonators. (f) Coupling to GP by compressing polaritons in tapered films of polar or semiconducting materials.

Figure 8.4(a) shows the conventional plasmon coupling configuration: a dielectric prism with a dielectric permittivity $\varepsilon_p > 1$ separated by a small gap from the graphene layer. This configuration is known as attenuated total internal reflection (ATR)[10,45] and is widely used in plasmonic sensing applications. Due to total internal reflection, occurring when the incident angle θ is larger than the critical one, $\theta_c = \arcsin(1/\sqrt{\varepsilon_p})$, the wave inside the prism is totally reflected, generating exponentially decaying fields at the bottom face of the prism. These evanescent fields have momentums, $k_x = \sqrt{\varepsilon_p} k_0 \sin\theta$, larger than those of free-space waves, $k_x > k_0$, and therefore for proper incident angle and refractive index of the prism, k_x can be matched with the wavevector of GPs, $k_x = k_p$:

$$k_p = \sqrt{\varepsilon_p} k_0 \sin\theta. \tag{8.16}$$

A big disadvantage of the coupling to GPs via prism is that the refractive index of the prism has to take enormous values. Recall that the refractive index of GPs, $q_p = k_p/k_0$, can be one or two orders of magnitude larger than 1 (see the GP dispersion curves in Fig. 8.2(b)). Therefore, using prisms made of existing natural materials, coupling to GPs due to the ATR is only possible at THz frequencies, where the GP dispersion is not so dramatically separated from the light line (compared to, for instance, mid-IR frequency range), and thus the values of refractive indices q_p are moderate.[46]

Figure 8.4(b) shows the diffraction grating, presenting another common structure for the excitation of GPs.[10,45,47] The scattering of an incident wave by a diffraction grating generates a set of plane waves (spatial Fourier harmonics of the scattered field) with discrete wavevectors, k_{nx}, along the grating vector $\boldsymbol{G}$ (reciprocal lattice vector). According to the Floquet theorem, the quantized wavevectors k_{nx} are given by $k_{nx} = k_0 \sin\theta + nG$ with $G = 2\pi/L$, where L is the period of the grating. The GP excitation condition on the diffraction grating, in a certain diffraction order n, reads as[48–49]

$$k_p = k_0 \sin\theta + nG. \tag{8.17}$$

The efficiency of the GP excitation by the grating can be very high (see Refs. 14 and 49 and Section 8.5 of this chapter), making the diffraction gratings promising candidates for GP-based technologies, particularly for light absorbers and sensors.[3] The grating provides the excitation of the GPs in the form of standing waves ("Bragg GPs"), rather than propagating GPs. On the other hand, the launching of propagating GPs, emanating from the local excitation region, can be crucial, for instance, in applications for communication, signal processing, coupling between quantum objects or sensing, based on absorption losses of GPs during the propagation across the sensed region.[3]

The propagating GPs can be locally excited by small emitters like, for instance, quantum dots, metallic nanoparticles or tips of the near-field microscopes.[35,36,43] Figure 8.4(c) shows the simplest example of a small emitter: a point dipolar source. An approximation

of a dipolar source is widely used to gain an insight into the physics behind the local excitation of GPs. The Fourier transform of the electric field radiated by a point dipole has a broad (plane) distribution of momenta. Therefore, when a dipole is placed in close proximity to the graphene sheet, the GPs are directly excited by the evanescent waves with momenta close to k_p.[27,51–55] For an appropriate distance between the dipole and graphene sheet (being of order of the GPs vertical confinement, δ_p), the efficiency of coupling to GPs can reach 100%.[54] In the conditions of the maximal coupling efficiency, the spontaneous emission rate of a point emitter is exclusively attributed to the decay via GPs.

The high efficiencies from point emitters are based on the assumption that the emitters play the role of electromagnetic energy sources. However, in real conditions, an external illumination usually presents the energy source for GPs. The external waves, polarizing a small object, placed close to the graphene sheet, induce an effective dipole (with the dipole moment depending upon the size and geometry of the object) and then the induced effective dipole couples to GPs. Due to the smallness of the object, its polarizability (proportional to the volume of the object) is also small. Therefore, the efficiencies of GP launching from the region of the local excitation are usually low.

The efficiency of local excitation of propagating GPs can be significantly improved by using resonance antennas, see Fig. 8.4(d). For instance, gold rods with lengths matching approximately a half of free-space wavelength can act as resonant dipole antennas. At resonance wavelengths, the antenna can provide a strong enhancement of the generated near-fields, as well as momenta increased orders of magnitude compared to k_0. The high-momentum components of the intense near-fields in the antenna match the GP wavevector, thus exciting propagating GPs.[34] Due to large sizes of the metallic antennas (roughly, being of order of $\lambda_0/2$) compared to GP wavelength, the wavefronts of the GPs can be controlled by shaping the antenna. For example, plane wavefronts can be generated by the flat regions of the antenna, while concave and convex shapes can generate

divergent and convergent GPs, respectively. The concave shapes can thus act as 2D lens for focusing GPs.

In appropriately structured graphene patches, reflection of the GPs from the edges can lead to constructive interference and create Fabry–Perot resonances. Therefore, shaped graphene structures can be considered as graphene (nano)resonators or graphene antennas.[56] The GP modes in graphene resonators can be directly excited by an external illumination,[11,21–23,57–60] see Fig. 8.4(e). The externally illuminated light scatters on strong discontinuities (edges) of graphene antennas, producing the near-fields with large momentums (similarly to metallic antennas), which then efficiently couple to GPs. Due to the uniform character of the electric field of plane waves, they can only couple to GP Fabry–Perot modes of certain symmetries (bright modes). For example, the dipolar modes with odd number of nodes can be excited, while those with even number of nodes are forbidden. Analogously to GPs excited on the diffraction grating, GP modes in graphene resonators can be beneficial for light absorbers[21] and ultra-sensitive actively controllable sensors.[4]

The last concept of GP launching we would like to mention in this section is compression of surface polaritons (SPs) on tapered slabs of bulk materials supporting surface polaritons (polaritonic waveguides). The latter can be either phonon–polaritons — in polar dielectrics — or plasmon–polaritons — in highly-doped simiconductors. The coupling mechanism is sketched in Fig. 8.4(f). First, a SP propagates along a slab with a negative dielectric permittivity. Note that the polariton in a relatively thick slab can be easily excited by the attenuated total internal reflection method, considered above. The polaritonic waveguide is tapered so that its thickness gradually decreases. Subsequent propagation of the polariton along a tapered slab yields a compression of its wavelength until it matches that of the GPs.[61] Finally, by placing the graphene above the thinned slab, the field of the SP couples efficiently to the GP. As has been shown theoretically, the polaritonic coupler can launch graphene plasmons with efficiencies larger than 25%.[61]

8.5. Excitation of Graphene Plasmons in Graphene Sheet with a Periodically Modulated Conductivity

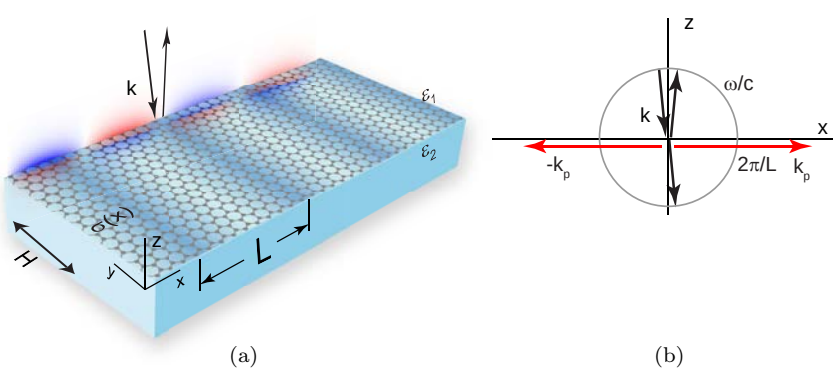
Let us consider the diffraction of a plane monochromatic electromagnetic wave on a graphene sheet with periodically modulated conductivity, α, see Fig. 8.5(a). For generality, we assume that α is an arbitrary periodic function of the coordinate x, so that $\alpha(x) = \alpha(x + L)$, where L is the period of the modulation.

As any periodic function, $\alpha(x)$ can be expanded into the Fourier series:

$$\alpha(x) = \sum_n \bar{\alpha}_n e^{inGx}, \quad G = \frac{2\pi}{L}, \tag{8.18}$$

where n is integer ($n = \cdots -2, -1, 0, 1, \ldots$). The Fourier harmonics $\bar{\alpha}_n$ are given by

$$\bar{\alpha}_n = \frac{1}{L} \int_{-L/2}^{L/2} \mathrm{d}x\, \alpha(x) e^{-inGx}. \tag{8.19}$$



(a) (b)

Fig. 8.5. Excitation of GPs on a graphene sheet with modulated conductivity. (a) Schematics of the structure. Nearly normally incident illuminating plane wave has the magnetic field parallel to the graphene sheet and perpendicular to the grating vector. The plane wave excites the graphene plasmon with the vertical electric field shown by a red–white–blue color plot. (b) k-vector diagram explaining the momentum matching. The GP k-vectors are shown by the red arrows, while the k-vectors of the incident, reflected and transmitted waves (in zeroth diffraction order) are shown by the black arrows. The circle represents the cross-section of the light cone, $k_x^2 + k_z^2 = k_0^2$, at a fixed frequency.

We assume that the dielectric permittivities of the upper half-space ($z > 0$) and of the lower half-space ($z < 0$) are ε_1 and ε_2, respectively, (see Fig. 8.5(a)). The x-components of the electric fields in the upper, E_{1x} and lower, E_{2x} half-spaces can be represented by the Fourier–Floquet expansion:

$$E_{1x} = e^{ik_x x - ik_z z} + \sum_n r_n e^{ik_{nx} x + ik_{1nz} z},$$

$$E_{2x} = \sum_n t_n e^{ik_{nx} x - ik_{2nz} z}, \tag{8.20}$$

where $k_{nx} = k_x + nG$ and $k_{1,2nz} = \sqrt{\varepsilon_{1,2} k_0^2 - k_{nx}^2}$ are the x- and z-components of the wavevectors for the diffracted (scattered) plane waves. The first term of the right-hand side in the first line of Eq. (8.20) presents the incident plane wave (which amplitude is set to 1). Both x- and z-components of the incident plane wave wavevector can be written through the angle of incidence θ (counting form the normal to the graphene sheet) as $k_x = \sqrt{\varepsilon}_1 k_0 \sin\theta$ and $k_z = \sqrt{\varepsilon}_1 k_0 \cos\theta$. The coefficients r_n and t_n present the amplitudes of the spatial Fourier harmonics (waves diffracted in the nth order) in the upper and lower half-spaces, respectively.

Using Maxwell's equations (see Section 8.2) we can find the y-component of the magnetic fields. Then, matching the fields at the graphene sheet ($z = 0$) according to Eq. (8.3), and using the Fourier series for $\alpha(x)$ in Eq. (8.18), we obtain

$$e^{ik_x x} + \sum_n (r_n - t_n) e^{ik_{nx} x} = 0,$$

$$-Y_i e^{ik_x x} + \sum_n (r_n Y_{1n} - t_n Y_{2n}) e^{ik_{nx} x} = -2 \sum_{n,m} \bar{\alpha}_m t_n e^{i(k_{nx} + mG)x}, \tag{8.21}$$

where $Y_i = \varepsilon_1 k_0 / k_z$ is the admittance (the inverse of the wave impedance) of the incident wave and $Y_{1,2n} = \varepsilon_{1,2} k_0 / k_{1,2nz}$ are the admittances of the diffracted waves. Taking into account that Eq. (8.21) must hold for any value of x, we have to equal the

coefficients at different exponentials. Then we obtain a compact system of equations for the amplitudes t_n:

$$\sum_m D_{nm} t_m = V_n,$$
$$D_{nm} = \delta_{nm}(Y_{1n} + Y_{2n}) + 2\bar{\alpha}_{n-m}, \quad V_n = 2Y_i\delta_{n0}, \tag{8.22}$$

where with δ_{nm} we mean the Kronecker symbol. The amplitudes, r_n, are related to the amplitudes, t_n, by the following relation

$$r_n = -\delta_{n0} + t_n. \tag{8.23}$$

The amplitudes r_n and t_n are related to the scattered power (normalized to the power in the incident wave) in the upper and lower half-spaces (R and T, respectively) according to

$$R = \sum_n \mathrm{Re}\left(\frac{Y_{1n}}{Y_i}\right)|r_n|^2, \quad T = \sum_n \mathrm{Re}\left(\frac{Y_{2n}}{Y_i}\right)|t_n|^2. \tag{8.24}$$

The power absorbed in the graphene layer is given by the following equation

$$A = 1 - R - T. \tag{8.25}$$

In short-period diffraction gratings (with $L \ll \lambda_0$), suitable for the excitation of GPs, the values of the wavevectors k_{nx} are large, $k_{nx} \gg k_0$. As a result, all the scattered waves, except the zeroth-order spatial field harmonic, are evanescent ($k_{1,nz}$ and $k_{2,nz}$ are imaginary for $n \neq 0$). Therefore, exclusively the zeroth-order term contributes to the sums in Eq. (8.24) ($k_{1,0z}$ and $k_{2,0z}$ are real).

The linear system of Eqs. (8.22) and (8.23) can be solved numerically for any type of periodic function $\alpha(x)$. For each type of the modulation, an appropriate number N of the diffraction orders, n, must be taken into account in order to achieve the convergence of the solution. In some cases, however (particularly, for periodic modulations with abrupt changes of $\alpha(x)$ as, for example, for graphene ribbons), the convergence with N can be very slow and the system of Eqs. (8.22) and (8.23) becomes unpractical. On the other hand, smooth profiles of $\alpha(x)$ provide a good convergence and the number

of required diffraction order is not large, so that even an analytical treatment of the system of equations is possible.

To illustrate how the system Eqs. (8.22) and (8.23) can be solved analytically, we will consider the simplest cosine profile of $\alpha(x)$:

$$\alpha(x) = \bar{\alpha}_0 + \Delta\alpha \cdot \cos(Gx) = \bar{\alpha}_0 + \frac{\Delta\alpha}{2} e^{-iGx} + \frac{\Delta\alpha}{2} e^{iGx}, \quad (8.26)$$

where $\bar{\alpha}_0$ (presenting the zeroth–order Fourier harmonic) is the average conductivity and $\Delta\alpha$ is the modulation amplitude. From equation (8.26), it follows that the amplitudes of the first- and second-order harmonics are

$$\bar{\alpha}_{-1} = \bar{\alpha}_1 = \frac{\Delta\alpha}{2}. \quad (8.27)$$

We will further simplify the solution of the diffraction problem by assuming the strictly normal incidence ($\theta = 0$, $k_x = 0$, $k_z = k_0$). With this assumption, the matrix D_{nm} of system of Eqs. (8.22) and (8.23) becomes three-diagonal and the amplitudes of the diffracted waves are symmetric with respect to the diffraction order n, i.e. $r_n = r_{-n}$ and $t_n = t_{-n}$.

Let us assume that the period of the grating L is adjusted in such a way that the k-vector of the 1st–order (and 2nd-order) spatial field harmonics matches with the k-vector of GP (see Fig. 8.5(b)): $k_{1x} = -k_{-1x} = k_p$. This condition corresponds to the so-called double resonance diffraction.[49,50,62] The amplitude of the First-order field harmonics (resonant harmonics) will resonantly increase and dominate over the rest of the field harmonics (non-resonant harmonics). Assuming small modulation amplitudes, $|\Delta\alpha| \ll |\bar{\alpha}_0|$, we can use the resonance perturbation theory. The main idea of the resonance perturbation theory consists in retention in the system of Eq. (8.22) the resonant field harmonics at the first place, and then the minimal number of non-resonant field harmonics (originating from the lowest-order scattering of the resonant field harmonics by the diffraction grating). The number of the non-resonant field harmonics (and the contributing scattering processes) are selected with respect to

the desired precision of the final solution (for more detailed description of the resonance perturbation theory in diffraction problems see [Refs. 48, 60, 61]). In our simplest example of the resonance diffraction, the first-order resonant field harmonic will mainly scatter into the neighboring zeroth- and second-order spatial field harmonic (via a single scattering process). Therefore, the system of equations (8.22) will only contain three equations for three variables: t_0, t_1, t_2. These equations take the following form:

$$\begin{aligned} (Y_{10} + Y_{20} + 2\bar{\alpha}_0)t_0 + 2\bar{\alpha}_1 t_1 &= 2Y_i \\ 2\bar{\alpha}_1 t_0 + (Y_{11} + Y_{21} + 2\bar{\alpha}_0)t_1 + 2\bar{\alpha}_1 t_2 &= 0 \\ 2\bar{\alpha}_1 t_1 + (Y_{12} + Y_{22} + 2\bar{\alpha}_0)t_2 &= 0 \end{aligned} \tag{8.28}$$

Solving the system of Eq. (8.28) by Gaussian elimination, and taking into account relation (8.23), we obtain the explicit analytical expressions for the amplitudes of the diffracted waves. The resonant amplitudes (First-order) read as

$$t_{\pm 1} = r_{\pm 1} = \frac{2Y_i}{b_0}\frac{\Delta\alpha}{D}, \tag{8.29}$$

where for compactness of writing, we have introduced the following notations:

$$\begin{aligned} b_n &= Y_{1n} + Y_{2n} + 2\bar{\alpha}_0, \\ D &= b_1 - \Delta\alpha^2\left(\frac{2}{b_0} + \frac{1}{b_2}\right). \end{aligned} \tag{8.30}$$

The second-order non-resonant amplitudes read as

$$t_{\pm 2} = r_{\pm 2} = \frac{2Y_i}{b_0 b_2}\frac{\Delta\alpha^2}{D}. \tag{8.31}$$

Finally (after some simple algebraic transformations), the zeroth-order amplitudes can be written in the form of the sum of the Fresnel coefficients for the unmodulated graphene sheet (r_F, t_F) and the term

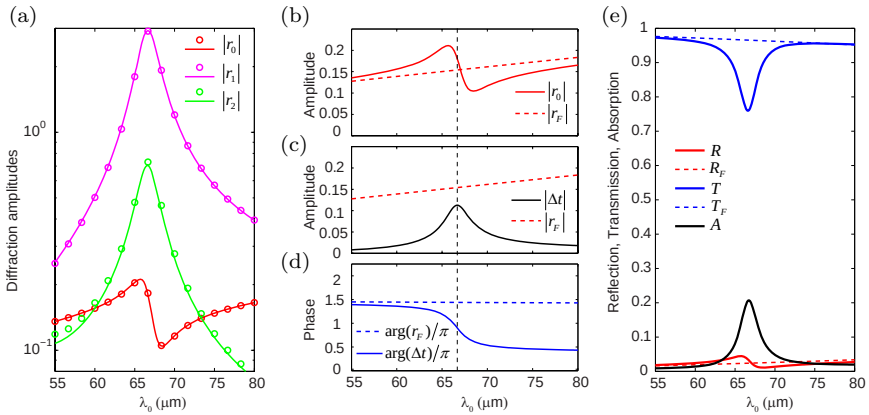


Fig. 8.6. GP resonance on a graphene sheet (surrounded by air, $\varepsilon_1 = \varepsilon_2 = 1$) with periodically modulated conductivity. (a) Amplitudes of the field Fourier harmonics (zeroth-, First- and second-order) as a function of the wavelength, λ_0. The curves correspond to the analytical solution given by equations (8.29)–(8.32), while the circles correspond to the numerical solution of the system of equations (8.22) and (8.23). (b)–(d) Explanation of the Fano-type profile in reflection: amplitudes and phases of r_F, r_0, and Δt as a function of λ_0. (e) Wavelength spectra of the reflection, transmission and absorption coefficients according to Eqs. (8.24) and (8.25). Fresnel transmission, $T_F = |t_F|^2$, and reflection, $R_F = |r_F|^2$, coefficients (according to equation (8.32)) are shown by the dashed curves. In the simulation of the conductivity grating, the following parameters have been used: grating period $L = 10\,\mu$m, modulation amplitude $\Delta\alpha = i\mathrm{Im}(\bar{\alpha}_0)/4$, Fermi energy $E_F = 0.2\,$eV, relaxation time $\tau = 1\,$ps, temperature $T = 300\,$K.

Δt, presenting the grating contribution:

$$r_0 = r_F + \Delta t, \quad t_0 = t_F + \Delta t,$$

$$r_F = \frac{Y_{10} - Y_{20} - 2\bar{\alpha}_0}{Y_{10} + Y_{20} + 2\bar{\alpha}_0}, \quad t_F = \frac{2Y_{10}}{Y_{10} + Y_{20} + 2\bar{\alpha}_0},$$

$$\Delta t = \frac{4Y_i}{b_0^2}\frac{\Delta\alpha^2}{D}. \tag{8.32}$$

The amplitudes r_0, $r_{\pm1}$ and $r_{\pm2}$ for normally incident plane wave onto a free-standing modulated graphene sheet are shown in Fig. 8.6(a). The parameters of the modulated graphene sheet (we assume that only the imaginary part of α is modulated) have been adjusted to have the GP resonance in the THz range, around

$\lambda_0 = 60\,\mu$m (the frequency of 4.5 THz). The analytical solution given by equations (8.29)–(8.32) (shown by the curves) perfectly matches with the numerical solution of the system of Eqs. (8.22) and (8.23) (shown by the circular symbols), in which 20 diffraction orders (from $n = -10$ to $n = 10$) have been taken into account. The amplitudes $r_{\pm 1}$ and $r_{\pm 2}$ clearly show the symmetric bell-shaped resonance profile, while the profile of r_0 is essentially asymmetric (this asymmetry will be discussed below in detail). The amplitude of the First-order amplitude dominates over the rest of the diffraction order amplitudes, justifying our initial assumption on the resonance in the first diffraction order. From the mathematical point of view, the resonance can be explained by small values of the denominator D in Eqs. (8.29)–(8.32), in the vicinity of the resonance wavelength. With the change of the wavelength, λ_0, the value of the GP wavevector k_p (having both real and imaginary parts) becomes very close to k_{1x} (purely real). In other words, with the change of λ_0, the pole of the amplitudes of the spatial field harmonics, $k_{1x} = k_p$, moves above the real axis in the complex plane of k_{1x}, in a close proximity to the point $k_{1x} = 2\pi/L$. In fact, the condition $D = 0$ (never reachable in Eqs. (8.29)–(8.32) due to purely real-valued k_{1x}) provides the dispersion relation of the GPs in the modulated graphene sheet. Neglecting the modulation ($\Delta\alpha = 0$), the equation $D = 0$ simply yields the GP dispersion relation in a graphene sheet placed on the boundary between two different dielectric half-spaces. It reads as $Y_{11} + Y_{21} + 2\bar{\alpha}_0 = 0$, or, more explicitly (changing k_{1x} to k_p)

$$\frac{\varepsilon_1 k_0}{\sqrt{\varepsilon_1 k_0^2 - k_p^2}} + \frac{\varepsilon_2 k_0}{\sqrt{\varepsilon_2 k_0^2 - k_p^2}} + 2\bar{\alpha}_0 = 0. \tag{8.33}$$

Recall that the simplified form of equation (8.33) is given in section 8.2 by Eq. (8.10) (valid when $|\alpha| \ll 1$).

According to Eqs. (8.29)–(8.32), $r_{\pm 1} \propto \Delta\alpha$, $r_{\pm 2} \propto \Delta\alpha^2$. As can be shown by a more detailed analysis of Eqs. (8.22) (being outside of the scope of this chapter), the amplitudes of the diffracted waves (except the zeroth-order one) decay with their diffraction

order as $r_{|n|} \propto \Delta\alpha^{|n|}$, and therefore the amplitudes of the high-order diffracted waves are negligibly small. In contrast, the amplitude of the zeroth-order diffracted waves, r_0 and t_0, are composed, according to Eq. (8.32), by the sum (interference) of the non-resonant terms, r_F and t_F, (reflection and transmission coefficients for the unmodulated graphene sheet, respectively) and the resonant term, Δt, arising due to the backscattering of the resonantly excited GP (in the $\pm 1^{\text{st}}$ diffraction orders) into the zeroth-order diffracted waves. Such an interference can be characterized as a Fano-resonance in which a non-resonant process interferes with the resonant one, producing an asymmetric shape of the resonance profile.[63] Indeed, as one can see in Fig. 8.6(b), the amplitude of the reflected zeroth-order wave, r_0, clearly has an asymmetric shape (solid red curve). The asymmetry is due to (i) similar amplitudes of r_F (dashed red curve in Fig. 8.6(c)) and Δt (solid black curve in Fig. 8.6(c)); and (ii) quick and strong change of the phase difference between r_F and Δt near the resonance wavelength (compare dashed and solid blue curves in Fig. 8.6(d)). Specifically, when the phase difference between r_F and Δt is smaller than $\pi/2$, the total amplitude $|r_0|$ is increased compared to $|r_F|$ (for wavelengths smaller than the resonance one, shown in Fig. 8.6(b)–(d) by a vertical dashed line). On the contrary, because the phase difference between r_F and Δt is larger than $\pi/2$, the total amplitude $|r_0|$ is decreased compared to $|r_F|$ (for wavelengths larger than the resonance one).

We would like to note that the asymmetric line-shape was originally found by Ugo Fano in his theoretical explanation of inelastic scattering of electrons from helium (1961).[64] Although the latter work is widely cited in the context of asymmetric line shapes, much less attention, however, is given to other important works of Fano related with asymmetry of resonances, and being of a similar relevance for Plasmonics. Namely, years earlier (1936–1941), Fano developed a theory which, for the first time, successfully explained strongly asymmetric resonance spectral features appearing due to the excitation of surface plasmon–polaritons in metallic diffraction gratings.[65,66] (Wood anomalies[48,67,68]). Afterwards, it

was realized that asymmetric-shape resonances (called "Fano resonances") present a general wave phenomenon, in which background and a resonant scattering processes interfere. At present, multiple examples of Fano resonances can be found in many areas of physics and engineering.[63]

Figure 8.6(e) shows reflection, R, transmission, T, and absorption, A, coefficients (given by Eqs. (8.24), (8.25)) for the periodically modulated graphene sheet. As a reference, both reflection and transmission coefficients for the unmodulated graphene sheet are also shown (dashed curves in Fig. 8.6(e)). For the chosen parameters of the graphene conductivity and its modulation amplitude, the absorption reaches 20%, while the transmission is well reduced to the value of 75% (compared to 96% for the unmodulated sheet). The reflection in GP resonance remains small, being 4% in its resonance maximum, versus 2% for the unmodulated sheet (at the same wavelength).

The GP resonance can be further enhanced (reducing the minimal transmission and enhancing maximal absorption or reflection) by optimizing the amplitude of the modulation. The optimal modulation amplitude can be estimated with the help of the analytical solution, given by Eqs. (8.29)–(8.32). Since the absorption is proportional to the intensity of the electric field in the graphene sheet, high values of GP-induced absorption (and thus, lower values of the transmission) require high GP field enhancement.

Let us thus find an optimal modulation amplitude yielding the maximum field enhancement of GP. The amplitude of the GP's electric field is given by the amplitude of the first-order spatial field harmonics, $r_{\pm 1}$ (or $t_{\pm 1}$). Let us simplify the expression for $r_{\pm 1}$ (Eq. (8.29)), assuming the air environment of graphene ($\varepsilon_1 = \varepsilon_2 = 1$). To simplify the derivations, it is convenient to introduce the following dimensionless notations for x- and z-components of the wavevectors of the diffracted waves

$$q_n = \frac{k_{nx}}{k_0}, \quad q_{nz} = \frac{k_{1,2nz}}{k_0}. \tag{8.34}$$

Specifically, for 0th-, 1st- and 2nd-diffraction orders, we have explicitly

$$q_0 = 0, \quad q_{\pm 1} = \pm\kappa, \quad q_{\pm 2} = \pm 2\kappa,$$
$$q_{0z} = 1, \quad q_{\pm 1z} = i\sqrt{\kappa^2 - 1}, \quad q_{\pm 2z} = \sqrt{\kappa^2 - 1}, \tag{8.35}$$

where $\kappa = G/k_0 = \lambda_0/L$. Then, according to the definition of the admittances, $Y_{1,2n}$, by using Eqs. (8.34), (8.35), we have the explicit expressions for them:

$$Y_{1,20} = 1, \quad Y_{1,2\pm 1} = \frac{1}{i\sqrt{\kappa^2 - 1}}, \quad Y_{1,2\pm 1} = \frac{1}{i\sqrt{4\kappa^2 - 1}}. \tag{8.36}$$

The coefficients b_n appearing in Eq. (8.30) can be now simplified to:

$$b_0 = 2(\bar{\alpha}_0 + 1), \;\; b_1 = 2\left(\bar{\alpha}_0 - \frac{i}{\sqrt{\kappa^2 - 1}}\right), \;\; b_2 = 2\left(\bar{\alpha}_0 - \frac{i}{\sqrt{4\kappa^2 - 1}}\right). \tag{8.37}$$

Taking into account Eqs. (8.35), (8.36), the amplitudes of the first-order spatial field harmonics, given by Eq. (8.29), become

$$r_{\pm 1} = \frac{-1}{b_0}\frac{\Delta\alpha q_{1z}}{1 + q_{1z}\Delta}, \quad \Delta = \bar{\alpha}_0 - \Delta\alpha^2\left(\frac{1}{b_0} + \frac{1}{2b_2}\right). \tag{8.38}$$

The maximal value of $|r_{\pm 1}|^2$ is achieved when the real part of the denominator in Eq. (8.38) takes zero value:

$$1 + \mathrm{Re}(q_{1z}\Delta) = 0. \tag{8.39}$$

Under the condition given by Eq. (8.39), taking into account that q_{1z} is purely imaginary for subwavelength periods of the grating, $L < \lambda_0$, the maximal value of $|r_{\pm 1}|^2$ becomes

$$|r_{\pm 1}|^2_{\max} = \left|\frac{\Delta\alpha}{b_0}\right|^2 \frac{1}{\mathrm{Re}(\Delta)^2}. \tag{8.40}$$

The maximal value of $|r_{\pm 1}|^2$ given by Eq. (8.40) corresponds to a certain fixed modulation amplitude, $\Delta\alpha$. Now, we can optimize the modulation amplitude (find the maximum of $|r_{\pm 1}|^2_{\max}$ considering it

as a function of $\Delta\alpha$). In order to optimize the modulation amplitude, let us further simplify Eq. (8.40), using the smallness of $\bar{\alpha}_0$ and assuming that κ approximately corresponds to the GP resonance in neglect of the modulation, i.e. $\kappa \simeq \mathrm{Re}(q_p) = 1/\mathrm{Im}(\bar{\alpha}_0)$ (see equation (8.7)). With these assumptions, taking into account that only the imaginary part of the conductivity is modulated, $\Delta\alpha = i|\Delta\alpha|$, we have $b_0 \simeq 2$, $b_2 = -i/\mathrm{Im}(\bar{\alpha}_0)$, so that $\mathrm{Re}(\Delta) \simeq \mathrm{Re}(\bar{\alpha}_0) + |\Delta\alpha|^2/2$, and Eq. (8.40) simplifies as

$$|r_{\pm 1}|^2_{\max} = \frac{|\Delta\alpha|^2}{(2\mathrm{Re}(\bar{\alpha}_0) + |\Delta\alpha|^2)^2}. \tag{8.41}$$

From Eq. (8.41), we see that $|r_{\pm 1}|^2_{\max}$ reaches its maximal (optimal) value at $|\Delta\alpha|^2 = 2\mathrm{Re}(\bar{\alpha}_0)$. Therefore, the optimal modulation amplitude, $\Delta\alpha_{\mathrm{opt}}$, reads:

$$\Delta\alpha_{\mathrm{opt}} = i\sqrt{2\mathrm{Re}(\bar{\alpha}_0)}. \tag{8.42}$$

At the modulation amplitude, $\Delta\alpha = \Delta\alpha_{\mathrm{opt}}$, the optimal value for $|r_{\pm 1}|^2$ depends exclusively upon the intrinsic losses in the unmodulated graphene sheet:

$$|r_{\pm 1}|^2_{\mathrm{opt}} = \frac{1}{8\mathrm{Re}(\bar{\alpha}_0)}, \tag{8.43}$$

while the resonant term in Eq. (8.32) takes approximately the constant value, $\Delta t \simeq -1/2$. Then, approximating the Fresnel coefficients by $r_F \simeq 0$ and $t_F \simeq 1$, from Eq. (8.32) we find the optimal values for transmission, reflection and absorption coefficients:

$$T_{\mathrm{opt}} = R_{\mathrm{opt}} = \frac{1}{4}, \quad A_{\mathrm{opt}} = \frac{1}{2}. \tag{8.44}$$

As can be shown by a more general mathematical analysis (see Refs. [21, 32]) the optimal values given by Eq. (8.44), correspond to the highest possible absorption by a symmetrically surrounded graphene sheet with an arbitrary spatial variation of the conductivity.

Note that the optimal modulation amplitude is attributed to the compromise between the radiation losses of the GP on the grating (given by the coupling to the zeroth–order diffracted wave)

and intrinsic damping of the unmodulated graphene sheet, given by the real part of its conductivity. Importantly, the optimal value of the modulation amplitude can easily violate the condition of the validity for the resonance perturbation theory ($\Delta\alpha < |\bar{\alpha}_0|$), since for moderate and, particularly, high intrinsic, losses we have $\sqrt{\mathrm{Re}(\bar{\alpha}_0)} \sim \mathrm{Im}(\bar{\alpha}_0)$. In this case, for modulation amplitudes close to $\Delta\alpha_{\mathrm{opt}}$, the analytical Eqs. (8.29)–(8.32) cannot be used for a quantitative description of the diffraction effects. Nevertheless, even in the regime when the validity of the analytical perturbation theory is restricted, optimal values given by Eqs. (8.42)–(8.44) can be a good estimation for the efficiency of the GP resonance in close to the optimal conditions.

8.6. Conclusions

In this chapter, we have tried to summarize recently developed concepts of graphene plasmonics. We have provided both a brief review on some of the most relevant physical effects in the topic and presented some background information. Therefore, we hope that the content of this chapter can be of interest both for specialists and for students or postdocs wishing to start working in the field.

Acknowledgments

We gratefully acknowledge Prof. Rainer Hillenbrand and Prof. Luis Martín-Moreno for reading this Chapter and making useful comments. We also acknowledge the financial support from European Commission under the Graphene Flagship (contract no. CNECTICT-604391) and the Spanish Ministry of Economy and Competitiveness (national project MAT2014-53432-C5-4-R).

References

1. Grigorenko, A. N., Polini, M. and Novoselov, K. S. (2012). *Nat. Photon* **6**, p. 749.

2. Cai, X., Sushkov, A. B., Jadidi, M. M., Nyakiti, L. O., Myers-Ward, R. L., Gaskill, D. K., Murphy, T. E., Fuhrer, M. S. and Drew, H. D. (2015). *Nano. Lett.* **15**, p. 4295.
3. Francescato, Y., Giannini, V., Yang, J., Huang, M. and Maier, S. A. (2014). *ACS Photonics* **1**, p. 437.
4. Rodrigo, D., Limaj, O., Janner, D., Etezadi, D., García de Abajo, F. J., Pruneri, V. and Altug, H. (2015). *Science* **349**, p. 165.
5. Hai Hu, Xiaoxia Yang, Feng Zhai, Debo Hu, Ruina Liu, Kaihui Liu, Zhipei Sun and Qing Dai (2016). *Nature Communications* **7**, p. 12334.
6. Andrea Marini, Iván Silveiro, and F. Javier García de Abajo (2015). *ACS Photonics* **2**(7), pp. 876–882.
7. Damon B. Farmer, Phaedon Avouris, Yilei Li, Tony F. Heinz and Shu-Jen Han (2016). *ACS Photonics* **3**(4), pp. 553–557.
8. Yilei Li, Hugen Yan, Damon B. Farmer, Xiang Meng, Wenjuan Zhu, Richard M. Osgood, Tony F. Heinz and Phaedon Avouris (2014). *Nano Lett.* **14**(3), pp. 1573–1577.
9. Vakil, A. and Engheta, N. (2011). *Science* **332**, p. 1291.
10. Raether, H. (ed.) (1988). *Surface Plasmons on Smooth and Rough Surfaces and on Gratings* (Springer Berlin, Heidelberg).
11. Yan, H., Low, T., Zhu, W., Wu, Y., Freitag, M., Li, X., Guinea, F., Avouris, P. and Xia, F. (2013). *Nat Photon* **7**, p. 394.
12. Brar, V. W., Jang, M. S., Sherrott, M., Kim, S., Lopez, J. J., Kim, L. B., Choi, M. and Atwater, H. (2014). *Nano Lett.* **14**, p. 3876.
13. Zhu, X., Wang, W., Yan, W., Larsen, M. B., Bøggild, P., Pedersen, T. G., Xiao, S., Zi, J. and Mortensen, N. A. (2014). *Nano Lett.* **14**, p. 2907.
14. Bludov, Y. V., Ferreira, A., Peres, N. M. R. and Vasilevskiy, M. I. (2013). *Int. J. Mod. Phys. B* **27**, p. 1341001.
15. García de Abajo, F. J. (2014). *ACS Photonics* **1**, p. 135.
16. Bao, Q. and Loh, K. P. (2012). *ACS Nano.* **6**, p. 3677.
17. Xia, F., Wang, H., Xiao, D., Dubey, M. and Ramasubramaniam, A. (2014). *Nat Photon* **8**, p. 899.
18. Low, T. and Avouris, P. (2014). *ACS Nano* **8**, p. 1086.
19. Jablan, M., Soljacic, M. and Buljan, H. (2013). *Proce. IEEE* **101**, p. 1689.
20. Politano, A. and Chiarello, G. (2014). *Nanoscale* **6**, p. 10927.
21. Thongrattanasiri, S., Koppens, F. H. L. and García de Abajo, F. J. (2012). *Phys. Rev. Lett.* **108**, p. 047401.
22. Nikitin, A. Y., Guinea, F., Garcia-Vidal, F. J. and Martin-Moreno, L. (2012). *Phys. Rev. B* **85**, p. 081405.
23. Nikitin, A. Y., Guinea, F., García-Vidal, F. J. and Martín-Moreno, L. (2011). *Phys. Rev. B* **84**, p. 161407.
24. Christensen, J., Manjavacas, A., Thongrattanasiri, S., Koppens, F. H. L. and García de Abajo, F. J. (2012). *ACS Nano* **6**, p. 431.
25. Smirnova, D. and Kivshar, Y. S. (2014). *Phys. Rev. B* **90**, p. 165433.
26. Cox, J. D. and Javier García de Abajo, F. (2014). *Nat. Commun.* **5**.

27. Koppens, F. H. L., Chang, D. E. and Garcìa de Abajo, F. J. (2011). *Nano. Lett.* **11**, p. 3370.
28. Garcia-Pomar, J. L., Nikitin, A. Y. and Martin-Moreno, L. (2013). *ACS Nano.* **7**, p. 4988.
29. Chen, J. *et al.* (2013). *Nano Lett.* **13**, p. 6210.
30. FeiZ *et al.* (2013). *Nat. Nano.* **8**, p. 821.
31. Crassee, I., Orlita, M., Potemski, M., Walter, A. L., Ostler, M., Seyller, T., Gaponenko, I., Chen, J. and Kuzmenko, A. B. (2012). *Nano Lett.* **12**, p. 2470.
32. Tymchenko, M., Nikitin, A. Y. and Martín-Moreno, L. (2013). *ACS Nano.* **7**, p. 9780.
33. Hadad, Y. and Steinberg, B. Z. (2013). *Phys. Rev. B* **88**, p. 075439.
34. Alonso-González, P. *et al.* (2014). *Science* **344**, p. 1369.
35. Chen, J. *et al.* (2012). *Nature* **487**, p. 77.
36. Fei, Z. *et al.* (2012). *Nature* **487**, p. 82.
37. Falkovsky, L. A. (2008). *Phys. Uspekhi* **51**, p. 887.
38. Hwang, E. H. and Das Sarma, S. (2007). *Phys. Rev. B* **75**, p. 205418.
39. Wunsch, B., Stauber, T., Sols, F. and Guinea, F. (2006). *New J. Phys.* **8**, p. 318.
40. Hanson, G. W. (2008). *J. Appl. Phys.* **103**, p. 064302.
41. Jablan, M., Buljan, H. and Soljačić, M. (2009). *Phys. Rev. B* **80**, p. 245435.
42. Wang, L. *et al.* (2013). *Science* **342**, p. 614.
43. Woessner, A. *et al.* (2015). *Nat. Mater.* **14**, p. 421.
44. Jackson, J. D. (1999). *Classical Electrodynamics* (Wiley, New York).
45. Agranovich, V. M. and Mills, D. L. (eds.) (1982). *Surface Polaritons : Electromagnetic Waves at Surfaces and Interfaces* (North-Holland Pub. Co., Amsterdam).
46. Bludov, Y. V., Vasilevskiy, M. I. and Peres, N. M. R. (2010). *Europhys. Lett.* **92**, p. 68001.
47. Gonçalves, P. A. D., Peres, N. M. R. (2016). *An Introduction to Graphene Plasmonics* (World Scientific Publishing Company).
48. Maystre, D., in *Plasmonics*, (2012). S. Enoch and N. Bonod (eds.) (Springer Berlin Heidelberg), Vol. 167, p. 39.
49. Slipchenko, T. M., Nesterov, M. L., Martin-Moreno, L. and Nikitin, A. Y. (2013). *J. Opti.* **15**, p. 114008.
50. Kats, A. V., Nesterov, M. L. and Nikitin, A. Y. (2007). *Phys. Rev. B* **76**, p. 045413.
51. Nikitin, A. Y., Guinea, F., Garcia-Vidal, F. J. and Martin-Moreno, L. (2011). *Phys. Rev. B* **84**, p. 195446.
52. Velizhanin, K. A. and Efimov, A. (2011). *Phys. Rev. B* **84**, p. 085401.
53. Gómez-Santos, G. and Stauber, T. (2011). *Phys. Rev. B* **84**, p. 165438.
54. Hanson, G. W., Yakovlev, A. B. and Mafi, A. (2011). *J. Appl. Phys.* **110**, p. 114305.

55. Huidobro, P. A., Nikitin, A. Y., González-Ballestero, C., Martín-Moreno, L. and García-Vidal, F. J. (2012). *Phys. Rev. B* **85**, p. 155438.
56. Nikitin, A. Y., Alonso-González, P., Vélez, Mastel, S., Centeno, A., Pesquera, A., Zurutuza, A., Casanova, F., Hueso, L. E., Koppens, F. H. L. and Hillenbrand, R. (2016). *Nature Photonics* **10**, pp. 239–243.
57. Ju, L. *et al.* (2011). *Nat. Nano.* **6**, p. 630.
58. Yan, H., Li, X., Chandra, B., Tulevski, G., Wu, Y., Freitag, M., Zhu, W., Avouris, P. and Xia, F. (2012). *Nat. Nano.* **7**, p. 330.
59. Fang, Z. *et al.* (2013). *ACS Nano.* **7**, p. 2388.
60. Fang, Z., Wang, Y., Schlather, A. E., Liu, Z., Ajayan, P. M., García de Abajo, F. J., Nordlander, P., Zhu, X. and Halas, N. J. (2014). *Nano Lett.* **14**, p. 299.
61. Nikitin, A. Y., Alonso-González, P. and Hillenbrand, R. (2014). *Nano Lett.* **14**, p. 2896.
62. Kats, A. V., Spevak, I. S. and Balakhonova, N. A. (2007). *Phys. Rev. B* **76**, p. 075407.
63. Miroshnichenko, A. E., Flach, S. and Kivshar, Y. S. (2010). *Rev. Mod. Phys.* **82**, p. 2257.
64. Fano, U. (1961). *Phys. Rev.* **124**, p. 1866.
65. Fano, U. (1936). *Phys. Rev.* **50**, p. 573.
66. Fano, U. (1941). *J. Opt. Soc. Am.* **31**, p. 213.
67. Wood, R. W. (1902). *Proc. Phys. Soc. London* **18**, p. 269.
68. Wood, R. W. (1935). *Phys. Rev.* **48**, p. 928.

CHAPTER 9

Plasmon Excitation by Fast Electrons

ULRICH HOHENESTER

University of Graz, Austria
ulrich.hohenester@uni-graz.at

9.1. Introduction

Plasmonics deals with light confinement at the nanoscale.[1] This is achieved by binding light to coherent electron oscillations at the surfaces of metallic nanoparticles, the so-called surface plasmons (SPs) or surface plasmon polaritons (SPPs). These SPs come together with large field enhancements and evanescent fields in the vicinity of metallic nanoparticles,[2] which allow light confinement to subdiffraction volumes.[3] Besides being of fundamental interest, this topic holds promise for a variety of photonics applications, such as optical communication and storage[4] or quantum optics.[5]

Although plasmonics forms a bridge between the micrometer scale of optics and the nanometer scale of nanostructures, the diffraction limit of light forbids in conventional optical microscopy the direct observation of light confinement with nanometer resolution. An approach to circumvent this shortcoming is to use instead of optical (photon) probes other probes, such as electrons, thereby entering the field of electron microscopy. As both photons and

electrons interact with plasmonic nanoparticles through the electrodynamic fields, both measurement techniques provide related information.

Electron energy loss spectroscopy (EELS) is a subfield of electron microscopy that has proven particularly useful in the context of plasmon field mapping. The basic principle is sketched in Fig. 9.1: Electrons with kinetic energies of typically 100 keV (corresponding to velocities comparable to the speed of light) pass by or penetrate through the metallic nanoparticle, excite particle plasmons, and lose energy. Finally, the energy loss of the electrons is detected and recorded. By raster-scanning the electron beam over the sample, one obtains a spatial map of the plasmon fields with nanometer resolution.

In the dawning of plasmonics, EELS was used for the first observation of bulk and SPs[6,7] (see also Ref. 8 for a brief historical sketch)

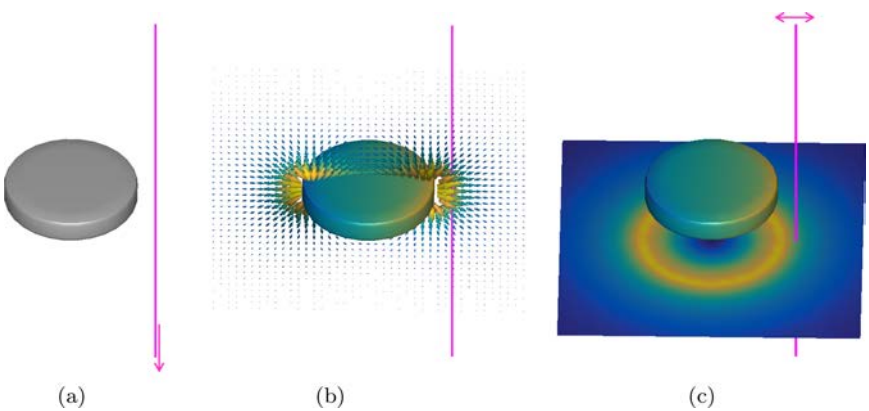


Fig. 9.1. Schematics of energy loss of fast electron. (a) A fast electron (see magenta line) passes by or penetrates through a metallic nanoparticle, here a silver nanodisk. (b) The electron excites a dipolar particle plasmon, which comes together with a localized electric field, and performs work against the induced field, thereby loosing a fraction of its kinetic energy. (c) By raster scanning the electron beam over the sample and recording the energy loss probability, one obtains a map of the plasmonic field distribution with nanometer spatial resolution and with sub-eV energy resolution.

and was later established as a unique characterization tool for the investigation of ensembles of plasmonic nanoparticles.[9] Initially, further progress was hindered by the lack of versatile fabrication tools for metallic nanoparticle and the limited energy resolution of EELS, which made it hard to resolve particle plasmons of a few eV in noble-metal nanoparticles such as gold or silver. The second generation of EELS measurements of plasmonic nanoparticles started with the pioneering studies of nanorods and nanotriangles,[10,11] which reported the mapping of plasmon mode distributions with nanometer spatial and sub-eV energy resolution, opening a most detailed view to the plasmonics world that was up to that point only accessible through computer simulations.

This chapter provides an overview of how such EELS mapping is performed, how it can be interpreted and simulated, and what has been achieved so far. Let me start with a disclaimer of what the paper is not about. Being a theoretical physicist, I have refrained from discussing experimental details. Excellent reviews about electron microscopy and EELS for plasmonic nanoparticles exist (see e.g. Refs. 12 and 13 and references therein), and the interested reader is referred to the pertinent literature. From the viewpoint of a theoretician, it is amazing to see how much the field of electron microscopy has matured in recent years. Electron microscopes nowadays come as true technical masterpieces, however, also hand in hand with enormous costs that force research groups to join efforts and make measurement times scarce and precious. In most of the joint studies where I have been involved, only a few measurement series were performed, and often it takes several weeks from the planning stage to the experiment. Yet, the true heroes of the trade remain those who actually perform the experiments. In particular with EELS for plasmonic nanoparticles, several critical points remain, such as sample preparation, sufficiently high-energy resolution, suppression of contamination during the measurement process, and post-processing of the recorded data. Thus, results that appear to come out of routine measurements are usually based on extremely hard work.

So this chapter deals more with theory, simulation, and results, although I had to realize with despair that there already exists an excellent review article by Javier García de Abajo[8] that covers almost everything. What I thus provide in the following is a short introduction to EELS with plasmonic nanoparticles, with a focus on theory and simulation, probably with more equations than needed. Section 9.2 introduces the basic equations. In section 9.3, I present the most common simulation approaches together with a few representative results to highlight the principles underlying EELS for plasmonic nanoparticles. The quasistatic approximation and an eigenmode expansion are discussed in section 9.4. Section 9.5 is the main part of this work and provides an overview of what has been done with EELS and what can be learned from EELS measurements. This section contains several references, which, however, are neither exhaustive nor complete. Finally, in section 9.6, I give a brief summary and an outlook to possible future developments.

9.2. Theory

In this section, we derive the theory for computing EELS probabilities. Our derivation closely follows the review article of García de Abajo.[8] To facilitate comparison with this work, in the following, we adopt a Gaussian unit system (conversion to SI units can be done along the guidelines given in Ref. 14) and we only consider nonmagnetic materials with a permeability $\mu = 1$. As shown in Fig. 9.1, in EELS a swift electron passing by or penetrating through a metallic nanoparticle (i) excites an SP, (ii) performs work against the induced SP field, and finally (iii) the electron's loss in kinetic energy is measured. We will next show how to account for the various steps within a semiclassical framework.

Within a small time interval $\mathrm{d}t$, an electron propagates the distance $\mathrm{d}\boldsymbol{\ell} = \boldsymbol{v}\,\mathrm{d}t$, where $\boldsymbol{v}$ is the electron velocity. Let $\boldsymbol{r}_e(t)$ denote the electron position, and $\boldsymbol{E}[\boldsymbol{r}_e(t)]$ and $\boldsymbol{B}[\boldsymbol{r}_e(t)]$ the electric and magnetic fields at the electron position, respectively. The work performed by

the electron against the electric field is then given by

$$\mathrm{d}W = -q\Big(\boldsymbol{E}[\boldsymbol{r}_e(t)] + \boldsymbol{v}\times\boldsymbol{B}[\boldsymbol{r}_e(t)]\Big)\cdot\boldsymbol{v}\,\mathrm{d}t = -q\boldsymbol{E}[\boldsymbol{r}_e(t)]\cdot\boldsymbol{v}\,\mathrm{d}t, \qquad (9.1)$$

where $q = -e$ is the charge of the electron. We have exploited the fact that magnetic fields cannot directly perform work[14] because of $(\boldsymbol{v}\times\boldsymbol{B})\cdot\boldsymbol{v} \equiv 0$. The total work performed by the electron, which corresponds to the energy loss ΔE, is then obtained by integrating over the entire electron trajectory,

$$\Delta E = e\int_{-\infty}^{\infty}\boldsymbol{E}[\boldsymbol{r}_e(t)]\cdot\boldsymbol{v}\,\mathrm{d}t. \qquad (9.2)$$

At this point, we have to be more specific about two points. First, the electric field entering Eq. (9.2) is the electric field felt by the electron, which, however, does not include the field produced by the electron itself.[a] Second, from now on we will assume that for swift electrons, with kinetic energies in the range from several tens to hundreds of keV, the electron's trajectory is not noticeably modified by the energy lost through plasmon excitation. We thus assume $\boldsymbol{r}_e(t) = \boldsymbol{r}_0 + \boldsymbol{v}\,t$. For an electron propagating along the positive z-direction, which will be considered if not noted differently, the electron trajectory becomes

$$\boldsymbol{r}_e(t) = \boldsymbol{R}_0 + \hat{\boldsymbol{z}}\,vt, \qquad (9.3)$$

with the *impact parameter* $\boldsymbol{R}_0 = (x_0, y_0)$. There is a subtle point about Eq. (9.3) that needs some clarification: although it is precisely the velocity change that allows to detect the electron's energy loss in experiment (through deflection in the magnetic field of the detector), in the theoretical approach the energy losses are already described through the integral expression of Eq. (9.2), accounting for the work

[a]The field produced by the electron and felt by the same electron would diverge for a point-like particle such as the electron. Self-interaction divergences are commonly treated in the field of quantum electrodynamics, but are always neglected in classical electrodynamics.

performed by the electron, and the consideration of minor trajectory changes would only lead to very small additional corrections. The energy loss ΔE of Eq. (9.2) can be spectrally decomposed into different frequency components using the Fourier transform

$$\boldsymbol{E}(\boldsymbol{r},t) = \int_{-\infty}^{\infty} e^{-i\omega t}\boldsymbol{E}(\boldsymbol{r},\omega)\,\frac{\mathrm{d}\omega}{2\pi}, \tag{9.4}$$

where ω is the angular frequency. Inserting this expression into Eq. (9.2) and using the definition $\Delta E = \int_0^\infty \hbar\omega\,\Gamma_{\mathrm{EELS}}(\boldsymbol{R}_0,\omega)\,\mathrm{d}\omega$ for the energy loss probability, with $\hbar\omega$ being the energy loss, we finally arrive at

$$\Gamma_{\mathrm{EELS}}(\boldsymbol{R}_0,\omega) = \frac{e}{\pi\hbar\omega}\int_{-\infty}^{\infty} \mathrm{Re}\left\{e^{-i\omega t}\boldsymbol{v}\cdot\boldsymbol{E}[\boldsymbol{r}_e(t),\omega]\right\}\,\mathrm{d}t, \tag{9.5}$$

where we have used $\boldsymbol{E}^*(\boldsymbol{r},\omega) = \boldsymbol{E}(\boldsymbol{r},-\omega)$, which directly follows from Eq. (9.4) for real electric fields $\boldsymbol{E}(\boldsymbol{r},t)$. Equation (9.5) is the central result of this section which allows us to compute the electron energy loss probability Γ_{EELS} once the electric field $\boldsymbol{E}(\boldsymbol{r},\omega)$ has been computed.

9.2.1. *Fields produced by swift electrons*

The charge distribution of a swift electron propagating along the trajectory of Eq. (9.3) is

$$\rho(\boldsymbol{r},t) = -e\delta(\boldsymbol{R}-\boldsymbol{R}_0)\delta(z-vt), \tag{9.6}$$

where $\boldsymbol{R} = (x,y)$ are the in-plane coordinates of the electron. The corresponding current distribution is $\boldsymbol{J}(\boldsymbol{r},t) = \boldsymbol{v}\rho(r,t)$. For the solution of Maxwell's equations in frequency space, we need the Fourier transform of Eq. (9.6), which reads

$$\rho(\boldsymbol{r},\omega) = \int e^{i\omega t}\rho(\boldsymbol{r},\omega)\,\mathrm{d}t = -\frac{e}{v}\delta(\boldsymbol{R}-\boldsymbol{R}_0)e^{iqz}, \tag{9.7}$$

where we have introduced the wavenumber $q = \omega/v$. For an unbounded medium with a background permittivity $\varepsilon(\omega)$, the electric field associated with $\rho(\boldsymbol{r},\omega)$ can be obtained from the solutions of Maxwell's equations. The calculation is explicitly worked out in

Ref. 14 (section 9.14) in the context of the Liénard–Wiechert potentials, by starting from the fields for an electron at rest and then boosting the electron to velocity v through a (relativistic) Lorentz transformation, and we get[8]

$$\boldsymbol{E}(\boldsymbol{r},\omega) = \frac{2e\omega}{v^2\gamma_\varepsilon\varepsilon} e^{iqz}\left[\frac{i}{\gamma_\varepsilon}K_0\left(\frac{\omega\rho}{v\gamma_\varepsilon}\right)\hat{\boldsymbol{z}} - K_1\left(\frac{\omega\rho}{v\gamma_\varepsilon}\right)\hat{\boldsymbol{\rho}}\right], \tag{9.8}$$

where $\gamma_\varepsilon = 1/\sqrt{1-\varepsilon v^2/c^2}$ is the Lorentz contraction factor, and the notation $\boldsymbol{\rho} = \boldsymbol{R} - \boldsymbol{R}_0$ has been employed. K_0 and K_1 are the modified Bessel functions of order 0 and 1, respectively. For completeness, we also give the expression for the scalar potential within the Lorentz gauge condition,

$$\phi(\boldsymbol{r},\omega) = -\frac{2e}{v\varepsilon} e^{iqz}\, K_0\left(\frac{\omega\rho}{v\gamma_\varepsilon}\right), \tag{9.9}$$

which will be used later in the context of the boundary element method (BEM) approach. The vector potential is given by $\boldsymbol{A}(\boldsymbol{r},\omega) = \varepsilon\frac{\boldsymbol{v}}{c}\phi(\boldsymbol{r},\omega)$.

Figure 9.2 shows the z and radial components of the electric field, Eq. (9.8). As can be seen, the radial component is much larger than the z component. For sufficiently small radial distances ρ, the radial component shows a $1/\rho$ dependence reminiscent of the electric field produced by a charged wire. At large distances, the electric field decays exponentially.

9.2.2. *Decomposition into bulk and surface losses*

It is convenient to split the energy loss probability of Eq. (9.5) into bulk and boundary contributions.[12] In general, if the electric field $\boldsymbol{E}(\boldsymbol{r},\omega)$ produced by the electron is at hand (e.g. as obtained from simulations) such splitting is not mandatory, but it helps to get more insight into the different loss channels. Bulk losses arise when the electron propagates through a medium described by a dielectric function $\varepsilon(\omega)$. They only depend on the dielectric properties of the material and the distance L propagated by the swift electron through

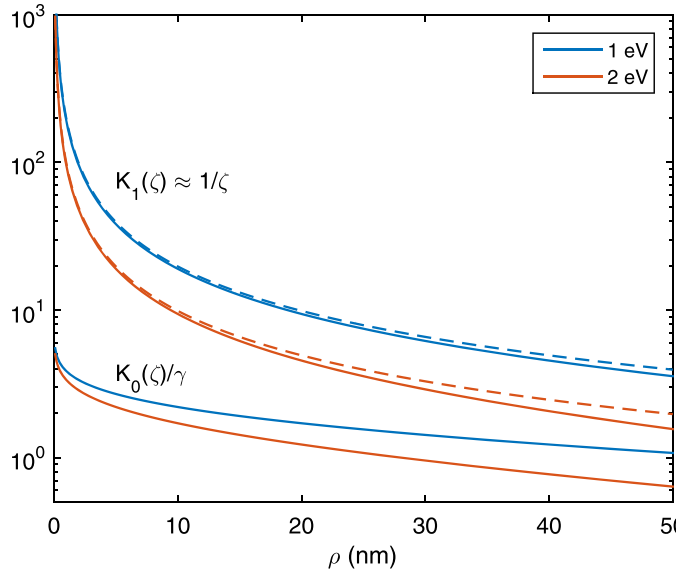


Fig. 9.2. Electric field components of Eq. (9.8) in z (K_0) and radial (K_1) directions, and for two selected loss energies. We use the abbreviation $\zeta = \omega\rho/(v\gamma_\varepsilon)$. The dashed lines report the $1/\rho$ dependence for the radial electric field of a charged wire, which agrees well with K_1 at sufficiently small distances.

this material, but not on the nanoparticle geometry[8]

$$\Gamma_{\text{bulk}}(\omega) = \frac{e^2 L}{\pi\hbar v^2}\,\text{Im}\left\{\left(\frac{v^2}{c^2} - \frac{1}{\varepsilon}\right)\ln\left(\frac{q_c^2 - k^2\varepsilon}{q^2 - k^2\varepsilon}\right)\right\}, \tag{9.10}$$

where $q_c \approx \sqrt{(mv\varphi_{\text{out}}/\hbar)^2 + q^2}$ is a cut-off frequency that is determined by the half-aperture collection angle of the microscope spectrometer φ_{out}. In the non-retarded limit $c \to \infty$, this expression reduces to

$$\Gamma_{\text{bulk}}^{\text{NR}}(\omega) = \frac{2e^2 L}{\pi\hbar v^2}\,\text{Im}\left\{-\frac{1}{\varepsilon}\right\}\ln\left(\frac{q_c}{q}\right). \tag{9.11}$$

As an example, we consider a Drude type permittivity $\varepsilon(\omega) = 1 - \omega_p^2/[\omega(\omega+i\gamma)]$ with ω_p and γ being the plasma frequency and collision frequency of the free-electron gas. In this case, the loss function can

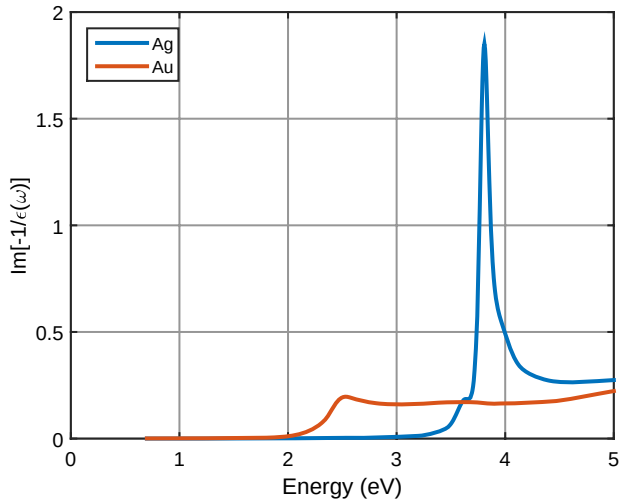


Fig. 9.3. Loss function $\mathrm{Im}[-1/\varepsilon(\omega)]$ for silver (Ag) and gold (Au), respectively, for the dielectric function tabulated in Ref. 15. For silver, the loss function is dominated by a single peak associated with bulk plasmon excitations, for gold one observes a broad distribution above 2 eV associated with d-band absorption.[1]

be evaluated explicitly and we obtain,

$$\mathrm{Im}\left\{-\frac{1}{\varepsilon}\right\} = \frac{2\omega\gamma\omega_p^2}{(\omega_p^2-\omega^2)^2+4\omega^2\gamma^2}.$$

This expression corresponds to a Lorentzian peaked at the plasma frequency $\omega \approx \omega_p$ and broadened by the collision frequency γ. For more realistic dielectric functions, such as those extracted from optical experiments,[15] we find similar shapes with practically no bulk losses in the low-energy regime relevant for particle plasmons, and a pronounced peak at the bulk plasmon energy, see Fig. 9.3. In case of gold, one observes a broad distribution above 2 eV associated with d-band absorption.[1]

We next consider the situation where the electron propagates through a system composed of different materials, described through homogeneous dielectric functions $\varepsilon_j(\omega)$, which are separated by sharp boundaries $\partial\Omega_j$. We split the different loss channels into bulk and boundary contributions, with $\Gamma_{\mathrm{bulk},j}(\omega)$ being given by Eq. (9.10),

and the boundary losses are often further separated into *begrenzung* (German expression for restriction) and *surface* losses.[12] The distinction between these contributions is somewhat subtle. In short, the begrenzung losses are due to the fact that dielectric effects entering Eq. (9.10) become modified in the vicinity of boundaries, for instance, due to a less efficient screening close to metal boundaries, and the modification of the loss probabilities is accounted for through an additional begrenzung term. All remaining contributions, in particular those originating from SPs, are denoted as surface losses. Below, we will show how these different losses can be accounted for within a simulation approach.

9.2.3. *Expressing EELS losses in terms of the dyadic Green function*

There is a formal and convenient connection between the EELS loss of Eq. (9.5) and the dyadic Green function, which was first stated and analyzed in Ref. 16. Let us first introduce the dyadic Green function, which rests on the following concept: as we are dealing with Maxwell's equations in the linear regime, it suffices to seek for the solution of delta-like current sources $\delta(\boldsymbol{r}-\boldsymbol{r}')\mathbb{1}$ subject to suitable boundary conditions, in general out-going waves at infinity.[14] Such a delta-like source is a mathematical idealization, since any meaningful current distribution must additionally comply with the continuity equation. Nevertheless, once we have obtained the solution of Maxwell's equations for a delta-like current distribution, we can immediately obtain the solution for a general current source $\boldsymbol{J}(\boldsymbol{r},\omega)$ through linear superposition of these delta-like sources.

To be more precise, we start from the wave equation for the electric field and for delta-like sources,

$$\nabla\times\nabla\times\boldsymbol{G}(\boldsymbol{r},\boldsymbol{r}',\omega)-k_0^2\varepsilon(\boldsymbol{r},\omega)\boldsymbol{G}(\boldsymbol{r},\boldsymbol{r}',\omega)=-\frac{1}{c^2}\delta(\boldsymbol{r}-\boldsymbol{r}')\mathbb{1}, \quad (9.12)$$

where $k_0=\omega/c$ is the wavenumber of light in vacuum and $\boldsymbol{G}(\boldsymbol{r},\boldsymbol{r}',\omega)$ is the *dyadic Green function*. It describes the electric field at position $\boldsymbol{r}$ for a delta-like current source at position $\boldsymbol{r}'$, oscillating with

frequency ω. Additionally, $\boldsymbol{G}$ is assumed to have the proper boundary conditions of out-going waves at infinity (retarded Green function[14]). With the dyadic Green function of Eq. (9.12), we can immediately write down a formal solution for a general current distribution $\boldsymbol{J}(\boldsymbol{r},\omega)$ as

$$\boldsymbol{E}(\boldsymbol{r},\omega) = -4\pi i\omega \int \boldsymbol{G}(\boldsymbol{r},\boldsymbol{r}',\omega)\cdot\boldsymbol{J}(\boldsymbol{r}',\omega)\,\mathrm{d}\boldsymbol{r}'. \tag{9.13}$$

Equation (9.13) follows from the defining Eq. (9.12) through a linear superposition of the delta-like sources, and thus also has the proper boundary conditions. For this reason, once we have determined the dyadic Green function of Eq. (9.12) for a given frequency ω and dielectric environment $\varepsilon(\boldsymbol{r},\omega)$, we have the solution of Maxwell's equation through Eq. (9.13) at hand.

The dyadic Green function can be immediately employed for the calculation of the energy loss probability of Eq. (9.5),

$$\Gamma_{\mathrm{EELS}}(\boldsymbol{R}_0,\omega) = -\frac{4ev^2}{\hbar}\int_{-\infty}^{\infty} \mathrm{Im}\left\{e^{-i\omega(t-t')}G_{zz}[\boldsymbol{r}_e(t),\boldsymbol{r}_e(t'),\omega]\right\}\mathrm{d}t\mathrm{d}t', \tag{9.14}$$

with $G_{zz} = \hat{\boldsymbol{z}}\cdot\boldsymbol{G}\cdot\hat{\boldsymbol{z}}$. Assuming a straight electron trajectory $\boldsymbol{r}_e(t) = \boldsymbol{R}_0 + \hat{\boldsymbol{z}}\,vt$ and indicating explicitly the dependence of $\boldsymbol{G}$ on the in-plane and z directions, this expression can be rewritten as

$$\Gamma_{\mathrm{EELS}}(\boldsymbol{R}_0,\omega) = -\frac{4e}{\hbar}\int_{-\infty}^{\infty} \mathrm{Im}\left\{e^{-iq(z-z')}G_{zz}[\boldsymbol{R}_0,z,\boldsymbol{R}_0,z',\omega]\right\}\mathrm{d}z\mathrm{d}z'. \tag{9.15}$$

Equation (9.15) expresses the EELS losses in terms of the dyadic Green function of Maxwell's theory. When the electron propagation path goes through several media, for instance for trajectories penetrating metallic nanoparticles, it is convenient to split the Green function into a bulk and boundary contribution,

$$\boldsymbol{G}(\boldsymbol{r},\boldsymbol{r}',\omega) = \boldsymbol{G}_{\mathrm{bulk}}(\boldsymbol{r},\boldsymbol{r}',\omega) + \boldsymbol{G}_{\mathrm{ind}}(\boldsymbol{r},\boldsymbol{r}',\omega). \tag{9.16}$$

Here, $-4\pi i\int \boldsymbol{G}_{\mathrm{bulk}}(\boldsymbol{r},\boldsymbol{r}',\omega)\cdot\boldsymbol{J}(\boldsymbol{r}',\omega)\,\mathrm{d}\boldsymbol{r}'$ gives the field for a current distribution inside an unbounded medium, and the induced part

$\boldsymbol{G}_{\mathrm{ind}}$ gives the begrenzung and surface contributions due to the particle boundaries. With this decomposition, we can split the losses of Eq. (9.15) into bulk losses, as previously discussed in section 9.2.2, and additional boundary losses viz.

$$\Gamma_{\mathrm{EELS}}(\boldsymbol{R}_0,\omega) = -\frac{4e}{\hbar}\int_{-\infty}^{\infty} \mathrm{Im}\left\{e^{-iq(z-z')}G_{\mathrm{ind},\,zz}[\boldsymbol{R}_0,z,\boldsymbol{R}_0,z',\omega]\right\} \times \mathrm{d}z\mathrm{d}z' + \Gamma_{\mathrm{bulk}}(\omega). \tag{9.17}$$

For simple geometries, such as layer structures or spheres, this expression can be evaluated analytically, as briefly discussed below. In the general case, one must employ numerical simulation approaches, as will be shown in section 9.3.

9.2.4. *Analytic expressions for simple systems*

For simple geometries, such as layer structures, infinitely long cylinders, or spheres, one can obtain analytic expressions which are often extremely useful, e.g. for testing the accuracy of simulation approaches. For spheres, one can resort to Mie theory and compute the EELS probabilities either within the quasistatic limit[17] (see section 9.4) or using the full Maxwell's equations.[18] Explicit expressions can be found in these papers as well as in Ref. 8.

9.3. Simulation Approaches

The simulation of EELS and plasmonic nanoparticles usually employs Eq. (9.17), which splits the loss probabilities into bulk and boundary contributions, and on generic Maxwell solvers. These solvers start with a given source, in our case the current distribution $\boldsymbol{J}(\boldsymbol{r},\omega) = \boldsymbol{v}\,\rho(\boldsymbol{r},\omega)$ of the swift electron, and compute the electromagnetic fields $\boldsymbol{E}(\boldsymbol{r},\omega)$ and $\boldsymbol{B}(\boldsymbol{r},\omega)$ using Maxwell's equations. Among the many available Maxwell solvers, in the following, we discuss three of them which have so far been predominantly used in the literature, namely, the discrete dipole approximation (DDA), the finite difference time domain (FDTD), and the BEM approaches.

9.3.1. *Discrete dipole approximation*

The conceptually probably most simple approach is based on the so-called discrete dipole approximation (DDA),[19,20] which approximates a metallic nanoparticle through a large collection of small polarizable particles ("discrete dipole"). A given external excitation $\boldsymbol{E}_{\text{ext}}$, such as the electric field of the swift electron, then polarizes the different dipoles according to

$$\boldsymbol{P}_j = \boldsymbol{\alpha}_j \cdot \boldsymbol{E}_j. \tag{9.18}$$

Here, $\boldsymbol{P}_j$ is the polarization of the j'th dipole with polarizability $\boldsymbol{\alpha}_j$, which in general is assumed to be isotropic, and $\boldsymbol{E}_j$ is the sum of $\boldsymbol{E}_{\text{ext}}$ and the polarization field produces by all dipoles. Equation (9.18) constitutes a matrix equation, which can be solved by inversion usually employing iterative schemes.[19,20] In Ref. 21, the authors presented a DDA implementation of EELS simulations and for electron trajectories that do not penetrate the nanoparticle.

9.3.2. *Finite difference time domain*

The FDTD approach is arguably the most successful and widely used simulation schemes in computational electrodynamics.[22,23] The basic idea is to discretize the computational domain by finite differences and to propagate the fields, starting with some initial conditions, in the time domain. Typically different spatial and time meshes are used for the electric and magnetic fields $\boldsymbol{E}$ and $\boldsymbol{B}$, which allows $\boldsymbol{E}$ and $\boldsymbol{B}$ to propagate in an efficient manner. A key ingredient of the FDTD approach are the perfectly matched layer (PML) absorbing boundary conditions which allow to restrict the computational domain, almost perfectly absorbing out-going waves at a user-defined boundary sufficiently far away from the simulated nanoparticles.

An FDTD implementation for EELS simulations was presented in Ref. 24. In a related scheme, EELS simulations with a discontinuous Galerkin time-domain method were presented.[25] Both simulation approaches work best for non-penetrating electron trajectories.

9.3.3. *Boundary element method*

The BEM approach was the first EELS simulation scheme used in the field of plasmonics. The development for BEM simulations of EELS was pioneered by Javier García de Abajo *et al.*[8,26] As we will present below primarily results from such BEM simulations, in the following we describe the main ingredients in more detail.

The BEM approach of Ref. 8, 26 uses as basic ingredients the scalar and vector potentials ϕ and $\boldsymbol{A}$ rather than the electromagnetic fields $\boldsymbol{E}$ and $\boldsymbol{B}$, and additionally employs the Lorentz gauge condition $\nabla \cdot \boldsymbol{A} = ik_0\varepsilon\phi$.[14] Potentials and fields are connected through $\boldsymbol{E} = ik_0\boldsymbol{A} - \nabla\phi$ and $\boldsymbol{B} = \nabla \times \boldsymbol{A}$. Both ϕ and $\boldsymbol{A}$ are solutions of the Helmholtz equations (rather than the wave equations for electromagnetic fields)

$$\left(\nabla^2 + k_0^2\varepsilon(\boldsymbol{r},\omega)\right)\phi(\boldsymbol{r},\omega) = -4\pi\,\frac{\rho(\boldsymbol{r},\omega)}{\varepsilon(\boldsymbol{r},\omega)} \tag{9.19a}$$

$$\left(\nabla^2 + k_0^2\varepsilon(\boldsymbol{r},\omega)\right)\boldsymbol{A}(\boldsymbol{r},\omega) = -\frac{4\pi}{c}\,\boldsymbol{J}(\boldsymbol{r},\omega). \tag{9.19b}$$

In what comes next, we foresee the separation of the EELS losses into bulk and boundary contributions. First, we split the scalar and vector potentials into external and induced contributions

$$\phi^j(\boldsymbol{r},\omega) = \phi^j_{\text{ext}}(\boldsymbol{r},\omega) + \phi^j_{\text{ind}}(\boldsymbol{r},\omega), \quad \phi^j_{\text{ext}}(\boldsymbol{r},\omega) = -\frac{2e}{v\varepsilon_j}e^{iqz}K_0\left(\frac{\omega\rho}{v\gamma_{\varepsilon_j}}\right) \tag{9.20a}$$

$$\boldsymbol{A}^j(\boldsymbol{r},\omega) = \boldsymbol{A}^j_{\text{ext}}(\boldsymbol{r},\omega) + \boldsymbol{A}^j_{\text{ind}}(\boldsymbol{r},\omega), \quad \boldsymbol{A}^j_{\text{ext}}(\boldsymbol{r},\omega) = \varepsilon_j\frac{\boldsymbol{v}}{c}\phi_{\text{ext}}(\boldsymbol{r},\omega). \tag{9.20b}$$

These potentials are to be used for the electron propagating within the medium with permittivity ε_j, i.e. $\boldsymbol{r} \in \Omega_j$. Due to the decomposition of Eqs. (9.20), the external and induced potentials give rise to bulk and boundary losses, respectively.

In order to obtain the induced potentials, we again make use of the concept of Green functions, which, however, in this case are not defined through the wave Eq. (9.12) but rather the Helmholtz

equation

$$\left(\nabla^2 + k_0^2 \varepsilon_j\right) G^j(\boldsymbol{r}, \boldsymbol{r}', \omega) = -4\pi\delta(\boldsymbol{r} - \boldsymbol{r}'), \quad G^j(\boldsymbol{r}, \boldsymbol{r}', \omega) = \frac{e^{ik_j|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r} - \boldsymbol{r}'|}, \tag{9.21}$$

with $k_j = \sqrt{\varepsilon_j}\, k_0$ being the wavenumber in medium j. We can now write down the solutions of the Helmholtz equations in the *ad hoc* form

$$\phi^j(\boldsymbol{r}, \omega) = \phi^j_{\text{ext}}(\boldsymbol{r}, \omega) + \oint_{\partial\Omega_j} G^j(\boldsymbol{r}, \boldsymbol{s}', \omega)\sigma^j(\boldsymbol{s}', \omega)\, \mathrm{d}\boldsymbol{s}' \tag{9.22a}$$

$$\boldsymbol{A}^j(\boldsymbol{r}, \omega) = \boldsymbol{A}^j_{\text{ext}}(\boldsymbol{r}, \omega) + \oint_{\partial\Omega_j} G^j(\boldsymbol{r}, \boldsymbol{s}', \omega)\boldsymbol{h}^j(\boldsymbol{s}', \omega)\, \mathrm{d}\boldsymbol{s}'. \tag{9.22b}$$

These expressions are constructed such that: (i) since ϕ^j_{ext}, $\boldsymbol{A}^j_{\text{ext}}$, and G^j fulfill the Helmholtz equations everywhere in the spatial domain Ω_j except *on* the boundary $\partial\Omega_j$, Eqs. (9.22) also fulfill the Helmholtz equations within Ω_j; (ii) the (artificial) surface charges and currents σ^j and $\boldsymbol{h}^j$ have to be chosen such that the boundary conditions of Maxwell's equations are fulfilled. As shown in Refs. 26, 27, this leads to a set of eight integral equations.

To render Eqs. (9.22) suitable for a numerical implementation within a BEM approach, we have to approximate the boundaries $\partial\Omega_j$ by boundary elements of finite size. The boundary integrals of Eqs. (9.22) are then represented by a sum over these boundary elements, and the eight constituent equations for the (now discretized) surface charges and currents become matrix equations, which can be solved through numerical matrix inversion.[26,27] In Refs. 27, 28, we presented details of a computational solution scheme for the BEM equations using the Matlab toolbox MNPBEM.

9.3.4. *Selected results of BEM simulations*

Figure 9.4 shows results of BEM simulations[28] for a swift electron with a kinetic energy of 200 keV (electron velocity $v/c \approx 0.7$) passing by a silver nanoparticle with a diameter of 80 nm. The dielectric function of Ag is taken from optical experiments[15] and the minimal

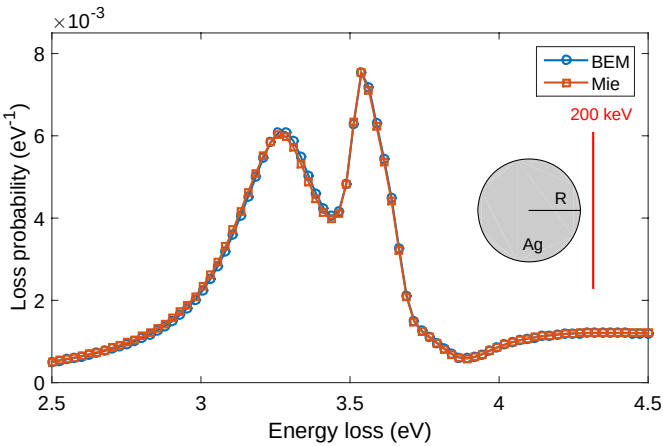


Fig. 9.4. Loss probability for electron trajectory passing by a silver nanosphere, as shown in inset. We compare the results of our BEM simulations with analytic results derived within Mie theory.[8] In the simulation, the nanosphere diameter is 80 nm, the silver dielectric function is extracted from optical experiments,[15] the background dielectric constant is one and the minimal distance between electron beam and nanosphere is 10 nm. We assume a kinetic electron energy of 200 keV.

distance between electron and nanosphere is 10 nm. For comparison, we also report results from Mie calculations[8,18] providing an analytic expression for Γ_{EELS}. As can be seen, the results of the BEM simulations and Mie calculations are in extremely good agreement, thus highlighting the accuracy of numerical simulation approaches. We tentatively assign the two peaks at loss energies of 3.26 eV and 3.54 eV to excitations of the dipolar and quadrupolar modes, respectively.

Figure 9.5 shows the induced electric field $\boldsymbol{E}_{\mathrm{ind}}$ at the dipole plasmon resonance, of 3.26 eV. On resonance, the induced field is approximately 90° phase delayed with respect to the driving field of the swift electron; for this reason, we plot the imaginary part of $\boldsymbol{E}_{\mathrm{ind}}$ only. According to Eq. (9.17), the energy loss is determined by the integrated work performed by the swift electron against the induced field.

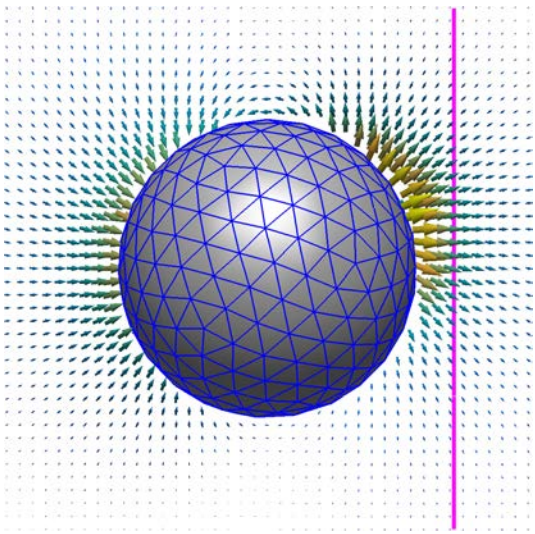


Fig. 9.5. Induced electric field $\boldsymbol{E}_{\mathrm{ind}}$ at the dipole resonance of 3.26 eV. All simulation parameters are identical to those given in Fig. 9.4. We show the discretization of the particle boundary as used in our BEM simulation approach.[28] The cones show the imaginary part of $\boldsymbol{E}_{\mathrm{ind}}$ in the plane of the trajectory of the swift electron (see magenta line).

9.4. Quasistatic Approximation and Modal Decomposition

9.4.1. *Quasistatic approximation*

Equations (9.5) and (9.15) are the general expressions that allow to compute the EEL probabilities from the solutions of the full Maxwell equations. In some cases, in particular for small nanoparticles and for a more intuitive understanding of EELS, it is advantageous to resort to the so-called quasistatic approximation.

In the quasistatic approximation, one assumes that the nanoparticle is much smaller than the wavelength of light, both inside and outside the metallic nanoparticle, such that in Eq. (9.19) we have $|\nabla\phi(\boldsymbol{r},\omega)| \gg |k\phi(\boldsymbol{r},\omega)|$. If this is the case, the vector potential is much smaller than the scalar potential (as can be inferred from

the Lorentz gauge condition), and is thus neglected. Instead of the Helmholtz equation, we then obtain the Poisson equation

$$\nabla^2\phi(\boldsymbol{r},\omega) = -4\pi\,\frac{\rho(\boldsymbol{r},\omega)}{\varepsilon(\boldsymbol{r},\omega)}. \tag{9.23}$$

The potentials inside and outside the nanoparticle have to be matched by employing the boundary conditions of Maxwell's equations at the particle boundary $\partial\Omega$, namely the continuity of the parallel component of the electric field and of the normal component of the dielectric displacement. Electric field $\boldsymbol{E}$ and scalar potential ϕ are related through $\boldsymbol{E} = -\nabla\phi$. The "quasi" of the quasistatic approximation refers to the fact that we keep in the solution of Eq. (9.23) and the consideration of the boundary conditions the full frequency content of $\varepsilon(\boldsymbol{r},\omega)$.

We next evaluate the energy loss probability of Eq. (9.5), which we rewrite in the form

$$\Gamma_{\mathrm{EELS}}(\boldsymbol{R}_0,\omega) = -\frac{1}{\pi\hbar\omega}\int \mathrm{Re}\left\{\boldsymbol{J}^*(\boldsymbol{r},\omega)\cdot\boldsymbol{E}_{\mathrm{ind}}(\boldsymbol{r},\omega)\right\}\,\mathrm{d}\boldsymbol{r} + \Gamma_{\mathrm{bulk}},$$

where $\boldsymbol{J}$ is the current distribution of the swift electron and $\boldsymbol{E}_{\mathrm{ind}}$ the induced field of the nanoparticle. Through repeated use of the divergence theorem, one can establish a number of useful expressions. We start with $\nabla\cdot(\boldsymbol{J}^*\phi_{\mathrm{ind}}) = (\nabla\cdot\boldsymbol{J}^*)\phi_{\mathrm{ind}} + \boldsymbol{J}^*\cdot\nabla\phi_{\mathrm{ind}}$, where, upon integration over the entire space and use of the divergence theorem, the term on the left-hand side vanishes because ϕ_{ind} approaches zero at infinity. Together with the continuity equation $i\omega\rho = \nabla\cdot\boldsymbol{J}$, we then find

$$\Gamma_{\mathrm{EELS}}(\boldsymbol{R}_0,\omega) = -\frac{1}{\pi\hbar}\int \mathrm{Im}\left\{\rho^*(\boldsymbol{r},\omega)\phi_{\mathrm{ind}}(\boldsymbol{r},\omega)\right\}\,\mathrm{d}\boldsymbol{r} + \Gamma_{\mathrm{bulk}}, \tag{9.24}$$

which is the equivalent form of Eq. (9.5) but within the quasistatic limit.

For an electron beam that does not penetrate the metallic nanoparticle, we can proceede even further. From $\nabla\cdot[(\nabla\phi^*)\phi_{\mathrm{ind}}] = (\nabla^2\phi^*)\phi_{\mathrm{ind}} + (\nabla\phi^*)\cdot(\nabla\phi_{\mathrm{ext}})$ together with $\rho = -\nabla^2\phi_{\mathrm{ind}}/(4\pi)$, we obtain after integration over the entire space and use of the

divergence theorem the intriguing expression

$$\Gamma_{\mathrm{EELS}}(\boldsymbol{R}_0,\omega) = \frac{1}{4\pi^2\hbar}\int \mathrm{Im}\left\{\boldsymbol{E}^*(\boldsymbol{r},\omega)\cdot\boldsymbol{E}_{\mathrm{ind}}(\boldsymbol{r},\omega)\right\}\,\mathrm{d}\boldsymbol{r}, \qquad (9.25)$$

stating that the loss probability is maximized when the integrated overlap between the field of the swift electron $\boldsymbol{E}^*$ (which is reminiscent of the field distribution of a charged wire) and of the "induced" particle plasmon field $\boldsymbol{E}_{\mathrm{ind}}$ is as large as possible.

9.4.2. *Solution of Maxwell's equations in the quasistatic approximation*

We will now show how to compute the scalar potential in case of a single boundary $\partial\Omega$, for a discussion of more general setups see Refs. 17, 27. First, we introduce the Green function for the Poisson equation,

$$\nabla^2 G(\boldsymbol{r},\boldsymbol{r}') = -4\pi\delta(\boldsymbol{r}-\boldsymbol{r}')\,, \quad G(\boldsymbol{r},\boldsymbol{r}') = \frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}. \qquad (9.26)$$

Similar to Eq. (9.22), we split the scalar potential into an external and induced part

$$\phi(\boldsymbol{r},\omega) = \phi_{\mathrm{ext}}(\boldsymbol{r},\omega) + \oint_{\partial\Omega} G(\boldsymbol{r},\boldsymbol{s}')\sigma(\boldsymbol{s}',\omega)\,\mathrm{d}\boldsymbol{s}', \qquad (9.27)$$

where ϕ_{ext} is the external potential associated with the charge distribution of the swift electron, see Eq. (9.20) and Ref. 8 for its quasistatic approximation. Again, we have introduced a surface charge distribution $\sigma(\boldsymbol{s},\omega)$ that has to be chosen such that the boundary conditions of Maxwell's equations are fulfilled. The continuity of the tangential electric field $\boldsymbol{E}^{\parallel}_{\mathrm{in}} = \boldsymbol{E}^{\parallel}_{\mathrm{out}}$ at the particle in- and outside is fulfilled when the potential is continuous, as is per construction the case in Eq. (9.22).[b] For the continuity of the dielectric displacement

[b]Note that in the quasistatic limit the external potential ϕ_{ext} does not depend on the material's permittivities,[8] similar to the Green function of the Poisson equation, and the surface charge distribution $\sigma(\boldsymbol{s},\omega)$ at the particle in- and outside is the same. In the case of the full Maxwell equations, see Eq. (9.22), ϕ^j_{ext}, $\boldsymbol{A}^j_{\mathrm{ext}}$

$D^{\perp}_{\text{in}} = D^{\perp}_{\text{out}}$, we have to perform in Eq. (9.27) a normal derivative $\frac{\partial}{\partial n} = \hat{\boldsymbol{n}} \cdot \nabla$ on both sides of the equation, and finally perform the limit $\boldsymbol{r} \to \boldsymbol{s}$. As discussed in more length in Ref. 26, some care has to be taken about this limit in the integral on the right-hand side. If we use a coordinate system with $\hat{\boldsymbol{n}}$ poynting in the z-direction and assume that the surface charge distribution σ is constant within a small circle of radius R, the contribution to the surface derivative of the integral becomes

$$\lim_{z\to\pm 0} \hat{\boldsymbol{n}} \cdot \int \frac{\boldsymbol{r}-\boldsymbol{s}'}{|\boldsymbol{r}-\boldsymbol{s}'|^3}\,\mathrm{d}\boldsymbol{s}' \to \lim_{z\to\pm 0} 2\pi z \int_0^R \frac{\rho \mathrm{d}\rho}{(\rho^2+z^2)^{\frac{3}{2}}} = \pm 2\pi, \quad (9.28)$$

where the sign depends on whether we approach the surface from the out- or inside of the particle boundary. Thus, we find for the surface derivatives of the dielectric displacement

$$\begin{aligned} &D^{\perp}_{\text{out}}(\boldsymbol{s},\omega) \\ &\quad = -\varepsilon_{\text{out}} \left(\frac{\partial \phi_{\text{ext}}(\boldsymbol{s},\omega)}{\partial n} + 2\pi\sigma(\boldsymbol{s},\omega) + \oint_{\partial\Omega} \frac{\partial G(\boldsymbol{s},\boldsymbol{s}')}{\partial n} \sigma(\boldsymbol{s}',\omega)\,\mathrm{d}\boldsymbol{s}' \right), \\ &D^{\perp}_{\text{in}}(\boldsymbol{s},\omega) \\ &\quad = -\varepsilon_{\text{in}} \left(\frac{\partial \phi_{\text{ext}}(\boldsymbol{s},\omega)}{\partial n} - 2\pi\sigma(\boldsymbol{s},\omega) + \oint_{\partial\Omega} \frac{\partial G(\boldsymbol{s},\boldsymbol{s}')}{\partial n} \sigma(\boldsymbol{s}',\omega)\,\mathrm{d}\boldsymbol{s}' \right). \end{aligned}$$

Subtracting the two expressions finally leads to the integral equation

$$\left[\Lambda(\omega)\sigma(\boldsymbol{s},\omega) + \oint_{\partial\Omega} \frac{\partial G(\boldsymbol{s},\boldsymbol{s}')}{\partial n} \sigma(\boldsymbol{s}',\omega)\,\mathrm{d}\boldsymbol{s}' \right] = -\frac{\partial \phi_{\text{ext}}(\boldsymbol{s},\omega)}{\partial n},$$

$$\Lambda(\omega) = 2\pi \frac{\varepsilon_{\text{out}} + \varepsilon_{\text{in}}}{\varepsilon_{\text{out}} - \varepsilon_{\text{in}}}, \quad (9.29)$$

which allows us to compute for a given external potential the surface charge distribution σ, and thus provides the solution of Maxwell's equations in the quasistatic limit. Again, Eq. (9.29) can be transformed from a boundary integral equation to a BEM equation by

and G^j dependent on ε_j, and one has to introduce different σ^j and $\boldsymbol{h}^j$ at the particle's in- and outside.

approximating the particle boundary through a set of boundary elements of finite size.[26,27] The solution for σ then involves a simple matrix inversion.

9.4.3. *Eigenmode expansion*

The neat thing about Eq. (9.29) is that all material parameters are embodied in the function $\Lambda(\omega)$. As we will discuss now, this allows for a convenient eigenmode expansion which was first introduced by Ouyang and Isaacson.[29] To this end, we introduce $\sigma_k(\boldsymbol{s})$ defined through

$$\oint_{\partial\Omega} \frac{\partial G(\boldsymbol{s},\boldsymbol{s}')}{\partial n}\sigma_k(\boldsymbol{s}')\,\mathrm{d}\boldsymbol{s}' = \lambda_k\sigma_k(\boldsymbol{s}), \tag{9.30}$$

which form a complete basis set satisfying the unusual orthogonality relations[c]

$$\oint_{\partial\Omega} \sigma_k(\boldsymbol{s})G(\boldsymbol{s},\boldsymbol{s}')\sigma_{k'}(\boldsymbol{s}')\,\mathrm{d}\boldsymbol{s}\mathrm{d}\boldsymbol{s}' = \delta_{kk'}.$$

In accordance with Ref. 31, we shall refer to these modes as *geometric eigenmode* as they only depend on the nanoparticle geometry. Through these modes, one can decompose the solution of Eq. (9.29) into contributions that only depend on material properties (described through Λ) and geometry (described through σ_k and λ_k). As worked out in Ref. 31, for a point charge at position $\boldsymbol{r}'$ the field at position $\boldsymbol{r}$ can be expressed through the induced Green function

$$G_{\mathrm{ind}}(\boldsymbol{r},\boldsymbol{r}',\omega) = -\frac{1}{\varepsilon}\sum_k \frac{\lambda_k+2\pi}{\Lambda(\omega)+\lambda_k}\phi_k(\boldsymbol{r})\phi_k(\boldsymbol{r}'),$$

$$\phi_k(\boldsymbol{r}) = \oint_{\partial\Omega} G(\boldsymbol{r},\boldsymbol{s})\sigma_k(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s}, \tag{9.31}$$

[c]There exists another definition for an eigenmode expansion,[30,31] where in addition to the right eigenmodes of Eq. (9.30) one introduces left eigenmodes $\tilde{\sigma}_k(\boldsymbol{s})$ defined through $\oint_{\partial\Omega}\tilde{\sigma}_k(\boldsymbol{s}')\frac{\partial G(\boldsymbol{s}',\boldsymbol{s})}{\partial n}\,d\boldsymbol{s}' = \lambda_k\tilde{\sigma}_k(\boldsymbol{s})$. The left and right eigenmodes form a biorthogonal set $\oint_{\partial\Omega}\sigma_k(\boldsymbol{s})\tilde{\sigma}_{k'}(\boldsymbol{s})\,d\boldsymbol{s} = \delta_{kk'}$ which is complete. This biorthogonal basis set is often advantageous for numerical implementations.[28,31]

where we have assumed that both $\boldsymbol{r}$ and $\boldsymbol{r}'$ are located outside the metallic nanoparticle in a medium with permittivity ε. Inserting the eigenmode expansion into Eq. (9.24) gives for an electron trajectory that does not penetrate the nanoparticle [see Eq. (9.17) for the corresponding expression using the full Maxwell's equations]

$$\Gamma_{\mathrm{EELS}}(\boldsymbol{R}_0,\omega) = -\frac{1}{\pi\hbar}\int \mathrm{Im}\left\{\rho^*(\boldsymbol{r},\omega)G_{\mathrm{ind}}(\boldsymbol{r},\boldsymbol{r}',\omega)\rho(\boldsymbol{r}',\omega)\right\}\,\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{r}',$$

which can be brought to the final form

$$\Gamma_{\mathrm{EELS}}(\boldsymbol{R}_0,\omega) = -\frac{e^2}{\pi\hbar v^2\varepsilon}\sum_k \mathrm{Im}\left\{\frac{\lambda_k+2\pi}{\Lambda(\omega)+\lambda_k}\right\}\left|\int_{-\infty}^{\infty} e^{iqz}\phi_k(\boldsymbol{R}_0,z)\,\mathrm{d}z\right|^2. \quad (9.32)$$

The term in curly brackets gives a Lorentzian lineshape at the resonance frequencies where $\mathrm{Re}[\Lambda(\omega)+\lambda_k] \approx 0$, the integral gives a form factor describing how well a given eigenmode can be excited by the electron beam. Equation (9.32) provides a quite transparent decomposition of the loss function into a lineshape function and an oscillator strength.

9.5. Results

References 10 and 11 were the first to investigate SP resonances of single metallic nanorods and nanotriangles, respectively. Figure 9.6 shows EEL spectra and maps for a silver nanotriangle, with a side length and height of approximately 80 nm and 10 nm, respectively. Panel (a) shows EEL spectra recorded for three different electron beam positions, as indicated in the inset which additionally shows a high-angle annular dark-field (HAADF) image of the nanotriangle. In the spectra, one can distinguish at least three peaks, whose maps are displayed in panels (c)–(e). Below, we will use a quasistatic mode expansion to label them as (c) dipole, (d) hexapole and (e) breathing modes.[32] From the figure it becomes apparent that, in comparison to optical spectroscopy, EELS can (i) map both modes that couple to light and those which are "dark",[33] and (ii) allows to map the

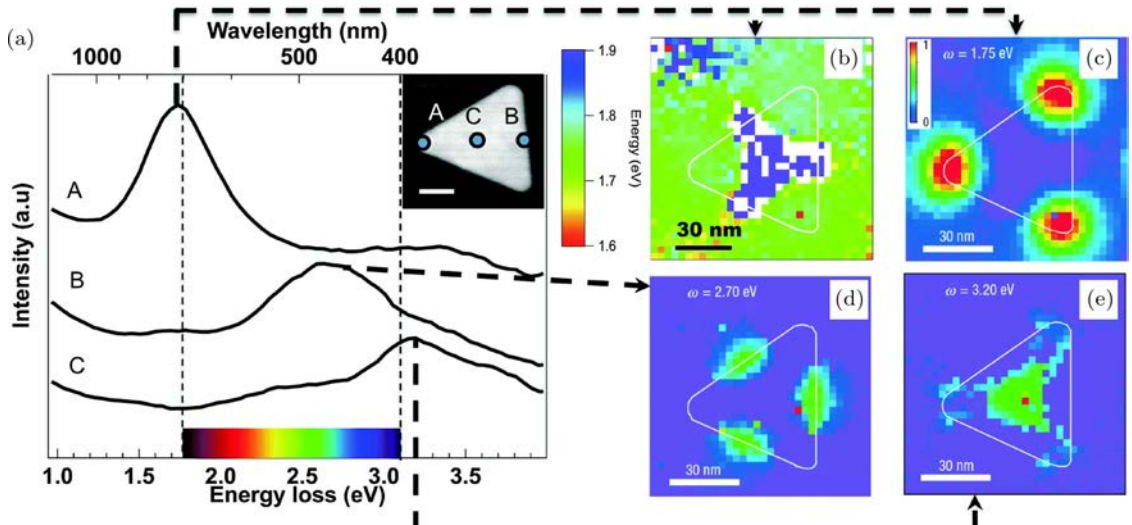


Fig. 9.6. EEL spectra and maps for a silver nanotriangle taken from Ref. 13. (a) EEL spectra taken at three different positions, as indicated in the inset. The scale bar is 20 nm. (b) Energy and (c) intensity map for the dipolar mode at 1.75 eV. (d, e) Intensity maps for second and third plasmon mode.

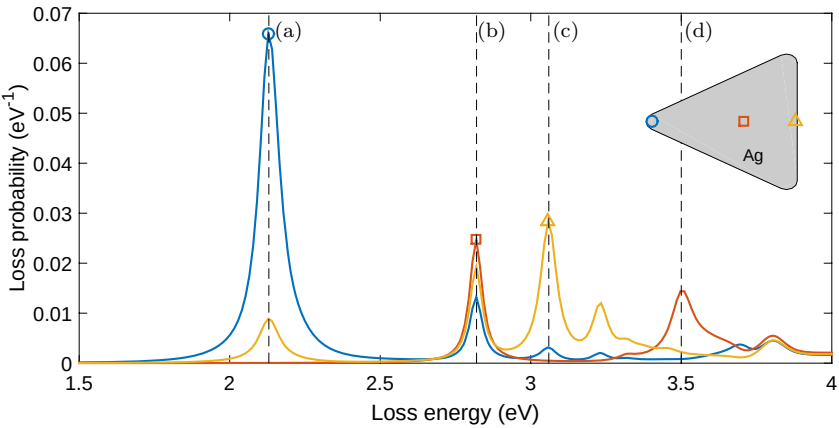


Fig. 9.7. EEL spectra at three different positions for the electron beam, indicated in the inset, as computed with the MNPBEM toolbox[27,28] for a silver triangle with a base length of approximately 80 nm and a height of 10 nm. In our simulations, we use dielectric functions extracted from optical experiments.[15] The dashed lines report the energetic positions of the plasmon resonances where the spatial EELS maps of Fig. 9.8 are computed.

EEL mode patterns with nanometer resolution. This renders EELS ideal for the investigation of the complete plasmonic mode spectrum. On the other hand, in comparison to optical spectroscopy the loss peaks are typically broader, due to the finite EELS energy resolution of ≈0.1 eV caused by the non-monochromatic energy distribution of the swift electrons, which, however, is usually not a serious limitation for the observation of SP peaks which are intrinsically broadened through ohmic losses of the metal and radiative damping. Another problem is the question of which quantity is measured in EELS. This point will be discussed in more detail in section 9.5.6.

Figure 9.7 shows simulated EEL spectra for a nanoparticle geometry similar to the experimental study shown in Fig. 9.6 (see also Ref. 11 for related simulation results). In comparison to the experiments, the simulated spectra show a richer peak structure, which is probably masked in the experiment by the finite spectral resolution, and the peaks are blueshifted. This shift is attributed to the neglect of the substrate in the simulations, as will be discussed in section 9.5.4.

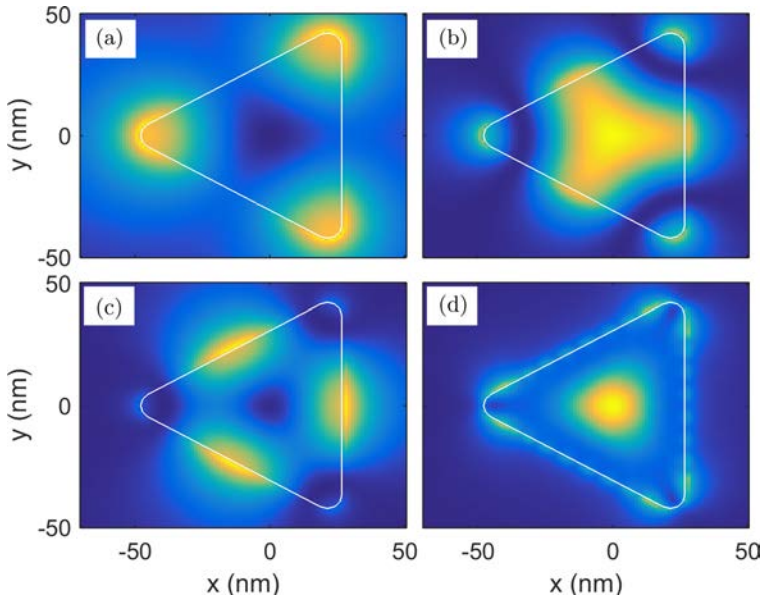


Fig. 9.8. Spatial EEL maps for the same nanotriangle as investigated in Fig. 9.7 at the plasmon resonances indicated by dashed lines. The modes can be assigned to (a) dipole, (b, c) hexapole and (d) breathing mode,[32,33] as discussed in the text.

We observe in the simulated spectra at least four pronounced peaks, labeled with (a)–(d), in addition to the bulk plasmon peak around 3.8 eV (see also Fig. 9.3). The EEL maps computed at the resonance energies are displayed in Fig. 9.8. In particular, the modes (a,b,d) compare well with the experimental maps of Fig. 9.6(c)–(e), showing that EELS experiments and simulations can be compared on par. We will return to a discussion of these modes in section 9.5.2.

9.5.1. *The example of a metallic nanodisk*

To get more insight to the plasmonic mode patterns and how they are measured in EELS, in the following we discuss the situation of a metallic nanodisk. A combined experimental and simulation study of silver nanodisks was presented in Refs. 32–35. In contrast to the nanotriangle discussed above, the nanodisk has cylinder symmetry

leading to the conservation of angular momentum,[26] which significantly facilitates the interpretation.

Figure 9.9 shows typical EEL spectra for a silver nanodisk with 60 nm diameter and 10 nm height. At the disk edge (triangular symbols), one observes a number of peaks, whereas in the center (circle) we only observe one pronounced peak. We again see a bulk loss peak at an energy of 3.8 eV. Since the nanodisk exhibits cylinder symmetry, the EELS probability only depends on the radial distance of the electron beam (measured with respect to the disk center) but not on the azimuthal angle. For this reason, in Fig. 9.10, we show instead of EEL maps a density plot of the EELS probabilities as a function of loss energy and impact parameter. In agreement to the spectra for selected impact parameters shown in Fig. 9.9, we observe (i) a series of peaks located at the disk boundaries, whose spatial extension decreases with increasing loss energy, (ii) a mode located in the disk center around 3.5 eV, and (iii) a featureless bulk plasmon peak at 3.8 eV.

In Fig. 9.11, we report the surface charge distributions for the geometric eigenmodes $\sigma_k(\boldsymbol{s})$ computed from Eq. (9.30). As the disk exhibits cylinder symmetry, the modes can be classified according to the angular momentum $\ell = 0, 1, 2, \ldots$.[26] The surface charge distributions thus depend through $\sigma_k \sim e^{\pm i\ell\varphi}$ on the azimuthal angle φ, with degenerate eigenvalues for $\pm\ell$. We can also form linear superpositions $\cos\ell\phi$ and $\sin\ell\phi$ from these degenerate modes. These (real-valued) eigenmodes naturally come out from BEM solvers and numerical diagonalization routines.[27] In EELS, the position of the electron beam determines the linear combination of the modes, and by rotating the electron beam around the nanodisk the mode pattern will rotate accordingly. For this reason, the loss probability, which accounts for mode excitations and the work of the electron performed against the electric field of the excited modes, does not depend on angle (see also Fig. 9.1(c)).

The dipolar, quadrupolar and hexapolar modes shown in Fig. 9.11 correspond to angular momentum numbers $\ell = \pm 1, \pm 2, \pm 3$, respectively. In principle, there exists an infinite number of such

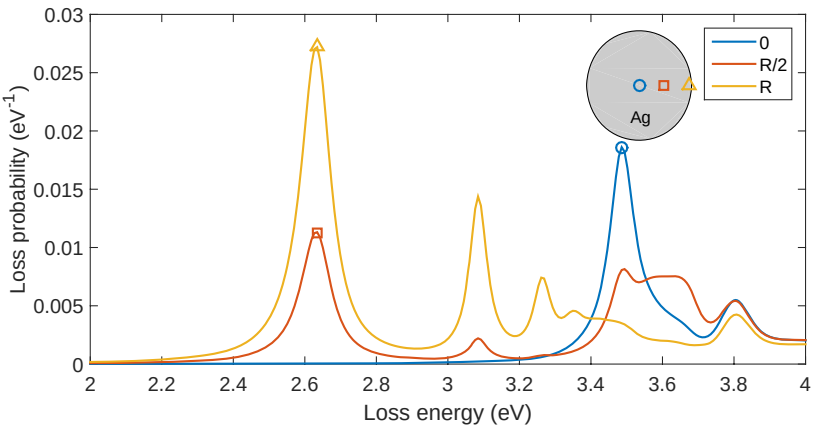


Fig. 9.9. The EEL spectra for a silver nanodisk with a diameter of 60 nm and a height of 10 nm. The impact parameters of the electron beams for the different spectra are reported in the inset, and the beam propagation direction is the z-direction perpendicular to the shaded disk.

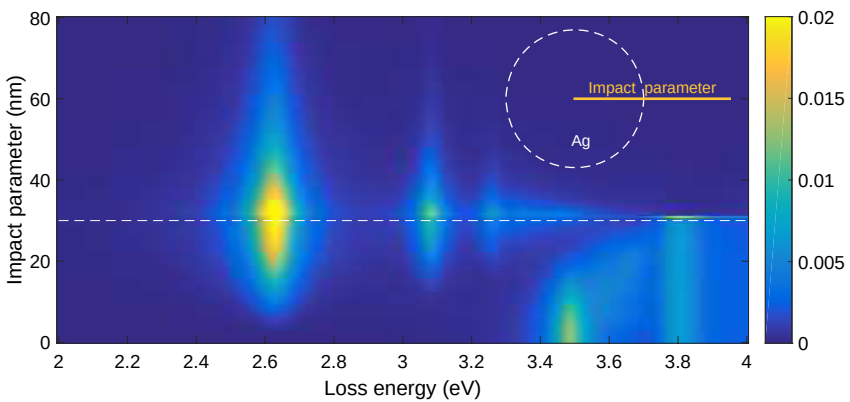


Fig. 9.10. Density map of EEL spectra for nanodisk and for different loss energies and impact parameters using the same geometry as investigated in Fig. 9.9. The dashed line at an impact parameter of 30 nm indicates the disk boundary.

modes. As can be seen in Fig. 9.10, the highest loss probabilities for these modes are at the disk edge. The spatial extension of the EELS signal decreases with increasing angular order ℓ, which is attributed to the tighter field confinement and the stronger localization of σ_k at

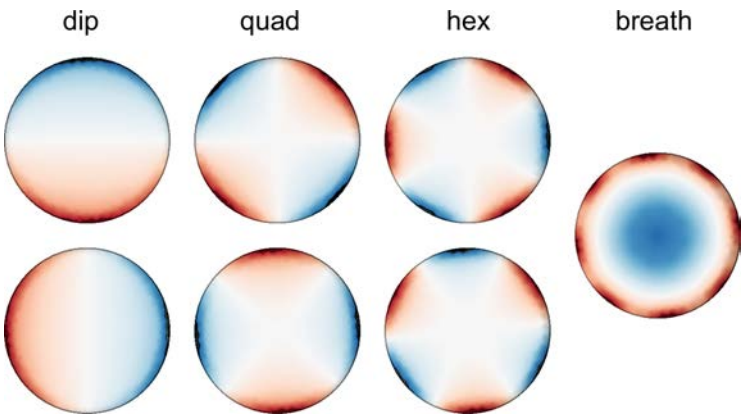


Fig. 9.11. Geometric eigenmodes [see Eq. (9.30)] for a nanodisk with a diameter to height ratio of 6:1. With increasing energy, we obtain dipole, quadrupole, hexapole and other multipole modes, whose surface charge distribution is largest at the disk edges. Additionally, we obtain a "breathing" mode with a charge distribution that oscillates in the radial direction. The slight deviations of the eigenmodes from cylinder symmetry are due to the non-spherical boundary discretization used in our simulations.

the disk edges for larger ℓ orders. The "breathing" mode in Fig. 9.11 corresponds to a radial oscillation mode with $\ell = 0$,[33] which shows up in the EEL maps of Fig. 9.10 at a loss energy of 3.5 eV. This mode is optically dark, at least for sufficiently small nanoparticles where retardation effects can be ignored.[35] As shown in Ref. 33, with increasing disk size additional radial "breathing" modes appear in the spectra, with mode patterns that have two to several zeros along the radial direction. In addition to these angular and radial plasmon excitations, there also exist modes with nodes along the vertical z-direction. However, for flat nanodisks these modes typically have very high plasmon energies and can be hardly observed in EELS. It is important to realize that this mode characterization in terms of angular, radial, and vertical modes is entirely dictated by symmetry.

In Ref. 34, it was shown that there exists an intriguing connection of radial and angular modes to SPs in planar film geometries and at the edges of such films. For films, the SP plasmon dispersion $\omega_{2D}(k_{\|})$

depends on the wavevector $\boldsymbol{k}_\parallel$ in the in-plane direction, whereas for film edges the plasmon dispersion $\omega_{1\mathrm{D}}(k_x)$ depends on the wavenumber k_x associated with the motion along the edge. The radial modes then can be interpreted as SP film modes, where the wavenumber $k_n \approx 2n\pi/d$ is determined by the disk diameter and the radial mode number $n = 1, 2, \ldots$.[d] As has been demonstrated both experimentally and theoretically,[33,34] the plasmon energies of the radial modes coincide almost perfectly with the SP energies $\hbar\omega_{2\mathrm{D}}(k_n)$. Similarly, the plasmon energies of the angular modes agree extremely well with the SP energies of edge modes $\hbar\omega_{1\mathrm{D}}(k_\ell)$, with the wavenumber associated with the angular order and the disk circumference according to $k_\ell \approx 2\ell\pi/(\pi d)$. Thus, the radial and angular disk modes can be mapped to the film and edge modes of planar films, where the wavenumbers are determined by the disk geometry (diameter and circumference). A similar characterization in terms of film and edge modes was also demonstrated for other particle geometries.[37,38]

9.5.2. *EELS mapping for nanotriangles and other geometries*

We briefly comment on the modes of a nanotriangle, with EEL spectra and maps already discussed at the beginning of section 9.5. Figures 9.12(a)–(c) show the eigenmodes for a nanotriangle, with increasing energy, as computed from Eq. (9.30). The modes of panel (a) have dipolar character, whereas modes (b, c) are more complicated.[32] Although not completely obvious from the figure, through linear combination of the two dipolar modes we can get modes where the dipole moment points into any direction within the (x, y) plane. Indeed, in optics the scattering or extinction spectra do not depend on the polarization direction within the triangle plane. Similarly, all EEL maps shown in Fig. 9.8 exhibit perfect triangular symmetry.

[d]Quite generally, one has to be careful about the reflection phase of the SP at the disk edge, which was neglected in Refs. 33,34. This phase was measured to be around 70° for a nanowire.[36]

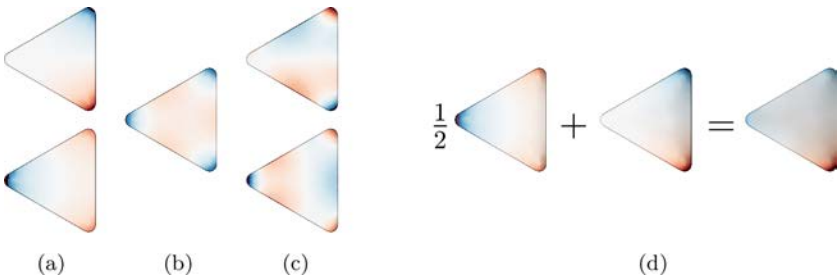


Fig. 9.12. Geometric eigenmodes for a nanotriangle with a base length to height ratio of 8:1. With increasing energy, we obtain (a) dipole and (b, c) second and third excited modes. (e) Through linear superposition of the degenerate dipole modes, one can rotate the triangular dipole mode into any direction.

In addition to the investigation of nanotriangles,[11,39–41] EELS was employed to a myriad of other nanoparticle geometries, such as nanorods,[10,42–47] nanocubes,[48–50] nanoholes,[51] nanodecahedra,[52] plasmonic tapers,[53] or split-ring resonators.[54] Additionally, for nanowires, EELS was used for probing the complex SP reflection coefficients at the wire ends[55,56] and for the investigation of plasmon modes in bent geometries.[57] All these studies underlined the great potential of EELS for nanometer imaging of plasmon modes with sub-eV energy resolution.

9.5.3. *Coupled particles*

Nanoparticle coupling allows for strong field enhancements in the gap regions, which can be exploited for extreme light confinement, surface enhanced Raman spectroscopy (SERS)[58] and sensorics, or for tailoring light-matter couplings.[1,3] Again, EELS can be used for the mapping of plasmon modes in coupled nanoparticles, but the interpretation of the signals from the hot-spot region requires some care. Consider, for example, the bowtie geometry shown in the inset of Fig. 9.13, which consists of two flipped nanotriangles separated by a narrow gap region. The modes of lowest energy correspond to bonding and antibonding dipolar modes, where the dipoles (see lower panel of Fig. 9.12) are either parallel (bonding mode) or antiparallel (antibonding mode). The dipole moments for these modes add up

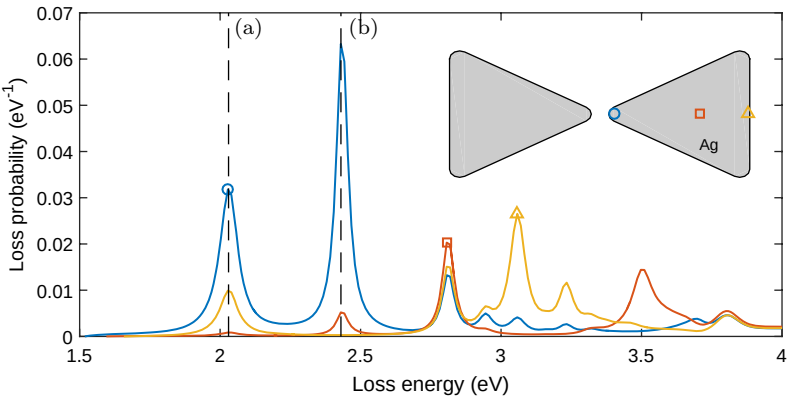


Fig. 9.13. The EEL spectra for coupled nanotriangles (bowtie geometry) and for selected electron beam positions, as indicated in the inset. The geometry of the individual nanotriangles is identical to those given in Fig. 9.7, and the gap size is 10 nm. Modes (a) and (b) correspond to the bonding and antibonding dipolar modes, whose EEL maps are shown in Fig. 9.14.

either constructively or destructively, resulting in optically bright and dark modes. For the bonding mode, the electric field in the gap region points from one nanotriangle to the other one and comes together with a strong near-field enhancement.

When an electron with propagation direction perpendicular to the bowtie plane shown in the inset of Fig. 9.14 moves through the hot-spot region, the velocity $\boldsymbol{v}$ is perpendicular to the induced electric field $\boldsymbol{E}_{\mathrm{ind}}$. Thus, the energy loss probability of Eq. (9.5) becomes small because of $\boldsymbol{v} \cdot \boldsymbol{E}_{\mathrm{ind}} \approx 0$, and EELS "is blind to hot spots".[59] Another way of understanding this effect is to recall Eq. (9.25) stating that the loss probability is given by the integrated overlap between the field of the swift electron $\boldsymbol{E}^*$ and of the "induced" particle plasmon field $\boldsymbol{E}_{\mathrm{ind}}$. Because of symmetry, $\boldsymbol{E}^*$ is symmetric with respect to reflection at the mirror plane, whereas $\boldsymbol{E}_{\mathrm{ind}}$ of the bonding mode is antisymmetric. Thus, the integral vanishes. Figure 9.14 shows EEL maps for the (a) bonding and (b) antibonding modes, showing indeed a strongly reduced EELS signal for the bonding mode in the gap region. In contrast, the antibonding mode has a high EELS signal, despite its low photonic local density of states (LDOS),[2] which is

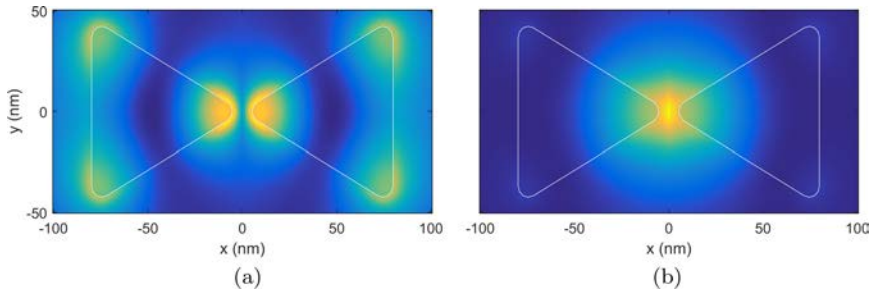


Fig. 9.14. The EEL maps for bowtie geometry, see Fig. 9.13, and for (a) bonding and (b) antibonding mode. For the bonding mode, the EEL intensity at the hot-spot in the gap region vanishes, as discussed in the text and in Ref. 59.

attributed to the strong components of $\boldsymbol{E}_{\mathrm{ind}}$ in z direction and the resulting large $\boldsymbol{v} \cdot \boldsymbol{E}_{\mathrm{ind}}$ contributions. Indeed, corresponding EEL maps were observed in Refs. 60, 61.

Other coupling phenomena have been observed for metal nanosphere dimers,[62,63] nanowire dimers,[64] oligomer-type nanocavities,[65] split-ring resonators,[66] or hybrid metal–semiconductor nanogap antennas.[67] The relation between SERS hot spots and EELS signals has been studied in Ref. 68. Non-locality[69] and quantum effects[61,70,71] have been investigated for coupled nanoparticles. Assembly of plasmonic nanoparticles through DNA strands has been demonstrated in Refs. 72,73.

9.5.4. *Substrate effects*

In EELS experiments, the plasmonic nanoparticles must be placed on some thin support, such as mica,[11] carbon grids,[42] or silicon nitride (SiN) membranes.[33] Large-area and homogeneous SiN membranes are available for various thicknesses, typically of the order of 10 nm. Figure 9.15 shows results of simulations performed with (solid lines) and without (dashed lines) a thin SiN membrane. Whereas the thin membrane leads to practically no additional electron losses, the energies of the plasmon modes are significantly redshifted due to the large permittivity $\varepsilon \approx 4$ of SiN and should be included in EELS simulations for quantitative comparison with

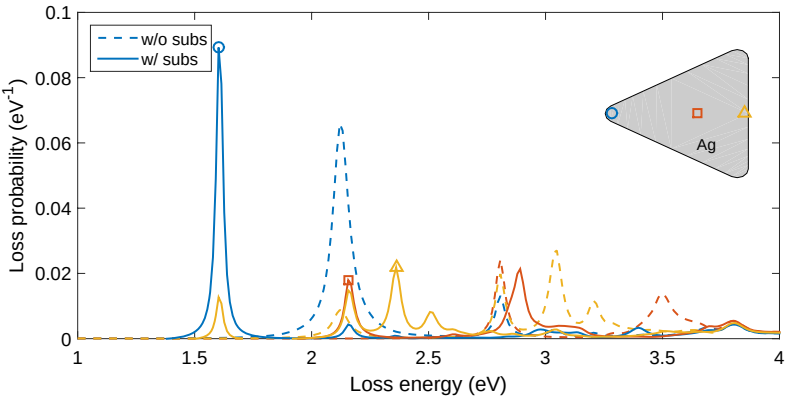


Fig. 9.15. The EEL spectra for nanotriangle and for selected electron beam positions, see also Fig. 9.7. The dashed lines correspond to simulations of a nanotriangle in free space, the solid lines show results of simulations where an additional 15 nm thick silicon nitride ($\varepsilon = 4$) membrane supporting the nanotriangle is considered. This membrane leads to a relatively large red-shift ≈ 0.5 eV of the plasmon peaks.

experiment.[74] Substrate-induced symmetry breaking of plasmon modes was reported in Refs. 34, 75.

9.5.5. *Combining electrons and photons*

When a fast electron excites a particle plasmon, the plasmon can decay by emitting a photon. This so-called cathodoluminescence (CL) provides additional information about plasmon modes.[12,13] See also Refs. 8, 76 for CL theory. CL was first observed for a silver nanosphere,[77] and has since then been applied in many experimental studies, such as for imaging of plasmonic modes in nanowires[78] or nanotriangles.[79] Correlated CL and EELS mapping has been studied in Refs. 44, 52, and the differences and similarities between these two techniques have been analyzed in Ref. 41.

Another interesting experimental technique based on electron microscopy is to combine photoexcitation and subsequent EELS imaging of SPs using the time-resolved photon-induced near-field electron microscopy (PINEM) technique,[80] which has been used for the investigation of spherical dimers[81] and for nanorods.[82]

Theoretical work has been concerned with plasmon electron energy-gain spectroscopy[83] and vortex electron beams.[84] Nonlinear inelastic electron scattering in EELS has been demonstrated in Ref. 85.

9.5.6. *Plasmon tomography*

A question that has attracted some interest is what quantity is measured in EELS. In Ref. 16, the authors established a link between EELS and the photonic LDOS, although this link is sometimes rather dubious.[59] An interesting recent development is the connection between EELS and plasmon field tomography. In Ref. 86, the authors showed that the full three-dimensional plasmon fields can be reconstructed from a series of rotated EELS maps, making various assumptions including the validity of the quasistatic approximation. A similar approach was developed independently in Ref. 49 and demonstrated experimentally for a silver nanocube. Angular-resolved EELS maps were presented for a split-ring resonator.[87] A tomography scheme using CL was shown in Ref. 88.

9.5.7. *Quantum effects*

Recent years have seen various efforts to investigate the limits of classical plasmonics, bringing EELS to the regime where quantum effects play a noticeable role. Quantum plasmon resonances of individual[70] and coupled[71] metallic nanoparticles were presented. In the latter approach, the electromagnetic forces on plasmonic nanoparticles induced by fast-electron beams were used to control the nanoparticle arrangements.[89] In a recent study,[90] the authors investigated silver nanocubes coupled through a molecular tunnel junction, which leads to plasmon tunneling and the emergence of charge-transfer plasmon peaks.

9.6. Summary and Outlook

To summarize, we have given a short introduction to EELS of plasmonic nanoparticles. In the first part, we have developed the

methodology for computing EEL signals and maps, including a generic description scheme within the framework of the Green's tensor of Maxwell's theory and an eigenmode expansion using the quasistatic approximation, and have discussed implementations for EELS simulations. In the second part, we have discussed typical EELS results for single and coupled nanoparticles, the impact of membranes needed to support the nanoparticles, as well as variants of EELS, such as CL, time-resolved PINEM, or plasmon field tomography. Although plasmonics remains an optics game, targeting at light confinement at the nanoscale through excitation of SPs, EELS has been established as an extremely efficient characterization tool that provides information about plasmonic field distributions with nanometer spatial and sub-eV energy resolution.

As usual, it is impossible to foresee future developments of the field, and predictions are doomed to failure. I will nevertheless try to speculate a little bit about future developments. First, the combination of electron microscopy and optics is particularly attractive for plasmonics, and CL will probably be employed more routinely in future studies. PINEM and variants which combine electron microscopy with a femtosecond time resolution have a great potential for unveiling coherent and incoherent plasmon dynamics at ultrashort time scales. Future EELS studies will also benefit from improved energy resolutions, providing spectra that have the intrinsic linewidths of the plasmonic peaks. Both from the experimental and theoretical side, plasmon tomography still needs further improvements, but holds a lot of promise for measuring plasmonic fields in full 3D and with unprecedented precision. By a similar token, electron holography[91] offers the possibility to retrieve phase information of plasmonic processes. One can also expect that novel materials will move into the focus of plasmonics, where EELS and electron microscopy provide an ideal platform for detailed investigations. Altogether, EELS of plasmonic nanoparticle and, more generally, the interaction of fast electrons with SPs will certainly play an important role in the (continuously) promising future of plasmonics.

References

1. Maier, S. A. (2007). *Plasmonics: Fundamentals and Applications* (Springer, Berlin).
2. Novotny, L. and Hecht, B. (2006). *Principles of Nano-Optics* (Cambridge University Press, Cambridge).
3. Schuller, J. A., Barnard, E. S., Cai, W., Jun, Y. C., White, J. S. and Brongersma, M. L. (2010). *Nat. Mat.* **9**, p. 193.
4. Zijlstra, P., Chon, J. W. M. and Gu, M. (2009). *Nature* **459**, p. 410.
5. Tame, M. S., McEnery, K. R., Ozdemir, S. K., Lee, J., Maier, S. A. and Kim, M. S. (2013). *Nature Phys.* **9**, p. 329.
6. Ritchie, R. H. (1957). *Phys. Rev.* **106**, p. 874.
7. Powell, C. J. and Swan, J. B. (1959). *Phys. Rev.* **115**, p. 869.
8. Garcia de Abajo, F. J. (2010). *Rev. Mod. Phys.* **82**, p. 209.
9. Kreibig, U. and Vollmer, M. (1995). *Optical Properties of Metal Clusters*, Vol. 25 of *Springer Series in Material Science* (Springer, Berlin).
10. Bosman, M., Keast, V. J., Watanabe, M., Maaroof, A. I. and Cortie, M. B. (2007). *Nanotechnology* **18**, p. 165505.
11. Nelayah, J., Kociak, M., Stephan, O., Garcia de Abajo, F. J., Tence, M., Henrard, L., Taverna, D., Pastoriza-Santos, I., Liz-Martin, L. M. and Colliex, C. (2007). *Nat. Phys.* **3**, p. 348.
12. Kociak, M. (2011), In *Scanning Transmission Electron Microscopy*, Pennycook, S. J. and Nellist, P. D. (Springer, Berlin), p. 163.
13. Kociak, M. and Stephan, O. (2014). *Chem. Soc. Rev.* **53**, p. 3865.
14. Jackson, J. D. (1999). *Classical Electrodynamics* (Wiley, New York).
15. Johnson, P. B. and Christy, R. W. (1972). *Phys. Rev. B* **6**, p. 4370.
16. Garcia de Abajo, F. J. and Kociak, M. (2008). *Phys. Rev. Lett.* **100**, p. 106804.
17. Garcia de Abajo, F. J. and Aizpurua, J. (1997). *Phys. Rev. B* **56**, p. 15873.
18. García de Abajo, F. J. (1999). *Phys. Rev. B* **59**, p. 3095.
19. Draine, B. T. (1988). *Astrophys. J.* **333**, p. 848.
20. Draine, B. T. and Flateau, P. J. (1994). *J. Opt. Soc. Am. A* **11**, p. 1491.
21. Bigelow, N. W., Vaschillo, A., Iberi, V., Camden, J. P. and Masiello, D. J. (2012). *ACS Nano.* **6**, p. 7497.
22. Yee, K. S. (1966). *IEEE Trans. Antennas Propag.* **14**, p. 302.
23. Taflove, A. and Hagness, S. C. (2005). *Computational Electrodynamics* (Artech House, Boston).
24. Cao, Y., Manjavacas, A., Large, N. and Nordlander, P. (2015). *ACS Photon.* **2**, p. 369.
25. Matyssek, C., Niegemann, J., Hergert, W. and Busch, K. (2011). *Photonics Nanostruct. Fundam. Appl.* **9**, p. 367.
26. Garcia de Abajo, F. J. and Howie, A. (2002). *Phys. Rev. B* **65**, p. 115418.
27. Hohenester, U. and Trügler, A. (2012). *Comp. Phys. Commun.* **183**, p. 370.
28. Hohenester, U. (2014). *Comp. Phys. Commun.* **185**, p. 1177.

29. Ouyang, F. and Isaacson, M. (1989). *Phil. Mag. B* **60**, p. 481.
30. Fuchs, R. (1975). *Phys. Rev. B* **11**, p. 1732.
31. Boudarham, G. and Kociak, M. (2012). *Phys. Rev. B* **85**, p. 245447.
32. Schmidt, F., Ditlbacher, H., Hofer, F., Krenn, J. R. and Hohenester, U. (2014). *Nano Lett.* **14**, p. 4810.
33. Schmidt, F.-P., Ditlbacher, H., Hohenester, U., Hohenau, A., Hofer, F. and Krenn, J. R. (2012). *Nano Lett.* **12**, p. 5780.
34. Schmidt, F.-P., Ditlbacher, H., Hohenester, U., Hohenau, A., Hofer, F. and Krenn, J. R. (2014b). *Nat. Commun.* **5**, p. 3604.
35. Krug, M., Reisecker, M., Hohenau, A., Ditlbacher, H., Trügler, A., Hohenester, U. and Krenn, J. R. (2014). *Appl. Phys. Lett.* **105**, p. 171103.
36. Gruber, C., Hirzer, A., Schmidt, V., Trügler, A., Hohenester, U., Ditlbacher, H., Hohenau, A. and Krenn, J. R. (2015). *Appl. Phys. Lett.* **106**, p. 081101.
37. Nelayah, J., Kociak, M., Stephan, O., Geuquet, N., Henrard, L., Garcia de Abajo, F. J., Pastoriza-Santos, I., Liz-Marzan L. M. and Colliex, C. (2010). *Nano Lett.* **10**, p. 902.
38. Gu, L., Sigle, W., Koch, C. T., Ögüt, B., van Aken, P. A., Talebi, N., Vogelgesang, R., Mu, J., Wen, X. and Mao, J. (2011). *Phys. Rev. B* **83**, p. 195433.
39. Nelayah, J., Gu, L., Sigle, W., Koch, C. T., Pastoriza-Santos, I., Liz-Martin, L. M. and van Anken, P. A. (2009). *Opt. Lett.* **34**, p. 1003.
40. Schaffer, B., Grogger, W., Kothleitner, G. and Hofer, F. (2010). *Ultramicroscopy* **110**, p. 1087.
41. Losquin, A., Zagonel, L. F., Myroshnychenko, V., Rodrguez-Gonzlez, B., Tence, M., Scarabelli, L., Förstner, J., Liz-Marzan, L. M., Garcia de Abajo, F. J., Stephan, O. *et al.* (2015). *Nano Lett.* **15**, p. 1229.
42. Schaffer, B., Hohenester, U., Trügler, A. and Hofer, F. (2009). *Phys. Rev. B* **79**, p. 041401(R).
43. Nicoletti, O., Wubs, M., Mortensen, N. A., Sigle, W., van Aken, P. A. and Midgley, P. A. (2011). *Opt. Exp.* **19**, p. 15371.
44. Guiton, B. S., Iberi, V., Li, S., Leonard, D. N., Parish, C. M., Kotula, P. G., Varela, M., Schatz, G. C., Pennycook, S. J. and Camden, J. P. (2011). *Nano Lett.* **11**, p. 3482.
45. Rossouw, D., Couillard, M., Vickery, J., Kumacheva, E. and Botton, G. A. (2011). *Nano Lett.* **11**, p. 1499.
46. Martin, J., Kociak, M., Mahfoud, Z., Proust, J., Gerard, D. and Plain, J. (2014). *Nano Lett.* **14**, p. 5517.
47. Collins, S. M., Nicoletti, O., Rossouw, D., Ostasevicius, T. and Midgley, P. A. (2014). *Phys. Rev. B* **90**, p. 155419.
48. Mazzucco Geuquet, S., Ye, J., Stephan, O., Van Roy, W., Van Dorpe, P., Henrard, L. and Kociak, M. (2012). *Nano Lett.* **12**, p. 1288.
49. Nicoletti, O., de la Pena, F., Leary, R. W., Holland, D. J., Ducati, C. and Midgley, P. A. (2013). *Nature* **502**, p. 80.
50. Goris, B., Guzzinati, G., Fernandez-Lopez, C., Perez-Juster, J., Liz-Marzan, L. M., Trügler, A., Hohenester, U., Verbeeck, J., Bals, S. and Van Tendeloo, G. (2014). *J. Chem. Phys. C* **118**, p. 15356.

51. Sigle, W., Nelayah, J., Koch, C. and van Aken, P. (2009). *Opt. Lett.* **34**, p. 2150.
52. Myroshnychenko, V., Nelayah, J., Adamo, G., Geuquet, N., Rodriguez-Fernandez, J., Pastoriza-Santos, I., MacDonald, K. F., Henrard, L., Liz-Marzan, L. M., Zheludev, N. I. *et al.* (2012). *Nano Lett.* **12**, p. 4172.
53. Talebi, N., Sigle, W., Vogelgesang, R., Esmann, M., Becker, S. F., Lienau, C. and van Aken, P. A. (2015). *ACS Nano* **9**, p. 7641.
54. Boudarham, G., Feth, N., Myroshnychenko, V., Linden, S., García de, Abajo, J., Wegener, M. and Kociak, M. (2010). *Phys. Rev. Lett.* **105**, p. 255501.
55. Zhou, X., Hörl, A., Trügler, A., Hohenester, U., Norris, T. B. and Herzing, A. A. (2014). *J. Appl. Phys.* **116**, p. 233101.
56. Schoen, D. T., Atre, A. C., Garcia-Etxarri, A., Dionne, J. A. and Brongersma, M. L. (2015). *Nano Lett.* **15**, p. 120.
57. Rossouw, D. and Botton, G. A. (2013). *Phys. Rev. Lett.* **110**, p. 066801.
58. Nie, S. and Emory, S. R. (1997). *Science* **275**, p. 1102.
59. Hohenester, U., Ditlbacher, H. and Krenn, J. (2009). *Phys. Rev. Lett.* **103**, p. 106801.
60. Koh, A. L., Fernandez-Dominguez, A. I., McComb, D. W., Maier, S. A. and Yang, J. K. W. (2011). *Nano Lett.* **11**, p. 1323.
61. Duan, H., Fernandez-Dominguez, A. I., Bosman, M., Maier, S. A. and Yang, J. K. W. (2012). *Nano Lett.* **12**, p. 1683.
62. Koh, A. L., Bao, K., Khan, I., Smith, W. E., Kothleitner, G., Nordlander, P., Maier, S. A. and McComb, D. W. (2009). *ACS Nano* **3**, p. 3015.
63. Barrow, S. J., Rossouw, D., Funston, A. M., Botton, G. A. and Mulvaney, P. (2014). *Nano Lett.* **14**, p. 3799.
64. Alber, I., Sigle, W., Demming-Janssen, F., Neumann, R., Trautmann, C., van Aken, P. A. and Toimil-Molares, M. E. (2012). *ACS Nano* **6**, p. 9711.
65. Ogüt, B., Talebi, N., Vogelgesang, R., Sigle, W. and van Aken, P. A. (2012). *Nano Lett.* **12**, p. 5239.
66. Liang, Q., Wen, Y., Mu, X., Reindl, T., Yu, W., Talebi, N. and van Aken, P. A. (2015). *Opt. Express* **23**, p. 20721.
67. Brintlinger, T., Herzing, A. A., Long, J. P., Vurgaftman, I., Stroud, R. and Simpkins, B. S. (2015). *ACS Nano* **9**, p. 6222.
68. Kneipp, K., Kneipp, H. and Kneipp, J. (2015). *Chem. Sci.* **6**, p. 2721.
69. Wiener, A., Duan, H., Bosman, M., Horsfield, A. P., Pendry, J. B., Yang, J. K. W., S. A. Maier and Fernndez-Domnguez, A. I. (2013). *ACS Nano* **7**, p. 6287.
70. Scholl, J. A., Koh, A. L. and Dionne, J. A. (2012). *Nature* **483**, p. 421.
71. Scholl, J. A., Garcia-Etxarri, A., Leen Koh, A. and Dionne, J. A. (2013). *Nano Lett.* **13**, p. 564.
72. Diaz-Egea, C., Sigle, W., van Aken, P. A. and Molina, S. I. (2013). *Nanosc. Res. Lett.* **8**, p. 337.
73. Sun, W., Boulais, E., Hakobyan, Y., Wang, W. L., Guan, A., Bathe, M. and Yin, P. (2014). *Science* **346**, p. 1258361.

74. Arbouet, A., Mlayah, A., Girard, C. and Colas des Francs, G. (2014). *New. J. Phys.* **16**, p. 113012.
75. Das, P. and Chini, T. K. (2014). *J. Phys. Chem. C* **118**, p. 26284.
76. Gomez-Medina, R., Yamamoto, N., Nakano, M. and Garcia de Abajo, F. J. (2008). *New J. Phys.* **10**, p. 105009.
77. Yamamoto, N., Araya, K. and García de Abajo, F. J. (2001). *Phys. Rev. B* **64**, p. 205419.
78. Vesseur, E. J. R., de Waele, R. Kuttge, M. and Polman, A. (2007). *Nano Lett.* **7**, p. 2843.
79. Chaturvedi, P., Hsu, K. H., Kumar, A., Fung, K. H., Mabon, J. C. and Fang, N. X. (2009). *ACS Nano* **3**, p. 2965.
80. Barwick, B., Flannigan, D. J. and Zewail, A. H. (2009). *Nature* **462**, p. 902.
81. Yurtsever, A., Baskin, J. S. and Zewail, A. H. (2012). *Nano Lett.* **12**, p. 5027.
82. Piazza, L., Lummen, T. T. A., Quinonez, E., Murooka, Y., Reed, B. W., Barwick, B. and Carbone, F. (2015). *Nat. Commun.* **6**, p. 6407.
83. Asenjo-Garcia, A. and Garcia de Abajo, F. J. (2013). *New J. Phys.* **15**, p. 103021.
84. Asenjo-Garcia, A. and García de Abajo, F. J. (2014). *Phys. Rev. Lett.* **113**, p. 066102.
85. Xu, C. K., Liu, W. J., Zhang, P. K., Li, M., Zhang, H. J., Xu, K. Z., Luo, Y. and Chen, X. J. (2014). *Nat. Phys.* **10**, p. 753.
86. Hörl, A., Trügler, A. and Hohenester, U. (2013). *Phys. Rev. Lett.* **111**, p. 086801.
87. von Cube, F., Niegemann, J., Irsen, S., Bell, D. C. and Linden, S. (2014). *Phys. Rev. B* **89**, p. 115434.
88. Atre, A. C., Brenny, B. J. M., Coenen, T., Garcia-Etxarri, A., Polman, A. and Dionne, J. A. (2015). *Nat. Nanotech.* **10**, p. 429.
89. Reyes-Coronado, A., Barrera, R. G., Batson, P. E., Echenique, P. M., Rivacoba, A. and Aizpurua, J. (2010). *Phys. Rev. B* **82**, p. 235429.
90. Tan, S. F., Wu, L., Yang, J. K. W., Bai, P., Bosman, M. and Nijhuis, C. A. (2014). *Science* **343**, p. 1496.
91. Midgley, P. A. and Dunin-Borkowski, R. E. (2009). *Nat. Mater.* **8**, p. 271.

CHAPTER 10

Thermoplasmonics

GUILLAUME BAFFOU

Institut Fresnel, Aix-Marseille Université, France
guillaume.baffou@fresnel.fr

ROMAIN QUIDANT

Institut de Ciencies Fotoniques (ICFO),
The Barcelona Institute of Science and Technology, Spain and
Institució Catalana de Recerca i Estudis Avançats (ICREA)
romain.quidant@icfo.es

Photothermal processes in plasmonic nanoparticles have been at the basis of a decade of burgeoning activities that set the field of thermoplasmonics. Here, we present an overview of this dynamic research area. Specifically, we detail the underlying physics of heat generation in plasmonic nanoparticles. Key aspects related to heat generation processes, determination of temperature increase (both numerically and experimentally), time scales, collective effects and heating under pulsed illumination are first addressed. Subsequently, we review the main targeted applications, namely photothermal cancer therapy, nanochemistry, thermophotovoltaics and heat-assisted magnetic recording.

10.1. Introduction

Under illumination, plasmonic nanoparticles act as sources of heat generation due to light absorption in the metal.[1] This process is

naturally enhanced at the plasmonic resonance wavelength and can become particularly strong, leading e.g., to nanoparticle damage, melting or boiling of the surrounding medium. Heat generation in plasmonic systems had been considered for long as a side effect, when the *optical* properties of plasmonic nanoparticles were the main subject of interest. In 2002 and 2003, two trendsetting advances evidenced possible benefits based on heat generation in plasmonics. First, the Orrit's group[2] introduced an optical technique for nanoparticle localization based on the detection of heating modulation of gold nanoparticles. Second, Pitsillides *et al.*[3] and Hirsch *et al.*[4] concomitantly proposed to use gold nanoparticles under illumination to locally destroy cancer cells by hyperthermia. These new concepts triggered the development of thermoplasmonics, which now denotes the branch of plasmonics involving photothermal processes.

Today, thermoplasmonics is arousing strong interest in the whole plasmonics community. Indeed, any experimental development in plasmonics may involve photothermal effects, even if it is not the original purpose. The temperature increase has to be quantified and its contribution to the experiment assessed. More generally, since any field of science features thermal effects, heating of plasmonic nanoparticles can lead to countless applications in a wide range of fields, including chemistry, biomedicine, magnetic recording, phase transition, polymer science or hydrodynamics.

The aim of this chapter is three-fold: (i) introducing the fundamentals of heat generation in plasmonics, (ii) detailing available experimental or numerical tools able to predict or measure local temperature increase in plasmonics and (iii) critically reviewing the most promising applications based on photothermal effects in plasmonics.

10.2. Fundamentals

This section presents the underlying theory of thermoplasmonics. We first address the notions of heat generation and absorption

cross-section, two important parameters to describe photothermal effects in plasmonics. We then focus on the magnitude of the temperature increase in plasmonic systems and how it can be predicted theoretically. This is a very common concern, which is not always easy to address, especially for non-spherical nanoparticles or under transient illumination or when collective thermal effects occur. A final part is dedicated to the theory of heating under pulsed illumination.

10.2.1. *Heat generation*

Heat generation in a metal nanoparticle under illumination arises from Joule effect. Let $q(\mathbf{r})$ be the heat source density inside the nanoparticle. It is related to the complex amplitude $\mathbf{E}(\mathbf{r})$ of the electric field by the relation[1]

$$q(\mathbf{r}) = \frac{1}{2}\mathrm{Re}[\mathbf{J}^*(\mathbf{r}).\mathbf{E}(\mathbf{r})] \tag{10.1}$$

where $\mathbf{J}^*$ denotes the complex conjugate of $\mathbf{J}$. Using the relations $\mathbf{J} = -i\omega\mathbf{P}$ and $\mathbf{P} = \varepsilon_0\varepsilon(\omega)\mathbf{E}$, one ends up with the expression of the heat source density as a function of the electric field amplitude inside the nanoparticle:

$$q(\mathbf{r}) = \frac{\omega}{2}\mathrm{Im}(\varepsilon(\omega))\varepsilon_0|\mathbf{E}(\mathbf{r})|^2 \tag{10.2}$$

Heat generation in plasmonics is thus related to the square of the electric field *inside* the metal nanoparticle. This result differs from the more common concern in plasmonics, where only the external optical near-field matters, for instance in applications related to enhanced light–matter interaction like surface enhanced Raman scattering or sensing (see Chapter 5). For this reason, nanoparticles acting as efficient near-field enhancers, such as gap dimers, do not necessarily behave as efficient photothermal transducers.[5] Conversely, elongated, thin morphologies enabling a better penetration of the electric field inside the metal are preferred in thermoplasmonics.[6]

10.2.2. *Absorption cross-section*

The absorption cross-section of a nanoparticle is defined as

$$P = \sigma_{\text{abs}} I \tag{10.3}$$

where P is the light power absorbed by the nanoparticle and I is the light intensity (power per unit area). The power P absorbed by the nanoparticle equals, to a good approximation, the power delivered by the nanoparticle. Other energy conversion pathways, such as fluorescence or thermal radiation, are fully negligible in general.

There exists a simple relation between the two important physical quantities introduced so far, $q(\mathbf{r})$ and σ_{abs}, that reads

$$P = \int q(\mathbf{r}) \mathrm{d}\mathbf{r} = \sigma_{\text{abs}} I, \tag{10.4}$$

where the integral runs over the nanoparticle volume. For small nanoparticles, the absorption cross-section is proportional to the volume of the nanoparticle. For gold, this proportionality is observed typically for nanoparticles smaller than 60 nm.[1]

10.2.3. *Temperature increase*

What the actual temperature increase is in plasmonics experiments is a widespread concern. Addressing this question is not straightforward, especially for non-spherical particles, or multiple nanoparticles in close proximity. The *experimental* measurement of the temperature distribution in plasmonics is the subject of a forthcoming section. In this part, we shall only focus on the numerical and analytical determination of a temperature distribution in plasmonics.

In most applications in plasmonics, nanoparticles are deposited on, or immersed in a dense medium (liquid or solid). This makes the temperature distribution in the system governed by the heat diffusion equation[7]:

$$c \frac{\partial T(\mathbf{r}, t)}{\partial t} - \kappa \nabla^2 T(\mathbf{r}, t) = q(\mathbf{r}, t) \tag{10.5}$$

where c and κ are the volumetric heat capacity and the thermal conductivity (of the metal or of the surrounding medium, depending

on $\mathbf{r}$). The heat source density q is defined in the previous section and naturally equals zero outside the nanoparticle. In general, the thermal conductivity of the metal is much larger than the thermal conductivity of the surrounding medium, $\kappa_{\mathrm{m}} \gg \kappa_{\mathrm{s}}$. Heat diffusion is consequently much faster inside the nanoparticle, which makes the *nanoparticle temperature uniform*. This is a general rule of thumb in thermoplasmonics. Non-uniform temperature in plasmonic structures can be only achieved with elongated structures and non-uniform illumination,[8] or with disconnected structures such as dimers.[9] Under the assumption of uniform nanoparticle temperature, one can easily calculate the temperature distribution in the simple case of a *spherical* particle of radius a immersed in a uniform medium. It yields

$$T(r) = \frac{\sigma_{\mathrm{abs}} I}{4\pi\kappa_{\mathrm{s}} r} \tag{10.6}$$

in the surrounding medium, i.e., for $r > a$. Importantly, the temperature decays as $1/r$ in the surrounding medium, where $r = |\mathbf{r}|$ denotes the radial coordinate. This rule also holds for non-spherical nanoparticles for r large compared to the dimensions of the nanoparticle. Inside the nanoparticle, the temperature increase is uniform and reads

$$T_{\mathrm{NP}} = \frac{\sigma_{\mathrm{abs}} I}{4\pi\kappa_{\mathrm{s}} a} \tag{10.7}$$

A plot of the overall temperature profile is represented in Fig. 10.1.

This description only holds for a spherical nanoparticle in a uniform surrounding medium. For non-spherical nanoparticles, no closed-form expressions exist and numerical simulations are in principle required. A simple approach was yet proposed[8] to retrieve the temperature increase using closed-form fitting expressions for a specific set of nanoparticle morphologies, namely rod, disc, ellipsoid and torus. The method consists in using Eq. 10.6 with an effective radius corrected by a factor β

$$T_{\mathrm{NP}} = \frac{\sigma_{\mathrm{abs}} I}{4\pi\kappa_s \beta a}, \tag{10.8}$$

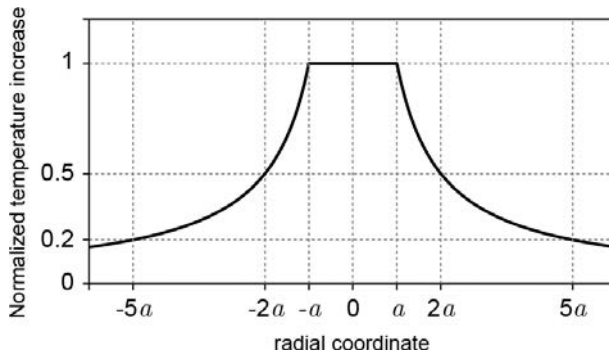


Fig. 10.1. Radial profile of the normalized temperature increase $T(r)/T_{\mathrm{NP}}$ around a spherical nanoparticle of radius a.

where a is now the equivalent radius of a sphere of the same volume. The fitting equations fixing the misprints of the original article[8] are reproduced here:

Ellipsoid	$\beta = \exp(\sqrt{1 + 0.0416\ln(D/d)} + 0.092^{*}\ln(D/d) - 1)$
Rod	$\beta = 1 + 0.09657\ln(D/d)$
Disk	$\beta = \exp[0.040 - 0.0124\ln(D/d) +$
	$0.0677\ln^2(D/d) - 0.00457\ln^3(D/d)]$
Ring	$\beta = 1.021 + 0.17442\ln^2(D/d - 0.625)$

Although this method provides a very simple way to retrieve a temperature increase in plasmonics for common nanoparticle morphologies, it involves the absorption cross-section, which has still to be calculated numerically *a priori*, using for instance the Boundary Element Method.[10]

For more specific nanoparticle geometries, it is still possible to use equation Eq. 10.6 by replacing the radius a by an effective radius coined the Laplace radius[9] a_{L}

$$T_{\mathrm{NP}} = \frac{\sigma_{\mathrm{abs}} I}{4\pi\kappa_s a_L}. \tag{10.9}$$

The numerical calculation only involves the inversion of a simple matrix. This formalism also enables the calculation of the temperature profile in the surrounding medium using a Green's function method.

A last common approach to compute temperature profiles in plasmonics consists in using Comsol. But this commercial software has to be used with care when considering thermodynamics calculations. Here are some precautions that have to be taken:

1. Do not consider a two-dimensional (2D) model even if one length of the nanoparticle is much larger than the other, for a rod or a wire for instance. This amounts to considering a nanoparticle with an infinite size (an infinite wire in this case), and changing the dimensionality changes the physics in thermodynamics. Even for an elongated system, numerical simulations have to be conducted using a 3D model, or possibly a 1D model in the case of a system with point symmetry, such as a sphere or a core-shell structure.
2. Mind the boundary conditions. In particular, do not use insulating walls, or the simulation is never going to converge. Heat has to be released somehow. Prefer walls that are far from the heat sources and set the walls to ambient temperature.

10.2.4. *Thermal collective effects in plasmonics*

The simple numerical approaches to determine a temperature increase in plasmonics as detailed in the previous paragraphs only hold for *single* nanoparticles (except Comsol). For assemblies of nanoparticles, thermal collective effects[11–13] are likely to occur, which it the subject of this paragraph.

When several nanoparticles are simultaneously illuminated, the temperature increase experienced by one specific nanoparticle results from its inner heat generation, but also from the contribution of all the other nanoparticles in its vicinity. This second contribution can be dominant or not, compared to the self-generation of heat, depending on the geometry of the system. This yields two regimes, the regime where thermal hot spots exist around each nanoparticles

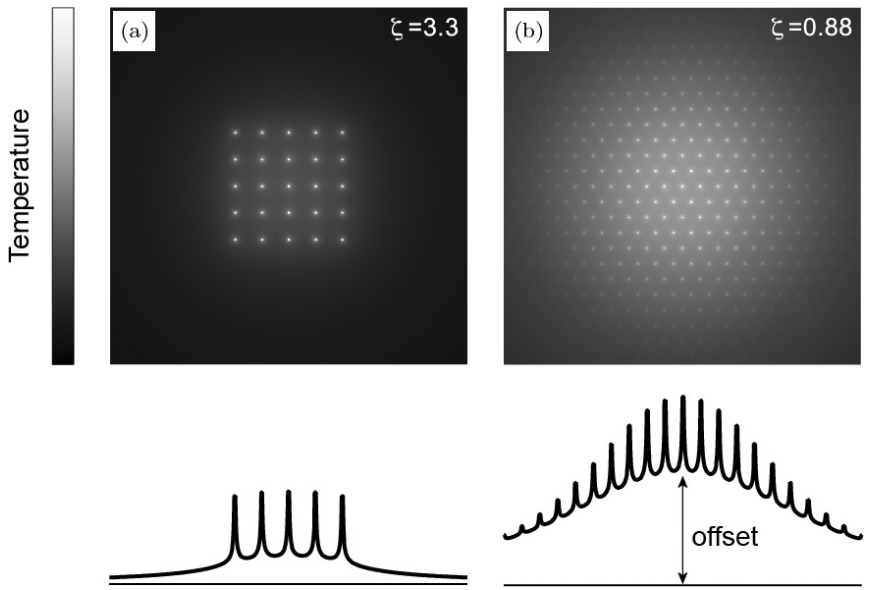


Fig. 10.2. Temperature maps of regular arrays of nanoparticles, and associated temperature profiles. (a) Case of large ζ value characterizing the temperature localization regime. (b) Case $\zeta < 1$ characterizing the presence of a dominant overall temperature offset resulting from thermal collective effects.

(temperature localization regime, Fig. 10.2(a)), and a regime where collective effects occur, characterized by an overall temperature offset, and a smooth temperature throughout the system despite the nanometric size of the heat sources (Fig. 10.2(b)). The occurrence of one regime or another can be easily determined by considering a dimensionless number than depends only on geometric properties of the system. For a two-dimensional distribution of nanoparticles (typically nanoparticles deposited on a substrate), this dimensionless number reads[13]

$$\zeta = \frac{p^2}{3La} \tag{10.10}$$

where p is the typical nanoparticle interdistance, L is the size of the nanoparticle distribution and $2a$ is the characteristic size of the nanoparticles. Many developments in plasmonics are based on the illumination of ensembles of nanoparticles where collective effects

are very likely to occur. Unfortunately, such a counterintuitive effect is sometimes discarded in reported works in plasmonics, leading to an underestimated effect of the temperature and misleading interpretations.

10.2.5. *Temperature dynamics*

A frequent concern in thermoplasmonics is to know how fast a temperature profile appears or disappears when switching on or off the illumination.[14] The most basic answer is: over the timescale

$$\tau_d = \frac{L^2}{D} \tag{10.11}$$

where L is the size of the heat source and D is the thermal diffusivity of the medium surrounding the source. L can be either the diameter of the nanoparticle in the case of a single nanoparticle illumination, or the size of an array of nanoparticles if collective thermal effects occur (see previous paragraph). For water, D is on the order of $1.4 \times 10^{-7}\,\mathrm{m^2\,s^{-1}}$ and for glass $3 \times 10^{-7}\,\mathrm{m^2\,s^{-1}}$. Plot of τ_D as a function of the size of the heat source is represented in Fig. 10.3 in the case of a water environment.

Equation (10.11) stands for the most rudimentary approach to get a rough estimation of the expected time scale of temperature dynamics. Reality may be slightly different from this simple estimation if the nanoparticle is endowed with a significant surface thermal resistivity (so-called Kapitza resistivity), which contributes to slow down any thermal exchange between the nanoparticle and its surroundings.[14] The value of the interface resistivity depends on the molecular coating of the nanoparticle and on the nature of the surrounding medium.

The actual temperature dynamics also depends on the nature of the illumination variations. For instance, when considering a point-like nanoparticle illuminated by a Heaviside intensity profile, the nanoparticle temperature varies according to the error function erf:

$$T(r,t) = \frac{P}{2\pi^{3/2} c_m \rho D}\left(1 - \mathrm{erf}\left(\frac{r}{\sqrt{4Dt}}\right)\right) \tag{10.12}$$

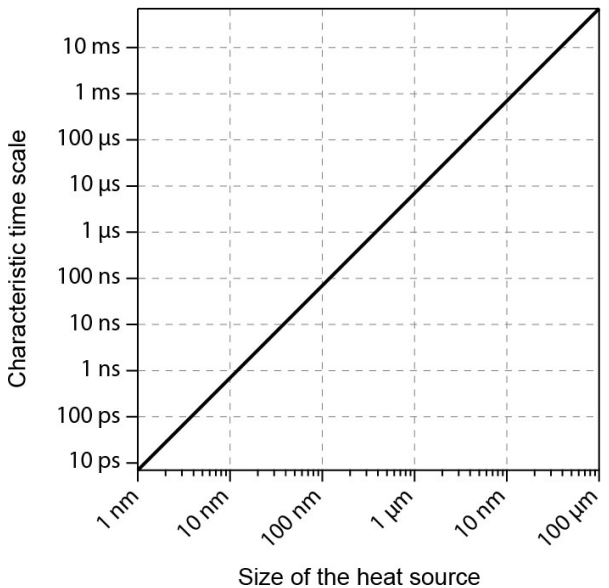


Fig. 10.3. Characteristic time scale of heat diffusion in water as a function of the size of the heat source.

However, if the nanoparticle is illuminated by a Dirac-like intensity profile, the nanoparticle temperature variation reads[14]

$$T(r,t) = \frac{\mathcal{E}_0}{c_m \rho} \frac{1}{(4\pi D\, t)^{3/2}} \exp\left(-\frac{r^2}{4\, Dt}\right) \tag{10.13}$$

Finally, if the nanoparticle is illuminated by a modulated light intensity at the angular frequency Ω, the nanoparticle temperature is[15]

$$T(r,t) = \frac{P}{4\pi\kappa_s r}\left[1 + e^{-(r-R)/\delta_{\text{th}}} \cos\left(\frac{r-R}{\delta_{\text{th}}} - \Omega t\right)\right] \tag{10.14}$$

where $\delta_{\text{th}} = \sqrt{2D/\Omega}$. Note that the decay length δ_{th} of the "thermal wave" is 2π times shorter than the period $2\pi\delta_{\text{th}}$ of the cosine. Consequently, no matter the angular frequency Ω of the excitation, it is not possible to observe thermal waves propagating from a modulated heat source. For a spherical nanoparticle, a more sophisticated expression can be derived.[16]

These three temperature distributions are represented in Fig. 10.4. Because they correspond to point-like sources of heat, they can be considered as Green's functions. Thus, they can be used as a basis to investigate more elaborated illumination conditions or particle distributions.

10.2.6. *Heating under pulsed illumination*

A large part of the current developments in thermoplasmonics are based on heating using femtosecond to nanosecond pulsed illumination. In this case, the purpose is often to induce a very fast and intense increase of temperature in the vicinity of single nanoparticles in order to generate physical effects such as bubble formation[17,18] or pressure wave generation.[19,20]

The physics of heat generation under pulsed illumination is slightly more complex than under continuous wave illumination and requires considering the following three-step process.[21,22]

1. The electronic gas of the metal nanoparticle primarily absorbs part of the incoming light energy.
2. This energy is transferred to the atomic lattice of the nanoparticle via electron–phonon interaction.
3. The nanoparticle cools down by releasing its thermal energy into the surrounding medium by thermal diffusion.

In this scenario, three time scales are involved. Step 1 occurs over a time scale that is the pulse duration τ_p. Step 2 occurs over a duration that corresponds to the time constant τ_ep of electron–phonon interaction, typically on the picosecond scale, (for gold,[23–25] $\tau_\mathrm{ep} = 1.7\,\mathrm{ps}$). And step 3 occurs over the cooling time scale of the nanoparticle that reads[14]

$$\tau_\mathrm{NP} = a^2 \frac{c_m}{3\,\kappa_s} \tag{10.15}$$

where a is the equivalent radius of the nanoparticle, c_m the volumetric heat capacity of the nanoparticle ($c_m = \rho c_{p,m}$ where ρ is the mass density and $c_{p,m}$ the specific heat capacity, more commonly found in constant tables) and κ_s the thermal conductivity of the surrounding

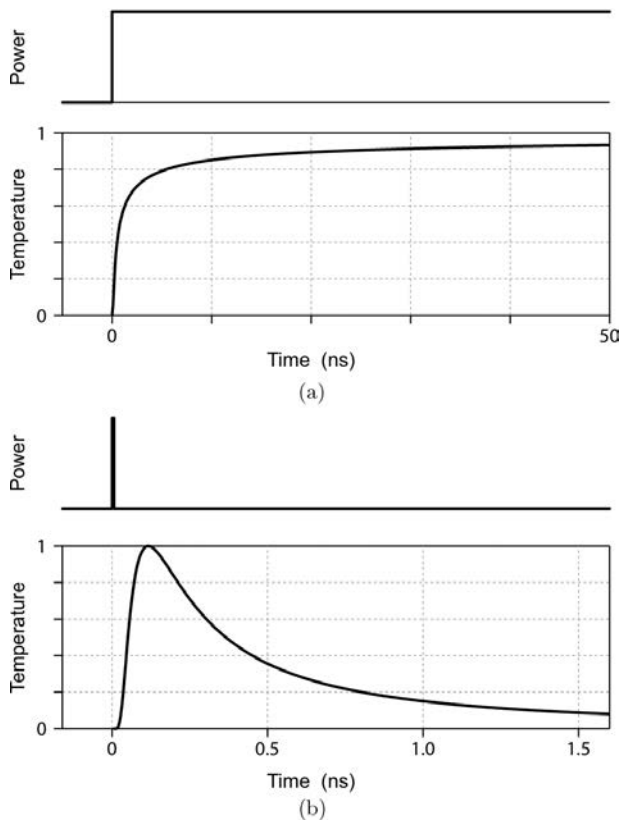


Fig. 10.4. (a) Absorbed power and temperature evolutions resulting from a Heaviside-like illumination. (b) Absorbed power and temperature evolutions resulting from a Dirac-like (pulsed) illumination.

medium. Depending on the relative values of these different time scales, different regimes can be observed. If $\tau_{\mathrm{NP}} \gg \tau_{\mathrm{ep}}$, i.e., if the heat release to the surroundings is much slower than the heating of the nanoparticle by the pulse of light, then a simple expression of the temperature increase in the nanoparticle can be derived from simple considerations of energy conservation. If F is the fluence of the illumination (energy per unit area), the total energy absorbed by the nanoparticle reads $e = \sigma_{\mathrm{abs}} F$. If the energy is transferred in totality to the nanoparticle before heat release to the surroundings,

which is our assumption, the temperature increase δT must satisfy the relation $e = V c_m \delta T$ where V is the nanoparticle volume and c_m the volumetric heat capacity of the metal. This yields an estimation of the maximum nanoparticle temperature increase achieved under pulsed illumination[14]:

$$\delta T = \frac{\sigma_{\text{abs}} F}{V c_m} \tag{10.16}$$

This expression does not assume any nanoparticle geometry and is useful to obtain an order of magnitude of the temperature increase under pulsed illumination. The assumption of no heat release during pulse absorption is valid mostly under femtosecond pulsed illumination and for nanoparticles that are larger than typically 40 nm. This ideal temperature increase appears as an upper limit. In general, the actual temperature increase is smaller than this value.[26]

Under pulsed illumination, the long-range $1/r$ temperature extension described in the previous paragraphs does not apply. On the contrary, the transient temperature increase remains confined at the vicinity of the nanoparticle in the case of small nanoparticles. More precisely, if one defines $T_{\max}(r) = \max_t[T(r,t)]$ as the maximum temperature increase achieved in the surrounding medium over time, $T_{\max}$ scales as $1/r^3$, a much steeper temperature decay compared to steady state.[14]

10.3. Thermal microscopy techniques

Measuring a temperature distribution on the nanometric and micrometric scales is not easy. First experiments related to temperature mapping in plasmonics only dates from 2009.[27] Since then, many other techniques have been developed, each of them with their own advantages and drawbacks. Let us review the most important developments in temperature mapping in plasmonics.

Most temperature mapping techniques developed in the context of thermoplasmonics are based on the use of fluorescent compounds dispersed in the nanoparticles' surroundings.[27–33] Many fluorescence properties are supposed to be temperature dependent. While the first

reported temperature mapping technique in plasmonics was based on the measurement of fluorescence polarization anisotropy,[5,27] other physical quantities such as fluorescence lifetime[31,32] and fluorescence spectra[28–30] have also proved to be efficient. Using this strategy, two approaches are usually adopted. The fluorescence compounds are either dispersed in a liquid medium surrounding the nanoparticles, or are embedded within the substrate.

An alternative temperature mapping technique is based on wavefront sensing.[33] In this label-free approach, which does not require the use of molecular probes, a planar wavefront crossing the medium of interest undergoes a thermal-induced distortion that is mapped using a quantitative wavefront sensor (based on quadriwave lateral shearing interferometry). This wavefront distortion is due to the temperature-induced variation of the refractive index of the medium surrounding the nanoparticles. As only liquids feature substantial temperature-induced variation of refractive index, this technique is restricted to applications involving a liquid environment, and where the nanoparticles are deposited on a planar substrate. The map of the wavefront distortion is subsequently post-processed using an inversion algorithm to retrieve the temperature distribution and the heat source density distribution (power per unit area) delivered in the system. This technique also enables the quantitative determination of absorption cross-sections of nanoparticles.[34]

Examples of temperature maps around plasmonic nanoparticles are presented in Fig. 10.5.

10.4. Applications

In this section, we review the main potential applications of thermoplasmonics and discuss their respective opportunities and challenges.

10.4.1. *Heat-assisted magnetic recording*

The maximum capacity of a hard disk, defined by its density of magnetic bits, is limited by the maximum magnetic field that can be produced around the recording head. Indeed, for the sake of reliability,

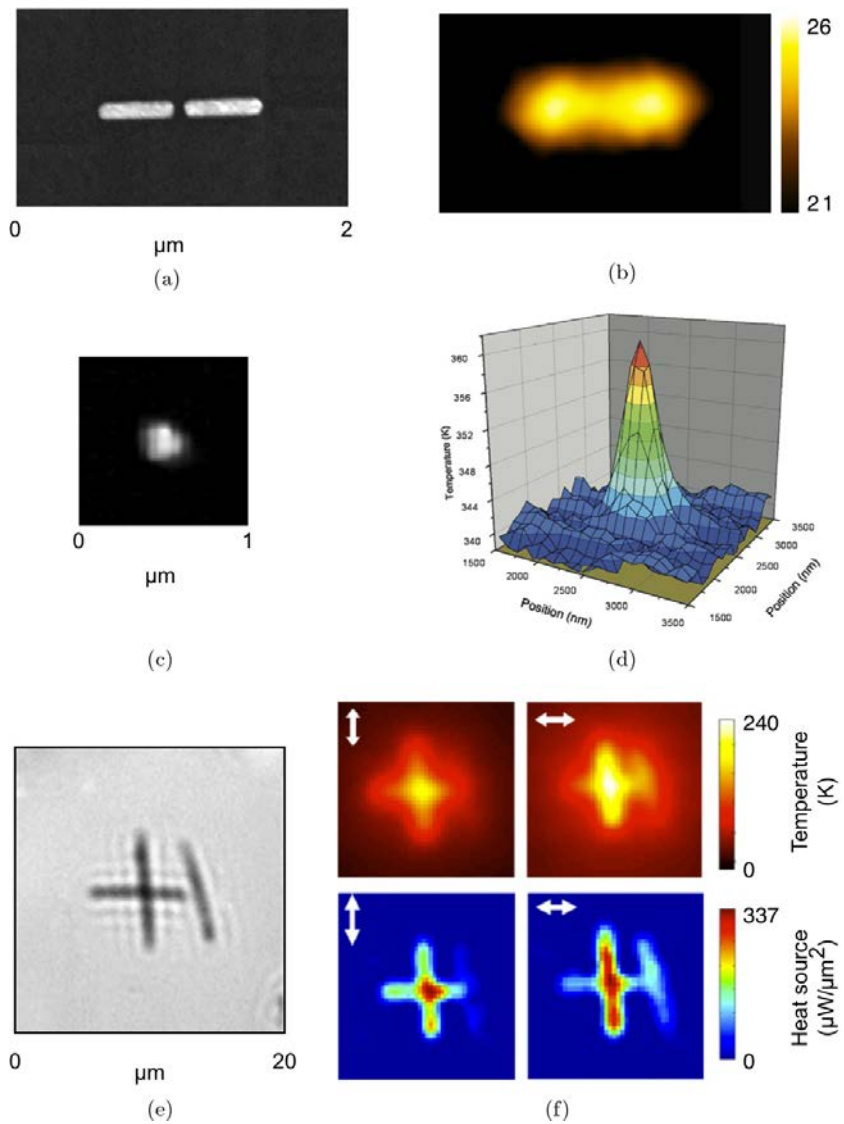


Fig. 10.5. (a,b) SEM image of a gold dimer antenna and the associated temperature distribution measured using fluorescent molecules and fluorescence polarization anisotropy (FPA) confocal microscopy.[5,27] (c,d) SEM image of a 40 nm gold dot and the associated temperature distribution measured using the fluorescence spectra of Erbium ionic centers embedded within the substrate.[29] (e,f) Optical image of gold nanowires and the associated temperature and heat source density distributions, measured using optical wavefront sensing.[33,35]

a decrease in the bit size requires an increase of the coercivity of the magnetic material (intensity of the applied magnetic field required to cancel the magnetization of that material, i.e., reach the Curie temperature). Consequently, there exists a minimum bit size beyond which the required magnetic field becomes prohibitive.[36]

Heat-assisted magnetic recording (HAMR) is one of the technologies that was proposed back to the 50s to overcome this technological problem and increase the capacity of hard drives. In the HAMR approach, a focused laser beam is used to locally heat up above the Curie point the portion of the disk that is being written. Such a local heating provisionally reduces the coercivity in such a way that writing requires a lower magnetic field.

While HAMR was originally demonstrated using illumination through a planar solid immersion lens (average optical spot size of $\lambda/4$), a much higher accuracy on local heating is required to substantially increase the current storage density. This is where nanoplasmonics is foreseen to play a major role. By exploiting the capability of metallic nanostructures to control light fields on the nanometer scale, it is possible to strongly reduce the dimension of the illumination spot and thus of the heated region.[36,38] Interestingly, unlike what one can first expect, direct heating of the metallic NP is not the dominant contribution. According to the papers published in this field, local heating results rather from the local absorption by the magnetic material of the optical near field bound to the nanoparticle. For this reason, the magnetic material and the plasmonic nanoparticle have to be as close as possible.

Plasmon-enhanced HAMR was pioneered in 2008 by Matsumoto *et al.* at Hitachi by exploiting the optical near field from a gold nanobeak.[39] One year later, Seagate reported on a experiment in which a plasmonic antenna was integrated at the extremity of a standard HAMR recording head.[37] In this implementation, the light coupled to the planar solid immersion lens was concentrated onto a 200 nm "lollipop" antenna (Fig. 10.6). The antenna was designed in such a way that it focused the intercepted light at its extremity, onto the magnetic material. Using a continuous laser power of 80 mW at

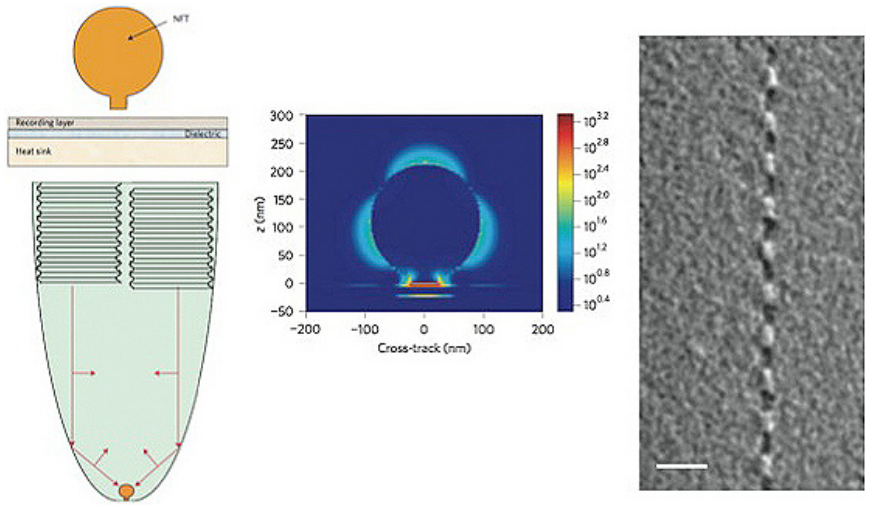


Fig. 10.6. (a) A planar solid immersion mirror with a dual offset grating used to focus a waveguide mode onto the lollipop NFT (bottom). The electric field for the TE mode is shown, which generates a longitudinal (vertical) field at the focus. Expended view of the NFT in front of the magnetic medium (top). (b) Profile of the field intensity $|\mathbf{E}|^2$ through a cross-section of the NFT. (c) MFM image of a recorded track. The track width is about 70 nm. Scale bar, 300 nm.[37]

a wavelength of 830 nm and a constant head-to-medium distance of about 15 nm, the authors demonstrated writing of a 75 nm (FWHM) track (versus around 250 nm with standard HAMR), which corresponds to an areal density of 375 $\mathrm{Tb/m^2}$. In 2010, Hitachi demonstrated 1.5$\mathrm{Pb/m^2}$ by combining a novel *E*-shape antenna design with a pre-patterned.[39]

Although local heating has been exclusively attributed to the optical absorption of the magnetic medium inducing the plasmonic hot spot, the actual mechanism of plasmon-based HAMR is expected to be more complex since the plasmonic nanoantenna acts itself as a nanosource of heat. Further studies are required to better understand and thus optimize this mechanism.

At that stage, the main hard drive companies seem to confirm that the new generation of storing devices will include

the plasmon-enhanced HAMR technology. Back in 2014, Seagate announced they planned to produce low quantities of 6–10 Tb capacity hard disks in the "near future" without confirming though they will use HAMR. As of 2015, no hard disks using HAMR are currently commercialized, but HAMR keeps progressing with demonstration drives produced by companies such as Seagate.[40] While TDK originally predicted that HAMR hard disks could be commercially released in 2015,[41] it seems they will rather enter the market in 2018.[42]

10.4.2. *Heat to electricity conversion (Thermophotovoltaics)*

In the framework of light to electricity conversion, it has been proposed that the unique optical properties of resonant metallic nanoparticles (MNPs) could benefit the field of photovoltaics.[43] MNPs integrated to a photovoltaic (PV) cell can act as efficient optical antennas capable of boosting solar light harvesting before concentrating collected photons into the photovoltaic medium. Yet, the presence of MNPs in a PV device also introduces additional loss channels along with technical limitations. All in all, it remains unclear whether plasmonics could substantially increase the efficiency of photovoltaic cells and if so, whether the technology would be industrially doable.

Conversely, enhanced absorption and heating in plasmonic MNPs could contribute to an alternative light-to-electricity conversion technology known as thermophotovoltaics (TPV).[44,45] Originally proposed in the 60s, TPV is the energy conversion process to convert heat into electricity via photons. A TPV device consists of a heat source that emits photons through blackbody radiation and a photovoltaic cell that converts these photons into current (see Fig. 10.7). The main technological challenge of developing TPV systems is matching the spectral sensitivity of the photocells with the radiation spectrum of the emitter, in order to achieve high conversion efficiency. In practice, this is done by shaping the emitter radiation spectrum by suppressing the emission of photons with energies

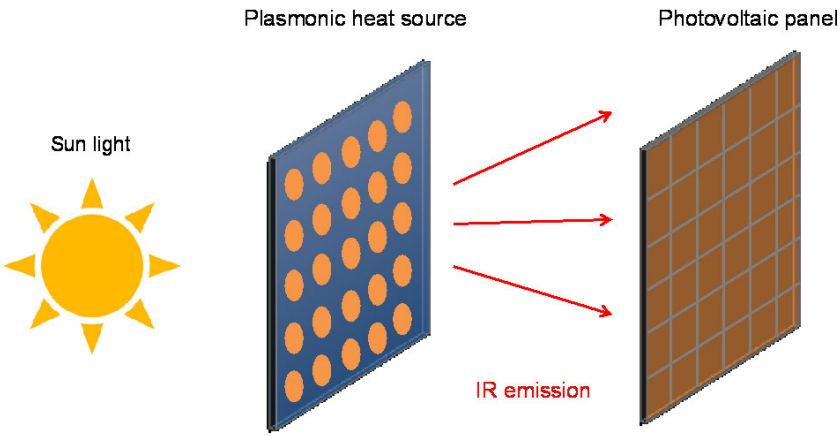


Fig. 10.7. Sketch of a TPV device based on photothermal conversion of plasmonic nanostructures.

below the bandgap of the photovoltaic cell. Spectral tunability can be achieved by using selectively emitting materials or combining a broadband emitter with a selective filter, which reflects low, energy photons back into the emitter. Alternatively, controlling the infrared photon density of states by structuring the emitter surface offers a narrow-band emission that is temperature independent.[46]

So far, most efforts have been focused on combustion driven systems to be implemented for instance in combination with gas, oil or wood burners. Alternatively, the heat source can be generated by absorption of solar light of an absorbing material. In a recent breakthrough experiment, Lenert *et al.* used a dense arrangement of vertical multiwall CNTs as a nearly ideal absorber (>0.99 absorbance over the all visible spectrum).[47] Selective emission is ensured by shaping the photon density of states with a Si/SiO_2 photonic crystal. This unique combination enabled the authors to reach a conversion efficiency of 3.2%, which outperforms the 1% efficiency of previous implementations.

The main advantage of plasmonic surfaces over other photoabsorbing materials is to enable both high absorption through the whole solar spectrum and narrow-band thermal emission as illustrated in Ref. [48]. In this configuration, the heating part is

formed by a periodic array of metallic nano-patches separated from a continuous metallic film by a thin dielectric spacer. Instead of gold or silver, the authors use high-loss refractory metals like tungsten and molybdenum that combine both high loss and high resistance to high temperatures. The absorbance spectra of the optimum design feature a broad peak over the solar spectrum that is used to collect and heat up and a sharp peak in the mid-IR that is used to feed the PV cell. The feature of the emission peak can be tuned by adjusting the structural parameters according to the Kirchhoff's law of thermal radiation. While solar TPV based on plasmonics seems to offer very interesting opportunities, this topic is only at its infancy and further studies will be necessary to confirm whether this could become a doable technology.

10.4.3. *Chemistry*

When noble metal nanoparticles are dispersed in a chemical reaction medium and illuminated at their plasmonic resonance, an increase of the chemical yield of the reaction can be observed. This observation gave birth to the research area that we name plasmon-induced nanochemistry (PINC) in this chapter. There are at least three possible mechanisms leading to the enhancement of chemical reactions using plasmonic nanoparticles[49]: (i) the optical near-field enhancement in the case of *photochemical* reactions, (ii) the local temperature increase due to light absorption and subsequent heat generation (named T-PINC, the subject that will be developed in this section) and (iii) hot-electron generation and transfer to surrounding oxidizing chemical species (see Chapter 12). In the context of the current chapter, devoted to thermoplasmonics, we will focus only on the second mechanism involving a local temperature increase.

Most chemical reactions are temperature dependent according to the Arrhenius law:

$$K(T) = A \exp(-E_a/RT),$$

where K is the equilibrium constant of the reaction, E_a is the activation energy and R is the ideal gas constant. According to this law, K is a strictly increasing function of T. Thus, nanoparticle heating can contribute to activate chemical reactions or boost reaction kinetics.

The benefits of using plasmonic nanoparticles compared with the use of a regular hot plate are the followings:

— Heating a small region makes it possible to make the thermal dynamics faster, due to a reduced thermal inertia (typically below the microsecond time scale).[1]
— Heating a micrometric area makes it possible to easily superheat the fluid above its boiling point, with possible application in solvothermal chemistry without using an autoclave.[50]
— Heating on the nanoscale enables the formation of products with a nanometer spatial resolution.

Consequently, T-PINC may not be so advantageous for producing industrial amounts of chemicals[51] and is mostly attractive upon localized heating. Indeed, when heating is performed by illuminating a macroscopic assembly of nanoparticles, the three above-mentioned benefits no longer hold and heating with a hot plate would be rather equivalent (except is the laser illumination is pulsed). The advantage of using nanometric particles is then lost, especially because of collective effects (temperature homogenization) that make the temperature completely uniform even at the nanoscale (see section 10.2).[13]

According to the literature, application of T-PINC can be divided into three main categories, all based on a two-dimensional geometry (plasmonic nanoparticles lying on a solid substrate):

— Plasmonic-assisted chemical vapor deposition.
— Thermoplasmonic-assisted catalysis in gas phase.
— Thermoplasmonic-assisted catalysis in liquids.

Here is a near exhaustive review of these three categories.

10.4.3.1. *Plasmon-assisted chemical vapor deposition (PACVD)*

In 2006, Goodwin's group[52] conducted the seminal experiment showing the concept of PINC. Chemical vapor deposition (CVD) is a well-known process suited to produce high-quality solid materials across extended regions. The authors introduced a new CVD process in which the local heating necessary to induce the deposition process was performed by local laser heating on the micrometric scale of gold nanoparticles deposited on a substrate. They named this technique PACVD (plasmon-assisted CVD). The author demonstrated microscale patterning of metal oxides such as PbO and TiO_2 on a glass substrate by local heating up to 150°C. Local laser heating for CVD, known as laser-assisted CVD (LACVD) was already known,[53] but limited by the size of the laser spot. The use of plasmonic nanoparticles strongly relaxes the requirements on the power of the heating laser and potentially enables pattering on the sub-diffraction limit. Sub-diffraction PACVD leading to the formation of silicon, germanium or carbon nanowires was later on achieved by Brongersma's group[54] in 2007, and more recently by Hwang *et al.* in 2011 in a more detailed study.[55] In 2008, Hung *et al.*[56] renamed this technique PRCVD (plasmon-resonant CVD), with little credit to the seminal work of Boyd *et al.*[52]

10.4.3.2. *Thermoplasmonic-assisted catalysis in gas phase*

Producing new catalytic compounds that are highly active under solar illumination is an important current challenge. With the pioneer works of Haruta starting in 1987,[57,58] gold is known to feature strong catalytic activity provided it is used in the form of nanoparticles supported by titanium dioxide (or other oxides), due to changes in the electronic structure, catalyst–substrate interactions, and morphological features, compared to its bulk counterpart. Today, oxidation of various volatile organic compounds, such as CO, CH_3OH, HCHO, have been demonstrated using gold nanoparticles on metal oxides, at moderately elevated temperatures (no light involved). The idea

of Chen *et al.*[59] in 2008 was to further enhance the heterogeneous catalytic activity of gold nanoparticles on oxides by heating the gold nanoparticles themselves under laser illumination. The authors used gold nanoparticles on ZrO_2 and SiO_2 (wide bandgap semiconductors to avoid any photocatalytic effect) to generate the oxidation of HCHO in CO_2. According to this work, the benefit of such an approach is that heating is limited to the catalytically active area of the system, hence a gain in energy consumption compared to a global heating approach. But the benefit compared to immersing a resistor inside the solution is not obvious. In 2010, the group of Cronin[60] borrowed the same idea to generate the exothermic oxidation of carbon monoxide in carbon dioxide, using gold nanoparticles on Fe_2O_3. The authors quantified the benefit of heating using a laser approach compared to a global heating. A gain of 2–3 orders of magnitude was evidenced.

10.4.3.3. *Thermoplasmonic-assisted catalysis in liquids*

Thermoplasmonic-assisted catalysis can also be adapted to liquid phase. This was first proposed first by Adleman *et al.* in 2009 and named PAC for plasmon-assisted catalysis.[51] The authors chose to investigate the thermal-induced reforming of a liquid mixture of ethanol and water, leading to the formation of CO_2, CO and H_2. The experiments were conducted in a microfluidic channel in order to more easily collect the gas products. The authors stressed the fact that their plasmon-assisted catalysis approach was general and could be used with a variety of endothermic catalytic processes. Along the same line, in 2011, Fasciani *et al.*[61] investigated the thermal decomposition of peroxides. The key parameter of this work was the use of a nanosecond-pulsed illumination (1 Hz repetition rate) to generate rapid and intense temperature increases, without boiling the fluid. This way, the authors evidenced that organic chemistry can be performed at relatively high-energy reaction at near-ambient temperature. In 2013, Vazquez-Vazquez *et al.*[62] presented an original approach where thermal-induced PAC was performed on the inside of hollow nanometric silica shells containing the reactants. The shell was

decorated with gold nanoparticles intended to generate the required heat upon illumination to generate the reaction. The main interest of this work was to demonstrate that local chemical reactions can be generated in a confined volume with a surrounding solvent at ambient pressure, and that the extent of reaction can be followed by SERS.

10.4.3.4. *Unidentified thermal effects in PINC*

In PINC developments, some observations have been demonstrated to originate from thermal effects, but some others are presumably attributed to optical effects or involve hot electron injection. In these two other schemes, it is often difficult to rule out any thermal contribution. This problem was pointed out by Yen and El Sayed in 2009.[63] The group studied the reduction of ferricyanide by thiosulfate and concluded that the origin of the observed enhanced chemical yield was purely due to a temperature increase of the solution, not due to a photocatalytic effect, questioning the interest of using gold nanoparticles compared to a regular hot plate. In the same spirit, the group of Cronin[60] monitored the temperature rise of their sample (simply using an infrared camera), which turned out to be significant. One can regret that all the reported works did not present such a rigorous and simple test to quantify the significance of thermal effects in PINC applications. For instance, the group of Linic conducted several studies in PINC[64,65] where a purely photocatalytic effect was claimed, while all the results are consistent with a purely thermal effect. In particular, Fig. 10.3b in Ref. [66] could be nicely fitted with the Arrhenius law, as well.

10.4.4. *Nanomedicine*

It is well established that under suitable conditions, heat can play a beneficial role in the treatment of diseases, and in particular of cancer. There exist two different regimes of hyperthermia. In the non-ablative regime, typically for temperature ranging between 41 and 45°C, heat is used as complement to radio- and chemotherapy

in order to increase their efficiency through enhanced tumor tissue oxygenation and immunological response. The second approach consists in using high enough temperature to ablate tissues (>45°C). In practice, a major challenge for both approaches is to deliver heat in a controlled, specific and non-invasive manner to the region of the patient body that needs to be treated.

A great achievement would be to deliver heat specifically and non-invasively, allowing higher temperatures locally without affecting surrounding healthy tissues. And this is where plasmonic nanoparticles can greatly contribute, acting as an artificial photothermal contrast agent upon illumination (nanoparticle-mediated hyperthermia). Note that similar principles are currently studied with both carbon nanotubes and magnetic nanoparticles.

In seminal works published almost simultaneously in 2003, the group of West,[66] on one hand, Pitsillides *et al.*[67] on the other, first demonstrated the feasibility of using plasmonic nanoparticles to ablate cancer cells upon illumination.

Shortly later, the group of West reported on longer survival of mice intravenously injected with gold nanoshells and treated with NIR light compared to the control (without nanoshells injection nor laser illumination) and the sham (laser illumination without nanoshells) groups.[68] Since then, many developments have contributed to increase the maturity of this approach to bring it closer to the clinic,[69–71] and several human trials have been performed, in particular by the company Nanospectra Biosciences Inc.[72]

Yet, despite these important advances, several challenges still prevent NP-assisted hyperthermia from becoming a widely used treatment:

(i) The first main challenge is related to the ability to specifically deliver the NPs to cancer tissues to maximize the absorption contrast with the rest of the healthy tissues. It is well established that intravenously injected NPs tend to accumulate in tumors in part due to their extra vasculature and less efficient lymphatic draining (known as enhanced permeability and

retention (EPR)). Yet, such passive targeting is inefficient, and better tumor-specific accumulation requires functionalizing the NPs surface with targeting molecules like antibodies.

(ii) A second major challenge is a consequence of the lack of control on the actual concentration of NPs delivered to the tumor that translates into a major uncertainty on the temperature increase achieved locally (similar problem as in photodynamic therapy (PDT)). A currently explored strategy to address this limitation is to create a feedback on the actual local temperature to properly adjust the illumination intensity.

(iii) Last but not least, the clinical viability of NP-assisted hyperthermia requires further knowledge about the clearance of NPs from the patient body and their possible toxicity due to accumulation in organs.

While lots of efforts are currently invested in this multidisciplinary research area that holds great promises, the entry of nanoparticle-mediated hyperthermia to the clinic will be conditioned by the possibility to overcome all the above-mentioned challenges.

References

1. Baffou, G. and Quidant, R. (2010). *Lasers Photon. Rev.* **4**, p. 864.
2. Boyer, D., Tamarat, P., Maali, A., Lounis, B. and Orrit, M. (2002). *Science* **202**, p. 1160.
3. Pitsillides, C. M., Joe, E. K., Wei, X., Anderson, R. R. and Lin, C. P. (2003). *Biophys. J.* **84**, p. 4023.
4. Hirsch, L. R., Stafford, R. J., Bankson, J. A., Sershen, S. R., Rivera, B., Price, R. E. Hazle, J. D., Halas, N. J. and West, J. L. (2003). *PNAS* **100**, p. 13549.
5. Baffou, G., Girard, C. and Quidant, R. (2010). *Phys. Rev. Lett.* **104**, p. 136805.
6. Baffou, G., Quidant, R. and Girard, C. (2009). *Appl. Phys. Lett.* **94**, p. 153109.
7. Carslaw, H. S. and Jaeger, J. C. (1986). *Conduction of Heat in Solids*, Oxford Science Publications,
8. Baffou, G., Quidant, R. and Garcia de Abajo, F. J. (2010). *ACS Nano* **4**, p. 709.
9. Baffou, G., Quidant, R. and Girard, C. (2010). *Phys. Rev. B* **82**, p. 165424.

10. Hohenester, U. and Trügler, A. (2012). *Comp. Phys. Commun.* **183**, p. 370.
11. Keblinski, P., Cahill, D. G., Bodapati, A., Sullivan, C. R. and Taton, T. A. (2006). *J. Appl. Phys.* **100**, p. 054305.
12. Govorov, A. O., Zhang, W., Skeini, T., Richardson, H., Lee, J. and Kotov, N. A. (2006). *Nanoscale Res. Lett.* **1**, p. 84.
13. Baffou, G., Berto, P., Ureña, E. B., Quidant, R., Monneret, S., Polleux, J. and Rigneault, H. (2013). *ACS Nano* **7**, p. 6478.
14. Baffou, G. and Rigneault, H. (2011). *Phys. Rev. B* **84**, p. 035415.
15. Berto, P., Mohamed, M. S. A., Rigneault, H. and Baffou, G. (2014). *Phys. Rev. B* **90**, p. 035439.
16. Heber, A., Selmke, M. and Cichos, F. (2015). *Phys. Chem. Chem. Phys.* **17**, p. 20868.
17. Kotaidis, V., Dahmen, C., von Plessen, G., Springer, F. and Plech, A. (2006). *J. Chem. Phys.* **124**, p. 184702.
18. Helb, E. Y. and Lapotko, D. O. (2008). *Nanotechnol.* **19**, p. 355702.
19. Egerev, S. *et al.* (2009). *Appl. Opt.* **48**, p. C38.
20. Vogel, A., Linz, N., Freidank, S. and Paltauf, G. (2008). *Phys. Rev. Lett.* **100**, p. 038102.
21. Inouye, H., Tanaka, K., Tanahashi, I. and Hirao, K. (1998). *Phys. Rev. B* **57**, p. 11334.
22. Grau, P., Morreeuw, J. P., Bercegol, H., Jonusauskas, G. and Vallée, F. (2003). *Phys. Rev. B.* **68**, p. 035424.
23. Huang, W., Qian, W., El-Sayed, M. A., Dong, Y. and Wang, Z. L. (2007). *J. Phys. Chem. C.* **111**, p. 10751.
24. Hodak, J. H., Henglein, A. and Hartland, G. V. (1999). *J. Chem. Phys.* **111**, p. 8613.
25. Link, S., Burda, C., Wang, Z. L. and El-Sayed, M. A. (1999). *J. Chem. Phys.* **111**, p. 1255.
26. Metwally, K., Mensah, S. and Baffou, G. (2015). *J. Phys. Chem. C* **119**, pp. 28586–28596.
27. Baffou, G., Kreuzer, M. P., Kulzer, F. and Quidant, R. (2009). *Opt. Exp.* **17**, p. 3291.
28. Clarke, M. L., Grace Chou, S. and Hwang, J. (2010). *J. Phys. Chem. Lett.* **1**, p. 1743.
29. Carlson, M. T., Khan, A. and Richardson, H. H. (2011). *Nano Lett.* **11**, p. 1061.
30. Mæstro, L. M., Haro-González, P., Coello, J. G. and Jaque, D. (2012). *Appl. Phys. Lett.* **100**, p. 201110.
31. Coppens, Z. J., Li, W., Walker, D. G. and Valentine, J. G. (2013). *Nano Lett.* **13**, p. 1023.
32. Freddi, *et al.* (2013). *Nano Lett.* **13**, p. 2004.
33. Baffou, G., Bon, P., Savatier, J., Polleux, J., Zhu, M., Merlin, M., Rigneault, H. and Monneret, S. (2012). *ACS Nano.* **6**, p. 2452.

34. Berto, P., Bermúdes Ureña, E., Bon, P., Quidant, R., Rigneault, H., Baffou, G. (2012). *Phys. Rev. B* **86**, p. 165417.
35. Bell, A., Fairfield, J. A., McCarthy, E. K., Mills, S., Boland, J. J., Baffou, G. and McCLoskey, D. (2015). *ACS Nano* **9**, p. 5551.
36. Kryder, M. H. *et al.* (2008). *Proceeding of the IEEE*, **96**, p. 1810.
37. Challener, W. A. *et al.* (2009). *Nat. Photon.* **3**, pp. 220–224.
38. Zhou, *et al.* (2014). *Nanophotonics* **3**, p. 141.
39. Stipe, *et al.* (2010). *Nat. Photon.* **4**, p. 484.
40. https://en.wikipedia.org/wiki/Seagate_Technology.
41. www.kitguru.net.
42. http://www.anandtech.com/show/9866/hard-disk-drives-with-hamr-technology-set-to-arrive-in-2018.
43. Atwater, H. A. and Polman, A. (2010). *Journal* **9**, pp. 205–213.
44. Bitnar, B., Durisch, W. and Holzner, R. (2013). *Appl. Ener.* **105**, pp. 430–438.
45. Ferrari, C., Melino, F., Pinelli, M. and Ruggero Spina, P. (2014). *Appl. Ener.* **113**, pp. 1717–1730.
46. Greffet, J. J., Carminati, R., Joulain, K., Mulet, J. P., Mainguy, S. and Chen, Y. (2002). *Nature* **416**, pp. 61–64.
47. Lenert, *et al.* (2014). *Nat. Nanotechnol.* **9**, p. 126.
48. Wu, C. (2012). *J. Opt.* **14**, p. 024005.
49. Baffou, G. and Quidant, R. (2014). *Chem. Soc. Rev.* **43**, pp. 3898–3907.
50. Baffou, G., Polleux, J., Rigneault, H. and Monneret, S. (2014). *J. Phys. Chem. C* **118**, p. 4890.
51. Adleman, J. R., Boyd, D. A., Goodwin, D. G. and Psaltis, D. (2009). *Nanolett.* **9**, pp. 4417–4423.
52. Boyd, D. A., Greengard, L., Brongersma, M., El-Naggar, M. Y., Goodwin, D. G. (2006). *Nanolett.* **6**, pp. 2592–2597.
53. Lehmann, O. and Stuke, M. (1995). *Science* **270**, pp. 1644–1646.
54. Cao, L., Barsic, D. N., Guichard, A. R. and Brongersma, M. L. (2007). *Nanolett.* **7**, p. 3523.
55. Hwang, D. J., Ryu, S. F. and Crigoropoulos, C. P. (2011). *Nanotechnol.* **22**, p. 385303.
56. Hsuang Hung, W., Hsu, I. K., Bushmaker, A., Kumar, R., Theiss, J. and Cronin, S. B. (2008). *Nanolett.* **8**, pp. 3278–3282.
57. Haruta, M., Kobayashi, T., Sano, H. and Yamada, N. (1987). *Chem. Lett.* **2**, pp. 405–408.
58. Cho, A. (2003). *Science* **299**, pp. 1684–1685.
59. Chen, X., Zhu, H. Y., Zhao, J. C., Zheng, Z. F. and Gao, X. P. (2008). *Angew. Chem. Int. Ed.* **47**, pp. 5353–5356.
60. Hsuang Hung, W., Aykol, M., Valley, D., Hou, W. and Cronin, S. B. (2010). *Nanolett* **10**, pp. 1314–1318.
61. Fasciani, C., Bueno Alejo, C. J., Grenier, M., Netto-Ferreira, J. C. and Scaniano, J. C. (2011). *Organic Lett.* **13**, pp. 204–207.

62. Vásquez-Vásquez, C., Vaz, B., Giannini, V., Pérez-Lorenzo, M., Alvarez-Puebla, R. A. and Correa-Duarte, M. A. (2013). *J. Am. Chem. Soc.* **135**, pp. 13616–13619.
63. Yen, C. W. and El-Sayed, M. A. (2009). *J. Phys. Chem. C* **113**, pp. 19585–19590.
64. Christopher, P., Ingram, D. B. and Linic, S. (2010). *J. Phys. Chem. C* **114**, pp. 9173–9177.
65. Christopher, P., Xin, H. and Linic, S. (2011). *Nat. Chem.* **3**, p. 467.
66. Hirsch, L. R. *et al.* (2003). *Proc. Natl. Acad. Sci. U.S.A.* 100, p. 13549.
67. Pitsillides, C. M., Joe, E. K., Anderson, R. R. and Lin, C. P. (2003). *Biophys. J.*, **84**, p. 4023.
68. O'Neal, D. P., Hirsch, L. R., Halas, N. J., Payne, J. D. and West, J. L. (2004). *Cancer Lett.* **25**, p. 209, pp. 171–176.
69. El-Sayed, M. A. *et al.* (2013). *PLOS* **8**, p. e76207.
70. Ayala-Orozco *et al.*, (2014). *ACS Nano* **8**, p. 6362.
71. Liu, X., Huang, N., Li, H., Wang, H., Jin, Q. and Ji, J. (2014). *Appl. Mater. & Interfaces* **6**, p. 5657.
72. www.nanospectra.com.

CHAPTER 11

Plasmon Polariton Field Mapping by Elastic Light Scattering from a Tip

EDWARD YOXALL* **and**
RAINER HILLENBRAND[†,‡,§]

*CIC NanoGUNE, Spain
†IKERBASQUE, Basque Foundation for Science, Spain
‡CIC NanoGUNE and UPV/EHU, Spain

11.1. Introduction

The most promising aspect of plasmonics is its ability to confine and manipulate light fields at nanometer length scales.[1–3] This is the basis for the development of ultrasensitive chemical sensing,[4] high-resolution microscopy[5] and optical circuitry,[6] to name just a few potential applications.[7] This confinement of fields, however, represents something of a doubled-edged sword. While it gives plasmonic devices and structures their unique properties, it also presents a number of practical challenges.

These challenges arise from the wave-like nature of light itself. With conventional, far-field optical techniques, light cannot be

*Corresponding author: §r.hillenbrand@nanogune.eu

squeezed into a space much smaller than its free-space wavelength λ. This lower bound is known as the diffraction limit and is usually approximated as $\lambda/2$. When plasmonic structures concentrate fields in very small volumes, the diffraction limit means that far-field optics lack the spatial resolution to study the field distributions created by the structures in any level of detail.

A number of sub-diffraction-limited techniques have thus been developed to fill this gap.[8,9] Some of them rely on electrons to provide the necessary resolution, as in the case of photoemission electron microscopy[10] (PEEM — photon-in/electron out), cathodoluminescence[11] (CL — electron-in/photon-out) and electron energy-loss spectroscopy (EELS — electron-in/electron out).[12] Others rely on nonlinear optical effects, including two-photon-induced luminescence[13,14] (TPL). Further techniques rely on directly probing the optical near field, which exists in the immediate vicinity of a structure (photon-in/photon-out). Among these are the two forms of scanning near-field optical microscopy (SNOM); aperture-type SNOM[15,16] (a-SNOM), which uses a tapered optical fiber to pick up fields on a surface, and scattering-type SNOM[17] (s-SNOM), which scatters near fields into the far field with a sharp tip so that they can be easily detected.

The focus of this chapter is exclusively on s-SNOM, which offers a number of advantages, including an extremely high spatial resolution (typically 10–20 nm), complete independence of the spatial resolution from the illuminating wavelength λ, allowing measurements from the visible to terahertz spectral regions, spectroscopic capabilities, discrimination between polarization states, time-resolved measurements and the ability to record both the magnitude *and* phase of plasmonic field distributions. Readers interested in the other listed techniques are directed to Ref. [9].

The sections that follow outline *how* s-SNOM is used to image plasmonic phenomena, but at this stage it is worth pointing out *why* one might want to do so in the first place. The first reason is as an experimental verification of theoretical predictions. Many facets of plasmonic behavior are first discovered with numerical simulations; imaging the field distribution around real-world objects caused by

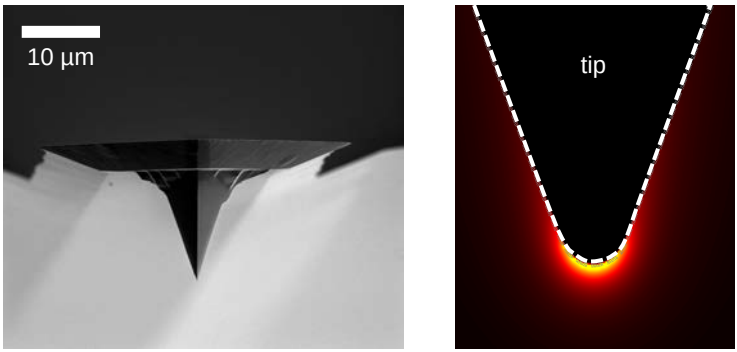
these newly discovered effects often provides conclusive evidence that the predictions are indeed correct. The second reason is that real-life materials and fabrication processes are imperfect. The performance of plasmonic devices is strongly linked to these imperfections, and the ability to examine the fields in real structures allows designers to check that their creations are working as well as they expect.

Before concluding this introduction, we offer a summary of the important nomenclature for this chapter. Strictly speaking, a *plasmon* is a collective oscillation of a free electron gas (such as exists inside a metal). A *surface plasmon* is the particular type of plasmon that is confined to a surface. When a surface plasmon interacts with light, a *surface plasmon polariton* (SPP) results. If the shape of the metallic structure supporting it does not restrict the SPP, it will propagate freely, before eventually decaying away. If the geometry of the structure is such that the SPP is confined, a resonance can result upon which the SPP is called a *localized* SPP (LSPP), or alternatively a *localized surface plasmon resonance* (LSPR). In this chapter, we adopt the more colloquial phrases *plasmon* and *plasmon resonance*, by which we really mean SPP and LSPP, respectively.

11.2. Overview of s-SNOM

The basis of s-SNOM is the atomic force microscope[18] (AFM). The AFM is itself based on the scanning tunneling microscope[19] (STM), which earned its inventors the Nobel Prize for Physics in 1986. In atomic force microscopy, a very sharp tip is raster scanned across a sample surface, and a feedback mechanism is used to regulate the tip–sample separation. In this way, an extremely high-resolution image of the sample topography is built up. It has even proven possible to image the chemical bonds of individual molecules using the technique.[20] AFM has become one of the key tools in imaging and manipulating matter at the nanoscale, and as such, a mature commercial sector exists to support its development and application. Figure 11.1(a) shows a typical AFM probe, comprising a sharp tip positioned at the end of a cantilever.

s-SNOM initially emerged in the mid-1990s[21–23] as a means of adding optical contrast to the plain topography image that AFM



(a) (b)

Fig. 11.1. (a) Electron micrograph of a typical AFM probe, made up of a sharp tip positioned at the end of a cantilever arm. The apex of the tip typically has a diameter of a few tens of nanometers. (b) The numerical simulation showing the enhancement and confinement (the "lightning rod" effect) of the electric field intensity at the apex of a nanoscale metallic tip, when the incident radiation is polarized parallel to the tip axis.

provides. This was achieved by focusing laser illumination onto the apex of a metallic AFM tip, and subsequently collecting the scattered light. Under the illumination, the metal tip acts like an optical antenna, both focusing and scattering the field in the immediate vicinity of its apex. This field enhancement at the tip apex is known as the "lightning rod" effect (see Fig. 11.1(b)), and has been exploited since the early 1980s, particularly in the field of tip-enhanced Raman spectroscopy.[24] Such strong fields at the apex allow the tip to interact with an extremely small volume of the sample in the vicinity of the strong fields, probing the sample's optical properties. The latter information is transmitted as a change in the tip-scattered light and can be detected in the far field by classical optical elements. Recording the tip-scattered field as the sample is scanned beneath the tip thus provides images of the sample optical properties with a spatial resolution defined by the size of the tip apex (typically just a few nanometers) and thus, well below the diffraction limit. Note that plasmon resonances supported by the s-SNOM probe itself have also been used to provide field enhancement at the tip apex.[25]

11.2.1. *Modes of operation*

Two main modes of operation can be identified for s-SNOM, depending on the strength of the interaction between the tip and the sample.

11.2.1.1. *Material contrast with metallic tips*

When s-SNOM is operated in the way outlined above i.e. with a metallic tip and the incident radiation (*p*-)polarized along the tip axis, the tip interacts strongly with the sample. The interaction can be qualitatively described by modelling the tip as a dipole and treating the sample as an image dipole.[26,27] In this case, differences in the near-field signal are fundamentally caused by changes in the dielectric constant across the sample surface, which we refer to as "material contrast".

As the technique has been refined (and the sensitivity increased), a large number of applications have been found for s-SNOM operated in this manner. The technique is particularly powerful when light from the mid-infrared part of the spectrum — the "chemical fingerprint" region — is used, where vibrational resonances in molecular bonds are especially good identifiers of different materials. Examples include polymer studies,[28–30] phonons,[31,32] charge-carrier density mapping,[33] and biological materials[34–37] to name just a few. Similarly, terahertz radiation has been used to map the conductivity of a transistor's cross-section.[38] Given that the spatial resolution in s-SNOM is independent of the wavelength and defined only by the tip's diameter, the gain in resolution compared to traditional far-field optics is particularly marked in the infrared and terahertz spectral regions.

11.2.1.2. *Near-field mapping with dielectric tips and s-polarized incident light*

The early attempts to image field distributions created by plasmon resonances with an s-SNOM followed a similar approach to what had worked so successfully for mapping material differences; a metal tip was used with *p*-polarized light to measure resonant field patterns

around and between gold nanoparticles.[39–41] In this case, the tip scattered the nanoparticles' local near fields into the far field, a process where the spatial resolution is again defined by the size of the tip's radius and thus sub-diffraction limited. It was quickly realized, however, that the large conductivity of the metallic tip could be perturbing the plasmon fields[42] (indeed, this perturbation can be so large that it has since been shown that the presence of a metallic tip can be used to precisely control an antenna resonance[43]).

Attempts to minimize any tip-induced distortion led to the search for alternative approaches. A carbon nanotube probe, for example, was shown to be capable of imaging a simple dipolar resonance on nanoscale gold islands.[44] The biggest breakthrough, however, came from using a silicon tip *and* rotating the polarization of the incident illumination from p to s, such that it lays perpendicular to the axis of the tip[45] (a full description of this experimental setup is given in the following section). The idea behind this approach was twofold: firstly, direct excitation of the tip would be avoided; and secondly, by choosing a dielectric over a metal, the tip would be a weak scatterer. These two advances greatly reduced the perturbation of the sample's near-fields by the tip, allowing for much more straightforward image interpretation, where even higher order resonances could easily be distinguished.[45–47]

When the incident light is s-polarized, the plasmon resonance in the sample structure can be strongly excited (while the tip is hardly excited at all — see Fig. 11.2). In this case, the tip only weakly interacts with the sample, and any light that is scattered by the tip can be directly attributed to the presence of the sample's near fields. Note that this is true for any sample structure that creates an electromagnetic (EM) field distribution (not just plasmon resonances): indeed, the near-field distribution around purely dielectric structures has even been mapped using s-SNOM[48]. Furthermore, it has been shown — to a first approximation — that the s and p-polarization states of the scattered light, which we label E^S and E^P, are related to the in-plane (E_x or E_y), and out-of-plane (E_z) components of the near-field distribution. Under this simplistic assumption, selection of

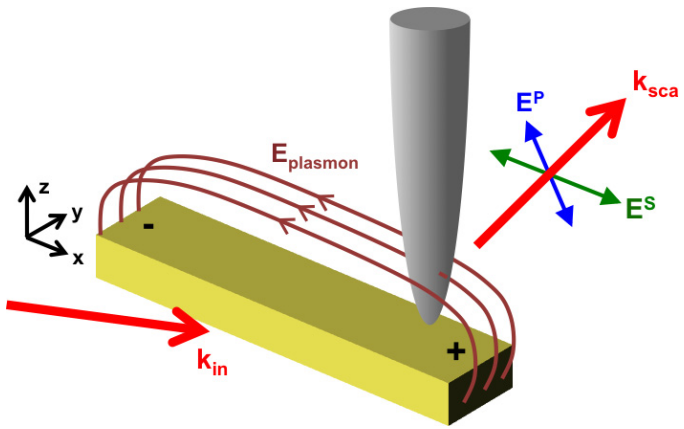


Fig. 11.2. Under a simplistic interpretation, incident radiation (with wavevector k_{in}) excites a plasmon resonance. The s-SNOM tip is subsequently illuminated by the near fields of the structure (field lines labeled E_{plasmon}), and scatters them into the far field (with wavevector k_{sca}). The *s* and *p*-polarized scattered fields, E^S and E^P, are then proportional to the in-plane, E_{x}, and out-of-plane, E_{z}, components of the near fields, respectively.

just the *p*-polarized scattered light E^P by means of a polarizer in the far field means that the out-of-plane component of the near-field distribution can be mapped. Likewise, the *s*-polarized scattered light E^S can be isolated to map the in-plane component of the resonance.[49] Note that the particular case where the *p*-polarized scattered light E^P is isolated i.e. *s*-polarization in, *p*-polarization out, is known as the cross-polarization scheme and is one of the more commonly used experimental setups.[50]

11.2.2. *A note regarding polarization sensitivity*

The interpretations that $E^S \sim E_x$ and $E^P \sim E_z$ are valid when the s-SNOM tip is the principle scatterer in a measurement.[41,51,49] Care must be taken, however, in analyzing images for which the sample structure also scatters strongly.[52] When this is the case, the tip not only scatters directly into the far field, but may also scatter via the structure due to a tip-structure coupling.[53] E^S and E^P can then depend on a much more complex combination (both

linear and nonlinear) of the near-field components, $E_{x,y,z}$. The image interpretation for measurements of the overall scattered field is therefore challenging, and developing a full understanding of this phenomenon is currently an active area of research.

11.3. Experimental Setups

11.3.1. *Basic AFM/s-SNOM principles*

The s-SNOM setup is based upon that of an AFM operating in intermittent contact (or "tapping") mode. For this mode of operation, the cantilever to which the sharp tip is attached (see Fig. 11.1(a)) is driven to oscillate in the vertical (z-)direction at its resonant frequency Ω, which is typically a few hundred kilohertz. The amplitude of oscillation (usually a few tens of nanometers) is monitored and is used to regulate the tip–sample distance at a constant value as the sample is scanned. By recording the vertical position of the sample as a function of its lateral position, the topography of the sample surface (with a resolution roughly equal to that of the tip's diameter) is acquired.

As mentioned in Section 11.2, what distinguishes s-SNOM from AFM is the light focused onto the tip–sample system. This focusing is often accomplished with a reflective parabolic mirror, rather than a dielectric lens, as this allows for easy transition between different spectral regions.

11.3.2. *s-SNOM with interferometry*

To extract both the amplitude and phase of the scattered field (i.e. the complex field), s-SNOM must be applied in conjunction with an interferometer. The knowledge of the phase is important in s-SNOM[54] regardless of whether it is being applied to measure material contrast (where the phase is often linked to the sample's absorption coefficient[55,56]) or in field distribution mapping (where the phase enables direct investigation of the modes in nanostructures, or the determination of the phase velocity direction in a propagating mode).

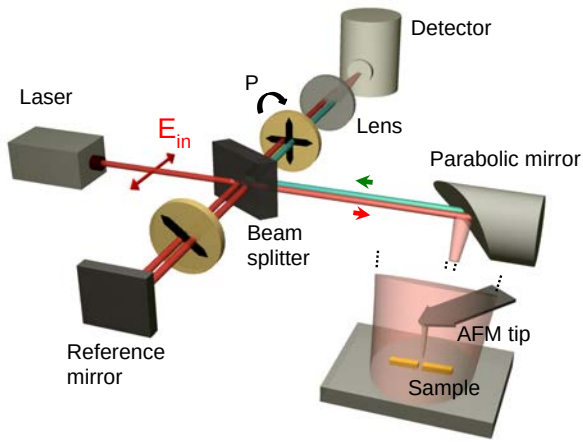


Fig. 11.3. Typical experimental setup of a side-illuminated cross-polarization scheme. Adapted from P. Alonso-González *et al.*[77]

Note that using an interferometer serves the additional purpose of amplifying the near-field signal. Such interferometric amplification has the special property that the signal can be amplified to a large amount without affecting the signal-to-shot-noise limit[57]; this is particularly important when the near-field signal is very weak.

11.3.2.1. *Side-illumination mode*

Figure 11.3 shows the typical experimental setup for mapping near-field distributions with an s-SNOM operated in side-illumination mode. The *s*-polarized light from a laser is first passed through a beamsplitter, where it splits into two arms. In the *signal arm* of the interferometer, the light is weakly focused (usually by a parabolic mirror) to the sample where the plasmon is excited. In the *reference arm*, light is reflected from a mirror. Note that if the *p*-polarized scattered field E^P is being detected, a polarizer at 45° must be placed in the reference arm to introduce a *p*-polarized component to the beam (without this addition, light from signal and reference arms would not constructively interfere). After recombination of the two beams at the beamsplitter, a final polarizer, labeled P, is positioned prior to the detector to select either the *s* or *p*-polarization.

The side-illumination scheme is a commonly used experimental setup, largely because it is very similar to the s-SNOM setup used for mapping material contrast across a sample surface with a metallic tip[17] (the only difference being the presence of the polarization optics). Care must be taken, however, with the effects of retardation due to the angled incident illumination. Retardation becomes significant when the phase of the illuminating beam varies across the area of the plasmon-supporting structure, as illustrated in Fig. 11.4. Side-illumination, therefore, is typically only suitable for narrow structures whose depth is considerably shorter than the illumination wavelength, such as simple rod antennas.

11.3.2.2. *Transmission mode*

Retardation effects can be overcome by operating an s-SNOM in transmission mode, where the illuminating light is normally incident upon the sample surface.[58,59] A typical setup is shown in Fig. 11.5. Light is weakly focused to the sample from beneath, thus providing a nearly homogenous sample illumination. This illumination scheme is similar to the experimental configuration in antenna-enhanced sensors and photodetectors, where light is typically normally incident on the sample. Illuminating the sample from beneath, however, means that the sample structures must be fabricated on a transparent substrate. Similar to the side-illumination mode, the tip-scattered field

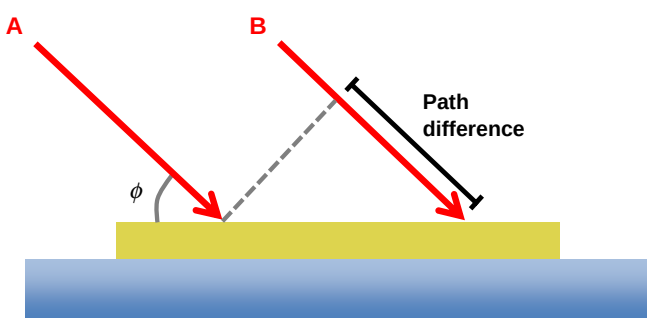


Fig. 11.4. The retardation effect. When two rays, labeled A and B, are obliquely incident upon a surface (at angle ϕ from the sample plane here), the phase of the electric field at the surface is non-uniform due to the path length difference.

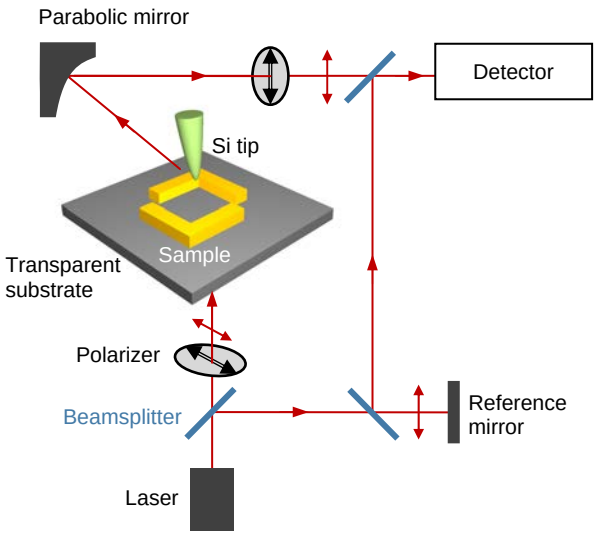


Fig. 11.5. Typical experimental setup of a transmission mode cross-polarization scheme. Adapted from Shang-Chi Jiang *et al.*[94]

is then combined with a reference beam at a beamsplitter, and the interference of the two measured at a detector.

The transmission mode setup allows the near-field distribution of much larger structures to be reliably mapped. However, it is technically more difficult to implement than the side-illumination mode. Transmission mode requires that two foci be overlapped — those of the focusing lens and the collection optics (the parabolic mirror in Fig. 11.5). Side-illumination, on the other hand, does not present this difficulty, as the same optical element acts to both focus light to the tip and sample and to recollimate the backscattered light.

11.4. Background Suppression

One of the general problems encountered in s-SNOM is how to distinguish the near-field scattering from the much larger background of light scattered from the body of the tip and directly from the sample itself. This can be seen by describing the overall scattered light as

the linear superposition of two scattering coefficients:

$$E_{\text{sca}} = \sigma E_{\text{in}} = (\sigma_{\text{NF}} + \sigma_{\text{B}})E_{\text{in}}, \tag{11.1}$$

where E_{sca} represents the overall scattered field, E_{in} the incident field, σ the overall scattering coefficient, σ_{NF} the near-field scattering coefficient and σ_B the background scattering coefficient.

11.4.1. *Distinguishing the near-field scattering and background scattering*

What allows the near-field scattering σ_{NF} and the background σ_B to be discerned is their very different sensitivities to the tip–sample separation distance: as σ_{NF} decays to negligible levels within a few nanometers of the sample surface, it is highly affected by the tip's vertical oscillations of a few tens of nanometers[17,60] (see Section 11.3.1, $\Omega \sim$ a few hundred kilohertz). The background scattering σ_B, on the other hand, only significantly varies over length scales equal to the illuminating wavelength, which is usually an order of magnitude or two larger than the tip oscillation amplitude. Most of the variation in the overall scattered field E_{sca} due to the tip's oscillation, therefore, comes from the near-field scattering coefficient σ_{NF}. By denoting the scattering coefficient of the overall scattering σ as a Fourier series, such that:

$$\sigma = \sum_{n=-\infty}^{\infty} \sigma_{\text{n}} e^{\text{in}\Omega t}, \tag{11.2}$$

it follows that there is a greater contribution of σ_{NF} in the higher harmonics of σ_n. Correspondingly, the ratio $\sigma_{\text{NF},n}$:$\sigma_{B,n}$ is also larger for the higher harmonics of σ_n. For sufficiently high harmonics of σ_n, therefore, the background scattering $\sigma_{B,n}$ has a negligibly small influence on the overall scattering σ_n, and the only contribution comes from the near-field coefficient $\sigma_{\text{NF},n}$. For the visible and near-infrared parts of the spectrum, this typically occurs at $n = 3$, whereas in the mid-infrared spectral range we find $n = 2$, being usually sufficient to suppress the background.[61]

11.4.2. *Removing the multiplicative background*

Simply demodulating the detector output at a frequency $n\Omega$, however, is not enough to completely remove the effects of the background scattering. Photodetectors register a voltage, U, proportional to the light intensity that falls upon them, and this intensity is itself related to the square of the scattering coefficient, σ^2. As a result, the detected voltage contains multiplicative cross-terms between the near-field and background Fourier coefficients, $\sigma_{\mathrm{NF},n}$ and $\sigma_{B,n}$. The demodulated detector signal at a given harmonic n is thus contaminated by the DC background scattering:

$$U_n \sim \sigma^*_{B,0}\sigma_{\mathrm{NF},n} \tag{11.3}$$

The influence of this multiplicative background can be completely removed by making use of the reference beam in the experimental setup's interferometer. When the reference beam, E_R, is included, the demodulated photodetector signal is given by:

$$U_n \sim E^*_R\sigma_{\mathrm{NF},n} + \sigma^*_{B,0}\sigma_{\mathrm{NF},n} \tag{11.4}$$

If the reference beam E_R is now modulated, a modulation of the $E_R\sigma_{\mathrm{NF},n}$ term is introduced, allowing for this term to be isolated. Several variations on this theme have been used as s-SNOM detection techniques, including two-step homodyne with sufficiently strong reference beam,[62] heterodyne,[63] pseudo-heterodyne,[64] phase-shifting,[65] holographic[66] and generalized lock-in detection.[67] The case where a reference beam is not used, and the background scattering alone is relied upon to amplify the near-field signal as shown in Eq. (11.3), is known as self-homodyne detection.[68] Although experimentally the most simple, it is not a preferred technique as the background scattering can vary depending on the position of the probe on the sample, often leading to difficulties in image interpretation (although it is still occasionally used for rapid point spectroscopy[69]). The benefit of using a reference beam, on the other hand, is that the factor E_R is constant, regardless of the probe's position.

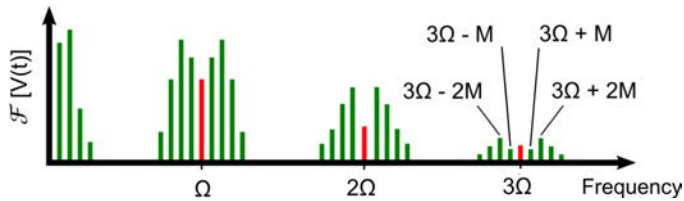


Fig. 11.6. Schematic for the Fourier transform of the pseudo-heterodyne techniques detector signal $V(t)$. The harmonics of the tip's oscillation frequency Ω are split into sidebands by the phase modulation of the reference beam, separated by multiples of the reference mirror oscillation frequency M. Adapted from M. Ocelic *et al.*[64]

11.4.3. *The pseudo-heterodyne technique*

Foremost amongst these detection methodologies for most of the past decade is the pseudo-heterodyne technique. First demonstrated in 2006,[64] it has risen to its position of prominence thanks to its ability to simultaneously measure both the amplitude and the phase of the near-field signal, and also for the simplicity of the optical components it requires, making it easy to use in any spectral range.

Pseudo-heterodyne detection works by sinusoidally phase-modulating the reference beam. This is achieved by oscillating the mirror in the reference arm of the interferometric setup (see Fig. 11.3) with a frequency M, which is typically a few hundred Hz, and thus, much smaller than Ω. As outlined above, the effect of the reference beam's phase modulation is to split the harmonics of the tip's oscillation frequency $n\Omega$ into sidebands, which are separated in frequency by the value of M. Figure 11.6 shows this principle schematically.

It can be shown that the odd and even sidebands are related to the imaginary and real parts of the near-field scattering $\sigma_{\mathrm{NF},n}$, respectively, in case the mirror oscillation amplitude is 21% of the illumination wavelength.[64] If we label the value of a given sideband as $C_{n\Omega+mM}$, the amplitude s_n and phase ϕ_n of the complex near-field coefficient are given by:

$$s_n = \sqrt{|C_{n\Omega+M}|^2 + |C_{n\Omega+2M}|^2}, \tag{11.5}$$

$$\phi_n = \tan^{-1} \frac{|C_{n\Omega+M}|}{|C_{n\Omega+2M}|}. \tag{11.6}$$

These experimentally measured values of s_n and ϕ_n thus correspond to the amplitude and phase of the near field (e.g. from an antenna or propagating polariton) at any given point on the sample surface.

11.5. Examples

For the efficient excitation of a polariton, the wavevector of the exciting radiation must be matched to that of the polariton[70] (known as phase or momentum matching). This concept is easily understood by examining a typical dispersion curve of a surface plasmon polariton, as seen in Fig. 11.7, and noting that the light line and the polariton dispersion do not overlap.

At this stage, it is useful to point out a practical difference between localized polariton resonances and propagating polaritons. The former can be excited by direct light illumination (a consequence of the curved surfaces of their supporting structures), whereas it is impossible to excite the latter by simply shining light onto a metal–dielectric interface (due to the wavevector mismatch Δk). In far-field

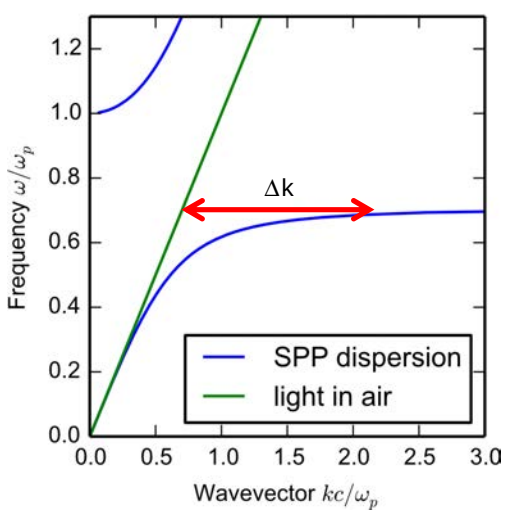


Fig. 11.7. SPP dispersion curve assuming zero loss. The wavevector mismatch Δk between light in free space and the SPP means that simply shining light on a metal-dielectric interface is not sufficient to excite the polariton.

experiments, propagating polaritons are usually launched in one of two ways; by using the evanescent fields that emanate from a prism in which light is totally internally reflected (in either the Kretschmann[71] or Otto[72] configuration), or by using a surface defect such as a slit or grating to provide the coupling.[73,74] Near-field experiments, on the other hand, typically launch polaritons using one of two approaches; either by fabricating antenna structures on the sample, or by making use of the antenna properties of the s-SNOM tip itself.

11.5.1. *Localized polaritons*

11.5.1.1. *Dipole antennas*

The simplest example of a near-field distribution measured by s-SNOM is the dipole antenna. Such antennas have been well studied in the literature, often as proof-of-principle experiments.[44–47,58,75,76] Here, we show the results of Schnell *et al.*,[58] where pseudo-heterodyne detection was used alongside the transmission mode to map the out-of-plane E_z component of a gold rod antenna via the cross-polarization scheme (Fig. 11.8).

The near-field images show "hot spots" at each end of the antenna, typical of the supported dipole mode. The optical phase shows that these areas of strong field enhancement have opposite phases, as would be expected for the out-of-plane E_z component of the antenna's near field (represented schematically in Fig. 11.2).

11.5.1.2. *Coupled infrared antennas*

As a straightforward demonstration of s-SNOM's ability to map the different near-field components of plasmonic nanoantennas by measuring the scattered light's different polarization states E^S and E^P, we present the key results of a paper published by Alonso-González *et al.* in 2013 on the coupling between two infrared nanoatennas.[77] Note that coupled antenna fields have been extensively studied with s-SNOM.[46,58,59,78,79] For the paper in question here, the authors used pseudo-heterodyne detection in conjunction with the side-illumination mode (as shown in Fig. 11.3) to image two gold rod

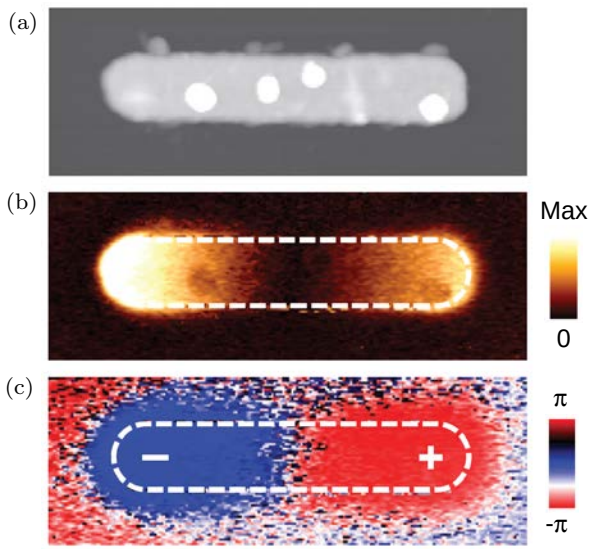


Fig. 11.8. (a) Topography showing a 1550 nm long gold rod with a 230 nm diameter. (b) Optical amplitude s_3 of the p-polarized scattered light, showing the E_z component of the rods near field. (c) Optical phase ϕ_3. The illumination wavelength was 9.6 μm. Reprinted from M. Schnell *et al.*[58]

antennas separated by a small gap on a calcium fluoride substrate (topography shown in Fig. 11.9a).

The p- and s-polarization scattered fields, demodulated at the 4th harmonic of the tip's oscillation frequency, are shown in Fig. 11.9(b) and 11.9(c). Note that the phase images ϕ_4 have been grayed out away from the antennas: when the signal is weak, the phase is essentially a random value between π and $-\pi$. For the p-polarized scattered field (revealing the out-of-plane near-field component), each nanoantenna shows a large amplitude signal at its extremities that oscillates out-of-phase by π radians (180°). In the gap, no field is observed, as there is no out-of-plane component in this location. For the s-polarized scattered field, an intense "hot spot" is observed in the gap (indicating a capacitive coupling between the antennas), whereas no phase shift is observed between the extremities. It reveals the strongly concentrated in-plane fields inside the gap.[93]

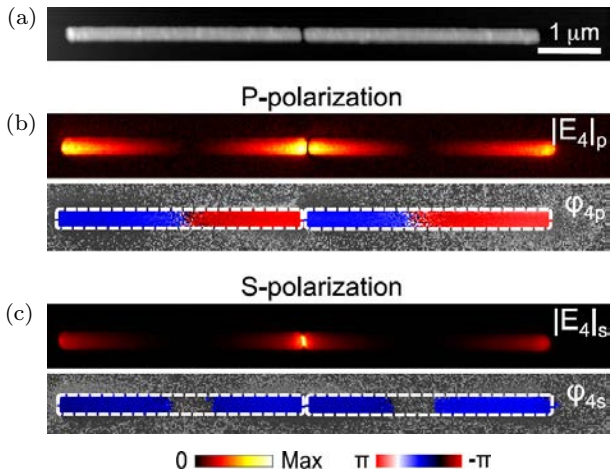


Fig. 11.9. (a) Topography of coupled nanoantennas. (b) *P*-polarized scattered light's amplitude and phase demodulated at the 4th harmonic of the tip oscillation frequency. (c) As for (b), but for *s*-polarized scattered light. The illumination wavelength is 11.1 μm. Reprinted from P. Alonso-González *et al.*[77]

11.5.2. *Propagating polaritons*

Propagating polaritons in near-field experiments are often launched with antennas, and these fall into two general categories: resonant and non-resonant. Resonant antennas (for example, a bar antenna supporting a dipole mode) are used when the coupling strength between the incident light and the polariton needs to be maximized.[80] Non-resonant structures (such as a simple metal edges,[81] disks[82] or slits[83]) have also been demonstrated and have the benefit of working across broad spectral ranges.

Alternatively, propagating polaritons can be excited directly by an s-SNOM tip.[84] Such polaritons propagate radially away from the tip, and can only be detected if they reflect from an object or an edge to create a standing wave. As the sample is raster scanned, the tip thus passes through areas of constructive and destructive interference, creating a series of fringes.[66,85–87] The spacing of these fringes is approximately $\lambda_p/2$, where λ_p is the polariton wavelength. One useful application of tip-launched polaritons is that by tuning

the frequency of the incident light and measuring the fringe spacing, the dispersion curve can be experimentally deduced.[32,88]

11.5.2.1. *Transmission lines and waveguides*

The study of transmission lines and waveguides is another research field in which s-SNOM has proved a useful tool for mapping near-field distributions.[89–91] A recent example is the work by Andryieuski *et al.*[92] in which mode propagation for plasmonic slot waveguides was studied for telecommunication wavelengths (1.55 μm). The topography of the waveguide is shown in Fig. 11.10a: a rectangular slot of width 300 nm was cut into a gold film along with two dipole antennas that would act as nanocouplers for the waveguide mode. For the experimental setup, this study used pseudo-heterodyne detection

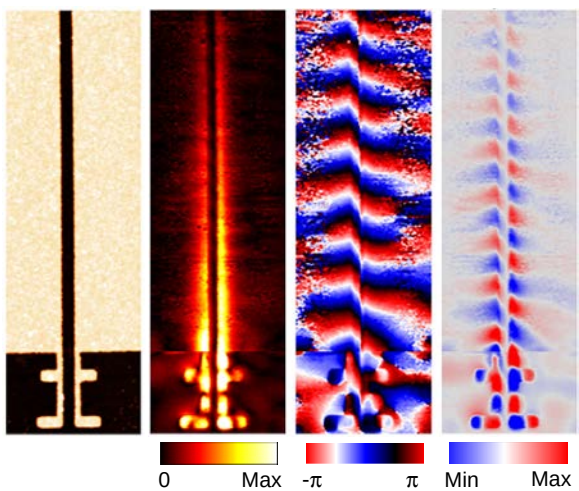


Fig. 11.10. (a) Slot-waveguide topography. (b) Amplitude, (c) phase and (d) real part of the slot mode fields, as extracted from the 2nd harmonic of the *p*-polarized tip-scattered light. The illumination wavelength is 1.55 μm. Reprinted with permission from A. Andryieuski *et al.*[92] Copyright 2014, American Chemical Society.

with transmission mode illumination (see Fig. 11.5). In terms of polarization sensitivity, the cross-polarization scheme was used, i.e. only the *p*-polarized scattered light was detected.

The amplitude, phase and real part of the scattered light stemming from the propagating waveguide mode (demodulated at the second harmonic of the tip's oscillation frequency) are shown in Fig. 11.10(b)–(d), respectively. Note that these images do not represent the raw recorded data — the raw data contain a background SPP field, which is simultaneously excited by the incident light. The images in Fig. 11.10 represent the data after this background has been removed.

The optical images reveal a propagating slot mode with decreasing amplitude and a linearly rotating phase. There is no out-of-plane field component in the middle of the slot (similar to the coupled antenna case of Section 11.5.1), and the field has opposite signs on each side of the slot at any given position.

The s-SNOM characterization of the near-field distribution of waveguide allows the authors to determine the efficiencies of different coupling antennas as well as determine the effective index and propagation lengths of the waveguide mode.

11.5.2.2. *Plasmon focusing*

As an example of mapping propagating SPPs, we summarize here work by Gjonaj *et al.*,[83] published in 2014. In this paper, the authors demonstrated the focusing of plasmons into a sub-100 nm area. The plasmons themselves were supported by a silver–silicon nitride layered structure and were launched by curved slits (Fig. 11.11(a)). Experimentally, the s-SNOM used to record the plasmon fields was operated in transmission mode with an illumination wavelength 532 nm. Pseudo-heterodyne detection was used to map both the amplitude and phase of the SPP (Fig. 11.11(b),(c)). Note that no polarization (i.e. *s* or *p*) was selected in the far-field: as a result, the amplitude and phase maps primarily represent the out-of-plane (E_z) component of the plasmon's field due to the tip's stronger scattering for fields along its principle axis.

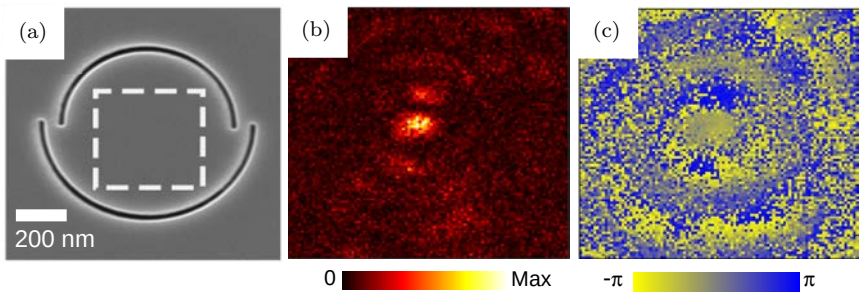


Fig. 11.11. (a) SEM image of the "plasmonic lens" used for focusing the SPPs. Top arc has a radius of 1 μm. The dashed box outlines area shown in (b) and (c). (b) Measured optical amplitude s_3 and (c) phase ϕ_3, representing the out-of-plane (E_z) component of the propagating SPPs. The incident wavelength is 532 nm. Reprinted with permission from B. Gjonaj *et al.*[83] Copyright 2014, American Chemical Society.

The optical images show curved wavefronts launched by the slits, which propagate to a small (∼70 nm) area of high confinement and optical power. This focus is the result of the constructive interference between two counter-propagating SPPs — the authors envision that such a plasmon focusing system might be used for *in vivo* biological imaging.

11.5.2.3. *Graphene plasmons*

One of the more promising candidates in the search for plasmonic materials that exhibit lower optical losses is graphene. Graphene is advantageous for several reasons; its 2D nature, high carrier mobility and ability to control the carrier concentration by electrical gating are among the most important. The wavelength of graphene plasmons, however, is particularly short compared to the wavelength of light in free space at the same frequency — for graphene plasmons to be launched, therefore, this large momentum mismatch must be overcome.

This was first demonstrated by Chen *et al.*[86] and Fei *et al.*[87] in 2012, who used a metallic s-SNOM tip to launch, and subsequently detect, graphene plasmons (with pseudo-heterodyne detection, and *p*-polarized incident light to maximize the near fields at the tip's

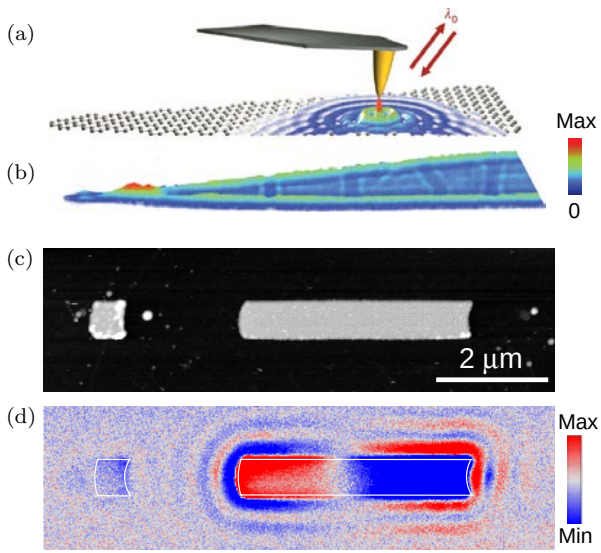


Fig. 11.12. (a) Schematic for tip-launching graphene plasmons in a graphene ribbon (12 μm long, 1 μm wide). Reprinted from J. Chen *et al.*[86] (b) Optical amplitude s_3, with an incident wavelength of 9.7 μm. (c) Topography of the sample showing off-resonance (left) and on-resonance (right) gold antennas on graphene. (d) Real part of the *p*-polarized tip-scattered light, representing the out-of-plane (E_z) component of the antenna/graphene plasmon fields. The illumination wavelength was 11.06 μm. From P. Alonso-González *et al.*[80] Reprinted with permission from AAAS.

apex). Figure 11.12(a) shows a schematic of the experiment, where the standing waves created by reflection of the tip-launched graphene plasmons from the edges of a tapered graphene ribbon are mapped (Fig. 11.12b). A series images at different incident frequencies allowed the authors to experimentally measure the graphene plasmon dispersion curve.

In 2014, Alonso-González *et al.* showed that this was also possible to launch graphene plasmons using resonant gold antennas.[80] These antennas provided strong near-fields at their extremities with large enough momenta to launch the plasmons. The topography of the sample is shown in Fig. 11.12(c), where both an off-resonance (left) and on-resonance antenna (right) are seen. The real part of the out-of-plane component (E_z) of the graphene

plasmons is shown in Fig. 11.12(d). Experimentally, this was mapped by using pseudo-heterodyne detection in conjunction with the side-illumination mode setup operated using the cross-polarization scheme (i.e. *s*-polarization in, *p*-polarization out).

The distance between the field maxima of the same polarity yields the graphene plasmon wavelength (380 nm), which is smaller than the free-space illumination wavelength (11.06 μm) by a factor of 27. Note that the off-resonance antenna, which does not exhibit a resonant field enhancement at its extremities, barely launches the plasmons at all.

11.6. Conclusion

s-SNOM is a powerful tool for characterizing plasmonic phenomena with nanoscale spatial resolution and within a broad spectral range including visible and infrared frequencies. It is based on AFM, where the probing tip acts simultaneously as mechanical and optical probe. Nanoscale optical imaging is achieved by recording the elastically scattered light simultaneously to topography. Interferometric detection thereby yields both, near-field optical amplitude and phase images. Imaging of the near-field distribution and wavefronts of localized and propagating surface plasmon polaritons on metal films, antennas and waveguide structures is often performed with dielectric Si tips. They essentially act as passive scattering probes, and thus minimally disturb the plasmonic structures. Plasmon mapping can be also performed with metallic tips, but in this case the strong field enhancement at the very tip apex might disturb plasmonic field distribution. On the other hand, the strongly confined near fields at the tip apex can provide the necessary momenta to launch strongly confined plasmons, for example in doped graphene. The tip-launched plasmons can be imaged when they are reflected at edges or defects, thus yielding plasmon interference patterns in the near-field images.

Plasmon imaging by s-SNOM can be applied for studying fundamental aspects of plasmon localization and propagation, but also for verifying novel design concepts and to perform quality control of emerging plasmonic devices. With the emergence of 2D materials

such as graphene, s-SNOM encounters a novel and exciting field, where exotic polariton phenomena are waiting to be explored.

References

1. Atwater, H. A. (2007). The Promise of Plasmonics. *Sci. Amer.* **17**, pp. 56–63.
2. Brongersma, M. L. and Shalaev, V. M. (2010). The case for plasmonics. *Science* **328**, pp. 440–441.
3. Schuller, J. A. *et al.* (2010). Plasmonics for extreme light concentration and manipulation. *Nat. Mater.* **9**, pp. 193–204.
4. Valsecchi, C. and Brolo, A. G. (2013). Periodic metallic nanostructures as plasmonic chemical sensors. *Langmuir* **29**, pp. 5638–5649.
5. Kawata, S., Inouye, Y. and Verma, P. (2009). Plasmonics for near-field nano-imaging and superlensing, *Nat. Photonics* **3**, pp. 388–394.
6. Engheta, N. (2007). Circuits with light at nanoscales: Optical nanocircuits inspired by metamaterials, *Science* **317**, pp. 1698–1702.
7. Stockman, M. I. (2011). Nanoplasmonics: The physics behind the applications, *Phys. Today* **64**, pp. 39–44.
8. Zheludev, N. I. (2008). What diffraction limit? *Nat. Mater.* **7**, pp. 420–422.
9. Vogelgesang, R. and Dmitriev, A. (2010). Real-space imaging of nanoplasmonic resonances. *Analyst* **135**, pp. 1175–1181.
10. Aeschlimann, M. *et al.* (2007). Adaptive subwavelength control of nano-optical fields, *Nature* **446**, pp. 301–304.
11. Bashevoy, M. V., Jonsson, F., MacDonald, K. F., Chen, Y. and Zheludev, N. I. (2007). Hyperspectral imaging of plasmonic nanostructures with nanoscale resolution, *Opt. Express* **15**, p. 11313.
12. Nelayah, J. *et al.* (2007). Mapping surface plasmons on a single metallic nanoparticle. *Nat. Phys.* **3**, pp. 348–353.
13. Muhlschlegel, P. (2005). Resonant optical antennas, *Science* **308**, pp. 1607–1609.
14. Ghenuche, P., Cherukulappurath, S., Taminiau, T. H., van Hulst, N. F. and Quidant, R. (2008). Spectroscopic mode mapping of resonant plasmon nanoantennas. *Phys. Rev. Lett.* **101**,
15. Ash, E. A. and Nicholls, G. (1972). Super-resolution aperture scanning Microscope, *Nature* **237**, pp. 510–512.
16. Babayan, Y. *et al.* (2009). Confining standing waves in optical corrals, *ACS. Nano* **3**, pp. 615–620.
17. Keilmann, F. and Hillenbrand, R. (2004). Near-field microscopy by elastic light scattering from a tip. *Philosophical Transactions of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences* **362**, pp. 787–805.
18. Binnig, G., Quate, C. F. and Gerber, C. (1986). Atomic Force Microscope. *Phys. Rev. Lett.* **56**, pp. 930–933.

19. Binnig, G., Rohrer, H., Gerber, C. and Weibel, E. (1982). Surface studies by scanning tunneling microscopy, *Phys. Rev. Lett.* **49**, pp. 57–61.
20. Gross, L., Mohn, F., Moll, N., Liljeroth, P. and Meyer, G. (2009). The chemical structure of a molecule resolved by atomic force microscopy. *Science* **325**, pp. 1110–1114.
21. Kawata, S and Inouye, Y. (1995). Scanning probe optical microscopy using a metallic probe tip. *Ultramicroscopy* **57**, pp. 313–317.
22. Bachelot, P. Gleyzes, and A. C. Boccara, (1995). Near-field optical microscope based on local perturbation of a diffraction spot. *Opt. Lett.* **20**, pp. 1924-1926.
23. Zenhausern, F., Martin, Y. and Wickramasinghe, H. K. (1995). Scanning interferometric apertureless microscopy — optical imaging at 10 angtrom resolution. *Science* **269**, pp. 1083–1085.
24. Stöckle, R. M., Suh, Y. D. Deckert, V., Zenobi, R. (2000). Nanoscale chemical analysis by tip-enhanced Raman spectroscopy. *Chem. Phys. Lett.* **318**, pp. 131–136.
25. Berweger, S., Atkin, J. M., Olmon, R. L. and Raschke, M. B. (2012). Light on the tip of a needle: Plasmonic nanofocusing for spectroscopy on the Nanoscale. *J. Phys. Chem. Lett.* **3**, pp. 945–952.
26. Knoll, B. and Keilmann, F. (2000). Enhanced dielectric contrast in scattering-type scanning near-field optical microscopy. *Optics Communications* **182**, pp. 321–328.
27. Cvitkovic, A., Ocelic, N. and Hillenbrand, R. (2007). Analytical model for quantitative prediction of material contrasts in scattering-type near-field optical microscopy. *Opt. Exp.* **15**, pp. 8550–8565.
28. Knoll, B. and Keilmann, F. (1999). Near-field probing of vibrational absorption for chemical microscopy, *Nature* **399**, pp. 134–137.
29. Taubner, T., Hillenbrand, R. and Keilmann, F. (2004). Nanoscale polymer recognition by spectral signature in scattering infrared near-field microscopy, *Appl. Phys. Lett.* **85**, pp. 5064–5066.
30. Raschke, M. B. *et al.* (2005). Apertureless near-field vibrational imaging of block-copolymer nanostructures with ultrahigh spatial resolution. *Chem. Phys. Chem.* **6**, pp. 2197–2203.
31. Hillenbrand, R., Taubner, T. and Keilmann, F. (2002). Phonon-enhanced light-matter interaction at the nanometre scale, *Nature* **418**, pp. 159–162.
32. Dai, S. *et al.* (2014). Tunable phonon polaritons in atomically thin van der waals crystals of boron nitride, *Science* **343**, pp. 1125–1129.
33. Stiegler, J. M. *et al.* (2010). Nanoscale free-carrier profiling of individual semiconductor nanowires by infrared near-field nanoscopy, *Nano Lett.* **10**, pp. 1387–1392.
34. Brehm, M., Taubner, T., Hillenbrand, R. and Keilmann, F. (2006). Infrared spectroscopic mapping of single nanoparticles and viruses at nanoscale resolution. *Nano Lett.* **6**, pp. 1307–1310.
35. Ballout, F. *et al.* (2011). Scanning near-field IR microscopy of proteins in lipid bilayers. *Phys. Chem. Chem. Phys.* **13**, pp. 21432–21436.

36. Paulite, M. *et al.* (2011). Imaging secondary structure of individual amyloid fibrils of a beta(2)-microglobulin fragment using near-field infrared spectroscopy, *J. Am. Chem. Soc.* **133**, pp. 7376–7383.
37. Amenabar, I. *et al.* (2013). Structural analysis and mapping of individual protein complexes by infrared nanospectroscopy, *Nat. Commun.* **4**,
38. Huber, A. J., Keilmann, F., Wittborn, J., Aizpurua, J. and Hillenbrand, R. (2008). Terahertz Near-field nanoscopy of mobile carriers in single semiconductor nanodevices, *Nano Lett.* **8**, pp. 3766–3770.
39. Hillenbrand, R. and Keilmann, F. (2001). Optical oscillation modes of plasmon particles observed in direct space by phase-contrast near-field microscopy, *Appl. Phys. B-Las. Opt.* **73**, pp. 239–243.
40. Wiederrecht, G. P. (2004). Near-field optical imaging of noble metal nanoparticles, *The Euro. Phys. J. Appl. Phys.* **28**, pp. 3–18.
41. Kim, Z. H. and Leone, S. R. (2008). Polarization-selective mapping of near-field intensity and phase around gold nanoparticles using apertureless near-field microscopy, *Opt. Express* **16**, pp. 1733–1741.
42. Kreibig, U. and Vollmer, M. (1995). *Optical properties of Metal Clusters.* (Springer).
43. García-Etxarri, A., Romero, I., García de Abajo, F. J., Hillenbrand, R. and Aizpurua, J. (2009). Influence of the tip in near-field imaging of nanoparticle plasmonic modes: Weak and strong coupling regimes. *Phys. Rev. B* **79**.
44. Hillenbrand, R., Keilmann, F., Hanarp, P., Sutherland, D. S. and Aizpurua, J. (2003). Coherent imaging of nanoscale plasmon patterns with a carbon nanotube optical probe. *Appl. Phys. Lett.* **83**, pp. 368–370.
45. Esteban, R. *et al.* (2008). Direct near-field optical imaging of higher order plasmonic resonances, *Nano Lett.* **8**, pp. 3155–3159.
46. Olmon, R. L., Krenz, P. M., Jones, A. C., Boreman, G. D. and Raschke, M. B. (2008). Near-field imaging of optical antenna modes in the mid-infrared, *Opt. Express* **16**, pp. 20295–20305.
47. Dorfmüller, J. *et al.* (2009). Fabry-Pérot resonances in one-dimensional plasmonic nanostructures. *Nano Lett.* **9**, pp. 2372–2377.
48. Habteyes, T. G. *et al.* (2014). Near-field mapping of optical modes on All-dielectric silicon nanodisks, *ACS Photonics* **1**, pp. 794–798.
49. Rahmani, M. *et al.* (2013). Plasmonic nanoclusters with rotational symmetry: Polarization-invariant far-field response vs changing near-field distribution. *ACS Nano.* **7**, pp. 11138–11146.
50. Esslinger, M., Dorfmuller, J., Khunsin, W., Vogelgesang, R. and Kern, K. (2012). Background-free imaging of plasmonic structures with cross-polarized apertureless scanning near-field optical microscopy. *Revi. Sci. Instru.* **83**, p. 033704.
51. Olmon, R. L. *et al.* (2010). Determination of electric-field, magnetic-field, and electric-current distributions of infrared optical antennas: A near-field optical vector network analyzer, *Phys. Rev. Lett.* **105**, p.

52. Kim, D.-S. and Kim, Z. H. (2012). Role of in-plane polarizability of the tip in scattering near-field microscopy of a plasmonic nanoparticle. *Opt. Express* **20**, pp. 8689–8699.
53. Alonso-González, P. *et al.* (2012). Resolving the electromagnetic mechanism of surface-enhanced light scattering at single hot spots. *Nat. Commun.* **3**, p. 684.
54. Carney, P. S., Deutsch, B., Govyadinov, A. A. and Hillenbrand, R. (2012). Phase in nanooptics, *ACS Nano.* **6**, pp. 8–12.
55. Huth, F. *et al.* (2012). Nano-FTIR absorption spectroscopy of molecular fingerprints at 20 nm spatial resolution, *Nano Lett.* **12**, pp. 3973–3978.
56. Govyadinov, A. A., Amenabar, I., Huth, F., Carney, P. S. and Hillenbrand, R. (2013). Quantitative measurement of local infrared absorption and dielectric function with tip-enhanced near-field microscopy. *J. Phys. Chemi. Lett.* **4**, pp. 1526–1531.
57. Hobbs, P. C. D. (2009). *Building Electro-optical Systems: Making it All Work.* (Wiley-Blackwell).
58. Schnell, M. *et al.* (2009). Controlling the near-field oscillations of loaded plasmonic nanoantennas, *Nat. Photon* **3**, pp. 287–291.
59. Schnell, M. *et al.* (2010). Amplitude- and phase-resolved near-field mapping of infrared antenna modes by transmission-mode scattering-type near-field microscopy. *J. Phys. Chem. C* **114**, pp. 7341–7345.
60. Bek, A., Vogelgesang, R. and Kern, K. (2005). Optical nonlinearity versus mechanical anharmonicity contrast in dynamic mode apertureless scanning near-field optical microscopy. *Appl. Phys. Lett.* **87**, p. 3.
61. Gucciardi, P. G. and Bachelier, G. (2006). Far-field background suppression in tip-modulated apertureless near-field optical microscopy. *J. Appl. Phys.* **99**, p. 124309.
62. Taubner, T., Hillenbrand, R. and Keilmann, F. (2003). Performance of visible and mid-infrared scattering-type near-field optical microscopes. *J. Microscopy-Oxford* **210**, pp. 311–314.
63. Hillenbrand, R. and Keilmann, F. (2000). Complex optical constants on a subwavelength scale. *Phys. Rev. Lett.* **85**, pp. 3029–3032.
64. Ocelic, N., Huber, A. and Hillenbrand, R. (2006). Pseudoheterodyne detection for background-free near-field spectroscopy. *Appl. Phys. Lett.* **89**, p. 01124.
65. Deutsch, B., Hillenbrand, R. and Novotny, L. (2008). Near-field amplitude and phase recovery using phase-shifting interferometry. *Opt. Exp.* **16**, pp. 494–501.
66. Schnell, M., Carney, P. S. and Hillenbrand, R. (2014). Synthetic optical holography for rapid nanoimaging. *Nat. Communi.* **5**, p. 4499.
67. Al Mohtar, A. *et al.* (2014). Generalized lock-in detection for interferometry: application to phase sensitive spectroscopy and near-field nanoscopy. *Opt. Exp.* **22**, p. 22232.

68. Inouye, Y. and Kawata, S. (1994). Near-field scanning optical microscope with a metallic probe tip. *Opt. Lett.* **19**, p. 159.
69. Craig, I. M. *et al.* (2013). Infrared near-field spectroscopy of trace explosives using an external cavity quantum cascade laser. *Opt. Exp.* **21**, p. 30401.
70. Maier, S. A. (2007). *Plasmonics: Fundamentals and Applications.* (Springer, USA).
71. Kretschmann, E. and Raether, H. (1968). Radiative decay of non-radiative surface plasmons excited by light. *Z. Naturforschung* **23A**, pp. 2135–2136.
72. Otto, A. (1968). Excitation of nonradiative surface plasma waves in silver by the method of frustrated total reflection. *Z. Physik* **216**, pp. 398–410.
73. Park, S., Lee, G., Song, S. H., Oh, C. H. and Kim, P. S. (2003). Resonant coupling of surface plasmons to radiation modes by use of dielectric gratings, *Opt. Lett.* **28**, p. 1870.
74. Offerhaus, H. L. *et al.* (2005). Creating focused plasmons by noncollinear phasematching on functional gratings, *Nano. Lett.* **5**, pp. 2144–2148.
75. Rang, M. *et al.* (2008). Optical near-field mapping of plasmonic nanoprisms. *Nano. Lett.* **8**, pp. 3357–3363.
76. Chen, J. *et al.* (2011). Plasmonic nickel nanoantennas, *Small* **7**, pp. 2341–2347.
77. Alonso-González, P. *et al.* (2013). Visualizing the near-field coupling and interference of bonding and anti-bonding modes in infrared dimer nanoantennas. *Opt. Expr.* **21**, pp. 1270–1280.
78. Alonso-Gonzalez, P. *et al.* (2011). Real-space mapping of fano interference in plasmonic metamolecules. *Nano. Lett.* **11**, pp. 3922–3926.
79. Kim, D.-S. *et al.* (2009). Real-space mapping of the strongly coupled plasmons of nanoparticle dimers, *Nano. Lett.* **9**, pp. 3619–3625.
80. Alonso-Gonzalez, P. *et al.* (2014). Controlling graphene plasmons with resonant metal antennas and spatial conductivity patterns. *Science* **344**, pp. 1369–1373.
81. Huber, A., Ocelic, N., Kazantsev, D. and Hillenbrand, R. (2005). Near-field imaging of mid-infrared surface phonon polariton propagation. *Appl. Phys. Lett.* **87**, p. 3.
82. Huber, A. J., Ocelic, N. and Hillenbrand, R. (2008). Local excitation and interference of surface phonon polaritons studied by near-field infrared microscopy. *J. Micros.* **229**, pp. 389–395.
83. Gjonaj, B. *et al.* (2014). Sub-100 nm Focusing of Short Wavelength Plasmons in homogeneous 2D space. *Nano Lett.* **14**, pp. 5598–5602.
84. Hecht, B., Bielefeldt, H., Novotny, L., Inouye, Y. and Pohl, D. W. (1996). Local excitation, scattering, and interference of surface plasmons, *Phys. Rev. Lett.* **77**, pp. 1889–1892.
85. Chang, Y. C. *et al.* (2008). Fourier analysis of surface plasmon waves launched from single nanohole and nanohole arrays: unraveling tip-induced effects. *Opt. Expr.* **16**, pp. 740–747.

86. Chen, J. *et al.* (2012). Optical nano-imaging of gate-tunable graphene plasmons. *Nature* **487**, pp. 77–81.
87. Fei, Z. *et al.* (2012). Gate-tuning of graphene plasmons revealed by infrared nano-imaging, *Nature* **487**, pp. 82–85.
88. Woessner, A. *et al.* (2014). Highly confined low-loss plasmons in graphene–boron nitride heterostructures, *Nat. Mater.* **14**, pp. 421–425.
89. Stefanon, I. *et al.* (2005). Heterodyne detection of guided waves using a scattering-type scanning near-field optical microscope, *Opt. Exp.* **13**, pp. 5553–5564.
90. Krenz, P. M., Olmon, R. L., Lail, B. A., Raschke, M. B. and Boreman, G. D. (2010). Near-field measurement of infrared coplanar strip transmission line attenuation and propagation constants. *Opt. Expr.* **18**, p. 21678.
91. Schnell, M. *et al.* (2011). Nanofocusing of mid-infrared energy with tapered transmission lines. *Nat. Photon.* **5**, pp. 283–287.
92. Andryieuski, A. *et al.* (2014). Direct Characterization of plasmonic slot waveguides and nanocouplers, *Nano. Lett.* **14**, pp. 3925–3929.
93. Neuman, T., Alonso-González, P., Garcia-Etxarri, A., Schnell, M., Hillenbrand, R. and Aizpurua, J. (2015), Mapping the near fields of plasmonic nanoantennas by scattering-type scanning near-field optical microscopy. *Laser & Photonics Reviews* **9**, pp. 637–649.
94. Shang-Chi, J., Xiang, X., Sarriugarte, P., Sheng-Wei, Ji., Xiao-Bo, Y., Yuan, W., Ru-Wen, P., Di, Wu, Hillenbrand, R., Xiang, Z. and Mu, W. (2013). Tuning the polarization state of light via time retardation with a microstructured surface, *Phys. Rev. B*, **88**, p. 161104.

CHAPTER 12

Chemical Synthesis of Plasmonic Nanoparticles

DORLETA JIMÉNEZ DE ABERASTURI[*,†,§], **CYRILLE HAMON**[*,§] **and LUIS M. LIZ-MARZÁN**[*,‡]

[*]Bionanoplasmonics Laboratory, CIC biomaGUNE, Spain
[‡]Centro de Investigación Biomédica en Red en el área temática de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Spain
[‡]IKERBASQUE, Basque Foundation for Science, Spain

The rapidly increasing interest in plasmonic nanoparticles (NPs) has driven the scientific community to explore a variety of innovative procedures for their production, including advanced nanolithography techniques, sacrificial templates and many others. While these techniques provide high-quality nanomaterials, scale-up is hindered due to potentially complicated and expensive processes. In this respect, finding complementary and alternative approaches is required. Bench-top chemical strategies do not suffer in principle from such limitations and are thus of tremendous interest in plasmonic materials design. We address in this chapter, the most commonly used methods of preparing plasmonic NPs by wet chemistry approaches. Rather than giving a list of manufacturing routes, which would be nearly impossible taking into account the vast available literature, we aim at providing the reader with the necessary knowledge to apprehend this field.

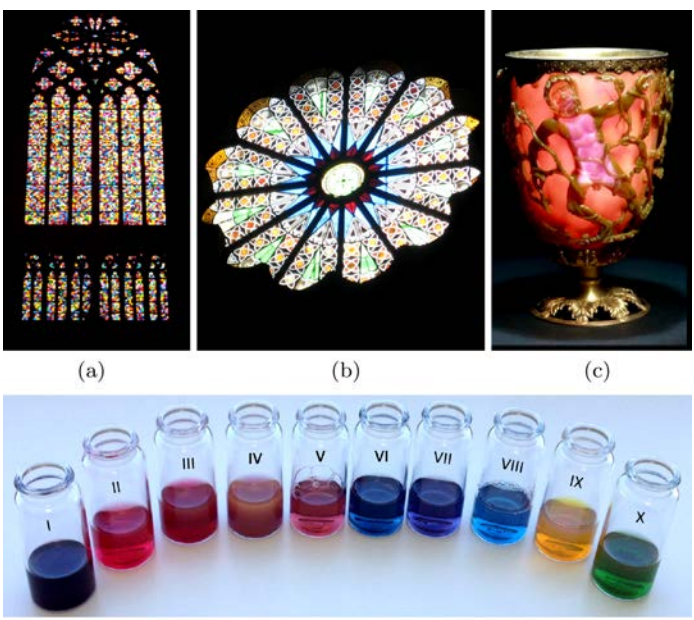
[*]Corresponding author: llizmarzan@cicbiomagune.es
[§]These authors contributed equally

12.1. Introduction

The exciting plasmonic properties of metal nanoparticles (NPs) such as gold and silver have been used since ancient times for the preparation of stained glasses. One of the most famous examples is the Lycurgus cup prepared by the Romans, which is still on display at the British Museum.[1] The utilization of metals as coloring agents was expanded in the middle ages thanks to alchemists like Cassius or Knuckel, who described the preparation of purple of Cassius and ruby red stained glass.[2] Scientific discussions on these phenomena started with Faraday in 1857 when he presented in a lecture at the Royal Institution the first systematic work on the interaction of finely divided metals with light.[3] He prepared aqueous dispersions of small gold particles by reducing a chloroaurate salt with white phosphorous and recognized that the resulting colors were due to the small size of the metal particles in the dispersion. In 1908, Gustav Mie[4] came forward to explain the red color of the gold NPs (Au NPs) in solution by applying a simplified resolution of Maxwell's equations for spheres; this model is still useful to describe the optical properties of spherical particles smaller than the wavelength of the incoming light. Since these pioneering works, many other protocols have been proposed that largely improved the level of control over the size and stability of NPs. Notably, the term "colloid science" gained great interest in the early years of the XXth century within the scientific community. In 1925, Richard Adolf Zsigmondy was awarded the Nobel Prize in chemistry for his work on the heterogeneous nature of colloidal solutions.[5,6] Later in 1950, Turkevich *et al.*[7] reduced gold salts with citrate anions forming monodisperse colloids in water, in the range of 10–50 nm. Frens[8] refined this protocol in 1973, accomplishing a better size tunability of the particles by adjusting the molar ratio of citrate to gold. This readily accessible synthesis, often denoted as citrate reduction method, is still popular nowadays, in part because the prepared colloids are stabilized by weak ligands (citrate ions) that can be readily exchanged to functionalize the particles with higher affinity ligands for different applications such as directed self-assembly,[9–11] biomedical activity[12,13] and many others.[14] Another

well-known synthesis is the Brust–Schiffrin method,[15] in which Au NPs are prepared in a two-phase system leading to nanocrystals as small as 1 nm, stabilized by alkanethiols that can be dispersed in organic solvents. Nowadays, many routes for the synthesis of a variety of nanocrystals made of different materials are well documented, even commercially available, but improvement of existing particles or preparation of new particles with different properties is still under the focus of many research groups. Additionally, the continuous development of novel microscopy techniques allows scientists to describe in detail the crystallinity and the precise morphology of the NPs, thereby helping to improve existing protocols. Although we focus the discussion on the preparation of plasmonic NPs made of noble metals such as gold and silver, some other examples of non-metallic NPs displaying plasmon resonances will also be provided.

The plasmonic properties of colloidal spherical Au NPs, typically characterized by a deep red color, can be modified by tuning particle



(a) (b) (c)

(d)

←

Fig. 12.1. Photographs of stained glasses from (a) Cologne Cathedral; (b) Santa Maria Sopra Minerva church in Rome; (c) Lycurgus cup; ©The trustees of the British museum, and (d) various NP solutions: (I) 5 nm Au NPs in organic solvent at high concentration[15]; (II) 18 nm Au NPs; (III) 55 nm Au NPs; (IV) 100 nm citrate Au NPs stabilized with citrate[16]; (V) gold nanorods[17]; (VI) nanostars[18]; (VII) branched gold AUNPs; (VIII) gold nanotriangles[19]; (IX) 40 nm AgNPs[20]; (X) silver-coated gold nanorods covered with silica.

size, between ∼3 and 200 nm, leading to plasmon shifts from ∼519 to 730 nm.[15,16] In the case of silver nanoparticles (Ag NPs), localized surface plasmon resonances (LSPRs) occur at higher energies around 400 nm but can also be tuned by increasing particle size.[20] It is also possible to synthesize nanoalloys made of gold and silver or other metals by co-reduction of the corresponding metallic salts in solution, thereby shifting the plasmon resonances within the range between 380 and 520 nm.[21] However, for applications that require a wider spectral range, alternative strategies have been developed to control not only the size but also the shape of the NPs. In contrast to the early development of spherical NPs, the synthesis of colloidal anisotropic NPs gained tremendous interest around the beginning of this century thanks to protocols that allowed high-yield preparation.[22] Research in this direction is thus motivated by the exciting optical properties of anisotropic NPs, which can accommodate additional plasmon modes related to various particle morphologies. A classification of the particles can, therefore, be made according to their dimensionality,[23] starting from 0D quasi-spherical particles, through 1D rod-like particles, bars and wires, 2D platelets and prisms to 3D cages and branched NPs, frequently characterized by concave surfaces displaying facets with high Miller indexes.[24]

Among the different available approaches for the preparation of anisotropic NPs, seed-mediated growth methods are most commonly used. The seed-mediated process comprises two steps: the preparation of a seed solution containing small crystallites (e.g. 1–5 nm), which is subsequently injected into a growth solution containing

metal salt, as well as reducing, stabilizing and shape directing agents. Other "one-step" synthetic methods also exist, albeit generally leading to smaller but more polydisperse NPs. Aiming to provide a clear overview to the readers, we describe here the synthesis of NPs according to their dimensions. Note that the field is so broad, as indicated by the huge amount of articles published in recent years, that it is difficult to cover all possible aspects. Although a universal view about the growth mechanism is still lacking in the literature, some of the reported protocols allow for a tight control over particle shape. Due to the need for increasingly sophisticated structures, Wang *et al.*[25] classified NP growth according to thermodynamically or kinetically controlled scenarios. We summarize here such general mechanisms, while more details on the preparation of specific anisotropic NPs can be found in the corresponding subsections.

In general terms, the growth of NPs is produced by a chemical reaction that provides free atoms or molecules that get involved in the growth. The interaction of these atoms with each other has been described in the classical theory of nucleation and growth. The different mechanisms that explain the formation of uniform particles are presented in Fig. 12.2.

In homogeneous precipitation, a short single nucleation burst occurs when the concentration of constituent species reaches critical supersaturation. Then, the so-obtained nuclei are allowed to grow uniformly by diffusion of solutes from the solution to their surface until the final size is attained. To obtain monodisperse particles, nucleation and growth stages must be separated. Kinetically speaking, this means that in order to achieve uniform NPs the rate at which the growth materials are produced should be slower than that of their consumption (heterogeneous deposition on existing nuclei), allowing all the nuclei to emerge at roughly the same time and to grow at the same rate. This corresponds to the model proposed by LaMer and Dinegar[26] (Curve I, Fig. 12.2). A self-sharpening growth process following multiple nucleation events is the so-called Ostwald ripening

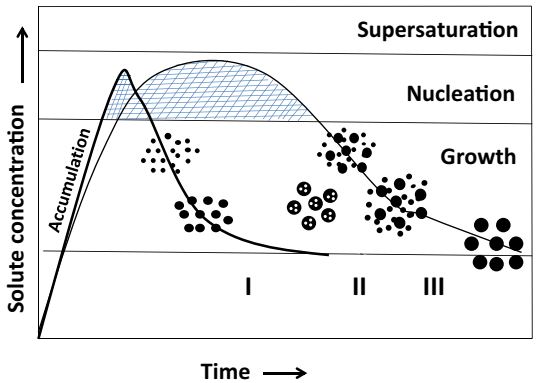


Fig. 12.2. Growth mechanisms explained by different models. Plot of solute concentration versus time, showing the initial accumulation of solute, the nucleation stage and the subsequent growth; curve I represents a single nucleation event and uniform growth by diffusion (classical model of LaMer and Dinegar)[26]; curve II shows nucleation, growth and aggregation of smaller subunits[27,28] and curve III represents multiple nucleation events and growth via Ostwald ripening.[29]

process, through which uniform particles can also be obtained (Curve III, Fig. 12.2). In some cases, however, it has been shown that uniform particles can also be obtained as a result of aggregation of much smaller subunits rather than continuous growth by diffusion,[30,31] which has been observed for example in the formation of magnetic NPs[27,28] (Curve II, Fig. 12.2).

Experimental studies of plasmonic nanocrystal growth in general invoke the LaMer mechanism, Ostwald ripening or both. In the seed-mediated growth method, nucleation and growth are temporally separated, thereby avoiding the overlap of nucleation and growth and the resulting polydispersity, as often occurs in "one pot" synthesis methods. Nucleation (i.e. formation of seeds) is generally induced by the reduction of metal salt by a strong reducing agent such as sodium borohydride ($NaBH_4$). In the second step, the injection of seeds into the growth solution initiates the slow reduction of the oxidized metal salt in solution at the surface of the seeds, which therefore also act as a catalyst. The metal salt would thus be the solute or monomer in the LaMer mechanism, as shown in Fig. 12.2. On the other hand,

Ostwald ripening corresponds to the growth of larger nanocrystals at the expense of smaller nanocrystals, the latter being progressively dissolved. It should, however, be noted that growth of larger nanocrystals may occur as well by aggregative growth, converting a suspension of non-stable small colloids into often more homogeneously distributed and stable aggregates, which can be either poly- or monocrystalline. These Ostwald ripening or aggregative growth processes may be avoided via a fine control of the reaction conditions involving strong stabilizing ligands that passivate the surface of the nanocrystals.

12.2. Controlling the Shape of Plasmonic Nanoparticles

12.2.1. *General considerations*

As introduced above, the initial work on colloidal plasmonic NPs dealt with spherical or quasi-spherical NPs, mainly due to the simpler synthetic procedures required for the reproducible formation of monodisperse colloids. Indeed, NPs tend to afford spherical geometry thereby reducing their surface free energy; in other words, it is the most thermodynamically favorable morphology. In order to access anisotropic nanostructures, it is necessary to find reaction conditions leading to the required symmetry break from the spherical seeds, as well as control the reaction kinetics. In the following, we guide the reader through the milestones in NP synthesis, including the key factors toward directing NP growth. Seed-mediated growth can be used to prepare a large library of particles having different shapes and crystallinity. Interestingly, slight changes in the reaction conditions may lead to different collective behavior of the reactants and produce completely different NPs. We focus here on the preparation of Au NPs, but the concepts can be translated to the synthesis of other metal particles. In this case, the growth solution is composed of a trivalent gold salt (tetrachloroauric acid, $HAuCl_4$) dissolved in an aqueous solution containing surfactants, which are typically hexacethyltrimethylammonium

bromide or chloride, denoted as CTA-Br or CTA-Cl, respectively. These surfactants contain quaternary ammonium headgroups, which can readily complex the metal ions in solution and ultimately the gold atoms on the particle surfaces, forming interdigitated bilayers in which the CTA^+ headgroups of the external layer face the solvent (water).[32,33] Homogenization of the solution at this point is important when shape-directing agents such as silver nitrate are added, as they are believed to interact synergistically with the other reactants. A weak reducing agent is also introduced, which can reduce Au(III) into Au(I), as reflected in a color change from yellowish to colorless. Ascorbic acid is the most popular reducing agent because its reduction strength can be easily tuned through adjustment of the solution pH, its electrochemical potential increasing at lower pH values. In a final step, the growth of the NPs is triggered by the fast injection of the seeds, and the typical duration ranges from 2 to 6 h.

Despite their apparent spherical shape when viewed in an electron microscope, gold seeds display crystalline facets. Seeds can be either mono- or poly-crystalline, the latter being characterized by the presence of stacking faults and/or twin boundaries in the lattice structure (Fig. 12.3). Importantly, as the seeds template the growth of larger particles, their crystallinity is a cornerstone to direct NPs morphology and size. Indeed, the addition of a larger amount of seeds to a given concentration of Au ions results in a larger number of smaller particles, whereas larger particles are obtained when less seeds are used. On the other hand, the crystalline habit is an important attribute of NPs because it will ultimately influence their catalytic, optical and electrical properties. It is thus significant to achieve control over the crystallinity of the seeds, not only to direct the NP shape but also to envision applications. Noble metals, gold and silver in particular, are characterized by a face-centered cubic (fcc) lattice. The addition of chemical species that can stabilize specific crystal facets has been proposed as a tool to fine-tune the growth rates in specific crystallographic directions, resulting in NPs with morphologies that may not always be those expected from thermodynamic considerations (i.e. octahedron and tetrahedron).[23,34,35]

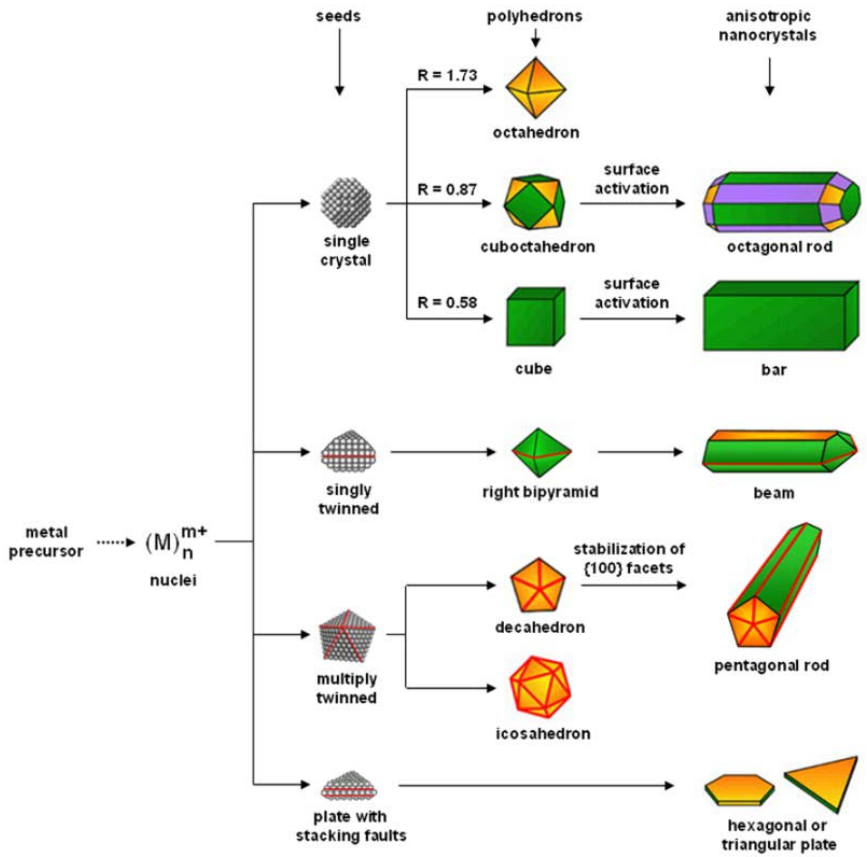


Fig. 12.3. Scheme illustrating the importance of the seed crystallinity to define particle shape. Seeds grow from the initial nuclei into either monocrystalline or polycrystalline structures, which in turn can be singly twinned or multiply twinned and/or contain stacking faults. The green, orange, and purple colors represent {100}, {111} and {110} facets, respectively. Twin planes are delineated in the drawing with red lines. The parameter R is defined as the ratio between the growth rates along ⟨100⟩ and ⟨111⟩ directions.[34] Reprinted with permission from Ref. 34. Copyright 2009 Wiley-VCH.

Capping agents preferentially adsorb onto certain types of crystal facets and can passivate them, so that the resulting particle overexpresses such facets, thereby directing particle shape. Recent optimization of some protocols has allowed drawing mechanistic

aspects that can be applied to most synthetic routes. It appears from experimental evidences that halide ions (Cl^-, Br^- and I^-) as well as silver ions are important to direct the particle morphology as they can preferentially adsorb on certain crystal facets. One interpretation of the mechanism underlying the action of silver ions on shape-directed growth is the so-called silver underpotential deposition, in which an Ag^0 adlayer would be deposited preferentially on high-index facets, which can thus be stabilized as the most abundant ones.[36] Iodide is known to stabilize {111} facets preferentially, promoting the growth of two-dimensional particles while preventing the growth of oblong particles.[19,37,38] On the contrary, an emerging method for the generation of anisotropic particles is the surface-selective etching of pre-formed nanocrystals.[39] In the following, the key ingredients for making and characterizing particles with different morphologies, including core@shell nanostructures, are described in the following subsections.

12.2.2. *The case of isotropic nanoparticles (0D)*

Crystallinity is a key parameter that determines the properties of NPs as it impacts their optical or electrical properties, thereby inducing differences between solids made of same material.

In general, in thermodynamic terms, the most favorable shapes are polyhedrons with convex surfaces, such as cubes, octahedrons or icosahedrons, depending on the number of twin defects in the seed and the ratio of growth rates along {111} and {100} directions. A nanocrystal will assume its equilibrium shape when its total Gibbs free energy reaches the global minimum.[40] Thus, it is known that for a metal nanocrystal, its equilibrium shape cannot be spherical because many high-index facets with high specific surface free energies would be required to obtain a perfect sphere.[23] However, this is more complex to explain as the types of facets of the NPs must be taken into account. Different facets have different densities of surface atoms, charges and often adsorbed ligands, resulting in different stabilities and, thus, different particle shapes. This is schematically shown in Fig. 12.4. However, most nanocrystals lie in between the

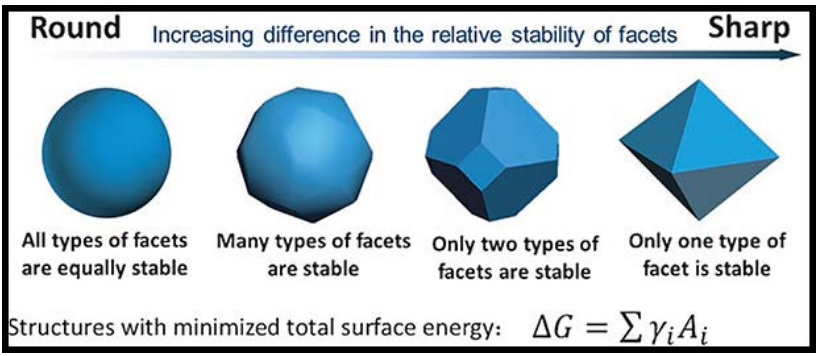
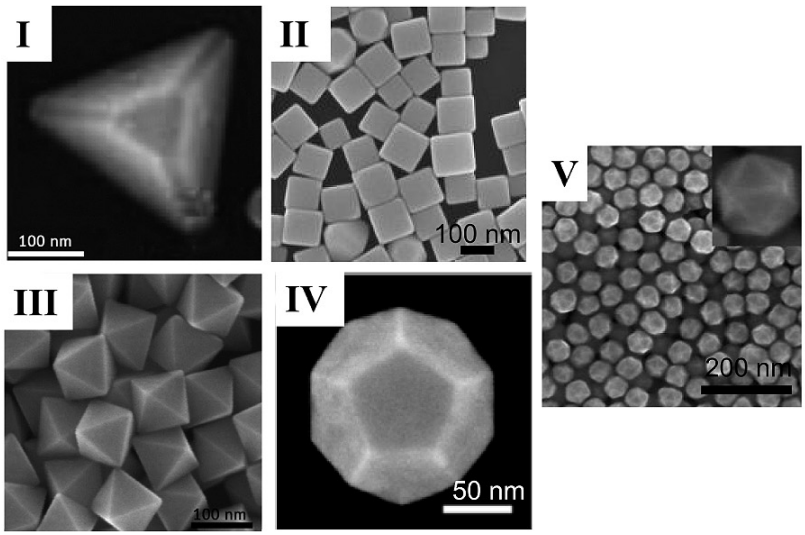


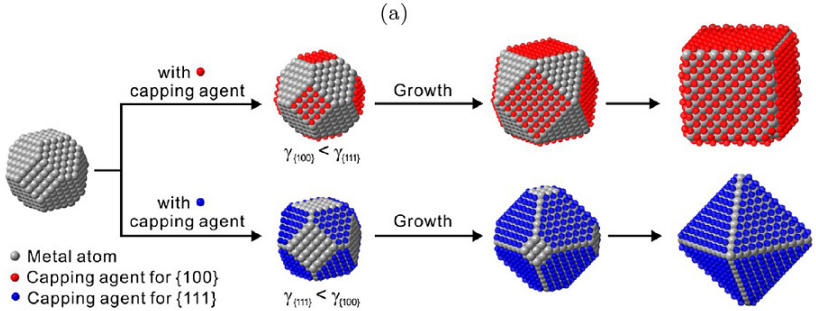
Fig. 12.4. NPs shape depending on the stability of its facets. Reprinted with permission from Ref. 25. Copyright 2014 Wiley-VCH.

two extremes; therefore, it can be said that polyhedral shapes are the most favorable ones.

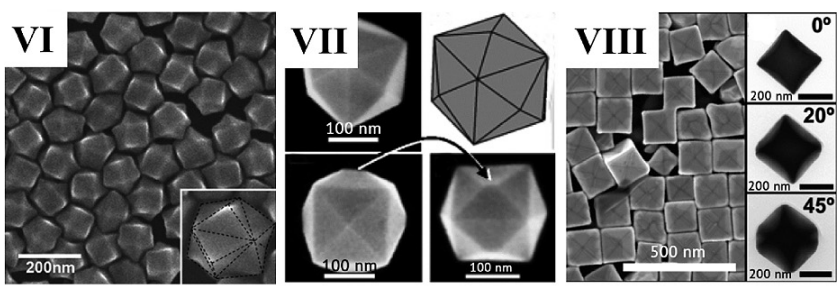
So-called Platonic NPs have aspect ratios near one and mainly differ in the number of crystal facets. In other words, they are polyhedrons containing regular polygonal facets with the same number of polygons meeting at each corner. The five platonic shapes are tetrahedron (4 triangular facets), hexahedron (cube; 6 square facets), octahedron (8 triangular facets), dodecahedron (12 pentagonal facets) and icosahedron (20 triangular facets) (Fig. 12.5(a)). In general, tetrahedrons, octahedrons and icosahedrons are dominated by {111} facets, whereas hexahedrons are bound by {100} and dodecahedrons by {110} facets. It should, however, be noted that often NPs are not perfect and some or all corners can be truncated, resulting in different structures (Fig. 12.5(c)). Thus, these platonic particles can also be used as seeds to form other different structures *via* seeded growth (Fig. 12.3). Note that capping ligands that bind selectively onto specific facets can direct the overgrowth, ultimately resulting in different particle shapes (Fig. 12.5(b)).[23] In addition, polyhedral NPs can also be used to obtain specific shapes via etching processes, which is commonly used for example to obtain concave particles as explained below.[24] Spherical monocrystalline gold particles have been obtained by oxidation of gold octahedrons with chloroauric acid.[41]


I
100 nm
II
100 nm
V
200 nm
III
100 nm
IV
50 nm

(a)


with
capping agent
Growth
$\gamma_{\{100\}} < \gamma_{\{111\}}$
with
capping agent
Growth
$\gamma_{\{111\}} < \gamma_{\{100\}}$
Metal atom
Capping agent for {100}
Capping agent for {111}

(b)


VI
200nm
VII
100 nm
100 nm
100 nm
VIII
0°
200 nm
20°
200 nm
45°
200 nm
500 nm

(c)

←

Fig. 12.5. (*figure on facing page*) (a) The five platonic shapes: (I) Au tetrahedron[42]; (II) Ag hexahedron (i.e. cube)[43]; (III) Au octahedron[44]; (IV) Au dodecahedron[45]; (V) Au icosahedron[42]: (b) Schematic illustrations showing the role of capping agents in directing the growth of a single-crystal seed made of an fcc metal. The shape of the resulting nanocrystals can be manipulated through the introduction of a capping agent (represented by red or blue dots) that selectively bind to a specific type of facet, altering the order of surface-free energies and eventually leading to the formation of a nanocube enclosed by {100} facets and an octahedron enclosed by {111} facets, respectively.[23] (c) Formed shapes from platonic NPs (VI) Au trisoctahedron NPs[46]; (VII) Pt tetrahexahedron NPs[47]; (VIII) Au concave hexahedron NPs.[48] (I) and (V) were adapted with permission from Ref. 42. Copyright 2004 Wiley-VCH. (II), (III), (IV), (VI), (VII) and (VIII) and (B) Reprinted with permission from Refs. [44–48], respectively. Copyright American Chemical Society.

12.2.3. *1D Elongated shapes: Rods and wires*

Although the first examples of efficient synthesis of rod-shaped Au NPs are probably electrochemical methods, either using alumina membranes as templates[49] or in solution,[50] the colloidal synthesis of gold nanorods with high yield was widely used after the modifications made by Murphy and El-Sayed groups, starting in 2001[51–54] Of all the possible NP shapes, nanorods are especially interesting as they offer strong plasmonic fields while exhibiting excellent tunability through adjustment of their aspect ratio.[55] Rather small variations in the synthesis of Au nanorods can, however, lead to either single crystals with octagonal cross-section or fivefold twinned crystals, with pentagonal cross-section. Other modifications have been reported to lead to e.g. rectangular or singly twinned crystals, but these will not be described here.[34,56] For the sake of clarity, the term nanowire will be used to refer to nanorods with high aspect ratio >10. Single crystal gold nanorods are elongated along the ⟨001⟩ direction and possess eight lateral {520} facets, whereas penta-twinned nanorods are elongated along the ⟨110⟩ direction and possess five {100} lateral facets.[36,57,58] Single crystal nanorods possess smooth ends as observed in electron microscopy due to a combination of {100}, {110} and {111} facets whereas penta-twinned rods possess sharp tips made of {111} facets.[36,57] The formation of single crystalline gold nanorods

is almost invariably mediated by silver ions, but penta-twinned rods do not contain silver. On the other hand, the synthesis of single crystalline silver rods has not been reported, but several examples exist for penta-twinned silver nanorods.[56,59,60]

Single crystalline gold rods are usually prepared in the presence of CTA-Br, ascorbic acid, $HAuCl_4$, $AgNO_3$ and single crystalline seeds (1.5 nm).[57,61] Therefore, in addition to the growth solution described above, silver ions are essential in this particular case to direct the rod shape, which has been ascribed to either underpotential deposition or to the formation of Ag-CTAB complexes in solution.[33,62–64] Interestingly, the aspect ratio of Au nanorods can be readily tailored between 1 and 5 by simply tuning the Au/Ag concentration ratio or the seed concentration. Additionally, recent improvements in gold nanorod synthesis have been reported which are based on the addition of a co-surfactant or other metal ions.[17,65–67] For instance, the addition of aromatic molecules provides further control on the reaction kinetics and the stabilization of crystal facets during the growth process (Fig. 12.6(a)).[65] As discussed above, pH can also be used to tune the rate of the reaction, by affecting the reduction power of ascorbic acid.

Penta-twinned gold nanorods are prepared in the presence of CTA-Br, ascorbic acid, $HAuCl_4$ and ca. 3.0 nm penta-twinned seeds.[57] In this case, the seeds are prepared in the presence of citrate anions instead of CTA-Br, which leads to a significant increase in the degree of twinning. The polycrystalline nature of the seed provides enough anisotropy to break the symmetry of the particles during growth. Indeed, the packing of the five tetrahedral subunits results in an incomplete filling of the space, inducing strain in the lattice structure.[68] Although this technique can be adapted to obtain rods with aspect ratios up to 10, post-synthesis overgrowth can be performed to increase the length and obtain wires with aspect ratios even greater than 20.[69] In a further development based on lowering both the pH and the amount of seeds, the growth of gold nanowires with aspect ratio over 200 has been reported.[70] On the other hand, flexible ultrathin gold nanowires can be synthesized in organic solvents,

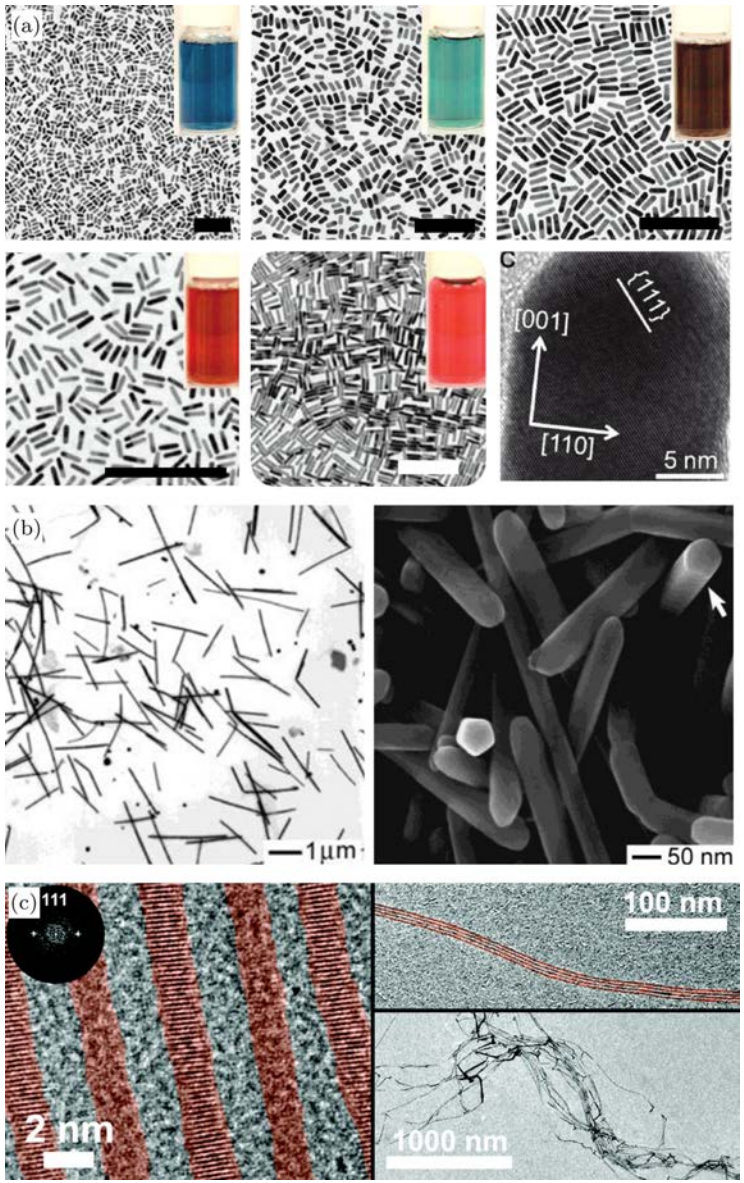


Fig. 12.6. Synthesis of different types of 1D nanocrystals. (a) TEM images of single crystalline gold nanorods of different aspect ratios displaying different optical properties as illustrated by the corresponding photographs of the colloidal suspensions. The scale bars on all images are 200 nm. (right) High-resolution TEM

Fig. 12.6. (*Caption Continued*) of a single crystalline nanorod.[67] (b) Silver nanowires synthesized by the polyol synthesis.[59] The pentagonal cross-section of the wires is visible on the right image.[72] (c) Ultra-thin gold nanowires prepared in organic solvent.[71] (a), (b) and (c) were adapted with permission from Refs. 65, 59, 71, 73, respectively. Copyright American Chemical Society.

in the presence of oleylamine, featuring a diameter of 1.6 nm but lengths up to 10 μm, thus expanding the aspect ratio control up to 2500 (Fig. 12.6(c)).[71] In this particular case, even if the solvent and ligands are different, the underlying growth mechanism may also be related to preferential adsorption of oleylamine on certain surface facets, thereby triggering anisotropic growth.

Regarding silver, the synthesis of penta-twinned nanowires has been reported by Sun *et al.*, based on the so-called polyol process,[59,72] which involves the reduction of silver nitrate by ethylene glycol in the presence of poly(vinylpyrrolidone) (PVP) and the introduction of seeds that can be either penta-twinned silver or Pt (Fig. 12.6(b)). Again, in this case, PVP acts as a molecular template by passivating {100} and {110} facets so that the addition of monomers occurs preferentially on {111} facets. A coalescence mechanism has been proposed for the formation of the nanowire in which small nanocrystals having high surface energy aggregate into a 1D structure.[60,73]

12.2.4. *2D platelets, triangles and disks*

During the last decade, several procedures have been reported to prepare plasmonic nanoplates, such as nanoprisms NPs used or nanotriangles (NTs), hexagons, pentagons, star shaped or disks, all of them based on the polyol process using a variety of polymers. In this way, the polymer can act as stabilizer, template and even reductant. Polymers such as poly(ethylene oxide), poly(propylene oxide) and poly(vinyl alcohol) are involved in the synthesis of gold nanoplates, often with no need of an additional reducing agent.[74] The polymer may, however, act only as stabilizer in which case an additional reducing agent such as L-ascorbic acid is required to reduce $HAuCl_4$, as reported by Yamamoto *et al.* for the formation of star-shaped

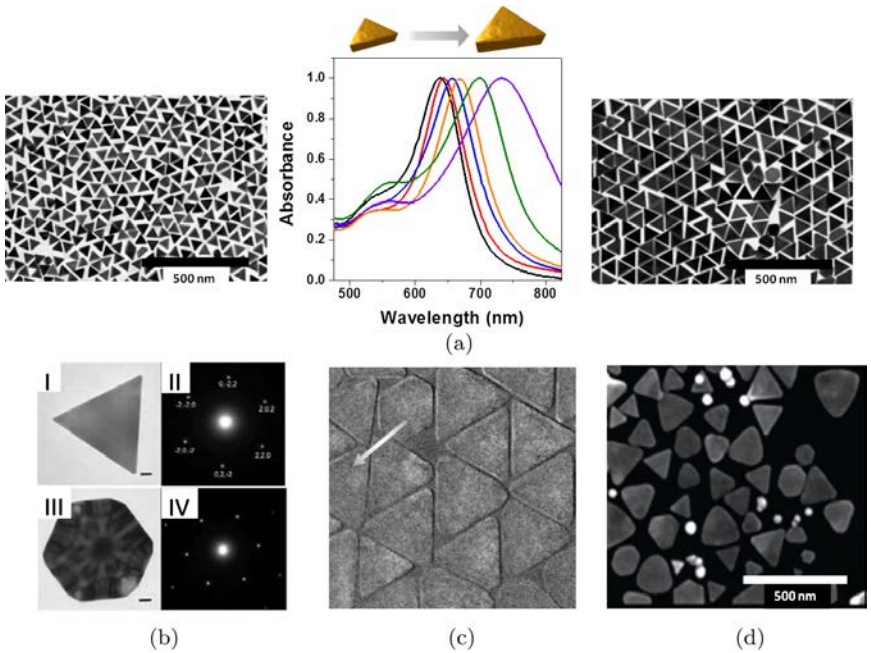


Fig. 12.7. Different examples of 2D nanocrystals. (a) TEM images of Au NTs with increasing edge lengths and the corresponding UV–vis spectra.[19] (b) (I) TEM image of a single triangular gold nanoplate and (II) the corresponding electron diffraction pattern. (III) TEM image of a truncated hexagonal gold nanoplate (with 6 long edges and 6 short edges) and (IV) the corresponding electron diffraction pattern. Scale bars: 50 nm.[77] (c) High-resolution SEM image of nanoprisms. The arrow highlights flat particles where contrast uniformity indicates thickness uniformity.[80] (d) SEM image of gold nanoprisms.[82] (a), (b) and (d) were adapted with permission from Refs. [19], [77], [82], respectively. Copyright American Chemical Society. (c) was adapted with permission from Ref. [80]. Copyright 2006 Wiley-VCH.

nanoplates coated with PVP.[75] In addition, this wet-chemical route has been extensively applied for other kinds of particles in which the morphology can be defined though the relative concentrations of the various reactants and the synthesis conditions. In 2004, a mild, large-scale synthesis of Au nanoplates was reported by Sun *et al.*[76] whereas Chu *et al.*[77] carried out the synthesis of triangular and hexagonal gold nanoplates in aqueous solution by thermal reduction of $HAuCl_4$ with trisodium citrate in the presence of CTAB.

Triangular nanoprisms of gold or silver are interesting due to their plasmonic features in the visible and IR regions and can be used for several applications.[78,79] The size of Au NTs can be tuned within a wide range by simply changing the experimental parameters. These methods usually yield particles with lateral dimensions above 100 nm.[80] Recently, Scarabelli *et al.*[19] Developed a method to synthesize Au NTs with edge lengths between 50 and 150 nm, which translated into LSPRs between 630 and 740 nm. The obtained crystals can be purified from by-products using a simple sedimentation method[81] comprising the addition of a high surfactant concentration, which induces morphology- and size-dependent depletion interactions, and leading to a shape purity as high as 95%. It has been confirmed that the amount of seeds and the addition of iodide ions to the growth solution are crucial to control the nanotriangle size. In addition, these particles are excellent candidates for plasmonic biological applications, so efforts have been made to develop particles that avoid the use of toxic capping agents such as CTAB or CTAC. Thus, CTAB-free, biocompatible single-crystalline nanoprisms have also been reported.[82]

12.2.5. *3D: Branched and concave morphologies: nanostars*

Metal nanocrystals with highly branched morphologies or with concave surfaces possessing high Miller indexes (>1) have become increasingly common during recent years. The sharp corners or edges and tips serve as efficient nanoantennas where electromagnetic fields are expected to be very high.[83] Dissolution and overgrowth are the two main synthetic approaches to prepare crystals with concave surfaces. The dissolution or corrosion of a metal has been shown to lead to surface degradation and structure failure, which can take place in different ways including etching, pitting, galvanic replacement and dealloying. Some of these mechanisms are explained in detail in the following section, and thus we restrict our discussion here to 3D structures prepared by seeded growth mechanisms. As mentioned in the section on 0D nanostructures, the seeded growth

mechanism generally favors polyhedral shapes with convex surfaces, such as cubes, octahedrons or icosahedrons, depending on the number of twin defects in the seed and the ratio between the growth rates along {111} and {100} directions. The formation of concave structures is however not thermodynamically favored. This limitation can be overcome by conducting the synthesis with an appropriate capping agent or under kinetically controlled conditions. For example, PVP passivates {100} rather than {111} facets in Ag, resulting in nanocubes or penta-twinned nanorods.[84] Additionally, under certain conditions, halide ions such as Br^- and I^-, preferentially adsorb onto {100} facets, thus promoting the formation of cubes, bars or rods enclosed by these facets.[85] Langille *et al.* recently reported a detailed study of NP growth on cetyltrimethylammonium chloride (CTAC)-capped seeds where the use of different halogen counter-anions was crucial in the synthesis of NPs of different shapes.[87] Although a number of examples have been published in which this approach allowed control over the growth of different metals, most likely it cannot be considered as a universal method, as each ligand interacts differently with the metal used. The kinetically controlled method is based on the manipulation of the growth rate at which atoms generated from the precursors can be added to the surface of a seed that may carry a capping agent or not, leading to different morphologies.[24] The formed atoms tend to be added to the edges, but the reaction kinetics can be manipulated by adjusting different parameters such as precursor and reductant concentrations, temperature, etc.[23] The presence or absence of silver ions is also crucial toward the formation of certain types of branched NPs such as gold nanostars (Au NSs),[86] as well as for others such as nanorods or high-index concave cubic Au nanocrystals with {720} surfaces.[87]

There are many examples of the synthesis of 3D structures obtained by optimization of the experimental conditions. Branched NPs including multipods.[88] and/or nanostars[18] (also known as nanoflowers and nanourchins[89]) have been extensively studied as they are excellent candidates for nanodevices and for *in vivo* applications due to characteristic LSPR peaks in the near IR

region, for which penetration of light in tissues is favored.[90] Chen and carroll reported in 2003 the synthesis of Au monopods, bipods, tripods and tetrapods through the reduction of Au(III) with ascorbic acid in the presence of CTAB and Ag platelets.[91] Sau and Murphy used the seed-mediated synthesis to modify the dimensions and the number of branches.[92] Hao *et al.* reported the synthesis of "branched" gold nanocrystals in high yield by a wet-chemical route where $HAuCl_4$ was reduced by sodium citrate in a solution of bis-(*p*-sulfonatophenyl)phenylphosphinedipotassium dehydrate (BSPP) and H_2O_2, at room temperature.[93] Although various other routes were subsequently used,[94] high yield production could not be achieved until 2008, when Kumar *et al.* proposed a seed-mediated reduction of $HAuCl_4$ in a concentrated solution of poly(vinylpyrrolidone) (PVP) in *N*,*N*-dimethylformamide (DMF) (Fig. 12.8(b)).[18] Another efficient, surfactant-free method was later developed by Yuan *et al.* in which naked NPs can be obtained, which is important for several applications such as SERS. By varying the amount of Ag^+ added to citrate-capped Au seeds in the presence of chloroauric acid and HCl, the shape and optical response can be tuned (Fig. 12.8(a)).[86] Similar procedures were subsequently employed to grow Au spikes onto different structures as for example Janus NPs[95] (Fig 12.8(c)) or nanodumbbells.[96]

Smooth nanodumbbells can also be prepared in high yield[97] by seeded growth of nanorods in the presence of iodide ions (Fig. 12.8(d)). Other morphologies such as gold nanotadpoles were synthesized by citrate reduction of chloroauric acid in the presence of sodium dodecyl sulfate as capping agent.[98] Recently, these structures were also synthesized through the standard seed-mediated process in the presence of Ag^+ ions and CTAB at room temperature.[99] Au dendrites are also interesting structures that possess a hierarchical tree-type architecture with trunks, branches, and leaf components (Fig. 12.8(e)). Huang *et al.* reported the first shape-controlled synthesis in the presence of supramolecular complexes formed from dodecyltrimethylammonium bromide (DTAB) and β-cyclodextrin (β-CD).[100]

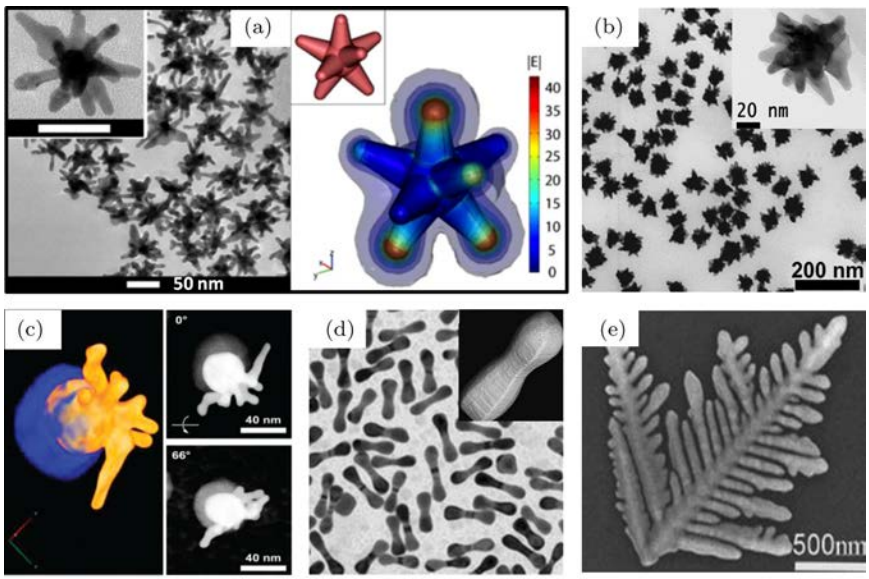


Fig. 12.8. Branched NPs. (a) TEM images of surfactant-free Au nanostars and simulation of electric field enhancement in their when excited by a z-polarized plane wave, propagating in the y-direction, with a wavelength of 800 nm.[86] (b) TEM images of PVP-coated Au nanostars.[18] (c) HAADF-STEM images and tomographic reconstruction of Au-silica Janus nanostars.[95] (d) TEM images of Au dumbbells.[101] (e) SEM image of an Au nanodendrite grown in mixed DTAB/β-CD solution.[100] (a) and (b) were adapted with permission from Refs. 86 and 18, respectively. Copyright IOP Science. (c) was adapted with permission from Ref. 95. Copyright Royal Society of Chemistry. (d) was adapted with permission from Ref. 97. Copyright 2008 Wiley-VCH. (e) was adapted with permission from Ref. 100 Copyright 2010 American Chemical Society.

12.2.6. *Core/shell nanoparticles (0D/1D/2D/3D)*

The discussion has been focused so far on plasmonic NPs made of a single metal. It is important, however, to remember that many other wet-chemistry synthesis methods exist for synthesizing particles of other compositions, e.g. with semiconductor or magnetic properties. In the scope of making heterostructures, synthetic approaches can be combined to fabricate particles with new functionalities. Indeed, growing one type of material on top of another by heteroepitaxy may affect charge carrier dynamics and lead to the emergence of unusual

properties, unattainable by any of the single components alone. For instance, heterostructured NPs based on semiconductors and noble metals show altered degree of quantum confinement and modified charge-carrier recombination or separation dynamics. Additionally, many examples exist on plasmonic NPs coated with inert materials such as silica, toward devising biocompatible NPs as the shell mediates the interaction with the surrounding media. Generally speaking, coating NPs with various organic or inorganic spacers can lead to a precise control over the interaction of plasmonic particles with their environment. The growth of platinum or palladium on gold or silver cores has also been widely used to tune or enhance their optical and catalytic properties, respectively. In the following, we focus on heterostructures made of gold and silver displaying novel morphology and optical properties, which allow us to complete the plasmonic toolbox. Gold and silver possess very similar lattice parameters (lattice mismatch below 0.3%), which allows us to envision growth by heteroepitaxy of one metal on the surface of the other with minimum mechanical stress at the interface between the two materials. Silver coated Au NPs (Au@Ag) can be easily prepared by seed-mediated growth in which gold seeds are injected in a growth solution containing dissolved silver salt and a mild reducing agent. On the other hand, gold-coated silver nanoparticles (Ag@Au) are more difficult to fabricate because of the oxidation of the silver core by the introduced gold ions, an effect that is known as galvanic replacement reaction and occurs because the silver core acts as a sacrificial anode for the cathodic reduction of gold ions. In other terms, galvanic replacement is a redox process occurring because gold has a higher reduction potential than silver. Even if this effect may be seen as a limitation, it actually enriches our plasmonic toolbox because hollow nanocrystals can be devised. In this section, we first introduce silver coating of Au NPs and subsequently describe the preparation of the hollow plasmonic nanocrystals.

In principle, no restrictions apply regarding the morphology of the gold particles used as seeds during the synthesis of Au@Ag NPs, which can be exploited to tune the morphology of the resulting

core-shell particles. Again, the growth solution comprises the dissolved salt, a weak reducing agent and shape directing and stabilizing agents. The manner in which the silver shell grows will depend on the surface energy of the exposed seed facets, which can be further tuned trough the addition of halides. Halides stabilize {100} silver facets, leading to preferential growth of other facets in the gold core. Consequently, a silver overgrowth reaction on either single crystalline or penta-twinned nanorods as seeds leads to preferential silver growth at the sides or the ends, respectively (Fig. 12.9).[57,101−103] Likewise, gold octahedral seeds can be converted into Au@Ag nanocubes enclosed by six {100} facets. It is interesting to note that in Fig. 12.9 the silver shell is readily discerned from the gold core by high-angle annular dark field scanning TEM (HAADF-STEM), in which the contrast is due to variation of the atomic number between the two components. Ag shells have thus been tuned into different morphologies such as cubic or truncated cubic, cuboctahedral, truncated octahedral and octahedral structures from rhombic dodecahedral seeds by adjusting the concentration of reactants.[104] Apart from the increase in size, Au@Ag NPs exhibit a higher extinction cross-section compared to the initial Au NPs, thereby resulting in enhanced optical properties. Au@Ag NPs show a more complex set of plasmon resonances than those in single metal particles. Notably, the plasmon resonance peaks of the gold cores are shifted toward higher energy as the silver shell thickness increases and additional higher energy peaks develop thanks to the better plasmonic performance of silver.[105]

The preparation of hollow nanocrystals via galvanic replacement has been widely used to engineer particles with novel structure and composition, which may even comprise more than two metals. As discussed above, the reaction occurs if the introduced metal ions possess a higher reduction potential than that of the metal composing the seeds, which is often referred as a sacrificial template. Typically, the template gets oxidized while the second metal is deposited on the outer surface by heteroepitaxy, forming complex structures. Various metals have been reported to promote galvanic replacement reactions such as Au, Pt and Pd.[106] reduced on Ag templates, and many

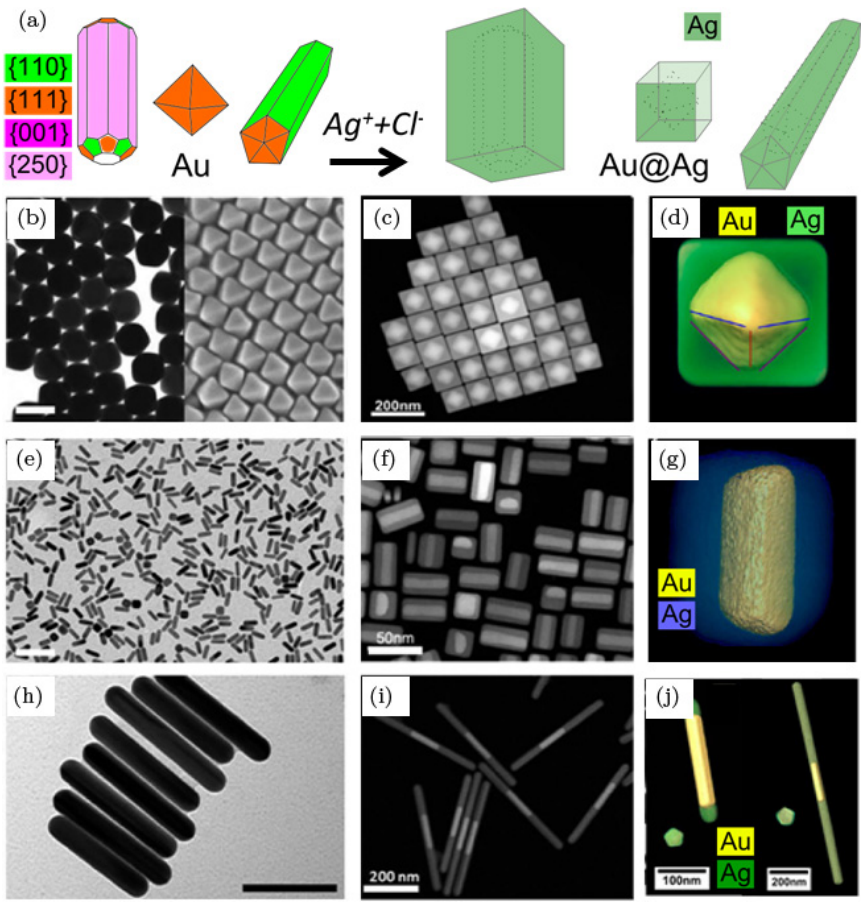


Fig. 12.9. Synthesis of Au@Ag NPs. (a) Scheme illustrating the importance of seed crystallinity to define the shape of the silver shell. Three types of gold seeds: single crystalline rods, octahedrons and penta-twinned rods are represented before and after silver deposition. (b)–(j) electron microscopy images characterizing the above-mentioned particles. Au@Ag NPs were prepared from gold seeds of different crystallinity (b)–(d) octahedrons. (e)–(g) single crystalline rods. (h)–(j) Pentagonally twinned rods. Three types of electron microscopy techniques were used: (b)–(e) TEM (and SEM in b), (c)–(i) STEM-HAADF, (d)–(j) Electron tomography. (b)–(i) and (j) were adapted with permission from Refs. 102 and 106, respectively. Copyright American Chemical Society.

other combinations of metals (or even oxides) can be used including Cu, Co, Mn_3O_4 and Fe_2O_3. Although most attention is directed toward the galvanic replacement of an Ag template by gold ions, the reaction mechanism may be generally translated to other cases. Colloidal synthesis of nanostructures involving galvanic replacement reactions was started by Sun *et al.* in 2002 and is nowadays widely used.[61] They prepared gold nanocages (Au NCs), hollow and porous gold nanostructures through galvanic replacement by reacting silver nanocubes with chloroauric acid. When gold ions are added to a suspension of Ag NPs, Ag surface atoms get quickly oxidized and dissolved, subsequently generating a small hole on one of the NP facets (Fig. 12.10(a)). Simultaneously, the generated electrons are captured by the gold ions to generate Au atoms by reduction. The deposition of gold on the surface of Ag NPs acts as an anticorrosion layer for the Ag underneath; as a result, the pinhole formed in the early stage of the process serves as an active site for the continuous dissolution of Ag inside the NP, ultimately forming hollow nanostructures.[106,107] Importantly, the oxidation of three Ag atoms from the NP is required for the reduction of one Au^{3+} ion on the NP when $HAuCl_4$ is used as gold precursor. Therefore, the stoichiometry of the two elements has to be carefully selected to control the level of void and porosity of the final product. During the titration of the silver particles by gold ions, the formation of an Au–Ag alloy is observed, but in the final stages, particles undergo a dealloying process where all the silver may be completely replaced. The optical properties can thus be tuned by controlling the amount of gold and porosity in the nanostructure. For instance, gold nanocages prepared from silver nanocubes may have plasmon modes in the NIR region.[107] By tuning the amount of $HAuCl_4$ in the reaction, the LSPR peak position of Au NCs can be precisely tuned to display LSPR peaks around 800 nm, while maintaining sizes around 40 nm and allowing specific targeting through bioconjugation with antibodies.[108] A wide range of silver nanostructures have been used, such as cubes, rods, wires, octahedrons, decahedrons and triangles, giving rise to the corresponding hollow nanostructures,

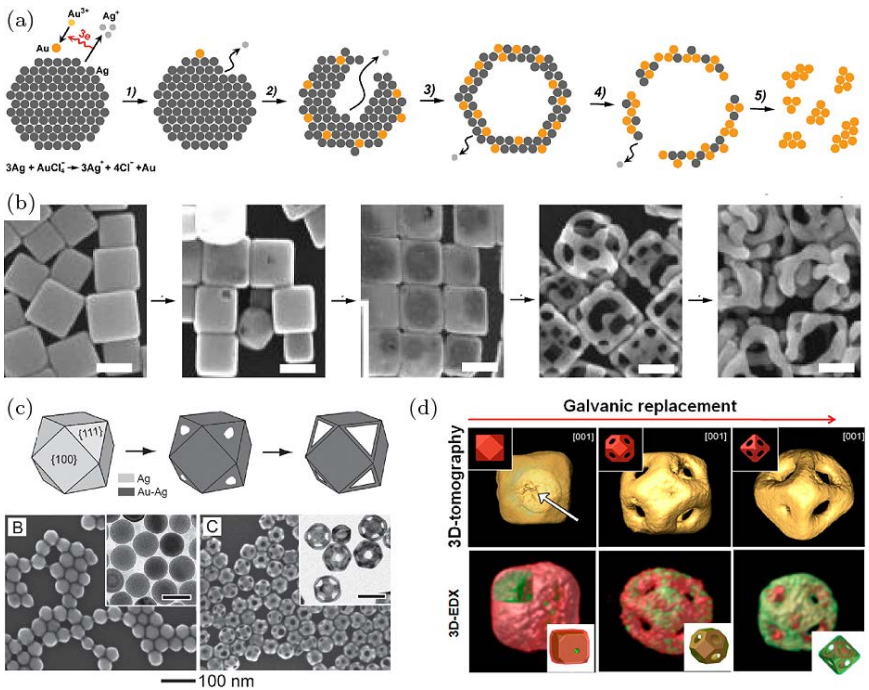


Fig. 12.10. Synthesis of hollow particles. (a) Schematic illustration of structural changes during the galvanic replacement reaction between an Ag NP and $HAuCl_4$ [106] (b) TEM images showing the structural and morphological changes during the titration of silver nanocubes and a gold salt ($HAuCl_4$) in aqueous solution. Scale bars on all images are 100 nm.[43] (c) Selective galvanic replacement on a silver cuboctahedron and $HAuCl_4$ solution, before (b) and after (c) galvanic replacement.[112] (d) 3D rendering of the morphology and chemical composition of Ag nanocubes and Au–Ag octahedral nanocages.[113] (a) and (c) were adapted with permission from Ref. [107], Copyright 2015 Wiley-VCH. (b) and (d) adapted with permission from Refs. [43] and [118]. Copyright 2015 American Chemical Society.

using gold as both an etchant and a coating agent.[109–118] Accordingly, various particle structures have been obtained including cages, tubes and frames. This synthetic approach for the preparation of hollow nanocrystals is simple, versatile and has been applied both in water and in organic solvents. The reaction is believed to occur at the same rate on the six {100} facets of Ag nanocubes, resulting in Au–Ag nanocages with random porosity (Fig. 12.10(b)).[43] On the

other hand, owing to the surface selectivity of the galvanic replacement reaction, the dissolution of Ag is specific on Ag cuboctahedrons which are enclosed by six {100} and eight {111} facets (Fig. 12.10(c)).[112] In this case, the dissolution of Ag atoms is initiated at the {111} facets, whereas the deposition of Au atoms occurs on the {100} facets covered by PVP. The nanocages produced from the Ag cuboctahedrons show uniform pores localized at the eight corners where the {111} facets were positioned (Fig. 12.10(d)).[113] Note that the surface selectivity of the reaction can be modified by the introduction of halides or surface-selective ligands such as PVP because they alter the surface free energy of the different facets. The galvanic replacement reaction can also be modulated by complexing the gold ions to tune the reduction potential. In most cases $HAuCl_4$ is used as a precursor (Au^{3+}/Au, 1.50 V versus standard hydrogen electrode (SHE); $AuCl^{4-}$/Au, 0.93 V versus SHE) with a higher potential than that for the silver template (Ag^{+}/Ag, 0.80 V versus SHE). However, the reduction potential of gold ions complexed with iodide becomes lower than that of the silver template[114,115] (AuI_4^-/Au, 0.56 V versus. SHE), which can be used to prevent the galvanic replacement reaction.[61,114] This strategy has been used to synthesize Ag@Au nanoplates in high yield.[114] In order to prevent the galvanic replacement reaction, a competing reaction can be introduced by means of ascorbic acid at high pH for fabricating Ag@Au nanocubes.[116] In general, the preparation of hollow nanocrystals is performed at 100°C to control the diffusion of Au and Ag atoms and obtain nearly atomically smooth nanocrystals. By performing the reaction at room temperature under particular conditions, other diffusive processes are allowed to take place like the Kirkendall effect. Gonźalez *et al.* combined galvanic and Kirkendall reactions to prepare double-walled nanocages at room temperature.[111]

12.3. Emerging Plasmonic Nanoparticles

Apart from noble metal NPs, certain non-metallic materials can exhibit LSPRs.[117,118] For example, nanocrystals "intrinsically"

doped by the formation of cation vacancies, also called self-doped, or "extrinsically" doped with heterogeneous dopant atoms present high concentrations of free charge carriers resulting in LSPR modes at NIR wavelengths, which are particularly useful not only for biomedical applications, but also for photonics, optoelectrics and nanoelectronics. The origins of LSPRs in doped semiconductors is the same as that in noble metal NPs, i.e. the resonant interaction between the electromagnetic field and free charge carriers. The charge carriers in noble metals are electrons, while in doped semiconductor or metal oxide nanocrystals can also be holes. Additionally, the density of these charge carriers can vary with the degree of doping or with the surface ligands, which cannot be realized in metals.

Intrinsically-doped NCs include doped semiconductor nanocrystals in which free holes result from the presence of cation vacancies. These are mainly composed by copper-deficient copper chalcogenide nanocrystals, including copper sulfide ($Cu_{2-x}S$),[119–121] copper selenide ($Cu_{2-x}Se$),[122,123] copper telluride ($Cu_{2-x}Te$).[124,125] and alloys $Cu_{2-x}SSe_{1-y}$.[126] The study and synthesis of these nanocrystals is in progress, and LSPR tuning is typically achieved by changing the amount of doping via different combinations or other strategies such as size and shape variations or tuning copper deficiency by oxidation. On the other hand, extrinsically-doped colloidal nanocrystals displaying LSPRs correspond basically to colloidal indium tin oxide (ITO),[127,128] ZnO doped with group III elements,[129] tungsten oxide ($WO_{3-\delta}$),[130] molybdenum trioxide (MO_{3-x}),[131] titanium dioxide (TiO_2)[132] or even silicon doped nanocrystals.[133] These novel materials are still emerging and further developments are expected in coming years, but the currently existing materials can be very attractive for applications where NIR and mid-IR light is involved.

12.4. Closing Remarks

In this chapter, the preparation of plasmonic NPs by wet-chemistry approaches has been discussed. We aimed at providing to the reader an overview of the past, the present and the future challenges in

the synthesis of plasmonic NPs through three sections. In particular, in the first section, the history of their preparation over the last century was introduced. In the second section, the recent literature on the preparation of nanocrystals having different dimensionalities and composition was discussed in detail. Finally, some examples depicting the synthesis of emerging plasmonic NPs were briefly presented in the third section.

Although spherical NPs have been prepared for more than a century, the synthesis of anisotropic NPs is still under extensive development in order to understand the mechanisms of formation and to prepare them with a high yield. The milestones in the synthesis processes invoke a fine control of the kinetics of the reaction and a careful addition of reactant. In particular, the roles of adsorbate such as halides as well as the role of the seed crystallinity and their influence to direct the final NP shape have been discussed. Even if significant advances have been made over the past decades, the comprehension of the processes is still empirical in many syntheses reported in the literature. Therefore, a unified theory encompassing the thermodynamics and the kinetics of the process as well as the description of the molecular interaction between the organic and inorganic moieties is anticipated. In this sense, the continuous development of novel microscopy and computational techniques will help scientists to improve knowledge in the field.

The preparation of nanoalloys and hollow particles has also been presented, illustrating the many faces of the chemical synthesis of plasmonic NPs. However, despite their exiting optical properties and the fine control of their morphology, plasmonic particles made of noble metals face a major problem for their implementation in devices, regarding the high cost of the raw materials. In this respect, devising plasmonic NPs made of more materials that are abundantly found on earth is required. Consequently, a short discussion presenting some of those emerging plasmonic NPs has been provided. In conclusion, NP synthesis is a broad field in constant rejuvenation along with the parallel development of new fields that aim to address their implementation in devices.

Acknowledgments

Funding from the European Commission (Grant #310445-2 SAVVY) and the European Research Council (ERC Advanced Grant #267867 Plasmaquo) are acknowledged.

References

1. Freestone, I., Meeks, N., Sax, M. and Higgitt, C. (2007). *Gold Bull.* **40**, p. 270.
2. Hunt, L. B. (1976). *Gold Bull.* **9**, p. 134.
3. Faraday, M. (1857). *Philos. Trans. R. Soc. Lond.* **147**, p. 145.
4. Mie, G. (1908). *Ann. Phys.* **25**, p. 377.
5. Zsigmondy, R. (1917). *The Chemistry of Colloids*, Wiley, New York.
6. *Nobel Lectures in Chemistry 1922–1941*, Elsevier, Amsterdam, 1966.
7. Turkevich, J., Stevenson, P. C. and Hillier, J. (1951). *Discuss. Faraday Soc.* p. 55.
8. Frens, G. (1973). *Nature*, **241**, p. 20.
9. Klinkova, A., Choueiri, R. M. and Kumacheva, E. (2014). *Chem. Soc. Rev.* **43**, p. 3976.
10. Xu, L., Ma, W., Wang, L., Xu, C., Kuang, H. and Kotov, N. A. (2013). *Chem. Soc. Rev.* **42**, p. 3114.
11. Hamon, C. and Liz-Marzán, L. M. (2015). *Chem. Eur. J.* **21**, p. 9956.
12. Sapsford, K. E., Algar, W. R., Berti, L., Gemmill, K. B., Casey, B. J., Oh, E., Stewart, M. H. and Medintz, I. L. (2013). *Chem. Rev.* **113**(3), p. 1904.
13. Dreaden, E. C., Alkilany, A. M., Huang, X., Murphy, C. J. and El-Sayed, M. A. (2012). *Chem. Soc. Rev.* **41**, p. 2740.
14. Tokel, O., Inci, F. and Demirci, U. (2014). *Chem. Rev.* **114**, p. 5728.
15. Brust, M., Walker, M., Bethell, D., Schiffrin, D. J. and Whyman, R. (1994). *J. Chem. Soc. Chem. Commun.* p. 801.
16. Bastus, N. G., Comenge, J. and Puntes, V. (2011). *Langmuir*, **27**, p. 11098.
17. Scarabelli, L., Grzelczak, M. and Liz-Marzán, L. M. (2013). *Chem. Mater.* **25**, p. 4232.
18. Kumar, P. S., Pastoriza-Santos, I., Rodríguez-González, B., García de Abajo, J. F. and Liz-Marzán, L. M. (2008). *Nanotechnology*, **19**, p. 015606.
19. Scarabelli, L., Coronado-Puchau, M., Giner-Casares, J. J., Langer, J. S. and Liz-Marzán, L. M. (2014). *ACS Nano* **8**, p. 5833.
20. Bastús, N. G., Merkoçi, F., Piella, J. and Puntes, V. (2014). *Chem. Mater.* **26**, p. 2836.
21. Cortie, M. B. and McDonagh, A. M. (2011). *Chem. Rev.* **111**, p. 3713.
22. Grzelczak, M., Perez-Juste, J., Mulvaney, P. and Liz- Marzán, L. M. (2008). *Chem. Soc. Rev.* **37**, p. 1783

23. Xia, Y., Xia, X. and Peng, H.-C. (2015). *J. Am. Chem. Soc.* **137**, p. 7947.
24. Zhang, H., Jin, M. and Xia, Y. (2012). *Angew. Chem. Int. Ed.* **51**, p. 7656.
25. Wang, Y., He, J., Liu, C., Chong, W. H. and Chen, H. (2014). *Angew. Chem. Int. Ed.* **54**, p. 2022.
26. LaMer, V. K. and Dinegar, R. H. (1950). *J. Am. Chem. Soc.* **72**, p. 4847.
27. Morales, M. P., González-Carreño, T. and Serna, C. J. (1992). *J. Mater. Res.* **7**, p. 2538.
28. Sugimoto, T. and Matijević, E. (1980). *J. Colloid. Interf. Sci.* **74**, p. 227.
29. Den Ouden, C. J. J. and Thompson, R. W. (1991). *J. Colloid. Interf. Sci.* **143**, p. 77.
30. Ocaña, M., Rodriguez-Clemente, R. and Serna, C. J. (1995). *Adv. Mater.* **7**, p. 212.
31. Brzeziński, M., Biedroń, T., Tracz, A., Kubisa, P. and Biela, T. (2014). *Macromol. Chem. Physic.* **215**, p. 27.
32. Murphy, C. J., Thompson, L. B., Alkilany, A. M., Sisco, P. N., Boulos, S. P., Sivapalan, S. T., Yang, J. A., Chernak, D. J. and Huang, J. (2010). *J. Phys. Chem. Lett.* p. 2867.
33. Gómez-Graña, S., Hubert, F., Testard, F., Guerrero-Martínez, A., Grillo, I., Liz-Marzán, L. M. and Spalla, O. (2011). *Langmuir*, **28**, p. 1453.
34. Xia, Y., Xiong, Y., Lim, B. and Skrabalak, S. E. (2009). *Angew. Chem. Int. Ed.* **48**, p. 60.
35. Barnard, A. S., Lin, X. M. and Curtiss, L. A. (2005). *J. Phys. Chem. B.* **109**, p. 24465.
36. Carbó-Argibay, E., Rodríguez-González, B., Gómez-Graña, S., Guerrero-Martínez, A., Pastoriza-Santos, I., Pérez-Juste, J. and Liz-Marzán, L. M. (2010). *Angew. Chem. Int. Ed.* **122**, p. 9587.
37. Personick, M. L. and Mirkin, C. A. (2013). *J. Am. Chem. Soc.* **135**, p. 18238.
38. Langille, M. R., Personick, M. L., Zhang, J. and Mirkin, C. A. (2012). *J. Am. Chem. Soc.* **134**, p. 14542.
39. O'Brien, M. N., Jones, M. R., Brown, K. A. and Mirkin, C. A. (2014). *J. Am. Chem. Soc.* **136**, p. 7603.
40. Markov, I. V. (1995). *Crystal growth for beginners: Fundaments of nucleation, crystal growth, and epitaxy*, (World Scientific, Singapore).
41. Lee, Y. J., Schade, N. B., Sun, L., Fan, J. A., Bae, D. R., Mariscal, M. M., Lee, G., Capasso, F., Sacanna, S., Manoharan, V. N. and Yi, G. R. (2013). *ACS Nano*, **7**, p. 11064.
42. Kim, F., Connor, S., Song, H., Kuykendall, T. and Yang, P. (2004). *Angew. Chem. Int. Ed.* **43**, p. 3673.
43. Sun, Y. and Xia, Y. (2004). *J. Am. Chem. Soc.* **126**, p. 3892.
44. Li, C., Shuford, K. L., Chen, M., Lee, E. J. and Cho, S. O. (2008). *ACS Nano*, **2**, p. 1760.
45. Niu, W., Zhang, W., Firdoz, S. and Lu, X. (2014). *J. Am. Chem. Soc.* **136**, p. 3010.
46. Yu, Y., Zhang, Q., Lu, X. and Lee, J. Y. (2010). *J. Phys. Chem. C*, **114**, p. 11119.

47. Tian, N., Zhou, Z. Y., Sun, S. G., Ding, Y. and Wang, Z. L. (2007). *Science*, **316**, p. 732.
48. Lohse, S. E., Burrows, N. D., Scarabelli, Liz-Marzán, L. M. and Murphy, C. J. (2014). *Chem. Mater.* **26**, p. 34.
49. Martin, C. R. (1991). *Adv. Mater.* **3**, p. 457.
50. Yu, Chang, S.-S., Lee, C.-L. and Wang, C. R. C. (1997). *J. Phys. Chem. B*, **101**, p. 6661.
51. Link, S. and El-Sayed, M. A. (1999). *J. Phys. Chem. B*, **103**, p. 8410.
52. Huang, X., Neretina, S. and El-Sayed, M. A. (2009). *Adv. Mater.* **21**, p. 4880.
53. Gole, A. and Murphy, C. J. (2004). *Chem. Mater.* **16**, p. 3633.
54. Burda, C., Chen, X., Narayanan, R. and El-Sayed, M. A. (2005). *Chem. Rev.* **105**, p. 1025.
55. Chen, H., Shao, L., Li, Q. and Wang, J. (2013). *Chem. Soc. Rev.* **42**, p. 2679.
56. Rycenga, M., Cobley, C. M., Zeng, J., Li, W., Moran, C. H., Zhang, Q., Qin, D. and Xia, Y. (2011). *Chem. Rev.* **111**, p. 3669.
57. Lohse, S. E., Burrows, N. D., Scarabelli, L., Liz-Marzán, L. M. and Murphy, C. J. (2014). *Chem. Mater.* **26**, p. 34.
58. Smith, D. K., Miller, N. R. and Korgel, B. A. (2009). *Langmuir*, **25**, p. 9518.
59. Sun, Y., Gates, B., Mayers, B. and Xia, Y. (2002). *Nano Lett.* **2**, p. 165.
60. Murph, S. E. H., Murphy, C. J., Leach, A. and Gall, K. (2015). *Cryst. Growth Des.* **15**, p. 1968.
61. Vigderman, L., Khanal, B. P. and Zubarev, E. R. (2012). *Adv. Mater.* **24**, p. 4811.
62. Almora-Barrios, N., Novell-Leruth, G., Whiting, P., Liz-Marzán, L. M. and López, N. (2014). *Nano Lett.* **14**, p. 871.
63. Hubert, F., Testard, F. and Spalla, O. (2008). *Langmuir*, **24**, p. 9219.
64. Liu and Guyot-Sionnest, P. (2005). *J. Phys. Chem. B*, **109**, p. 22192.
65. Ye, X., Jin, L., Caglayan, H., Chen, J., Xing, G., Zheng, C., Doan-Nguyen, V., Kang, Y., Engheta, N., Kagan, C. R. and Murray, C. B. (2012). *ACS Nano*, **6**, p. 2804.
66. Wadams, R. C., Fabris, L., Vaia, R. A. and Park, K. (2013). *Chem. Mater.* **25**, p. 4772.
67. Henkel, A., Jakab, A., Brunklaus, G. and Sönnichsen, C. (2009). *J. Phys. Chem. C*, **113**, p. 2200.
68. Johnson, C. L., Snoeck, E., Ezcurdia, M., Rodriguez-Gonzalez, B., Pastoriza-Santos, I., Liz-Marzán, L. M. and Hytch, M. J. (2008). *Nat. Mater.* **7**, p. 120.
69. Murphy, C. J. and Jana, N. R. (2002). *Adv. Mater.* **14**, p. 80.
70. Kim, F., Sohn, K., Wu, J. and Huang, J. (2008). *J. Am. Chem. Soc.* **130**, p. 14442.
71. Pazos-Pérez, N., Baranov, D., Irsen, S., Hilgendorff, M., Liz-Marzán, L. M. and Giersig, M. (2008). *Langmuir*, **24**, p. 9855.
72. Sun, Y., Mayers, B., Herricks, T. and Xia, Y. (2003). *Nano Lett.* **3**, p. 955.

73. Giersig, M., Pastoriza-Santos, I. and Liz-Marzán, L. M. (2004). *J. Mater. Chem.* **14**, p. 607.
74. Porel, S., Singh, S. and Radhakrishnan, T. P. (2005). *Chem. Comm.* p. 2387.
75. Yamamoto, M., Kashiwagi, Y., Sakata, T., Mori, H. and Nakamoto, M. (2005). *Chem. Mater.* **17**, p. 5391.
76. Sun, X., Dong, S. and Wang, E. (2004). *Angew. Chem. Int. Ed.* **43**, p. 6360.
77. Chu, H. C., Kuo, C. H. and Huang, M. H. (2006). *Inorg. Chem.* **45**, p. 808.
78. Millstone, J. E., Hurst, S. J., Métraux, G. S., Cutler, J. I. and Mirkin, C. A. (2009). *Small*, **5**, p. 646.
79. Bastys, V., Pastoriza-Santos, I., Rodríguez-González, B., Vaisnoras, R. and Liz-Marzán, L. M. (2006). *Adv. Funct. Mater.* **16**, p. 766.
80. Millstone, J. E., Métraux, G. S. and Mirkin, C. A. (2006). *Adv. Funct. Mater.* **16**, p. 1209.
81. Park, K., Koerner, H. and Vaia, R. A. (2010). *Nano Lett.* **10**, p. 1433.
82. Pelaz, B., Grazu, V., Ibarra, A., Magen, C., del Pino, P. and de la Fuente, J. M. (2012). *Langmuir*, **28**, p. 8965.
83. Guerrero-Martínez, A., Barbosa, S., Pastoriza-Santos, I. and Liz-Marzán, L. M. (2011). *Curr. Op. Colloid Inter. Sci.* **16**, p. 118.
84. Wiley, B., Sun, Y., Mayers, B. and Xia, Y. (2005). *Chem. Eur. J.* **11**, p. 454.
85. Tsung, C. K., Kuhn, J. N., Huang, W., Aliaga, C., Hung, L. I., Somorjai, G. A. and Yang, P. (2009). *J. Am. Chem. Soc.* **131**, p. 5816.
86. Yuan, H., Khoury, C. G., Hwang, H., Wilson, C. M., Grant, G. A. and Vo-Dinh, T. (2012). *Nanotechnology*, **23**, p. 075102.
87. Zhang, J., Langille, M. R., Personick, M. L., Zhang, K., Li, S. and Mirkin, C. A. (2010). *J. Am. Chem. Soc.* **132**, p. 14012.
88. Kim, D. Y., Yu, T., Cho, E. C., Ma, Y. F., Park, O. O. and Xia, Y. (2011). *Angew. Chem. Int. Ed.* **123**, p. 6452.
89. Lu, L. H., Ai, K. L. and Ozaki, Y. (2008). *Langmuir* **24**, p. 1058.
90. Jimenez de Aberasturi, D., Serrano-Montes, A. B. and Liz-Marzán, L. M. (2015). *Adv. Opt. Mater.* **3**, p. 602.
91. Chen, S. and Carroll, D. L. (2002). *Nano Lett.* **2**, p. 1003.
92. Sau, T. K. and Murphy, C. J. (2004). *J. Am. Chem. Soc.* **126**, p. 8648.
93. Hao, E., Bailey, R. C., Schatz, G. C., Hupp, J. T. and Li, S. (2004). *Nano Lett.* **4**, p. 327.
94. Nehl, C. L., Liao, H. and Hafner, J. H. (2006). *Nano Lett.* **6**, p. 683.
95. Rodriguez-Fernandez, D., Altantzis, T., Heidari, H., Bals, S. and Liz-Marzán, L. M. (2014). *Chem. Comm.* **50**, p. 79.
96. Novikov, S. M., Sánchez-Iglesias, A., Schmidt, M. K., Chuvilin, A., Aizpurua, J., Grzelczak, M. and Liz-Marzán, L. M. (2014). *Part. Part. Syst. Charact.* **31**, p. 77.
97. Grzelczak, M., Sanchez-Iglesias, A., Rodriguez-Gonzalez, B., Alvarez-Puebla, R., Perez-Juste, J. and Liz-Marzan, L. M. (2008). *Adv. Funct. Mater.* **18**, p. 3780.

98. Hu, J., Zhang, Y., Liu, B., Liu, J., Zhou, H., Xu, Y., Jiang, Y., Yang, Z. and Tian, Z. Q. (2004). *J. Am. Chem. Soc.* **126**, p. 9470.
99. Li, F., Tian, D. and Cui, H. (2012). *Luminescence* **28**, p. 7.
100. Huang, T., Meng, F. and Qi, L. (2010). *Langmuir* **26**, p. 7582.
101. Gómez-Graña, S., Goris, B., Altantzis, T., Fernández-López, C., Carbó-Argibay, E., Guerrero-Martínez, A., Almora-Barrios, N., López, N., Pastoriza-Santos, I., Pérez-Juste, J., Bals, S., Van Tendeloo, G. and Liz-Marzán, L. M. (2013). *J. Phys. Chem. Lett.* **4**, p. 2209.
102. Park, K., Drummy, L. F. and Vaia, R. A. (2011). *J. Mater. Chem.* **21**, p. 15608.
103. Polavarapu, L., Mourdikoudis, S., Pastoriza-Santos, I. and Perez-Juste, J. (2015). *Cryst. Eng. Comm.* **17**, p. 3727.
104. Tsao, Y.-C., Rej, S., Chiu, C.-Y. and Huang, M. H. (2014). *J. Am. Chem. Soc.* **136**, p. 396.
105. Mayer, M., Scarabelli, L., March, K., Altantzis, T., Tebbe, M., Kociak, M., Bals, S., Garcia de Abajo, J., Fery, A. and Liz-Marzán, L. M. (2015). *Nano Lett.* **15**, p. 5427.
106. Xia, X., Wang, Y., Ruditskiy, A. and Xia, Y. (2013). *Adv. Mater.* **25**, p. 6313.
107. Skrabalak, S. E., Au, L., Li, X. and Xia, Y. (2007) *Nat. Protoc.* **2**, p. 2182.
108. Chen, J., Saeki, F., Wiley, B. J., Cang, H., Cobb, M. J., Li, Z. Y., Au, L., Zhang, H., Kimmey, M. B., Li, X. and Xia, Y. (2005). *Nano Lett.* **5**, p. 473.
109. Kim, M. H., Lu, X., Wiley, B., Lee, E. P. and Xia, Y. (2008). *J. Phys. Chem. C* **112**, p. 7872.
110. Skrabalak, S. E., Chen, J., Sun, Y., Lu, X., Au, L., Cobley, C. M. and Xia, Y. (2008). *Acc. Chem. Res.* **41**, p. 1587.
111. González, E., Arbiol, J. and Puntes, V. F. (2011). *Science* **334**, p. 1377.
112. Cobley, C. M. and Xia, Y. (2010). *Mater. Sci. Eng. R* **70**, p. 44.
113. Goris, B., Polavarapu, L., Bals, S., Van Tendeloo, G. and Liz-Marzán, L. M. (2014). *Nano Lett.* **14**, p. 3220.
114. Gao, C., Lu, Z., Liu, Y., Zhang, Q., Chi, M., Cheng, Q. and Yin, Y. (2012). *Angew. Chem. Int. Ed.* **51**, p. 5629.
115. Gao, C., Zhang, Q., Lu, Z. and Yin, Y. (2011). *J. Am. Chem. Soc.* **133**, p. 19706.
116. Yang, Y., Liu, J., Fu, Z.-W. and Qin, D. (2014). *J. Am. Chem. Soc.* **136**, p. 8153.
117. Liu, X. and Swihart, M. T. (2014) *Chem. Soc. Rev.* **43**, p. 3908.
118. Comin, A. and Manna, L. (2014). *Chem. Soc. Rev.* **43**, p. 3957.
119. Luther, J. M., Jain, P. K., Ewers, T. and Alivisatos, A. P. (2011). *Nat. Mater.* **10**, p. 361.
120. Hsu, S. W., On, K. and Tao, A. R. (2011). *J. Am. Chem. Soc.* **133**, p. 19072.
121. Liu, X., Wang, X., Zhou, B., Law, W. C., Cartwright, A. N. and Swihart, M. T. (2013). *Adv. Funct. Mater.* **23**, p. 1256.

122. Riha, S. C., Johnson, D. C. and Prieto, A. L. (2011). *J. Am. Chem. Soc.* **133**, p. 1383.
123. Deka, S., Genovese, A., Zhang, Y., Miszta, K., Bertoni, G., Krahne, R., Giannini, C. and Manna, L. (2010). *J. Am. Chem. Soc.* **132**, p. 8912.
124. Kriegel, I., Jiang, C., Rodríguez-Fernández, J., Schaller, R. D., Talapin, D. V., da Como, E. and Feldmann, J. (2012). *J. Am. Chem. Soc.* **134**, p. 1583.
125. Li, W., Zamani, R., Rivera Gil, P., Pelaz, B., Ibañez, M., Cadavid, D., Shavel, A., Alvarez-Puebla, R. A., Parak, W. J., Arbiol, J. and Cabot, A. (2013). *J. Am. Chem. Soc.* **135**, p. 7098.
126. Dilena, E., Dorfs, D., George, C., Miszta, K., Povia, M., Genovese, A., Casu, A., Prato, M. and Manna, L. (2012). *J. Mater. Chem.* **22**, p. 13023.
127. Kanehara, M., Koike, H., Yoshinaga, T. and Teranishi, T. (2009). *J. Am. Chem. Soc.* **131**, p. 17736.
128. Garcia, G., Buonsanti, R., Runnerstrom, E. L., Mendelsberg, R. J., Llordes, A., Anders, A., Richardson, T. J. and Milliron, D. J. (2011). *Nano Lett.* **11**, p. 4415.
129. Buonsanti, R., Llordes, A., Aloni, S., Helms, B. A. and Milliron, D. J. (2011). *Nano Lett.* **11**, p. 4706.
130. Manthiram, K. and Alivisatos, A. P. (2012). *J. Am. Chem. Soc.* **134**, p. 3995.
131. Huang, Q., Hu, S., Zhuang, J. and Wang, X. (2012). *Chem. Eur. J.* **18**, p. 15283.
132. Gordon, T. R., Cargnello, M., Paik, T., Mangolini, F., Weber, R. T., Fornasiero, P. and Murray, C. B. (2012). *J. Am. Chem. Soc.* **134**, p. 6751.
133. Rowe, D. J., Jeong, J. S., Mkhoyan, K. A. and Kortshagen, U. R. (2013). *Nano Lett.* **13**, p. 1317.

Index

World Scientific Series in Nanoscience and Nanotechnology

(Continuation of series card page)

Vol. 13 *Nanoelectronics: A Molecular View*
by Avik Ghosh (University of Virginia, USA)

Vol. 12 *Nanomaterials for Photocatalytic Chemistry*
edited by Yugang Sun (Temple University, USA)

Vol. 11 *Molecular Bioelectronics: The 19 Years of Progress Second Edition*
by Nicolini Claudio (University of Genoa, Italy)

Vol. 10 *Pore Scale Phenomena: Frontiers in Energy and Environment*
edited by John Poate (Colorado School of Mines, USA), Tissa Illangasekare (Colorado School of Mines, USA), Hossein Kazemi (Colorado School of Mines, USA) and Robert Kee (Colorado School of Mines, USA)

Vol. 9 *Handbook of Biomimetics and Bioinspiration: Biologically-Driven Engineering of Materials, Processes, Devices, and Systems (In 3 Volumes)*
edited by Esmaiel Jabbari (University of South Carolina, USA), Deok-Ho Kim (University of Washington, USA), Luke P Lee (University of California, Berkeley, USA), Amir Ghaemmaghami (University of Nottingham, UK) and Ali Khademhosseini (Harvard University, USA & Massachusetts Institute of Technology, USA)

Vol. 8 *Polyoxometalate Chemistry: Some Recent Trends*
edited by Francis Sécheresse (Université de Versailles-St Quentin, France)

Vol. 7 *Scanning Probe Microscopy for Energy Research*
edited by Dawn A Bonnell (The University of Pennsylvania, USA) and Sergei V Kalinin (Oak Ridge National Laboratory, USA)

Vol. 6 *Plasmon Resonances in Nanoparticles*
by Isaak D Mayergoyz (University of Maryland, USA)

Vol. 5 *Inorganic Nanomaterials from Nanotubes to Fullerene-Like Nanoparticles: Fundamentals and Applications*
by Reshef Tenne (Weizmann Institute of Science, Israel)

*For the complete list of volumes in this series, please visit
www.worldscientific.com/series/wssnn