

UNIVERSITE MONTPELLIER II  
SCIENCES ET TECHNIQUES DU LANGUEDOC

T H E S E

pour obtenir le grade de

**DOCTEUR DE L'UNIVERSITE MONTPELLIER II**

***Discipline : Physique de la Matière Condensée***

Ecole Doctorale: Physique et Chimie

présentée et soutenue publiquement

par

*Alexandru Ioan Căbuz*

*alexcabuz@gmail.com*

le 19 juin 2007

Titre :

*Electromagnetic Metamaterials*

-

*From Photonic Crystals to Negative Index Composites*

**Jury:**

Prof. V. Lorman  
Prof. D. Felbacq  
Prof. A. Lakhtakia  
Prof. J.-M. Raimond  
Prof. J.-P. Vigneron  
Prof. D. Wiersma  
Prof. D. Cassagne

Université Montpellier II  
Université Montpellier II  
Penn State University, USA  
Ecole Normale Supérieure, Paris  
FUNDP, Namur Belgique  
LENS, Florence Italie  
Université Montpellier II

Président  
Directeur de Thèse  
Rapporteur (non présent)  
Rapporteur  
Rapporteur  
Examineur  
Examineur



# Contents

<b>Introduction</b>	<b>5</b>
<b>1 Electrons, dielectrics and beyond</b>	<b>9</b>
1.1 Spatial averaging as truncation . . . . .	9
1.2 Ensemble averaging . . . . .	13
1.3 Periodic media . . . . .	17
1.4 Polarizability, susceptibility – electric . . . . .	19
1.4.1 Small atoms – electric . . . . .	21
1.4.2 Big atoms – electric . . . . .	25
1.5 Polarizability, susceptibility – magnetic . . . . .	26
1.5.1 Small atoms – magnetic . . . . .	27
1.6 Permittivity and permeability – index and impedance . . . . .	28
1.7 Inhomogeneous models and local models . . . . .	30
1.8 Constructive versus holistic . . . . .	35
1.9 Beyond dielectrics - truncation revisited . . . . .	36
<b>2 Photonic crystals – the super-prism effect</b>	<b>39</b>
2.1 Introduction . . . . .	39
2.2 Numerical Methods . . . . .	40
2.2.1 The plane wave expansion method . . . . .	40
2.2.2 The finite difference time domain method . . . . .	43
2.3 Iso-frequency curves – the construction line . . . . .	44
2.3.1 The effective index approximation . . . . .	46
2.4 The super-prism effect . . . . .	47
2.5 Transmission efficiency . . . . .	49
2.6 The incident beam width . . . . .	52
<b>3 Negative Index Composite Metamaterials</b>	<b>55</b>
3.1 Prior art . . . . .	55
3.1.1 The Imaging Studies . . . . .	55
3.1.2 The Materials Studies . . . . .	57
3.2 Negative index of refraction – the superlens . . . . .	59
3.3 Flat lenses . . . . .	64
3.4 Thin wire medium . . . . .	66
3.4.1 Transfer matrix of thin metal wire grating . . . . .	66
3.4.2 Effective permittivity - analytical . . . . .	68
3.4.3 Effective permittivity - numerical . . . . .	70

3.5 Dielectric resonators . . . . .	71
3.6 Metal resonators . . . . .	72
3.7 Composite metamaterial – the 1D stack approach . . . . .	79
<b>Conclusion</b>	<b>87</b>
<b>Acknowledgements</b>	<b>89</b>
<b>Appendices</b>	
<b>A Multiscattering theory of circular rod gratings</b>	<b>91</b>
A.1 Single circular rod . . . . .	91
A.2 Infinite grating . . . . .	92
A.3 Plane wave - Bessel series representation . . . . .	92
A.4 Multi-scattering contribution to the local field . . . . .	93
A.5 Computing $A_0$ , $A_1$ and $A_2$ numerically . . . . .	94
A.6 Estimating $A_0$ , $A_1$ and $A_2$ analytically . . . . .	95
A.7 Transmission and reflection coefficients . . . . .	97
A.8 Domain of validity - comparison with rigorous simulations . . . . .	98
<b>B The parameter extraction procedure</b>	<b>101</b>
<b>Bibliography</b>	<b>105</b>

# Introduction

In the early 1980's classical electromagnetism was seen as a largely played out discipline. It had had well over a century in which to reach full maturity. After the exciting period during and immediately following the Second World War, when transmission line theory and antenna theory reached their final form, very few new ideas came to light. Research concentrated more and more on phenomena involving quantum effects, such as superconductivity, or on aspects related to exotic behavior such as bianisotropy or nonlinear effects. Classical electromagnetism appeared to have run out of tricks.

However, by the late eighties excitement was in the air once again, and once more, the Semiconductor Revolution had been the enabling agent. It was pointed out (independently by Yablonovitch [1] and John [2]) that, theoretically, artificial materials could be designed that would exhibit a *photonic band-gap* (PBG), by analogy to the electronic bandgaps responsible for the behavior of semiconductors and insulators. Whereby semiconductors owed their usefulness to the fact that the position of the Fermi level within the gap – and thereby their electric behavior – was malleable, it was hoped that the PBG would allow a control of *photons* analogous to the control of *electrons* achieved in semiconductors.

In order to understand these developments in the broader context of electromagnetic metamaterials research, we now consider the different methods available to control light, while emphasizing the notion of *scale*. We will divide phenomena and devices into categories following the relative scale of *four* characteristic lengths. The first, and most obvious is the wavelength of light,  $\lambda$ . The second is the typical size of atoms,  $a$ , where an “atom” is generally and somewhat arbitrarily understood as a collection of electrons and nuclei that can be seen as a scattering unit in the sense that its interaction with its environment may be treated perturbatively. This somewhat vague definition covers not only atoms in the chemical sense but also any electromagnetically convenient grouping thereof, including molecules, unit cells of crystals or macroscopic scatterers like dipole antennas. The third scale is the inter-atomic distance  $d$  while the fourth scale is that on which the medium is structured,  $s$ , where “structure” refers to any heterogeneity of the medium other than that associated with the internal structure of the atoms themselves: for instance, the presence of regions where their internal structure, density or arrangement is different, or where they are absent. Note that the inter-atomic distance and the atomic size can only be identified in the case of natural dielectric media. They may be different, however, in man made metamaterials, a liberty we use to full advantage in Chapter 3.

The most familiar case is where  $d \approx a \ll \lambda$  and  $s \rightarrow \infty$ . This is the case of an unbounded dielectric medium. The microscopic inhomogeneity on the  $a$  scale is averaged to give a *homogeneous* description of the medium as represented by the permittivity  $\epsilon$  and permeability  $\mu$ . A remarkable feature in the homogeneous description is that it is equivalent to treating each atom of the material as a point scatterer which acquires a *dipole moment* in an external applied field [3]. This *dipolar approximation* shall be very useful since it applies not only to physical atoms, but to any structure of a size much smaller than the wavelength of an applied field. While, as we shall see, it is not absolutely necessary in order to *define* the macroscopic parameters permittivity  $\epsilon$  and permeability  $\mu$ , if the dipolar approximation is not valid, then the *usefulness* of the homogeneous parameters  $\mu, \epsilon$  is drastically reduced (e.g. when quadrupolar

moments are non-negligible). These parameters are in general dyadic (rank 2 tensors) and depend on the frequency.

It is shown in Chapter 1 that in order for an analytic closed form model of the medium to be possible, it is not only necessary for the atoms to behave as point dipoles in interaction with the macroscopic field, but *also in interaction with each other*. They must therefore be not only smaller than the wavelength but also smaller than the distances separating them. In the case of a natural, solid dielectric this is clearly not the case, but when dealing with metamaterials composed of macroscopic scatterers this is not an obstacle. As an example, a multiscattering model is offered in Appendix A that can be used to determine frequency regions where a dielectric rod grating can be seen as a row of dipoles or not.

Another familiar case is where  $d \approx a \ll \lambda \ll s$ . In this category one finds devices such as lenses and mirrors, and phenomena explained using beam optics. Similarly, two- or multi-conductor transmission lines (parallel plate, coaxial, micro-strip) as well as dielectric waveguides in the single mode regime far below cutoff fall into this category. What all these devices have in common is that light propagation in the structure is characterized by only two (possibly frequency dependent) parameters: a phase parameter (the optical index in optics, the propagation constant in transmission lines) and an impedance parameter. The phase parameter governs the phase variation with propagation in a given medium (or section of transmission line), while the impedance parameter governs transmission at interfaces between media (except for total internal reflection, which is a phase-phenomenon). In the case of two or three dimensional structures (e.g. dielectric media) these parameters can take the form of tensors. When they are well defined, the index and impedance contain the same information as the permittivity and permeability mentioned above. Generally speaking, the former are more phenomenologically rooted, while the latter have more physical origins; there are situations, notably when  $\mu$  and  $\varepsilon$  are tensors, when the index and impedance are not well defined in the sense that they depend on the polarization of the field. In these cases the index and impedance can be seen as properties of the *field*, or *wave*, while the permittivity and permeability as properties of *matter*, or the *medium*, in which the wave is propagating. They are obviously closely related. Engineers tend to prefer the first set, while physicists the second. We will generally prefer using the second set whenever possible because it distinguishes naturally between electric and magnetic phenomena.

A somewhat less familiar case but which had also been studied extensively prior to the developments of the late 80's is the case of  $d \approx a \ll s \ll \lambda$ . This situation is encountered, for instance, when considering the propagation of microwaves through clouds or through metal-dielectric composites such as cermets. The first results in this case were obtained as early as 1904, with the work of Maxwell-Garnett [4] and the Wiener bounds several years later [5], and continued through the work of Bruggeman [6], Hashin and Shtrikman [7], Milton [8] and finally Tsang and Kong [9]. This approach was characterized by the *lack* of information about the detailed microscopic structure of the material under study; design was not seen as an option. Thus, a precise approach was not possible and bounds had to be obtained on the effective permittivity of the medium based on the information available, such as the filling fraction of the component materials, and the symmetry of the microscopic particles involved [10]. Indeed, the approach of Maxwell-Garnett, for instance, *relies* on the microscopic randomness of the medium.

Meanwhile, for cases when the microscopic structure of the medium was known in detail, more rigorous methods were developed starting in the 1970's [11]. These heavily mathematical methods would not seek to average the field within a given structure, but rather to obtain the (partial differential) equation which the asymptotic field would need to satisfy, when  $\lambda \rightarrow \infty$ . By comparing the equation obtained with the standard wave equation in a homogeneous material, effective medium parameters could be obtained. However these methods are often mathematically intense and provide little in the way of physical intuition. Moreover in certain cases this rigorous approach is not necessary, namely in cases where the medium can be considered as composed of *independent* scattering elements. In these

cases the same averaging methods could be applied to the scatterers as to the atoms themselves when introducing the macroscopic fields  $\mathbf{D}$  and  $\mathbf{H}$ . The scattering elements could therefore be treated as a kind of macroscopic "meta-atoms". Indeed, some steps in this direction had been taken by Kock as early as 1948 [12] but they generated relatively little interest before the late 1990's, see below.

This is where electromagnetic material science stood at the beginning of the 1980's. Meanwhile, it is interesting to note that the situation in the field of electron transport was quite different. At room temperature the wavelength of most electrons in a typical medium is *shorter* than the typical distance between atoms. In the case of conductors, some of the electrons may have wavelengths larger than the atomic distances, though not by much. We see, therefore that for electrons propagating in a material we have  $d \approx a \approx \lambda_e$  which means that the approximations we need to define homogeneous parameters such as the index are no longer justified. An altogether different approach is required. The electronic "field" propagating in a periodic lattice is no longer sinusoidal, nor even strictly periodic, but rather a quasiperiodic wave known as a *Bloch wave*. The propagation of a Bloch wave in a lattice is considerably more complicated than that of a plane wave in a homogeneous medium. Its phase parameter depends not only on the frequency but also on the direction of propagation. It is possible to have waves with different phase parameters propagating in the same direction at the same frequency and even to find frequencies at which no propagation is allowed *at all*. These phenomena are lumped together under the name of *spatial dispersion* and they appear when the scale of the wavelength is comparable with one or the other two length scales: the atomic scale  $a$ , or the structural scale  $s$ . The notions of index or impedance then become unwieldy or outright useless. In general, in order to characterize the propagation, one needs to perform a separate calculation for each frequency and phase parameter that one is interested in, see Chapter 2.

In this context we can say that Yablonovitch's original idea basically comes down to a realization that the behavior of electrons when  $\lambda_e \approx a$  may be reproduced by photons when  $\lambda_p \approx s$ . In this way a pattern of the relative permittivity as a function of position may play the same role for photons as the periodic spatial distribution of atoms in a crystal does for electrons. The difference is that with modern technology the macroscopic scatterers can be designed and tailored to our convenience whereas we have only limited liberty of controlling the behavior of atoms or their precise assembly.

In all fairness we must also mention that a considerable amount of work had been done before the 80's on one type of structure for which  $s \approx \lambda$ : the diffraction grating. These structures, made of thin, long indentations on flat dielectric or metallic surfaces, had been under study since the late 19th century. A series of high performance mathematical and numerical methods had been painstakingly developed over the years, and by the early 80's the field had reached full maturity. However, in spite of the similarities with the physics of electron transport in crystalline solids, the crosstalk between the two fields was minimal. For instance the classic 1980 text edited by R. Petit on the "Electromagnetic Theory of Gratings" makes a single passing reference to the Floquet-Bloch theorem, a major pillar of solid state physics, with no further development. At the time gratings were seen as structures more akin to a scattering *obstacle* rather than an extreme case (a monolayer) of a spatially dispersive *medium*. It took the shift in perspective brought about by the work of Yablonovitch to realize how deep the analogy goes.

But the realization that by designing the medium on the  $s$  scale we can tailor the (complicated) dispersion relation when  $\lambda \approx s$  naturally led to considering the - in principle simpler - situation of designing media with  $\lambda \gg s$ , as suggested by Kock in 1948. Research in the design of artificial dielectrics did begin, though timidly at first and it is only in the late 90's that the work of Pendry [13, 14, 15] can finally be considered to have launched the field of metamaterials. Pendry pointed out that there is a class of extremely exotic materials which do not exist in nature, but which could be obtained by a careful design of the material structure on the  $s$  scale: media with a negative permittivity and a negative

permeability, or, equivalently, with a negative index of refraction. We will refer to these artificial media as negative index metamaterials or double negative media.

Since all the media we shall be interested in will be periodic, then for the rest of this text we will take both  $s$  and  $d$  equal to the structure period. Thus what we mean by “atom” will depend on the context: in real dielectric media we will mean real atoms, while in artificial metamaterials we will mean meta-atom, or scatterer. We therefore have three important scales,  $a$ ,  $d$  and  $\lambda$ . We will be concerned with the study of electromagnetic materials for which the wavelength is either about the same size as  $d$ , or for which the wavelength is larger, though not by much. *We will be working either within, or bordering on, the intermediate grey area between heterogeneous ( $\lambda < d$ ) and homogeneous ( $\lambda \gg d$ ).* This grey area is characterized by a phenomenon rarely alluded to in established electromagnetics texts, *spatial dispersion*. It appears whenever  $\lambda$  nears either  $d$  or  $a$  (remember we have set  $s = d$ ).

In the first chapter, we will provide a presentation of classic homogenization theory that does not ignore spatial dispersion as is done in most formulations, but on the contrary, highlights it at each step. In this way we shall be building up a physical intuition about its origins, the main factors that shape it, and the conditions under which it can or cannot be ignored. *Whereas classical expositions such as those available in most texts rely on holistic arguments of mainly academic interest, our approach is a constructive one, with an explicit emphasis on **design**.*

The chapter begins with the *microscopic* Maxwell’s equations in a region of space populated with point charges, and defines the fields and parameters that enter into the well known *macroscopic* Maxwell equations and the associated constitutive relations. This first part of the chapter corresponds to the homogenization of *natural* dielectrics. This case is then compared and contrasted with the case of *artificial* metamaterials, whereby we shall be laying the ground work and introducing the main ideas of later chapters. In particular, the conceptually rich Section 1.7 introduces the novel notions of *custom-made effective medium models*, and of *meta-photonic crystals*. The rest of the text draws on the physical ideas introduced in the first chapter by considering first photonic crystals in Chapter 2 and then negative index metamaterials in Chapter 3.

In Chapter 2 we will show how the spatial dispersion relation can be calculated exactly in dielectric structures and in this context it will prove to be an ally, since it will lead to surprising new phenomena. In Chapter 3, however, the situation is reversed due to the fact that the most interesting phenomena and applications rely on non-spatially dispersive structures. Spatial dispersion becomes an enemy and the goal is to find ways to diagnose and avoid it.

# Chapter 1

## Electrons, dielectrics and beyond

### 1.1 Spatial averaging as truncation

This section presents the first step leading from the microscopic description of matter, involving point charges moving in empty space, to the macroscopic description, involving parameters such as the permittivity and the permeability. It is a way of averaging the microscopic fields  $\underline{\mathbf{e}}$  and  $\underline{\mathbf{b}}$  and the distribution of charge and of current  $\underline{\eta}(\mathbf{x}, t)$  and  $\underline{\mathbf{j}}(\mathbf{x}, t)$  to obtain the macroscopic fields  $\underline{\mathbf{E}}$ ,  $\underline{\mathbf{B}}$ ,  $\underline{\mathbf{P}}$  and  $\underline{\mathbf{M}}$  and the macroscopic charge and current densities  $\underline{\rho}(\mathbf{x}, t)$  and  $\underline{\mathbf{J}}(\mathbf{x}, t)$ . The meaning of the underlines is discussed below. If one assumes the medium behaves as a linear time invariant system, then one can define permittivity and permeability tensors through the relations

$$\begin{aligned}\bar{\varepsilon}\mathbf{E} &= \varepsilon_0\mathbf{E} + \mathbf{P} \\ \bar{\mu}\mathbf{B} &= \mu_0(\mathbf{B} + \bar{\mu}\mathbf{M})\end{aligned}$$

where the underlines are absent for reasons discussed below. The formal asymmetry between these definitions has the benefit of leading to a highly symmetrical formulation of the *macroscopic* Maxwell equations, which we will obtain in Section 1.6. It also results in a simple relationship between the permittivity and the permeability on one hand, and the phenomenological parameters of refractive index  $n$  and the impedance  $Z$  on the other hand, which in an isotropic medium are defined as

$$\begin{aligned}n &= \sqrt{\mu\bar{\varepsilon}} \\ Z &= \sqrt{\frac{\mu}{\bar{\varepsilon}}}.\end{aligned}$$

$\mathbf{P}$  and  $\mathbf{M}$  are called the macroscopic polarization and magnetization respectively, and they represent the overall macroscopic effect of the microscopically complicated distribution of *bound* charges and *bound* currents *within* the atoms.

We start, therefore, with the microscopic Maxwell's equations:

$$\begin{aligned}\nabla \cdot \underline{\mathbf{b}} &= 0 & \nabla \times \underline{\mathbf{e}} + \frac{\partial \underline{\mathbf{b}}}{\partial t} &= 0 \\ \nabla \cdot \underline{\mathbf{e}} &= \underline{\eta}/\varepsilon_0 & \frac{1}{\mu_0}\nabla \times \underline{\mathbf{b}} - \varepsilon_0\frac{\partial \underline{\mathbf{e}}}{\partial t} &= \underline{\mathbf{j}}\end{aligned}\tag{1.1}$$

and with an averaging procedure. Various approaches to the averaging have been put forward: spatial, temporal or ensemble averaging. It was argued by Russakoff [16] that only the spatial averaging is truly necessary in order to consistently define the macroscopic *fields*. However, as we shall see, in order to define macroscopic *parameters* such as the relative permittivity, an additional, ensemble average, is

required. The usual macroscopic quantities we are familiar with are therefore *both* spatial and ensemble averages. Since in this section we are concerned only with the consequences of spatial averaging all quantities will be underlined as a reminder that they have *not yet* been ensemble averaged. When the ensemble average is taken, the underlines are removed, and some of the relationships in this section may have to be reconsidered.

The spatial averaging can be seen from two points of view: as a spatial “sliding average”, or as a lowpass filter in reciprocal space, or  $\mathbf{k}$ -space. We explain by considering a generic space and time dependent quantity  $\underline{\xi}(\mathbf{x}, t)$  though in what follows the time is fixed and we will omit it to avoid cluttering the equations.

In the moving average view the macroscopic quantity  $[\underline{\xi}(\mathbf{x})]$  is defined at each point by taking the average of the original  $\underline{\xi}(\mathbf{x})$  over a small region centered at  $\mathbf{x}$ . We write

$$[\underline{\xi}(\mathbf{x})] = \int d^3x' f(\mathbf{x}') \underline{\xi}(\mathbf{x} - \mathbf{x}') \quad (1.2)$$

where the function  $f(\mathbf{x})$  is real, its support is microscopically large but macroscopically small, it contains the origin where it is nonzero, is normalized to 1:  $\oint f(\mathbf{x}) dV = 1$  and is radially symmetric in order to preserve the symmetry properties of  $\underline{\xi}$ :  $f = f(r)$ . This corresponds to a generalized version of our intuitive notion of a sliding average. The form of the integral above is also known as a *convolution* and we can rewrite the equation as

$$[\underline{\xi}(\mathbf{x})] = f(\mathbf{x}) \circ \underline{\xi}(\mathbf{x}) \quad (1.3)$$

where the small circle denotes convolution.

In the lowpass filter view the average is seen as a truncation of the spatial Fourier transform of the quantity  $\underline{\xi}(\mathbf{x})$  whereby all components with  $|\mathbf{k}| > k_0$  are excluded. We apply the convolution theorem to Eq. (1.2):

$$\begin{aligned} [\underline{\xi}(\mathbf{x})] &= f(\mathbf{x}) \circ \underline{\xi}(\mathbf{x}) \\ &= \mathcal{F}^{-1}(\mathcal{F}(f(\mathbf{x}))\mathcal{F}(\underline{\xi}(\mathbf{x}))) \\ &= \mathcal{F}^{-1}(\tilde{f}(\mathbf{k})\tilde{\xi}(\mathbf{k})) \end{aligned} \quad (1.4)$$

where the Fourier transform of  $\underline{\xi}$  is denoted as  $\mathcal{F}(\underline{\xi}(\mathbf{x})) = \tilde{\xi}(\mathbf{k})$  and has the specific form

$$\mathcal{F}(\underline{\xi}(\mathbf{x})) = \int \underline{\xi}(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} d^3x \quad \mathcal{F}^{-1}(\tilde{\xi}(\mathbf{k})) = \frac{1}{(2\pi)^3} \int \tilde{\xi}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} d^3k. \quad (1.5)$$

It is clear that  $\tilde{f}(\mathbf{k})$  plays the role of a filter on the frequency components of  $\underline{\xi}(\mathbf{x})$ . In our case we want to remove the microscopic features of  $\underline{\xi}$  which is equivalent to removing its high frequency components.  $\tilde{f}(\mathbf{k})$  must then be a low pass filter, a point of view emphasized by Robinson [17]. Moreover, from the well known general properties of the Fourier transform we know that if  $f$  is well-behaved, normalized to 1 and symmetrical, then  $\tilde{f} \approx 1$  and  $\nabla_{\mathbf{k}} \tilde{f} \approx \mathbf{0}$  in some neighborhood of  $\mathbf{k} = \mathbf{0}$  and the approximations can be made arbitrarily good in the right neighborhood. The importance of these facts will become clear below.

Since the convolution commutes with space and time differentiation, when we apply the brackets to Maxwell's equations we obtain directly

$$\begin{aligned} \nabla \cdot [\mathbf{b}] &= 0 & \nabla \times [\mathbf{e}] + \frac{\partial [\mathbf{b}]}{\partial t} &= 0 \\ \nabla \cdot [\mathbf{e}] &= [\eta]/\varepsilon_0 & \frac{1}{\mu_0} \nabla \times [\mathbf{b}] - \varepsilon_0 \frac{\partial [\mathbf{e}]}{\partial t} &= [\mathbf{j}] \end{aligned} \quad (1.6)$$

The macroscopic fields  $\underline{\mathbf{E}}$  and  $\underline{\mathbf{B}}$  are then defined as  $\underline{\mathbf{E}} = [\underline{\mathbf{e}}]$  and  $\underline{\mathbf{B}} = [\underline{\mathbf{b}}]$  and in order to obtain the macroscopic equations we need to write out the average charge and current densities,  $[\underline{\eta}]$  and  $[\underline{\mathbf{j}}]$ . We will write out only the charge density in detail.

We now make two simplifying assumptions.

The first, and relatively innocuous one, is that the medium as a whole is neutral. This assumption is due to the fact that electromagnetic interactions are so strong compared to the masses of the objects involved that electrical charges will quickly pair up, such that even over microscopic distances (say, several unit cells) most media of interest are all but almost perfectly neutral.

The second assumption, which we call the *atomic assumption* (the medium is composed of *stable atoms*), is far more consequential and deserves a serious discussion. However since it is not required in this section or the next, and its central importance and impact will become clear only later, we leave this discussion to the section on “Ensemble averaging” below.

For the time being we simply group charges in the medium in atoms. Thus the whole charge distribution of the medium can be considered as a sum over the charge distributions of individual atoms  $\underline{\eta} = \sum_n \underline{\eta}_n(\mathbf{x} - \mathbf{x}_n)$ . In other words there are no free or surplus charges. Note that the individual atoms need not be neutral, only collectively.

We now apply Eq. (1.4) to  $[\underline{\eta}(\mathbf{x})]$ :

$$\begin{aligned} [\underline{\eta}(\mathbf{x})] &= \mathcal{F}^{-1}(\tilde{f}(\mathbf{k})\tilde{\eta}(\mathbf{k})) \\ &= \mathcal{F}^{-1}(\tilde{f}(\mathbf{k})\sum_n \tilde{\eta}_n(\mathbf{k})) \\ &= \sum_n \mathcal{F}^{-1}(\tilde{f}(\mathbf{k})\tilde{\eta}_n(\mathbf{k})). \end{aligned} \quad (1.7)$$

Since we have seen that the multiplication by  $\tilde{f}(\mathbf{k})$  has the role of a filter which passes only frequency components with  $\mathbf{k}$  close to  $\mathbf{0}$  it is reasonable to attempt to represent  $\tilde{\eta}_n(\mathbf{k})$  as a Taylor series around  $\mathbf{k} = \mathbf{0}$  and hope that we may only need to keep a few terms. We have

$$\begin{aligned} \tilde{\eta}_n(\mathbf{k}) &= \tilde{\eta}_n(\mathbf{k})|_{\mathbf{k}=\mathbf{0}} + \mathbf{k} \cdot \nabla_{\mathbf{k}} \tilde{\eta}_n(\mathbf{k})|_{\mathbf{k}=\mathbf{0}} + \mathbf{k} \cdot \tilde{\mathbf{R}}_n(\mathbf{k}) \\ &= \tilde{\eta}_n(\mathbf{k})|_{\mathbf{k}=\mathbf{0}} + \mathbf{k} \cdot \left( \nabla_{\mathbf{k}} \tilde{\eta}_n(\mathbf{k})|_{\mathbf{k}=\mathbf{0}} + \tilde{\mathbf{R}}_n(\mathbf{k}) \right) \end{aligned} \quad (1.8)$$

where the first term is easily seen as the total net charge  $q_n$ , the second term is the dipolar term, while the  $\mathbf{k}$  dependent term  $\tilde{\mathbf{R}}_n$  collects all the rest of the higher order multipolar terms, which we hope are small; the above equation is therefore not an approximation but a true equality. Before going any further, let us try to get a feel for the physical meaning of the quantity in parentheses. Let us assume the  $\tilde{\mathbf{R}}_n$  term is negligible, and write out the gradient term. The interpretation is facilitated if we take the Fourier transform around  $\mathbf{x}_n$ . From Eq. (1.5) we have

$$\begin{aligned} \nabla_{\mathbf{k}} \tilde{\eta}_n(\mathbf{k})|_{\mathbf{k}=\mathbf{0}} &= \int \underline{\eta}_n(\mathbf{x}) e^{-i\mathbf{k} \cdot (\mathbf{x} - \mathbf{x}_n)} (-i(\mathbf{x} - \mathbf{x}_n)) d^3x \Big|_{\mathbf{k}=\mathbf{0}} \\ &= -i \int (\mathbf{x} - \mathbf{x}_n) \underline{\eta}_n(\mathbf{x}) d^3x \\ &= -i \underline{\mathbf{p}}_n \end{aligned}$$

where we have introduced  $\underline{\mathbf{p}}_n$ , the *equivalent point dipole moment* of the atom, in the limit of  $\mathbf{k} \rightarrow \mathbf{0}$ . Assuming  $\tilde{\mathbf{R}}_n$  to be negligible is therefore equivalent to what in the Introduction was referred to as the *dipolar approximation*. We now define the *generalized electric moment*

$$\tilde{\underline{\mathbf{p}}}_n(\mathbf{k}) = i \nabla_{\mathbf{k}} \tilde{\eta}_n(\mathbf{k})|_{\mathbf{k}=\mathbf{0}} + i \tilde{\mathbf{R}}_n(\mathbf{k}) = \underline{\mathbf{p}}_n + i \tilde{\mathbf{R}}_n(\mathbf{k}) \quad (1.9)$$

Note that the charge distribution of a given atom at any given time  $t$  need not be symmetrical, even when there is no external field applied. The dipolar term  $\nabla_{\mathbf{k}}\tilde{\eta}_n(\mathbf{k})|_{\mathbf{k}=\mathbf{0}}$  therefore need not be zero. Summarizing:

$$\tilde{\eta}_n(\mathbf{k}) = \underline{q}_n - i\mathbf{k} \cdot \tilde{\mathbf{p}}_n(\mathbf{k}) \quad (1.10)$$

We now write the Taylor expansion of  $\tilde{f}(\mathbf{k})$ :

$$\begin{aligned} \tilde{f}(\mathbf{k}) &= 1 + \mathbf{k} \cdot \nabla_{\mathbf{k}}\tilde{f}_n(\mathbf{k})|_{\mathbf{k}=\mathbf{0}} + \mathbf{k} \cdot \tilde{\mathbf{R}}_f(\mathbf{k}) \\ &= 1 + \mathbf{k} \cdot \tilde{\mathbf{R}}_f(\mathbf{k}) \end{aligned} \quad (1.11)$$

where we have used the symmetry of  $f$  as mentioned above. The remainder terms  $\tilde{\mathbf{R}}_n$  and  $\tilde{\mathbf{R}}_f$  are by definition null at the origin:  $\tilde{\mathbf{R}}_n(\mathbf{k})|_{\mathbf{k}=\mathbf{0}} = \tilde{\mathbf{R}}_f(\mathbf{k})|_{\mathbf{k}=\mathbf{0}} = 0$  and continuous there. Moreover it is important to note that the  $\mathbf{k}$  dependent rest terms  $\tilde{\mathbf{R}}_n(\mathbf{k})$  and  $\tilde{\mathbf{R}}_f(\mathbf{k})$  are not on the same footing, from a physical point of view. While the  $\tilde{\mathbf{R}}_n$  term is related to the microscopic configuration of the medium at the given time, the  $\tilde{\mathbf{R}}_f$  term is related to the properties of the function  $f$  which is a mathematical construct that we can choose as suits our needs. We can therefore constrain  $f$  to be such that  $\tilde{\mathbf{R}}_f(\mathbf{k})$  be arbitrarily small compared to the other terms in Eq. (1.7). We shall see what this constraint entails in the next section.

The generic term of Eq. (1.7) takes the form:

$$\begin{aligned} \mathcal{F}^{-1}(\tilde{f}(\mathbf{k})\tilde{\eta}_n(\mathbf{k})) &= \mathcal{F}^{-1}\left(\underline{q}_n\tilde{f}(\mathbf{k}) - i\mathbf{k} \cdot \tilde{\mathbf{p}}_n(\mathbf{k})\tilde{f}(\mathbf{k})\right) \\ &= \underline{q}_n\delta(\mathbf{x} - \mathbf{x}_n) \circ f(\mathbf{x}) - \delta'(\mathbf{x} - \mathbf{x}_n) \circ \mathcal{F}^{-1}\left(\tilde{f}(\mathbf{k})\tilde{\mathbf{p}}_n(\mathbf{k})\right) \\ &= \underline{q}_nf(\mathbf{x} - \mathbf{x}_n) - \nabla \cdot (f(\mathbf{x}) \circ \underline{\mathbf{p}}_n(\mathbf{x} - \mathbf{x}_n)) \end{aligned} \quad (1.12)$$

What is the physical meaning of this result ? For the interpretation of the first term it is sufficient to look at the definition of the smoothing process, Eq. (1.3). We have

$$\underline{q}_nf(\mathbf{x} - \mathbf{x}_n) = \underline{q}_n\delta(\mathbf{x} - \mathbf{x}_n) \circ f(\mathbf{x}) = [\underline{q}_n\delta(\mathbf{x} - \mathbf{x}_n)]$$

so from a macroscopic point of view the net charge of the atom is seen as if the atom consisted of a single point charge  $\underline{q}_n$  localized at the center of the atom,  $\mathbf{x}_n$ . Even though the actual charge distribution within the atom may be complicated, with many individual point charges spread over a finite volume, the smoothing process wipes out all the detailed information leaving only two aspects: the net charge  $\underline{q}_n$  and the mean position  $\mathbf{x}_n$ .

The interpretation of the second term is not quite as straightforward. We write out the position dependent polarization vector:

$$\begin{aligned} \underline{\mathbf{p}}_n(\mathbf{x} - \mathbf{x}_n) &= \mathcal{F}^{-1}(\tilde{\mathbf{p}}_n(\mathbf{k})) \\ &= \mathcal{F}^{-1}(\underline{\mathbf{p}}_n + i\tilde{\mathbf{R}}_n(\mathbf{k})) \\ &= \underline{\mathbf{p}}_n\delta(\mathbf{x} - \mathbf{x}_n) + i\underline{\mathbf{R}}_n(\mathbf{x} - \mathbf{x}_n) \end{aligned} \quad (1.13)$$

Note that since  $\tilde{\mathbf{R}}_n(\mathbf{k})$  is null at the origin by definition, this means that  $\underline{\mathbf{R}}_n(\mathbf{x} - \mathbf{x}_n)$  integrates to zero over all space. The electric polarization of the atom therefore has two components. One of them is singular, the ideal dipole localized at the center of the atom, while the other is regular and decreases to zero quickly with distance. When the  $\mathbf{k}$  dependence of  $\tilde{\mathbf{p}}_n(\mathbf{k})$  is negligible the homogenization process reduces the atom to a smoothed version of a point dipole  $\underline{\mathbf{p}}_n$  localized at  $\mathbf{x}_n$ . In the more general case, however, we must write  $\nabla \cdot [\underline{\mathbf{p}}_n(\mathbf{x} - \mathbf{x}_n)]$  where the electric moment of the atom cannot be idealized as a

point dipole but is smeared out, in a sense, over a finite region of space. We now sum over all the atoms to obtain the total smoothed charge density

$$[\underline{\eta}(\mathbf{x})] = -\nabla \cdot \underline{\mathbf{P}}(\mathbf{x}) \quad (1.14)$$

where the macroscopic polarization  $\underline{\mathbf{P}}(\mathbf{x})$  is defined

$$\underline{\mathbf{P}}(\mathbf{x}) = \left[ \sum_n \underline{\mathbf{p}}_n(\mathbf{x} - \mathbf{x}_n) \right] \quad (1.15)$$

which in the limit of  $\mathbf{k} \rightarrow \Gamma$  (the origin in reciprocal space) becomes

$$\underline{\mathbf{P}}^\Gamma(\mathbf{x}) = \left[ \sum_n \underline{\mathbf{p}}_n \delta(\mathbf{x} - \mathbf{x}_n) \right].$$

It is important to point out that the main shortcoming of the results of this section is the fact that the spatial averaging is done at *one instant in time*. Consequently, none of the quantities defined in this section (marked with underlines) contain any information about the actual properties of the medium, about its behavior, or about its response to an applied external stimulus. All underlined quantities only give information about the *state* of the system at one instant, not about its dynamics, or evolution. In order to obtain a new set of quantities which do contain such information it is necessary to eliminate the corrupting effects of the random statistical microscopic fluctuations of the medium. This requires an ensemble averaging procedure, which is discussed in the next section.

Before moving on, note that the way we have defined the macroscopic polarization differs from the way it is defined in texts such as Jackson's [3]. In our case, the  $\underline{\mathbf{P}}(\mathbf{x})$  includes the quadrupolar and higher terms of the multipolar expansion of the microscopic charge distribution  $\eta$ , such that Eq. (1.14) is an exact equality in neutral media with no free charges. In Jackson's definition the macroscopic polarization is defined as what we would call the macroscopic *dipolar* polarization and in that case the Eq. (1.14) would be only an approximation. We have here a first glimpse of one of the two main physical origins of the phenomenon of spatial dispersion: ***the proximity of  $\lambda$  and  $a$***  (the other, the proximity of  $\lambda$  and  $d$ , is discussed below in Sections 1.3 and 1.4). When higher multipolar terms of the polarization are not negligible, we say the medium exhibits spatial dispersion, which means that its response at a given point depends not only on the instantaneous *intensity* of an applied field at that point but also on its *instantaneous phase and direction of propagation at that point*. However, since for the time being we cannot rigorously speak of "response to applied fields" for the reason explained above, we leave this discussion to Section 1.4 where we introduce the susceptibility.

## 1.2 Ensemble averaging

If one wants to characterize a given system, then one must find a way to specify some kind of correlation between stimuli and responses. In the case of linear systems this connection is encapsulated in the *transfer function*. In the case of the dielectric media which we want to study, the electric divergence equation  $\nabla \cdot \underline{\mathbf{e}} = \underline{\eta}/\epsilon_0$  indicates that there is a connection between the electric field and the charge distribution. It therefore seems reasonable to attempt to describe the medium in terms of the response of the charge distribution to an external applied field. Since we have assumed the medium is neutral and has no free charges, then the most important term in the charge distribution is the polarization, see Eq. (1.10). A way to characterize the system would be to specify a relationship between the macroscopic electric field and the polarization of the medium. In this sense the electric field would be the input signal

or stimulus of our linear medium, and the polarization the output, or response. They would be related by a transfer function. However, in order for the transfer function to be a useful tool, the system must be not only linear, but also *time invariant*.

A glance over the arguments of the previous section makes it clear that the model is not time invariant if only the spatial averaging is used. In fact the quantities defined above contain no information about the *behavior* of the medium, or its response to a given stimulus, but only about the *configuration* at a given fixed time  $t$ . It is not possible to make any clear correlation between these quantities and external applied fields or even between these quantities and general properties of the medium itself such as its periodicity. The forces acting between charges in the medium and the electric and magnetic fields (due to sources both external and internal to the medium) have not been accounted for, not even approximately or indirectly. In fact we have only made use of the fact that there is a region with a (singular) distribution of charges and a continuous field, that high spatial frequencies in *both* are unobservable, and of the existence (though not the nature) of an interaction between field and charges. The force between charges need not even be central, for instance. If the Lorentz force did not exist, the macroscopic polarization field of Eq. (1.15) would still be a well defined quantity. The macroscopic fields as defined above are therefore of no use in describing the *behavior* of the medium. For instance, one undesirable feature of the macroscopic quantities as defined above is that even in a structure composed of atoms arranged periodically, the microscopic quantities would not be periodical. In fact, their spatial Fourier spectrum would contain a certain amount of noise (of thermal origin). The arguments of this section, however, are general, and a detailed discussion of periodic media is left to the next section.

A means of rendering the medium time invariant is required. At first it might seem like a time averaging may be required but this is not quite the correct answer. Firstly, many oscillating quantities including the electric field would average to zero rendering the average useless. Also, averaging over periods of time shorter than a full oscillation does not render the system time invariant either due to the corrupting effect of correlations in time between the positions of particles. A way to eliminate these effects and to obtain a truly time invariant description is the *ensemble average*. It could be seen as a very special type of time averaging which excludes correlation effects.

Going into the arcane and subtle aspects of the precise nature of the averaging procedure required is beyond the scope of this chapter. We will therefore postulate that there exists a procedure, called ensemble averaging, that has the following key features:

1. Makes it possible to treat the structure as a linear time invariant system, in particular allowing us to define inputs, outputs and transfer functions. Transfer functions are particularly useful since it is they that contain the macroscopic time invariant information we seek about the *behavior* (as opposed to simply the *state*) of the system.
2. Incorporates the symmetry properties of the medium into the ensemble distribution function by eliminating random spatial fluctuations so as to bring the structure within the reach of Bloch's theorem. In other words the ensemble averaged charge distribution of a periodic medium must be periodic.
3. Includes the effect of the interactions between the particles and the microscopic fields as well as the interactions (electromagnetic *and* quantum) among the particles themselves, on the sub-atomic scale. Both interactions within the same atom and with particles in neighboring atoms must be accounted for. These effects would be reflected in the ensemble distribution function.

An introduction to the details of ensemble averaging is available in Chapters 6 and 7 of Ref. [17] and references therein while a quantum mechanical treatment can be found in Refs. [18, 19]. The consequences of point 2 will be discussed in the next section. The rest of this section is devoted to

a discussion of the 3'd point, particularly its close connection to the *atomic assumption* which was introduced without justification in the previous section.

The atomic assumption consists of assuming the medium is composed of atoms, where an atom is understood as a stable collection of charges with an existence independent of the material of which it is a part. The word “stable” has a particular significance in a text such as this, which explicitly limits itself to purely classical considerations, because stable atoms do not exist in classical electromagnetism. As a matter of fact, this failure of Maxwell’s theory was one of the major motivations for the later development of quantum theory. We therefore have two options. Either we stick to purely classical considerations, or we attempt to include, even if only phenomenologically, quantum effects. If we want to remain strictly within the classical domain, then, since atoms are outside its scope, no study of lossless dielectric media is possible, and we are limited to the study of collisionless plasmas, since they can be treated completely using only Maxwell’s equations [20]. Otherwise, we must include quantum effects, even if only phenomenologically, through an ad-hoc relation introduced at the appropriate moment.

Obviously, we must choose the latter option. This immediately implies that, by definition, our ensemble average is a procedure which does not leave Maxwell’s equations unchanged. The equations are not invariant with respect to ensemble averaging. We therefore emphasize that one should not expect the divergence equation  $\nabla \cdot \underline{\mathbf{e}} = \underline{\eta}/\varepsilon_0$  to suffice in describing the interaction between the microscopic field  $\mathbf{e}$  and the charge distribution  $\eta$ , even when magnetic effects are negligible. An independent equation containing additional information is required. This equation is the equation which establishes the relationship between the ensemble averaged microscopic electric field and the ensemble averaged atomic polarization vector:

$$\mathbf{p}_n(\mathbf{x} - \mathbf{x}_n) = \varepsilon_0 \gamma_n^e(\mathbf{x} - \mathbf{x}_n) \mathbf{e}(\mathbf{x}) \quad (1.16)$$

Note that this equation would be meaningless within the framework of the previous section, where the ensemble average had *not* been taken, because we want  $\gamma_n^e$  to be a time invariant *property* of the internal structure of atom  $n$ , not just the (randomly fluctuating in time) factor of proportionality between the polarization and the electric field at some time  $t$ . Thus each atom will be treated as a black box, and the interactions between the particles inside will be inaccessible to us. Some of these interactions may be electrical in nature, but some will clearly not be, and “for all we are concerned atoms could equally well be held together with glue and rubber bands”, as Robinson correctly observes [17]. Thus the internal dynamics of the atom, as well as the modification in its *internal* dynamics due to the presence of neighboring atoms’ electron clouds, are encapsulated within the parameter  $\gamma_n^e$  such that the total polarization can be written as a sum over the atoms, and Eq. (1.15) remains valid (though only formally) without the underlines:

$$\mathbf{P}(\mathbf{x}) = \left[ \sum_n \mathbf{p}_n(\mathbf{x} - \mathbf{x}_n) \right] = [\mathbf{p}(\mathbf{x})] = [\varepsilon_0 \gamma^e(\mathbf{x}) \mathbf{e}(\mathbf{x})] \quad (1.17)$$

where

$$\mathbf{p}(\mathbf{x}) = \sum_n \mathbf{p}_n(\mathbf{x} - \mathbf{x}_n) \quad \text{and} \quad \gamma^e(\mathbf{x}) = \sum_n \gamma_n^e(\mathbf{x} - \mathbf{x}_n)$$

We see that the effect of taking an ensemble average in an atomic medium is to replace the two fundamental interacting quantities  $\underline{\mathbf{e}}$  and  $\underline{\eta}$  with the quantities  $\mathbf{e}$  and  $\mathbf{p}$ . The charges are displaced by an electric field according to Eq. (1.16) while the polarization charge density  $\eta(\mathbf{x}) = -\nabla \cdot \mathbf{p}(\mathbf{x})$  produces

an electric field according to Coulomb's law,

$$\mathbf{e}_{\text{pol}}(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \nabla_{\mathbf{x}} \int \frac{\nabla_{\mathbf{x}'} \cdot \mathbf{p}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (1.18)$$

The net result of ensemble averaging is therefore that it is now possible to define a time independent polarizability, such that the time dependence of the polarization  $\mathbf{p}$  (and therefore the charge distribution  $\eta$ ) is tied directly to the time dependence of the electric field. The two equations  $\nabla \cdot \underline{\mathbf{e}}(\mathbf{x}, t) = \underline{\eta}(\mathbf{x}, t)/\epsilon_0$  and  $\mathbf{F}_j(t) = q_j \underline{\mathbf{e}}(\mathbf{x}_j, t)$  have been replaced by equations (1.18) and (1.16) respectively. There are no more forces, and no more point charges, only two position and time dependent continuous fields,  $\mathbf{e}$  and  $\mathbf{p}$  whose time dependence is synchronous if the atomic polarizability is real. The total electric field then satisfies the equation

$$\mathbf{e}(\mathbf{x}) = \mathbf{E}_{\text{ext}}(\mathbf{x}) + \mathbf{e}_{\text{pol}}(\mathbf{x}) = \mathbf{E}_{\text{ext}}(\mathbf{x}) + \frac{1}{4\pi\epsilon_0} \nabla_{\mathbf{x}} \int \frac{\nabla_{\mathbf{x}'} \cdot \gamma^e(\mathbf{x}') \mathbf{e}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x'$$

The order in which the ensemble average and the spatial average are taken is indifferent for our purposes. However this may not be the case if one were undertaking a study of the statistical fluctuations of the system. Such a study would require a closer look at the details of the ensemble averaging process, which is beyond the scope of this work.

In the above we have assumed the polarizability is a scalar, meaning that the atoms are perfectly isotropic. However, a diagonal dyadic polarizability can be included with minimal effort. As long as the polarizability is diagonal the equations above, as well as those of the rest of this chapter retain their form.

We note that Eq. (1.17) indicates that if the smoothing function  $f$  is much wider than the support of the atomic polarizability  $\gamma_n^e(\mathbf{x} - \mathbf{x}_n)$  then the result of the smoothing is equivalent to the smoothing of a point dipole with dipole moment  $\mathbf{p}_n = \gamma_n^e(\mathbf{x} - \mathbf{x}_n) \mathbf{e}(\mathbf{x})$ . However, if the smoothing function is not sufficiently large then the point dipole approximation is not satisfactory, and the fact that the atoms are not negligibly small but have a finite size begins to make itself felt. In this case the macroscopic polarization has contributions from higher order multipolar terms, which are, in essence, manifestations of the internal structure of atoms. When noticeable, these contributions lead to a dependence of the polarization on the phase and direction of propagation of the electric field, and therefore, to spatial dispersion. This contribution, which is due to the nearness of the  $a$  and  $\lambda$  scales, may be termed **atomic spatial dispersion** to distinguish it from the **lattice spatial dispersion** which is due to the nearness of the  $d$  and  $\lambda$  scales. The latter is discussed in the next two sections.

One may wonder, however, why we would ever want to choose an averaging volume so small that it starts to be comparable to the size of the atoms. Until now there has been no indication that there might be some upper limit or constraint to the size of the averaging volume. But we will see in the next section that this constraint is imposed by the size of the wavelength of the macroscopic electric field inside the medium. Consequently, for short wavelengths, the averaging volume may have to be reduced to the point where the internal structure of the atoms starts to play a role, through the quadrupolar and higher terms in the macroscopic polarization.

The second of the three properties we have postulated for the ensemble average brings periodic media within the reach of our theory, and the next section is devoted to exploring the consequences.

## 1.3 Periodic media

In this section we explore the consequences of the fact that the truncation (spatial averaging) discussed above must be *simultaneously* applied to the electric field and the charge distribution, since they appear in the same equation, the divergence equation. We also take a closer look at the Taylor expansions of  $\tilde{\eta}(\mathbf{k})$  and  $\tilde{f}(\mathbf{k})$  in relation to the behavior of the molecular polarization vector  $\mathbf{p}_n(\mathbf{x} - \mathbf{x}_n)$ .

The Bloch theorem tells us that a wave at a single temporal frequency  $\omega$  propagating in a periodic lattice has a space dependent part of the form

$$\mathbf{e}(\mathbf{x}) = \mathbf{U}(\mathbf{x})e^{i\mathbf{k}_B\mathbf{x}} \quad (1.19)$$

where  $\mathbf{U}(\mathbf{x})$  is a function with the periodicity of the lattice and where the time dependence has been ignored as above. If we consider a 1D lattice of period  $a$  then the function  $\mathbf{U}(\mathbf{x})$  contains spatial frequencies no lower than  $K = 2\pi/a$  (without counting the null frequency). The local field  $\mathbf{e}$  which results when a wave at a *single* temporal frequency  $\omega$  propagates in a periodic lattice contains *more than one* spatial frequency; specifically it contains the frequencies:  $\mathbf{k}_B + n\mathbf{K}$ ,  $n \in \mathbb{Z}$ . However in a homogeneous material only one spatial frequency is present. The smoothing procedure must therefore remove the harmonics which are due to the periodic structure, namely  $\mathbf{k}_B + n\mathbf{K}$ ,  $n \in \mathbb{Z}^*$  leaving only the Bloch phase harmonic  $\mathbf{k}_B$ . The cutoff beyond which the filter  $\tilde{f}(\mathbf{k})$  must strongly attenuate the spatial Fourier components must therefore be below  $K$  (actually  $K/2$ ). The definition of “strong attenuation” depends on the precision we require of our effective medium model.

More precisely, the region in reciprocal space that the lowpass filter may admit is called *the first Brillouin zone*. For a definition and detailed discussion see Refs. [21, 22]. The filter does not have to allow all  $\mathbf{k}$  in the Zone but only those  $\mathbf{k}$  necessary for the description of the problem at hand (only  $\mathbf{k}_B$  for a single plane wave propagating in an infinite medium), while of course satisfying all the rest of the restrictions already imposed on it (symmetry, normalization, smoothness).

Since this filter is applied to both sides of the electric divergence equation simultaneously this means that the frequencies filtered out of  $\mathbf{E}(\mathbf{k})$  must also be filtered out of  $\tilde{\eta}(\mathbf{k})$ . The cutoffs for the two are the same. The  $\mathbf{k}$ 's of the previous section are therefore identified with the wavevectors (spatial harmonics) of the electric field. We can now take a closer look at the physical implications of the mathematical results of the previous section. In what follows we will refer to  $K/2$  or the “edge of the BZ” interchangeably, with the first being the 1D version of the second, and useful for purposes of illustration. We begin by considering the various options for the choice of filter function  $\tilde{f}(\mathbf{k})$ .

A reasonable first try for  $\tilde{f}(\mathbf{k})$  is the Gaussian function, whose Fourier transform is also a Gaussian. This function removes all frequencies outside the first BZ, but also partially filters out some frequencies just inside it. This makes no difference, of course, if the wavevector  $\mathbf{k}_B$  in Eq. (1.19) is very close to  $\Gamma$  such that  $\tilde{f}(\mathbf{k}_B) \approx 1$ . However, if  $\mathbf{k}_B$  is closer to the BZ edge a Gaussian will no longer give a good representation of the medium. A possible improvement is, for instance, a function of the form (in 1D)

$$\text{real}(\tilde{f}(\mathbf{k})) = \exp\left(-\frac{1}{2\nu}\left(\frac{2k}{K}\right)^{2\nu}\right), \quad (1.20)$$

see Fig. 1.1. When  $\nu = 1$  this is a Gaussian, but by increasing it the function can be made arbitrarily close to 1 over the interval  $k_B \in (-K/2, K/2)$ . Its imaginary part is defined by the requirement that  $f$  be real. Thus we can always find a suitable function  $f$  such that the variation of the field at the  $a$  scale is averaged out, while the variation at the  $\lambda_B = 2\pi/k_B$  scale is preserved to arbitrary accuracy. The price paid, however, is that of using filtering functions  $\tilde{f}(\mathbf{k})$  with a sharper and sharper cutoff. The effect this has on the spatial smoothing function  $f$  can be seen in Fig. 1.1 where we compare the absolute values

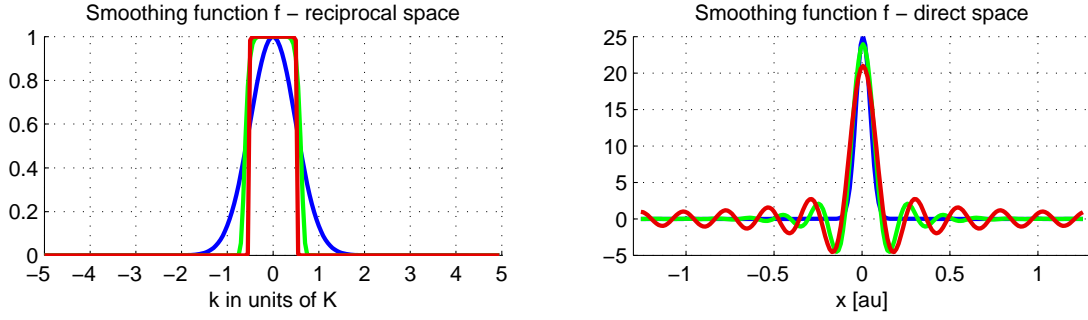


Figure 1.1: Comparison of smoothing functions  $f$  with sharper and sharper cutoffs, in direct and reciprocal space. The blue curves are Gaussians while the green and red curves correspond to  $\nu$  of 5 and 30 respectively. The sharper cutoff leads to broader smoothing functions in direct space.

of  $f$  and  $\tilde{f}$  for  $\nu = 1$  (Gaussian) and  $\nu$  very large (square filter).

When the filter function  $\tilde{f}$  is Gaussian, then for a cutoff around  $K/2$  the averaging volume is of a size around  $a$ . However, as the wavelength approaches  $a$ , the filter must become sharper, the averaging volume becomes considerably larger than  $a$ . Firstly, this implies that the value of a generic macroscopic quantity  $[\xi]$  at a point  $\mathbf{x}$  depends on the microscopic details of  $\xi$  two, three, or more unit cells away from  $x$ . Secondly, it means that the averaging volume may then become comparable to the wavelength. This in turn means that the value of  $[\xi]$  at the point  $\mathbf{x}$  depends on values of the electric field *other than* the value of the electric field at  $\mathbf{x}$ .

This is not an intuitive phenomenon. One way of approaching it is to think of temporal rather than spatial frequencies. The job of the function  $\tilde{f}$  is to tell apart two frequencies,  $k_B$  and  $K/2$ . Imagine the two as “sounds” starting in phase. If the two frequencies are quite different, it takes a small fraction of a wavelength for the two signals to go out of phase. Consequently it takes only a short time to tell them apart, or in the case of spatial frequencies, a short distance. When the two frequencies are closer, however, they will stay in phase for longer, perhaps several wavelengths. Consequently it takes a longer time to tell them apart, or in the case of spatial frequencies, a longer distance. This is why a sharper filter  $\tilde{f}$  in  $k$ -space requires a more sprawled out smoothing function  $f$  in real space.

The size of the smoothing volume, in turn, is important because it is responsible for the macroscopically non-local behavior of the medium. In fact, the spatial averaging implicitly results in non-local macroscopic quantities in the sense that the macroscopic polarization at any given point does not depend only on the electric field at that point but over a whole region surrounding it, the smoothing volume  $f$ . In most cases of interest, however, the wavelength within the medium is much larger than this volume, the electric field being roughly constant over it. This in effect hides the non-locality of the macroscopic description, making the medium response seem local.

As the wavelength of the external field becomes smaller, the wavevector moves farther away from  $\Gamma$ , the origin of the reciprocal space. This in turn requires smoothing functions with sharper cutoffs in reciprocal space, corresponding to broader averaging volumes in direct space. There comes a point where the variation of the macroscopic field over the size of the averaging volume is non negligible, in which case the macroscopic polarization field depends not only on the intensity of the applied electric field, but also on its phase and direction of propagation. The phase and direction of propagation in turn are functions of the wavevector  $\mathbf{k}$  of the EM wave in the medium, a situation which is known under the name of spatial dispersion. Nonlocality and spatial dispersion are therefore seen to be two sides of the same coin.

Yet another way to think about it is to consider the product in  $k$ -space between a very wide, smooth function ( $\tilde{\eta}_n(\mathbf{k})$ ) and an almost square filter ( $\tilde{f}(\mathbf{k})$  of Eq. (1.20) with a very large  $\nu$ ). By multiplying

them one could say that the sharpness of the filter “introduces” higher frequencies into the original signal, with the result that while the original  $\tilde{\eta}(\mathbf{k})$  may have been almost constant (independent of  $\mathbf{k}$  over the region of interest), the smoothed version may have a very strong  $\mathbf{k}$  dependence, or in other words, a strong *spatial dispersion*. The need for sharp filtering could then be said to “introduce” spatial dispersion into physical quantities such as the charge density  $\eta_n$ .

This duality is the direct consequence of the appearance in the same equation (the electric divergence equation), of the field  $\mathbf{E}$  and the charge distribution. When the smoothing is applied, this equation imposes a constraint in that one must keep the macroscopic oscillations of the *field*, but filter out the microscopic oscillations of the *charge density*. As these two spatial frequencies become closer the smoothing volume  $f$  becomes wider. These ideas are clarified and rendered more explicit in the next section where we introduce the *susceptibility*.

But before moving on we must clarify an aspect related to the interactions between neighboring atoms. In real dielectric media, composed of atoms with overlapping electron clouds, the interactions between neighboring atoms are governed by both quantum mechanics and classical electromagnetics. The black box inside which we hide the quantum effects needs to include not only interactions within atoms, but also with neighboring atoms, due to the non-negligible electron cloud overlap. The atomic polarization field,  $\mathbf{p}_n(\mathbf{x})$  will therefore overlap with those of neighboring atoms, which means that the electromagnetic interaction between them cannot be expressed in terms of multipole expansions, due to the fact that these are valid only *outside* a given charge distribution. This has the potential to complicate considerably the analysis of the interaction between the  $\mathbf{e}$  and  $\mathbf{p}$  vectors. However, since the goal of this work is to construct an analytic model of metamaterials obeying the *macroscopic* Maxwell equations, we are only interested in the case of clearly separated atoms whose interactions are mediated by multipolar fields. In this case the charge distributions (the scattering objects) do not overlap, and they interact with each other through both a). dipolar fields, and b). higher multipolar fields. The latter we lump together under the label “near fields” since they attenuate rapidly with distance.

## 1.4 Polarizability, susceptibility – electric

Dielectric media are linear systems when the field intensity is not too large. As with any linear system, there are a number of degrees of freedom, some of which are of interest and some of which are either not of interest or in any case inaccessible or unobservable. In the case of dielectric media, which are composed of extremely large numbers of extremely small particles, the unobservable parameters are those related to the microscopic degrees of freedom of the particles. The external description of dielectric media, also known as the macroscopic description, therefore must be obtained by an averaging over a large number of inaccessible microscopic degrees of freedom. Distances on this microscopic scale are therefore meaningless from the point of view of the macroscopic description. This leads to a spatial smearing which implies that the observable properties of the material at a given position  $x$  are in fact the result of a large number of individual microscopic interactions over a whole region surrounding point  $x$ . What happens at  $x$ , therefore, depends to some extent on the conditions prevailing in a certain volume  $v$  *surrounding*  $x$ . This is what we refer to below as the “electromagnetic neighborhood” of  $x$ . If we apply, therefore, some position dependent stimulus to the material, the macroscopic response of the medium at  $x$  will depend on the stimulus over the whole volume  $v$ . If the stimulus is the macroscopic electric field noted  $\mathbf{E}$  and the response the polarization  $\mathbf{P}$  then we can define the transfer function of the medium through the relation

$$\mathbf{P}(\mathbf{x}) = \varepsilon_0 \chi^e(\mathbf{x}) \circ \mathbf{E}(\mathbf{x}) = \varepsilon_0 \chi^e(\mathbf{x}) \circ [\mathbf{e}(\mathbf{x})]. \quad (1.21)$$

The microscopic field inside the structure is in turn composed of two contributions, one external, and one internal, due to the polarization charge density in the medium, given by  $\mathbf{e}_{\text{pol}}$  of Eq. (1.18):

$$\mathbf{e}(\mathbf{x}) = \mathbf{E}_{\text{ext}}(\mathbf{x}) + \mathbf{e}_{\text{pol}}(\mathbf{x}).$$

The nonlocal nature of the transfer function might seem peculiar given that the interaction between charged particles and electromagnetic fields is local, according to the electrostatic Lorentz force equation  $\mathbf{F}_n = q_n \mathbf{E}(\mathbf{x}_n)$ , where  $\mathbf{F}_n$  is the force on particle  $n$  located at  $\mathbf{x}_n$ . The force on a particle at  $\mathbf{x}_n$  depends only on the electromagnetic fields at  $\mathbf{x}_n$ .

This however, is a purely classical view, a view that, as we have seen in the section on ensemble averaging, is incompatible with any consideration of a medium composed of stable atoms. The ensemble averaging is the procedure that accounts for these quantum effects, resulting in a continuous charge distribution and rendering the notion of position of any given charge henceforth meaningless. The spatial averaging required to eliminate the oscillations of the electric field due to the periodicity of the lattice contributes even further to this blurring. Since both averages involve a loss of information about the positions and velocities of particles, they implicitly render the description nonlocal.

From a historical point of view, the definition of the response function of a dielectric medium as nonlocal in direct space (and therefore local in reciprocal space) can be understood by considering the fact that from a classical perspective, where light is seen as a wave, it does not make much sense to insist on the notion of position of the wave, and rather more on the frequency (temporal or spatial). It was far easier to fix the wavelength of a light wave than its exact position, indeed, the very notion of the position of a wave seemed meaningless. Transfer functions local in direct space are typical of particle like behavior, while transfer functions local in reciprocal space are typical of wave like behavior. Since the dual wave/particle nature of light became known rather late, *after* the work on the photoelectric effect in the early years of the 20th century, the non-local wave-like description was (and for most purposes remains) the most natural.

We now compare Eq. (1.21) with Eq. (1.17) obtaining the master equation of the dielectric medium

$$[\mathbf{p}(\mathbf{x})] = [\varepsilon_0 \gamma^e(\mathbf{x}) \mathbf{e}(\mathbf{x})] = \varepsilon_0 f(\mathbf{x}) \circ (\gamma^e(\mathbf{x}) \mathbf{e}(\mathbf{x})) = \varepsilon_0 \chi^e(\mathbf{x}) \circ [\mathbf{e}(\mathbf{x})] \quad (1.22)$$

where we have assumed zero intrinsic polarization at zero field (no ferro-electricity). The susceptibility  $\chi^e$  is therefore a macroscopic quantity *defined as a relationship between two macroscopic quantities* rather than as an average of some microscopic quantity. As such it is a macroscopic property whose relationship to the microscopic description is indirect and intuitively slippery. While  $\chi^e$  is obviously fundamentally dependent on  $\gamma^e$ , it is far from clear in what way (if any) this dependence may be made more explicit or straightforward in the general case. The difficulty resides in the fact that in this equation we see an intricate interplay of the micro- and macro-, of classical and quantum, worlds that are intuitively and physically apart and no straightforward or smooth transition is possible.

Eq. (1.22) provides the ab-initio starting point for calculating the susceptibility of a given medium. However, in order to obtain the position dependent polarizability of atoms their internal dynamics must be considered, and one cannot avoid a detailed quantum mechanical analysis. In the case of *naturally occurring* dielectric media the microscopic and the macroscopic are worlds governed by different rules. The microscopic is governed by Schrodinger's equation, while the macroscopic is governed by Maxwell's equations. A detailed understanding is therefore a very ambitious enterprise and often the phenomenological approach is the most pragmatic. As we shall see in the following chapters, however, this will not be the case for artificial materials, since in that case, both the unit cell and the macroscopic description are governed by the same equations, the macroscopic Maxwell's equations. An exact description is

therefore much more easily accessible.

If we consider Eq. (1.22) then it is clear that spatial dispersion can be ignored whenever the right hand side can be written as a product, that is, whenever the left hand side is proportional to  $[\mathbf{e}(\mathbf{x})]$ , meaning that the susceptibility is singular. As a first, and rather trivial case, this is possible when the polarizability  $\gamma^e(\mathbf{x})$  is simply a constant. Then it can be taken out of the convolution and we would have  $\chi^e(\mathbf{x}) = \varepsilon_0 \gamma^e \delta(\mathbf{x})$ . However, a constant polarizability means either there are no atoms, or they are much larger than the period (see subsection 1.4.2), neither case being very interesting or relevant. A second, far less obvious and more relevant case, is that where the wavelength is sufficiently large that the field can be considered constant over a volume the size of the electromagnetic neighborhood of any given lattice site. This case is discussed in detail in the next subsection.

In the following discussion we will assume the medium is periodic with cubic symmetry and a single atom per unit cell, for purposes of illustration. We consider two simplified but important cases. The first is the case where the atom is much smaller than the unit cell, and the unit cell is much smaller than the averaging volume  $f$ . The second is where the size of the atom is similar to or greater than the unit cell, but both are much smaller than the averaging volume. We will refer to the two situations as the “small atoms” and “large atoms” cases.

### 1.4.1 Small atoms – electric

Our small-atoms assumption has two components. First, since the atoms are small, their electron clouds do not overlap, and it is therefore possible to expand the electric field due to any particular atom into a multipole series. Second, the distances separating the atoms from each other must be larger than the distance over which the quadrupole and higher moments are negligible. This implies that the field seen by the neighbors of any particular atom is given, to an arbitrarily good approximation, *only* by the dipolar component of the field of said atom. From the point of view of any particular lattice site the rest of the lattice can be treated as a collection of point dipoles. The validity of this approximation in a simple geometry such as that of a circular dielectric rod grating can be verified using the multiscattering model of Appendix A.

In this form the small-atoms assumption has a very important consequence. The fact that the atoms are isolated means that the left side of Eq. (1.22) must be modified. By assuming the atoms are isolated we have eliminated the effect that the internal fields, or the higher multipole fields of any given atom could have on the others. In fact, the field seen by any given atom is no longer the *total* field due to all the other atoms, but the dipolar *radiated* field due to all the other atoms. The distinction between the total field and the radiated field of a distribution of charges (an atom) is important because the second excludes the fields over the region *occupied* by said distribution. A multipole expansion is only given with respect to some closed surface which must completely enclose the charge and is valid only *outside* of it. For instance if we consider the field of a point dipole placed at the origin we have

$$\mathbf{e}_{dipole}(\mathbf{x}) = \frac{1}{4\pi\varepsilon_0} \nabla \left( \mathbf{p} \cdot \nabla \frac{1}{\|\mathbf{x}\|} \right) = \frac{1}{4\pi\varepsilon_0} \frac{(3\hat{\mathbf{x}}(\hat{\mathbf{x}} \cdot \mathbf{p}) - \mathbf{p})}{\|\mathbf{x}\|^3} - \frac{\mathbf{p}}{3\varepsilon_0} \delta(\mathbf{x}) \quad (1.23)$$

whereas the radiated field (in the static limit) of the same dipole is

$$\mathbf{e}_{dipole}^*(\mathbf{x}) = \frac{1}{4\pi\varepsilon_0} \frac{(3\hat{\mathbf{x}}(\hat{\mathbf{x}} \cdot \mathbf{p}) - \mathbf{p})}{\|\mathbf{x}\|^3} \quad (1.24)$$

since the singular term encapsulates the localized fields which are not seen by neighboring scatterers.  $\hat{\mathbf{x}}$  is the unit vector in the direction of  $\mathbf{x}$ . In what follows we will distinguish radiated fields from total fields with a star in the superscript. The field  $\mathbf{e}^*(\mathbf{x})$  seen by any particular dipole can therefore be written in

terms of the total field as

$$\mathbf{e}^*(\mathbf{x}) = \mathbf{e}(\mathbf{x}) + \sum_n \frac{\mathbf{p}_n \delta(\mathbf{x} - \mathbf{x}_n)}{3\epsilon_0} \quad (1.25)$$

Eq. (1.22) for this case must therefore be written as

$$f(\mathbf{x}) \circ (\gamma^e(\mathbf{x})\mathbf{e}^*(\mathbf{x})) = \chi^e(\mathbf{x}) \circ [\mathbf{e}(\mathbf{x})] \quad (1.26)$$

We begin by writing the total polarizability  $\gamma^e(\mathbf{x})$  as a sum over the atomic polarizabilities  $\gamma_n^e(\mathbf{x})$

$$f(\mathbf{x}) \circ (\gamma^e(\mathbf{x})\mathbf{e}^*(\mathbf{x})) = f(\mathbf{x}) \circ \left( \sum_n \gamma_n^e(\mathbf{x})\mathbf{e}^*(\mathbf{x}) \right) \quad (1.27)$$

Since we have assumed that  $\gamma_n^e(\mathbf{x})$  is very localized, then its Fourier transform is well represented by the zeroth term of its Taylor expansion, such that  $\tilde{\gamma}_n^e(\mathbf{k}) \approx \tilde{\gamma}_n^e(\mathbf{0}) \equiv \gamma_s^e$  (all atoms are identical) and we have  $\gamma_n^e(\mathbf{x}) \approx \gamma_s^e \delta(\mathbf{x} - \mathbf{x}_n)$  where the s subscript stands for the singular, or DC component of the polarizability. The left side of Eq. (1.22) becomes

$$f(\mathbf{x}) \circ (\gamma^e(\mathbf{x})\mathbf{e}^*(\mathbf{x})) = f(\mathbf{x}) \circ \left( \mathbf{e}^*(\mathbf{x}) \sum_n \gamma_s^e \delta(\mathbf{x} - \mathbf{x}_n) \right)$$

We take the Fourier transform of the above equation to obtain

$$\tilde{f}(\mathbf{k}) \left( \tilde{\mathbf{e}}^*(\mathbf{k}) \circ \frac{\gamma_s^e}{V_{\text{uc}}} \sum_n \delta(\mathbf{k} - \mathbf{G}_n) \right) \quad (1.28)$$

where the Fourier transform of a Dirac comb is also a Dirac comb in reciprocal space, the  $\mathbf{G}_n$  are the reciprocal lattice vectors, and  $V_{\text{uc}}$  is the volume of the unit cell in real space (see Eq. 2-12 of Kittel [21]). We note  $N = 1/V_{\text{uc}}$  the atomic number density. The electric field is quasiperiodic as per Bloch's theorem, such that its Fourier transform can be written

$$\tilde{\mathbf{e}}^*(\mathbf{k}) = \sum_m \mathbf{E}_m^* \delta(\mathbf{k} - \mathbf{k}_B - \mathbf{G}_m)$$

and we can rewrite expression (1.28) as

$$N\gamma_s^e \tilde{f}(\mathbf{k}) \left( \sum_{n,m} \mathbf{E}_m^* \delta(\mathbf{k} - \mathbf{k}_B - \mathbf{G}_n - \mathbf{G}_m) \right)$$

We must now recall that the function  $\tilde{f}(\mathbf{k})$  has been *designed* in order to filter out all spatial frequencies which are not in the first Brillouin zone. Consequently, in the above sum, only those terms will survive where  $\mathbf{G}_n + \mathbf{G}_m = \mathbf{0}$  such that the delta function is at  $\mathbf{k}_B$  and therefore within the first BZ. This relation is satisfied when  $m = -n$  and therefore for any  $m$  nonzero, there will be exactly one  $n$  such that the term survives. The  $\mathbf{E}_0$  term is not filtered out because it is already by default in the first BZ. We can therefore rewrite the above expression

$$N\gamma_s^e \tilde{f}(\mathbf{k}) \left( \mathbf{E}_0^* \delta(\mathbf{k} - \mathbf{k}_B) + \sum_{n \neq 0} \mathbf{E}_n^* \delta(\mathbf{k} - \mathbf{k}_B) \right).$$

Now, the  $\mathbf{E}_n$  coefficients are the Fourier coefficients of a Bloch wave, and as such they have some dependence on the Bloch vector  $\mathbf{k}_B$ . When this vector is large (approaching the edges of the BZ)

this dependence is strong, the quantity in parentheses becomes a function of  $\mathbf{k}_B$ , leading to a spatially dispersive medium. However, when this vector approaches  $\mathbf{0}$  the dependence diminishes and the Fourier coefficients become, to a good approximation, independent of  $\mathbf{k}_B$ . Moreover, when the medium is highly symmetrical (cubic symmetry) one can show [3] that the sum over  $\mathbf{E}_n^*$  reduces to zero and we obtain

$$\begin{aligned} f(\mathbf{x}) \circ (\gamma^e(\mathbf{x})\mathbf{e}^*(\mathbf{x})) &= \mathcal{F}^{-1} \left( N\gamma_s^e \tilde{f}(\mathbf{k}) (\mathbf{E}_0^* \delta(\mathbf{k} - \mathbf{k}_B)) \right) \\ &= N\gamma_s^e [\mathbf{e}^*(\mathbf{x})] \end{aligned}$$

This only holds for cubic crystals. The averaged radiation field can be written in terms of the averaged total field by using Eqs. (1.23) and (1.24) and the fact that the unit cell contains only one atom:

$$\begin{aligned} N\gamma_s^e [\mathbf{e}^*(\mathbf{x})] &= N\gamma_s^e \left[ \mathbf{e}(\mathbf{x}) + \sum_n \frac{\mathbf{p}_n \delta(\mathbf{x} - \mathbf{x}_n)}{3\varepsilon_0} \right] \\ &= N\gamma_s^e \left( [\mathbf{e}(\mathbf{x})] + \frac{1}{3\varepsilon_0} [\mathbf{p}(\mathbf{x})] \right) \\ &= N\gamma_s^e \left( [\mathbf{e}(\mathbf{x})] + \frac{1}{3} \chi^e(\mathbf{x}) \circ [\mathbf{e}(\mathbf{x})] \right) \end{aligned}$$

Putting this back into the definition of the susceptibility we have

$$N\gamma_s^e \left( [\mathbf{e}(\mathbf{x})] + \frac{1}{3} \chi^e(\mathbf{x}) \circ [\mathbf{e}(\mathbf{x})] \right) = \chi^e(\mathbf{x}) \circ [\mathbf{e}(\mathbf{x})] \quad (1.29)$$

and by grouping the terms containing  $\chi$  we obtain

$$\chi^e(\mathbf{x}) \circ [\mathbf{e}(\mathbf{x})] = \frac{N\gamma_s^e}{1 - N\gamma_s^e/3} [\mathbf{e}(\mathbf{x})].$$

The susceptibility is therefore singular:

$$\chi^e(\mathbf{x}) = \frac{N\gamma_s^e}{1 - N\gamma_s^e/3} \delta(\mathbf{x}) \quad (1.30)$$

and we recognize the Mossotti-Clausius relation. When the polarizability is anisotropic but diagonal this relation can be considered to hold separately for each coordinate component, or may equivalently be rewritten as

$$\bar{\chi}^e(\mathbf{x}) = N\bar{\gamma}_s^e (\mathbf{I} - N\bar{\gamma}_s^e/3)^{-1} \delta(\mathbf{x})$$

We see that the difference between the total field of a dipole and its radiated field is essential. The assumption that the atoms are non-overlapping, which enables us to use the multipole expansion, is equivalent to assuming the fields localized on the atoms do not play a role in the behavior of the other atoms. If we had subsequently used the total field average  $[\mathbf{e}(\mathbf{x})]$  on the left side of Eq. (1.36) instead of the radiated field  $[\mathbf{e}^*(\mathbf{x})]$  that would have amounted to a contradiction of the small atoms assumption.

Before going on to discuss the case of strongly coupled atoms we must point out an aspect of detail, which is consistently neglected in most discussions of the Mossotti-Clausius result. The above discussion makes no distinction between fields incident on the medium under study from outside, and fields originating within the medium. In Eq. (1.25), for instance, the term  $\mathbf{e}(\mathbf{x})$  contains, in principle, contributions from both the radiation fields of the scatterers within the medium, as well as external incident fields presumably due to charge distributions outside the medium. This is the distinction between the internal and the external field. In the interest of clarity this distinction was not made explicitly above, but it must be mentioned because it becomes important when one considers energy conservation.

Strictly speaking the above discussion is inconsistent from a point of view of energy conservation. If we consider a single scattering atom exposed to an incident plane wave, then it is easy to see that it will accept some energy from the wave, and it will scatter some energy. If the incident field stops furnishing energy, then the dipole will eventually radiate its energy and gradually settle down in a rest state (classically speaking). The dipole acts like a damped oscillator. As long as it is driven it oscillates, but when the driving field stops, it radiates away its remaining energy and eventually comes to rest. However this means that the polarizability cannot be real. The radiation mechanism must be accounted for through an imaginary *radiative damping* term. This may seem surprising if one considers Eq. (1.30) because it seems to imply that even in a medium composed of non-absorbing scatterers, the susceptibility must be imaginary, which is clearly false.

The answer is that the polarizability which appears in this section is in fact not exactly the same as if the atom had been alone in space, even if electron cloud overlap and near field effects are accounted for. The difference is that whereas for the free-space atom scattered (or radiated) energy is forever lost, for the atom in the bulk of a material this energy is not lost because each of the atoms surrounding it sends *back* a small part. In fact, since the medium is assumed to be infinite (the atom is very far from any interface), and *regular* (cubic or rectangular symmetry), *all* of the scattered energy is scattered back by the other atoms in the crystal. This energy balance can be accounted for by replacing the free-space imaginary polarizability  $\gamma$  with a real polarizability  $\gamma'$  given by the relation (see Section III of Ref. [23])

$$\gamma' = \frac{1}{\text{Real}\left(\frac{1}{\gamma}\right)}.$$

It is assumed the scatterers exhibit no magneto-electric coupling, in other words that the electric polarization  $\mathbf{p}$  depends *only* on the electric field  $\mathbf{e}$ , and the magnetic polarization  $\mathbf{m}$  depends *only* on the magnetic flux density  $\mathbf{b}$ . In this section and the following only real polarizabilities are considered, and the primes are omitted for readability.

We have seen in this section that when  $\mathbf{k}_B$  is sufficiently close to  $\Gamma$  the medium is non-spatially dispersive, and the Mossotti-Clausius relation is obeyed. It is not clear however, how close it needs to be. This can be qualitatively explained using the notion of electromagnetic neighborhood. In fact, since the interaction between two atoms decreases quickly with distance, it follows that any given atom in the lattice only interacts to any considerable extent with a small number of other atoms in its immediate vicinity. The volume containing these neighbors we call the “electromagnetic neighborhood” of the given atom. The internal field at any lattice site is mostly due to the atoms in its electromagnetic neighborhood. Now, the derivation of the Mossotti-Clausius relation makes use of the fact that the applied field is uniform throughout the lattice. In reality this is only an approximation, because we are not working at  $\omega = 0$ , and the field is not uniform. However, if the field is uniform *over a volume the size of the EM neighborhood* of a given atom, then from the point of view of the atom the result is the same. The electric field *looks* uniform.

We should also point out that in the Small Atoms case, since the  $a$  scale is smaller than the  $d$  scale, the atomic and lattice contributions are expected to be unequal. In fact, in this case the lattice effects dominates. This is why, in Chapter 3, we will treat the resonators as point dipole scatterers. The multipole coupling between neighboring cells (atomic dispersion) is a weak effect compared to the effect of the nonuniformity of the smoothed field over these cells (lattice dispersion).

In fact, as we will see in the next section, from the point of view of the *formal* validity of the Mossotti-Clausius relation, it is not even necessary that the interactions between neighbors be dipolar in nature, since the polarizability can simply be redefined to include non-dipolar effects. What is important is that the neighbors see the same applied field. When this is no longer the case, then different neighbors see

different fields, which depend on the phase and direction of propagation of the applied field, resulting in spatial dispersion.

### 1.4.2 Big atoms – electric

In this case we can no longer consider the atoms isolated, and we must take into account the *total* fields at the positions of the atoms. The total position dependent polarizability of the medium is a function with the periodicity of the medium,  $\gamma^e(\mathbf{x})$ . Its Fourier transform can be written as

$$\tilde{\gamma}^e(\mathbf{k}) = \sum_i \gamma_i^e \delta(\mathbf{k} - \mathbf{G}_i)$$

where  $\gamma_i^e$  are the Fourier coefficients, and  $\gamma_0^e = \gamma_s^e/V_{uc}$ . The case of the small atoms amounts to assuming  $\gamma_i^e \cong \gamma_0^e$ . We now drop this simplification, and in addition assume that the quantum aspects of the electron cloud overlap of neighboring atoms is already contained in the polarizability. Eq. (1.28) becomes

$$\begin{aligned} \tilde{f}(\mathbf{k}) (\tilde{\gamma}(\mathbf{k}) \circ \tilde{\mathbf{e}}(\mathbf{k})) &= \tilde{f}(\mathbf{k}) \left( \sum_i \gamma_i^e \delta(\mathbf{k} - \mathbf{G}_i) \right) \circ \left( \sum_j \mathbf{E}_j \delta(\mathbf{k} - \mathbf{G}_j - \mathbf{k}_B) \right) \\ &= \tilde{f}(\mathbf{k}) \sum_{i,j} \gamma_i^e \mathbf{E}_j \delta(\mathbf{k} - \mathbf{G}_i - \mathbf{G}_j - \mathbf{k}_B) \end{aligned} \quad (1.31)$$

The field and the polarization are combined in a more complicated way than before. However, we know, from the first section, that when the wavelength is large enough, the medium is equivalent to a lattice of point dipoles, *whether the atoms overlap or not*. Macroscopically the behavior of each atom depends on only two quantities: its effective polarizability, and the effective field that acts on it. However, since all that is visible on the macroscopic scale is the product of these two quantities, and the microscopic information has been discarded, then the meaning that may be attached to these quantities is open to discussion, a discussion we leave to Section 1.8. For now, we have two options.

1. We can consider that the effective atomic polarizability is the same as for non-overlapping atoms, but that the effective field seen by the atoms is not the same as before.
2. We can consider that the field seen by the atoms is the same as before, but that the effective atomic polarizability seen on the macroscopic scale is different due to the overlap.

We first consider option 1.

In the previous section, we showed that Small Atoms do not see each other's internal fields, and we wrote

$$[\mathbf{e}^*(\mathbf{x})] = \left[ \mathbf{e}(\mathbf{x}) + \sum_n \frac{\mathbf{p}_n \delta(\mathbf{x} - \mathbf{x}_n)}{3\epsilon_0} \right]$$

When the atoms are much larger than the distance separating them, then we can consider that they see the total field in the medium, and we have

$$[\mathbf{e}^*(\mathbf{x})] = [\mathbf{e}(\mathbf{x})]$$

Intermediate situations may therefore be accounted for phenomenologically by writing

$$[\mathbf{e}^*(\mathbf{x})] = \left[ \mathbf{e}(\mathbf{x}) + \alpha \sum_n \frac{\mathbf{p}_n \delta(\mathbf{x} - \mathbf{x}_n)}{3\epsilon_0} \right]$$

where the parameter  $\alpha$  tells us to what extent the atoms can be considered as isolated. When  $\alpha = 1$  the atoms are perfectly isolated and when  $\alpha = 0$  the atoms are so large compared to the period that the polarizability is simply a constant, a situation also mentioned on page 21 above. This leads to a Mossotti-Clausius-*like* relation of the form

$$\chi(\mathbf{x}) = \frac{N\gamma_s^e}{1 - \alpha N\gamma_s^e/3} \delta(\mathbf{x})$$

We now consider option 2. In this case we basically include the phenomenological parameter  $\alpha$  into the effective polarizability, which was a phenomenological parameter to start with. We define

$$\gamma_{\text{eff}}^e = \frac{3\gamma_s^e}{3 + N\gamma_s^e(1 - \alpha)} \quad (1.32)$$

and the Mossotti-Clausius relation remains formally unchanged

$$\chi^e(\mathbf{x}) = \frac{N\gamma_{\text{eff}}^e}{1 - N\gamma_{\text{eff}}^e/3} \delta(\mathbf{x})$$

## 1.5 Polarizability, susceptibility – magnetic

We have so far said nothing about the magnetic activity of the material under study. As mentioned above, the mathematical details involved in the spatial averaging are somewhat more involved in the case of the microscopic current density  $\mathbf{j}$ . However, when there is no net or free charge (and therefore no free current), the space and ensemble averaged bound current can be written

$$[\mathbf{j}(\mathbf{x})] = \frac{\partial[\mathbf{p}(\mathbf{x})]}{\partial t} + \nabla \times [\mathbf{m}(\mathbf{x})]$$

Here  $\mathbf{m}(\mathbf{x})$  stands for

$$\mathbf{m}(\mathbf{x}) = \sum_n \mathbf{m}_n(\mathbf{x} - \mathbf{x}_n)$$

and  $\mathbf{m}_n$  is the atomic magnetic moment. By noting the macroscopic magnetization  $\mathbf{M}(\mathbf{x}) = [\mathbf{m}(\mathbf{x})]$  and introducing the magnetic field

$$\mathbf{H}(\mathbf{x}) = \frac{\mathbf{B}(\mathbf{x})}{\mu_0} - \mathbf{M}(\mathbf{x}) \quad (1.33)$$

the inhomogeneous curl equation from Eq. (1.6) becomes

$$\nabla \times \mathbf{H}(\mathbf{x}) - \frac{\partial}{\partial t} (\varepsilon_0 \mathbf{E}(\mathbf{x}) + \mathbf{P}(\mathbf{x})) = 0 \quad (1.34)$$

The reason we have introduced the magnetic field  $\mathbf{H}$  is that for historical reasons the magnetic susceptibility is defined as relating  $\mathbf{M}$  and  $\mathbf{H}$  (rather than  $\mathbf{M}$  and  $\mathbf{B}$ )

$$\mathbf{M}(\mathbf{x}) = \chi^m(\mathbf{x}) \circ \mathbf{H}(\mathbf{x})$$

The atomic polarization, however, cannot be defined with respect to  $\mathbf{h}$  because there is no such thing as the microscopic magnetic field. The magnetic field is a derived macroscopic quantity that is *not* the smoothed version of some microscopic field. The magnetic polarizability is defined by the relation

$$\mathbf{m}(\mathbf{x}) = \sum_n \mathbf{m}_n(\mathbf{x}) = \sum_n \frac{1}{\mu_0} \gamma_n^m(\mathbf{x}) \mathbf{b}(\mathbf{x}) = \frac{1}{\mu_0} \gamma^m(\mathbf{x}) \mathbf{b}(\mathbf{x})$$

where the  $1/\mu_0$  is for later convenience. The analog of Eq. (1.22) is then obtained by combining the last two equations

$$f(\mathbf{x}) \circ (\gamma^m(\mathbf{x})\mathbf{b}(\mathbf{x})) = \chi^m(\mathbf{x}) \circ \mathbf{H}(\mathbf{x}) \quad (1.35)$$

In the following section we consider the small atoms case.

### 1.5.1 Small atoms – magnetic

As before, when the atoms are small, we must distinguish between the total and the radiated field. Since on the microscopic scale only  $\mathbf{b}$  is defined, then the magnetic flux density of a dipole at the origin is

$$\mathbf{b}_{dipole}(\mathbf{x}) = \frac{\mu_0}{4\pi} \nabla \times \left( \frac{\mathbf{m} \times \mathbf{x}}{\|\mathbf{x}\|^3} \right) = \frac{\mu_0}{4\pi} \frac{(3\hat{\mathbf{x}}(\hat{\mathbf{x}} \cdot \mathbf{m}) - \mathbf{m})}{\|\mathbf{x}\|^3} + \frac{2\mu_0\mathbf{m}}{3} \delta(\mathbf{x})$$

whereas the radiated field (in the static limit) of the same dipole is

$$\mathbf{b}_{dipole}^*(\mathbf{x}) = \frac{\mu_0}{4\pi} \frac{(3\hat{\mathbf{x}}(\hat{\mathbf{x}} \cdot \mathbf{m}) - \mathbf{m})}{\|\mathbf{x}\|^3}$$

Compare and contrast with Eqs. (1.23) and (1.24). Consequently we can write the microscopic radiated field in the medium in terms of the total field as

$$\mathbf{b}^*(\mathbf{x}) = \mathbf{b}(\mathbf{x}) - \sum_n \frac{2\mu_0\mathbf{m}_n\delta(\mathbf{x} - \mathbf{x}_n)}{3}$$

The development then proceeds by analogy with section 1.4.1 until we obtain

$$\begin{aligned} f(\mathbf{x}) \circ (\gamma(\mathbf{x})\mathbf{b}^*(\mathbf{x})) = N\gamma_s^m [\mathbf{b}^*(\mathbf{x})] &= N\gamma_s^m \left[ \mathbf{b}(\mathbf{x}) - \sum_n \frac{2\mu_0\mathbf{m}_n\delta(\mathbf{x} - \mathbf{x}_n)}{3} \right] \\ &= N\gamma_s^m \left( [\mathbf{b}(\mathbf{x})] - \frac{2\mu_0}{3} [\mathbf{m}(\mathbf{x})] \right) \\ &= N\gamma_s^m \left( [\mathbf{b}(\mathbf{x})] - \frac{2\mu_0}{3} \chi^m(\mathbf{x}) \circ \mathbf{H}(\mathbf{x}) \right) \end{aligned}$$

Putting this back into the definition of the susceptibility we have

$$\mathbf{M}(\mathbf{x}) = N\gamma_s^m \left( [\mathbf{b}(\mathbf{x})] - \frac{2\mu_0}{3} \chi^m(\mathbf{x}) \circ \mathbf{H}(\mathbf{x}) \right) = \chi^m(\mathbf{x}) \circ \mathbf{H}(\mathbf{x}) \quad (1.36)$$

We now replace

$$[\mathbf{b}(\mathbf{x})] = \mathbf{B}(\mathbf{x}) = \mu_0\mathbf{H}(\mathbf{x}) + \mu_0\mathbf{M}(\mathbf{x}) = \mu_0\mathbf{H}(\mathbf{x}) + \mu_0\chi^m(\mathbf{x}) \circ \mathbf{H}(\mathbf{x})$$

and group terms containing  $\chi^m(\mathbf{x})$  to obtain

$$\chi^m(\mathbf{x}) \circ \mathbf{H}(\mathbf{x}) = \frac{N\gamma_s^m}{1 - N\gamma_s^m/3} \mathbf{H}(\mathbf{x})$$

which results in a singular susceptibility given by the magnetic version of the Mossotti-Clausius relation

$$\chi^m(\mathbf{x}) = \frac{N\gamma_s^m}{1 - N\gamma_s^m/3} \delta(\mathbf{x})$$

The path was somewhat different but we have arrived at a very similar result as for the electric

susceptibility. Within this formulation electric and magnetic phenomena are formally symmetrical on the *macroscopic* scale in the absence of free charges.

## 1.6 Permittivity and permeability – index and impedance

Now that we have obtained macroscopic parameters which characterize the behavior of the material we would like to write the full macroscopic Maxwell's equations in order to obtain the plane wave solutions. The spatial and ensemble averaged divergence equation

$$\nabla \cdot [\mathbf{e}(\mathbf{x})] = [\eta(\mathbf{x})]/\varepsilon_0$$

becomes (using Eq. (1.14))

$$\varepsilon_0 \nabla \cdot \mathbf{E}(\mathbf{x}) = -\nabla \cdot \mathbf{P}(\mathbf{x})$$

We now introduce the electric flux density, or the electric displacement vector

$$\mathbf{D}(\mathbf{x}) = \varepsilon_0 \mathbf{E}(\mathbf{x}) + \mathbf{P}(\mathbf{x}) = \varepsilon_0 (\delta(\mathbf{x}) + \chi^e(\mathbf{x})) \circ \mathbf{E}(\mathbf{x}) = \varepsilon_0 \varepsilon(\mathbf{x}) \circ \mathbf{E}(\mathbf{x}).$$

where the relative permittivity (also often referred to abusively as the dielectric “constant”) is defined in general by  $\varepsilon(\mathbf{x}) = \delta(\mathbf{x}) + \chi^e(\mathbf{x})$  which reduces to  $\varepsilon(\mathbf{x}) = 1 + \chi^e(\mathbf{x})$  for local media. The electric divergence equation in a medium with no free charges now takes the simple form

$$\nabla \cdot \mathbf{D}(\mathbf{x}) = 0$$

If we use the newly introduced electric displacement field the macroscopic Maxwell-Ampère equation from Eq. (1.34) is then written

$$\nabla \times \mathbf{H}(\mathbf{x}) - \frac{\partial \mathbf{D}(\mathbf{x})}{\partial t} = 0$$

By rewriting Eq. (1.33) we have

$$\mathbf{B}(\mathbf{x}) = \mu_0 \mathbf{H}(\mathbf{x}) + \mu_0 \chi^m(\mathbf{x}) \circ \mathbf{H}(\mathbf{x}) = \mu_0 \mu(\mathbf{x}) \circ \mathbf{H}(\mathbf{x})$$

where the relative permeability is defined in general by  $\mu(\mathbf{x}) = \delta(\mathbf{x}) + \chi^m(\mathbf{x})$  which reduces to  $\mu(\mathbf{x}) = (1 + \chi^m(\mathbf{x}))\delta(\mathbf{x})$  in local media.

We are now able to write the complete source free macroscopic Maxwell's equations:

$$\nabla \times \mathbf{H}(\mathbf{x}) + i\omega \mathbf{D}(\mathbf{x}) = 0 \tag{1.37}$$

$$\nabla \cdot \mathbf{B}(\mathbf{x}) = 0 \tag{1.38}$$

$$\nabla \times \mathbf{E}(\mathbf{x}) - i\omega \mathbf{B}(\mathbf{x}) = 0 \tag{1.39}$$

$$\nabla \cdot \mathbf{D}(\mathbf{x}) = 0 \tag{1.40}$$

and the corresponding constitutive relations

$$\mathbf{B}(\mathbf{x}) = \mu_0 \mu(\mathbf{x}) \circ \mathbf{H}(\mathbf{x})$$

$$\mathbf{D}(\mathbf{x}) = \varepsilon_0 \varepsilon(\mathbf{x}) \circ \mathbf{E}(\mathbf{x})$$

We see immediately one of the benefits of the homogenization procedure. Whereas on a microscopic scale the electric and magnetic phenomena are inevitably asymmetric, from a macroscopic point of view

electric and magnetic phenomena in source free regions are, *at least formally*, perfectly symmetric, as can be seen by inspection of the above equations.

In general the solutions to these equations are complicated and require a numerical computational approach. However, there are some simple cases where the solutions can be written explicitly. An example is that of local homogeneous isotropic media, for which the relative permittivity and permeability are purely singular and scalar  $\varepsilon(\mathbf{x}) = \varepsilon\delta(\mathbf{x})$  and  $\mu(\mathbf{x}) = \mu\delta(\mathbf{x})$ . In this case by eliminating  $\mathbf{H}$  from Eqs. (1.37) and (1.39) we obtain the Helmholtz wave equation

$$(\nabla^2 + \mu_0\varepsilon_0\mu\varepsilon\omega^2)\mathbf{E}(\mathbf{x}, t) = \mathbf{0}.$$

One possible solution is a field of the form

$$\mathbf{E}(\mathbf{x}) = \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} e^{ikx - i\omega t}. \quad (1.41)$$

This represents a plane wave of frequency  $\omega$  propagating in the positive  $x$  direction.  $k$  is the wavevector and by the Helmholtz equation it is related to  $\omega$  through the relation

$$k^2 = \mu_0\varepsilon_0\mu\varepsilon\omega^2$$

where if we denote the speed of light in a vacuum  $c = 1/\sqrt{\mu_0\varepsilon_0}$  and we introduce the index  $n^2 = \mu\varepsilon$  then we can rewrite it as

$$k^2 = n^2 \frac{\omega^2}{c^2}$$

It is interesting to consider the lossy case. If the electric field and the polarization of the medium are not exactly in phase, then the permittivity and/or the permeability must have a non-zero imaginary part. It is easily seen that this imaginary part is positive for lossy media, and negative for gain media when the time dependence is of the form  $e^{-i\omega t}$ , because if  $k$  has a positive imaginary part the wave is attenuated as it propagates along the  $x$  axis. We write  $k = \beta + i\alpha$  with  $\alpha, \beta$  real and  $\alpha > 0$ . We have

$$\begin{aligned} \mu &= \mu' + i\mu'' \\ \varepsilon &= \varepsilon' + i\varepsilon'' \\ n^2 &= \mu'\varepsilon' - \mu''\varepsilon'' + i(\mu'\varepsilon'' + \varepsilon'\mu'') \end{aligned}$$

with  $\mu'', \varepsilon'' > 0$ . Once the permittivity and the permeability have been determined for a given material then the wave propagates with a wavevector that can be determined from the equations

$$\beta^2 - \alpha^2 = \mu'\varepsilon' - \mu''\varepsilon'' \quad (1.42)$$

$$2\alpha\beta = \mu'\varepsilon'' + \varepsilon'\mu'' \quad (1.43)$$

In particular, if the imaginary parts are much smaller than the real parts of the permeability and permittivity but the real parts of both are negative then  $\beta$  must be negative also, as can be seen from the second equation. Consequently, *a material with negative real parts of permittivity and permeability will exhibit a negative real part of the index*. It is important to note, however, that this condition is sufficient but not necessary, at least not in lossy media. In other words the real part of the index can be negative even when one of the real parts of either the permittivity or permeability is positive. This can be seen from Eq. 1.43 if one keeps in mind that  $\alpha$ ,  $\varepsilon''$  and  $\mu''$  must all be positive in a passive medium

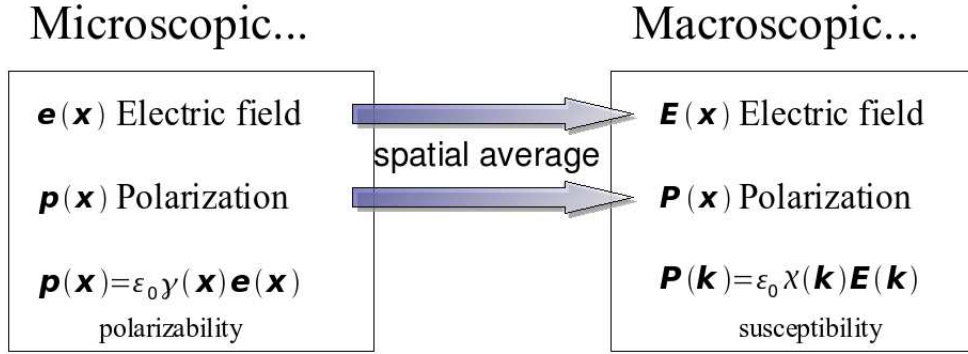


Figure 1.2: Schematic representation of the homogenization procedure.

with the sign conventions chosen in this work.

These concepts can be extended in a straightforward way to media for which the permittivity and permeability are  $3 \times 3$  diagonal tensors. In those cases the index of a given wave depends on the polarization of the electric and magnetic field vectors. For instance if the electric field is polarized purely along  $x$ , and the magnetic field is polarized purely along  $y$  then  $n = \mu_y \epsilon_x$ . However note that there is no such thing as the “tensor index” in the anisotropic case. The definition  $n^2 = \mu \epsilon$  is only valid in the scalar case. In the anisotropic case one must begin once more with Eqs. (1.37) and (1.39) (where the permittivity and the permeability no longer commute with the curl operator) and it is then possible to define an index ellipsoid, which specifies the index corresponding to any given field polarization. When the ellipsoid is an ellipsoid of revolution the medium is termed *uniaxial*, the most common type, both in natural and in artificial media.

## 1.7 Inhomogeneous models and local models

Homogenization is the process whereby the complicated and cumbersome microscopic fields existing in a heterogeneous medium are replaced by smooth varying (a.k.a. macroscopic) fields which, though ignoring the detailed aspects on the scale of heterogeneity, are very useful for characterizing the behavior of the medium. In essence, at a given wavelength, the field propagation in a given medium is independent of the microscopic details of the structure, being sensitive only to its macroscopic, average properties. Homogenization can therefore be seen as the process whereby all the (presumably useless) microscopic information is discarded, leaving only the useful, macroscopic information.

As we will see below, an important consequence of this process is that the blurring of the microscopic details of the field and the charge distribution implicitly makes the description nonlocal. In other words, once in the macroscopic world, the notion of “point” must be blurred into the notion of “ball”. One can no longer locate any point in space to within less than a certain distance which is related to the amount of blurring that has been done, or to the amount of microscopic information that has been discarded in the homogenization process. We therefore expect macroscopic fields to be related to each other in a non-local way related to how blurred the microscopic picture is by the homogenization process.

In this section we take a step back and take a broader look at the theory developed in the previous sections. We have outlined a procedure whereby one starts with a set of microscopic quantities ( $\mathbf{e}(\mathbf{x})$ ,  $\mathbf{p}(\mathbf{x})$ ,  $\gamma(\mathbf{x})$ ) and proceeds to obtain a set of macroscopic quantities ( $\mathbf{E}(\mathbf{x})$ ,  $\mathbf{P}(\mathbf{x})$ ,  $\chi(\mathbf{x})$ ) via a series of steps involving, in particular, spatial averaging. The spatial averaging takes the form of a convolution by a smoothing function  $f(\mathbf{x})$  and the macroscopic parameter of electric susceptibility is defined by the

following relation between the macroscopic field and polarization

$$\mathbf{P}(\mathbf{k}) = \chi(\mathbf{k})\mathbf{E}(\mathbf{k}).$$

By writing out the above equation in detail, such that only microscopic quantities appear, we obtain the *master equation* of the effective medium *model*:

$$f(\mathbf{x}) \circ (\gamma^e(\mathbf{x})\mathbf{e}(\mathbf{x})) = \chi^e(\mathbf{x}) \circ (f(\mathbf{x}) \circ \mathbf{e}(\mathbf{x}))$$

This equation is quite complex and non-intuitive due to the fact the the averaging function  $f(\mathbf{x})$  has the role of erasing microscopic information. Consequently, the susceptibility defined by this equation is not unique. Any amount of microscopic spatial jitter can be added, and the equation will remain correct. A detailed discussion of this equation is beyond the scope of this work but I must emphasize several aspects related to the role of the function  $f(\mathbf{x})$ .

In most formulations of the derivation of the macroscopic Maxwell equations [16, 17, 3], this function appears as a purely academic construct, and the sole concern of the authors is to state the required properties and to show that all of them can, in principle be satisfied. Thus, the jury is rigged, in the sense that one already knows the susceptibility that a given medium has, and the only concern is to show that a smoothing function  $f(\mathbf{x})$  can be found that will result in the expected value of the susceptibility. These arguments are focused above all on the coherence and consistency of the theoretical construction, rather than on the notion of model-building. In a sense, the model (the susceptibility) was already built and all there was left to do was to show that it was consistent with all the other assumptions that are commonly made about the microscopic behavior of dielectric media.

The point of view I would like to emphasize here is different. In this approach, we do not know in advance the model that we must obtain. Instead, we make full use of the fact that  $f(\mathbf{x})$  is a mathematical construct which we can choose as suits us. We treat the smoothing function  $f(\mathbf{x})$  as a *dial which we can tune* to obtain different types of models: we hereby introduce the novel concept of a ***custom-made effective medium model***. In order to understand the usefulness of this mathematical dial, we must take a closer look at the relation defining the susceptibility.

Let us consider the impact of the spatial averaging on the definition of the susceptibility:  $\mathbf{P}(\mathbf{k}) = \chi(\mathbf{k})\mathbf{E}(\mathbf{k})$ . In direct space this takes the form of a convolution integral:

$$\mathbf{P}(\mathbf{x}) = \chi(\mathbf{x}) \circ \mathbf{E}(\mathbf{x}) = \int d^3x' \chi(\mathbf{x} - \mathbf{x}')\mathbf{E}(\mathbf{x}').$$

Since, as argued above, the macroscopic description has been blurred by the spatial averaging, we expect the susceptibility to have a size similar to the averaging volume. In other words, the *macroscopic* polarization at some point in space is expected to depend on the *macroscopic* field over a region of similar size to the averaging volume  $f(\mathbf{x})$ . But let us consider what happens in the case of very large wavelength. In that case the macroscopic electric field can be considered constant over a region as large as an averaging volume, and the electric field can be taken out of the above integral. The macroscopic polarization becomes

$$\mathbf{P}(\mathbf{x}) \cong \mathbf{E}(\mathbf{x}) \int d^3x' \chi(\mathbf{x} - \mathbf{x}') = \chi_{\text{DC}}\mathbf{E}(\mathbf{x})$$

where  $\chi_{\text{DC}}$  is the DC component of the susceptibility. In this case, for all intents and purposes, the susceptibility acts as if it was singular:  $\chi_{\text{eff}} \approx \chi_{\text{DC}}\delta(\mathbf{x})$  and the macroscopic model acts as if it was local. In essence the implicit non-locality of the macroscopic description is *hidden by the size of the wavelength*. The electric field acts as a lowpass filter, in a sense, on the observed susceptibility and the blurring due to the spatial averaging is not detectable.

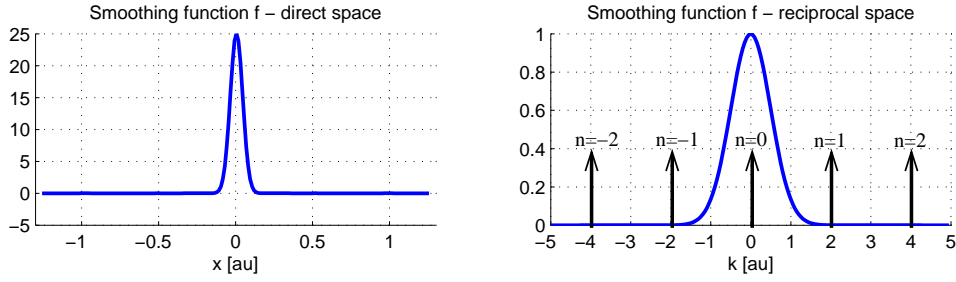


Figure 1.3: Smoothing function  $f(\mathbf{x})$  in one dimension - case of very large wavelength.

This works well as long as the wavelength of the macroscopic field in the medium is much larger than the averaging volume such that the field can be considered constant over the extent of the averaging volume. In this case the macroscopic model acts as if it was local, and therefore exhibits no spatial dispersion. This is the situation most widely considered in textbooks or other contexts.

However, as the wavelength becomes shorter, the non-locality must inevitably come into its rights. Indeed, as soon as the electric field is no longer approximately constant over volumes the size of  $f(\mathbf{x})$  the finite size of the susceptibility starts to make itself felt, with the result that one must once again write  $\mathbf{P}(\mathbf{x}) = \chi(\mathbf{x}) \circ \mathbf{E}(\mathbf{x})$ . Note that the above discussion implies that what is local or non-local is not the *medium* as is commonly thought, but the *model* we have made of it, which depends on the position of the  $f(\mathbf{x})$  dial. By choosing a sufficiently small  $f(\mathbf{x})$  we can always obtain a local model. Non-locality (known as spatial dispersion in reciprocal space) is a direct consequence, not of the properties of the medium, but of the neglect of microscopic details of the field and charge distributions which spatial averaging implies.

In this way, it becomes clear why a local model may be preferable to a non-local one if one is interested in understanding the physics of a given medium. Non-local models are in a sense incomplete descriptions of the underlying media through their neglect of important microscopic features of the field and charge distribution in the media.

In order to understand where the missing microscopic information resides it is simplest to consider the homogenization of a one-dimensional periodic system. The left side of Fig. 1.3 shows a Gaussian smoothing function in real space. If one recalls that spatial averaging is a convolution integral this figure shows that the value of some macroscopic smoothed quantity (for instance the macroscopic electric field  $\mathbf{E}(\mathbf{x})$ ) at  $\mathbf{x} = 0$  depends on the corresponding microscopic quantity (the microscopic electric field  $\mathbf{e}(\mathbf{x})$ ) over a region the size of the support of  $f(\mathbf{x})$ , which in this case is about  $0.2\text{au}$  across. In reciprocal space, however, a convolution integral transforms to a simple product and the smoothing function  $f(\mathbf{k})$  can be interpreted as a lowpass filter on the spatial frequency harmonics of the quantity being filtered. In our case, the microscopic electric field propagates in a periodic lattice of period  $d$  and Bloch's theorem tells us it is composed of a series of Bloch harmonics at spatial frequencies  $k_B + nK$  where  $k_B = 2\pi/\lambda$  and  $K = 2\pi/d$ . When the wavelength is very large the  $k_B$  term can be virtually ignored and the Bloch harmonics coincide with the reciprocal lattice vectors of the periodic medium. The homogenization process can then be seen (Fig. 1.3) to consist of filtering out all but the lowest harmonic, corresponding to  $n = 0$ .

This is relatively devoid of complications as long as the wavelength is very large. However, as the wavelength becomes smaller, the  $k_B$  term can no longer be ignored, and the Bloch harmonics no longer coincide with the reciprocal lattice vectors of the medium. This is illustrated on the right side of Fig. 1.4. Notice that in order for the filter  $f(\mathbf{k})$  to continue to serve its purpose, that is, to smooth over all higher harmonics except the lowest one, its shape must change. This shape change has a very important consequence in real space: the averaging volume becomes significantly larger as can be seen in the left

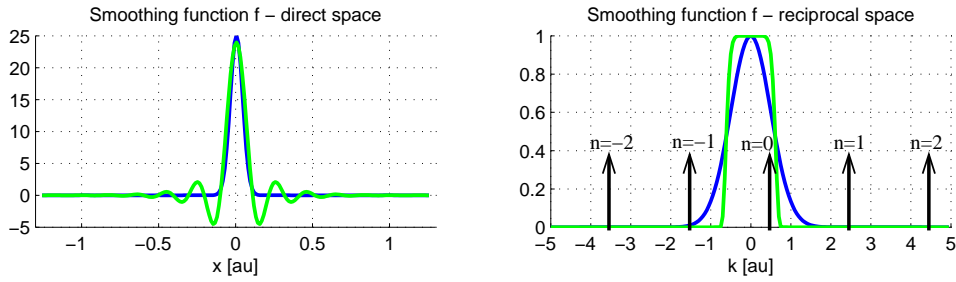


Figure 1.4: Smoothing function  $f(\mathbf{x})$  in one dimension - case of shorter wavelength (green), case of very large wavelength (blue).

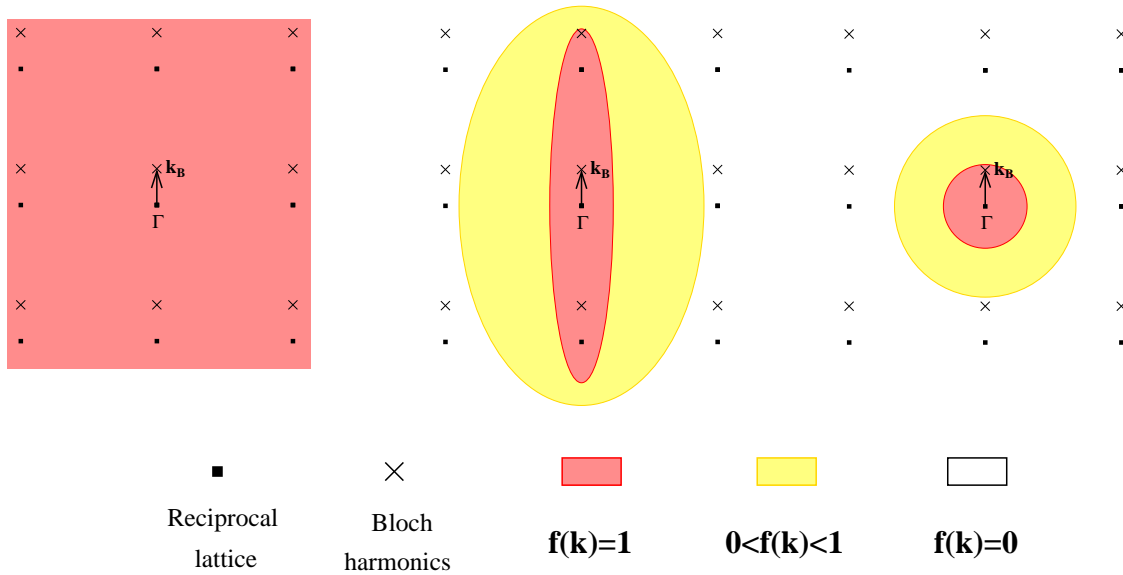


Figure 1.5: Three possible models corresponding to three  $f(\mathbf{k})$  functions are illustrated. The first is the identity model, the one where all harmonics are kept, corresponding in real space to  $f(\mathbf{x}) = \delta(\mathbf{x})$  the Dirac delta. The model on the right side is the homogeneous model, only the lowest harmonic is kept, the rest are averaged over. The case in the middle is an intermediate model, where two of the  $y$  harmonics are kept. All three models are effective medium models, but only the one on the right is homogeneous. But is it also local?

plot.

Eventually, as the wavelength becomes smaller and the averaging volume larger one arrives at a situation where the *macroscopic* electric field can no longer be considered constant over the size of the averaging volume. The homogeneous model has become nonlocal, or spatially dispersive. This is a result of the fact that the information contained in the higher Bloch harmonics which are filtered out in the averaging process is in fact non-negligible, and a model *must* include these higher harmonics if it is to be complete, and therefore local. It is these higher spatial harmonics filtered out in the spatial averaging process which contain the “missing information” discussed above.

These arguments can be extended in a straightforward way to the two-dimensional case. This is illustrated in Fig. 1.5. From left to right we have the identity model ( $f(\mathbf{x}) = \delta(\mathbf{x})$  the Dirac delta), a 1D model and the homogeneous model. As claimed above, the smoothing function  $f(\mathbf{x})$  can be tuned to obtain different types of models. In the frequency region where non-local effects appear this idea can be used to trade off non-locality against in-homogeneity. If the above arguments are correct, then the non-locality of the homogeneous model is related to the importance of the information contained in the higher Bloch harmonics of the field. This, in turn, leads to two predictions, which we test numerically

in Chapter 3:

1. If the homogeneous model and the 1D model disagree (give different transmission and reflection coefficients for normal incidence, for instance) at a given frequency, then the homogeneous model is non-local, and it also disagrees with the identity model. We expect this because if the 1D intermediate model disagrees with the homogeneous model, this implies the  $y$ -harmonics which are included in the 1D model contain important information about the medium, and cannot be ignored, meaning that the homogeneous model is incomplete and therefore nonlocal.
2. If the homogeneous model agrees with the identity model, then it is local, and it also agrees with the 1D intermediate model. We expect this because if the homogeneous and the identity model agree, then this implies that the information contained in *all* the higher harmonics is irrelevant so that the homogeneous model is complete and therefore local.

It is clear that this approach can be used to determine in a straightforward way whether a homogeneous model is non-local or not, without going through tedious off-normal transmission and reflection calculations. But this is only one of the applications of the new approach to homogenization discussed in this section. We discuss two others in particular.

The first idea starts from the fact that the metamaterial structures which are most easily analyzed using analytical methods [13, 14], being the most physically intuitive, do not always coincide with the structures that are most easily fabricated. In particular, a number of structures have been proposed recently [24, 25] which experiments have shown to exhibit interesting behavior. However, in such structures it is not always clear what is the important region of the unit cell, for instance, related to the magnetic or the electric activity of the material. This question is important for reasons of design optimisation. One would like to know, for instance, whether it is the thickness of the metal components, their length, their shape, or the distance separating them that dictates whether a strong magnetic response will be observed at a given frequency. This question can be settled by using the above approach to create inhomogeneous models of these structures which enable to designer to look *inside* the unit cell and gain a better understanding of the physical phenomena giving rise to the macroscopic behavior observed in the experiment (real or numerical), even if the wavelength is sufficiently large that the homogeneous behavior of the structure is not in question.

The second, and perhaps most interesting, possibility opened up by the ideas of this section is the idea that composite metamaterials may be useful even for wavelengths which are *not* large enough for the medium to behave as a homogeneous local medium. For wavelengths which are in the intermediate regime where spatial dispersion holds sway, the structure may be effectively modeled as an *inhomogeneous* effective medium, which may be seen as a *meta-photonic crystal*. This is an interesting new concept, because it makes use of a frequency region hitherto considered useless, and moreover, because this approach may allow the design of photonic crystals previously impossible to realize on the same scale. For instance, one may create novel new effective media, made up of interlocking negative permittivity and negative permeability regions, a possibility which goes far beyond current photonic crystal structures which mainly consist of regions of naturally occurring dielectric with a positive index alternating with regions of index equal to 1 (air holes). Classic numerical tools for the study of photonic crystals such as the MIT Photonic Bands software package will have to be updated to take into account the possibility of *meta-photonic crystals*, that is photonic crystals made of combinations of truly arbitrary alternating media within the unit cell.

## 1.8 Constructive versus holistic

This chapter has been mainly concerned with establishing connections between two main theaters: the micro- and the macro-. In any study of this kind, in which one studies a system which is composed of a large number of smaller and simpler systems, there are two main paradigms, the constructive and the holistic.

The constructive approach endeavors to start with the atoms (in the very general sense of “basic units”), and deduce the properties of the whole. There are generally two steps involved. The first characterizes the behavior of the atoms by analyzing their internal structure. The second step accounts for the interactions between the basic units and establishes the dependence between the parameters characterizing the whole and the parameters characterizing the atoms. The manipulations involved in calculating the parameters of the whole from a knowledge of the internal structure of the atoms and their mutual interaction can be more or less complicated depending on the type of structure. When the relationships connecting the micro- and the macro- parameters are sufficiently simple to be written down in closed form we will speak of a *physical* description of the system because it is ideal as a guide for the intuition. The phenomena involved therefore lend themselves readily to *design*. However when the relationships are more complicated a purely *computational* approach is required. The system still falls within the limits of the theoretical tools available, but it is somewhat removed from the realm of intuition. Design is still possible though it may involve some trial and error.

The *holistic* approach considers the structure as a monolithic whole, which cannot be described in a simple or straightforward way in terms of its components. It becomes useful when the behavior of the whole is sufficiently different from the behavior of the parts that no direct or intuitively useful connection can be drawn. This approach is at best *intensely* computational, and at worst purely phenomenological. Design is properly speaking not possible, except through the groping use of purely trial and error methods.

In view of the preceding discussion we see that the dielectric medium model presented in this chapter is unfortunately a combination of all three of these approaches. The reason for this is that even though the dielectric media can be considered as linear time invariant systems, the subatomic domain is not governed by the same rules as the macroscopic domain. Dielectric media straddle two physical worlds, the classical and the quantum, and if one wishes to remain in the classical domain, then one is forced to account for quantum effects *phenomenologically*. Once the appropriate black box parameters have been introduced (in our case the atomic polarizability  $\gamma_n(\mathbf{x})$ ) the modeling can continue *computationally*, in general. There is also a particular situation where we can dispense with the computation, and where once the quantum effects have been locked away inside the polarizability, the rest of the description simplifies considerably, resulting in the famous *physical* relation known as the Mossotti-Clausius formula. This physical description, however, only applies in the case of highly symmetrical crystals composed of well isolated atoms. If the atoms are larger (or, equivalently, closer to each other) we are forced to revert to the computational description, which would take into account cumbersome higher multipole interaction terms, in order to evaluate the parameter  $\alpha$ .

The non-dipolar coupling parameter  $\alpha$  therefore has purely electromagnetic roots, while the polarizability  $\gamma(x)$  is a parameter that is related to the internal structure of the atoms and is therefore of quantum mechanical origin. The definition of the effective polarizability  $\gamma_{\text{eff}}$  (Eq. 1.32) is therefore conceptually hybrid and *its role is not to express our knowledge of the system, but to hide our ignorance*. The notion of effective polarizability can be extended to all materials when the wavelength is sufficiently large even though in some cases, such as for example that of covalent crystals, where the electron clouds of neighboring atoms overlap considerably, its usefulness or physical meaning is dubious. Macroscopically all that is observed is the polarization per unit volume, and therefore each unit cell can be assigned a dipole moment. But the exact relationship between this dipole moment and the atomic internal struc-

ture and the microscopic field is lost in the homogenization process. This is why the Mossotti-Clausius relation is in reality far less general than it seems. Its generality is implemented artificially by defining the quantities appearing in it in such a way as to satisfy the relation. In many cases, consequently, it is a *reverse-engineered* relation and cannot be seen as a properly *physical* description in the sense in which we use the term here.

In summary, we have seen above that the main reason why a phenomenological parameter had to be introduced was in order to avoid delving into the quantum realm. This was to some extent useful as long as the wavelength was sufficiently large such that the electric field could be considered constant over the size of the support of  $f$  which we refer to as the averaging volume. In the following chapters however we will consider only structures where the atoms, *defined* as the basic scattering units, will be described purely classically. In this case both the subatomic and the macroscopic domains will be governed by the same equations (the macroscopic Maxwell equations) and a complete, though perhaps computational, description will be possible. The need for black box parameters will be eliminated and the rest of this work endeavors to remain within the *constructive* paradigm, trying, wherever possible, to stick to as physical a description as possible.

The next chapter presents a study of dielectric photonic crystals, structures where the wavelength is comparable to the crystal period. The atoms (dielectric rods or holes) are large and interact strongly through higher multipolar fields. A computational approach is required, which we describe and illustrate on several concrete examples.

The final chapter considers the effective medium theory of metallic metamaterials particularly in view of the possibility of designing media which on the macroscopic scale exhibit a negative index of refraction, that is, a negative permeability and a negative permittivity simultaneously. Since we seek to design such media, we take a physical approach, in order to obtain a description of the effective medium that is as explicit as possible. Since the materials required are inevitably three dimensional, as we shall see, the computational approach, though possible, is often impractical for purposes of design.

## 1.9 Beyond dielectrics - truncation revisited

The novelty of the last two decades of research in the field of electromagnetic research springs from the realization that the situations where  $\lambda \approx s$  and  $\lambda > s$  are analogous to the situations where  $\lambda \approx a$  or  $\lambda > a$  with the very important difference that modern technology allows us to *design* the geometry on the  $s$  scale in ways that are impossible on the  $a$  scale. As a result we can control and tailor the behavior of the EM field in ways that were previously impossible. Moreover, working on the  $s$  scale rather than the  $a$  scale has the major advantage that all phenomena can be understood completely using only one set of equations and physical intuitions: the macroscopic Maxwell's equations, whereas, as we saw above, working on the  $a$  scale has the major inconvenient of the inevitability of quantum mechanics for a complete description. We are now standing safely on the solid ground of the macroscopic Maxwell's equations.

Methods for solving them numerically in heterogeneous dielectric media are discussed in the next chapter, while metallic media will be discussed in chapter 3. We first give a general prescription for homogenization of macroscopic structures, and in the rest of this section we consider 1D dielectric structures.

If we are presented with a periodic macroscopic structure which responds to an incident field with a wavelength far larger than the periodicity then it is possible to define effective permittivity and permeability parameters similarly to the prescription of Ref. [26]. Namely, one averages the constitutive relations component-wise:

$$\begin{aligned}\mu_{\text{eff}}^{i,j} &= \frac{\langle B_i \rangle}{\langle H_j \rangle} \\ \varepsilon_{\text{eff}}^{i,j} &= \frac{\langle D_i \rangle}{\langle E_j \rangle}\end{aligned}$$

The brackets indicate spatial averaging, which as discussed above can be seen as the truncation of the spatial Fourier spectrum of the concerned quantity. This is theoretically straightforward, though experimentally cumbersome, since the values of the fields must be known everywhere. The above definitions will become useful in relation to the 1D stack model of composite metamaterials described in Section 3.7.

Let us now attempt to obtain a homogeneous description of a 1D dielectric stack. We consider the constitutive relation:

$$\mathbf{D}(x) = \varepsilon(x)\mathbf{E}(x)$$

We would like to obtain a relationship between the averaged electric field and the averaged displacement,  $[\mathbf{E}(x)]$  and  $[\mathbf{D}(x)]$ . However, convolution is not associative with respect to multiplication except when the multiplication is by a scalar, or a constant. In a 1D medium there are two cases of polarization. When  $\mathbf{E}$  is parallel to the layers then the continuity of  $\mathbf{E}_{\parallel}$  at the interfaces insures that  $\mathbf{E}$  is constant over the whole structure, if the wavelength is large. In this case we can write

$$[\mathbf{D}(x)] = [\varepsilon(x)\mathbf{E}(x)] = [\varepsilon(x)][\mathbf{E}(x)]$$

The effective permittivity is just the average permittivity. However, if the electric field is polarized normal to the layers this is no longer the case.  $\mathbf{E}$  is no longer constant, but  $\mathbf{D}$  is. Therefore, if we divide by  $\varepsilon(x)$  before averaging we can write

$$[\mathbf{E}(x)] = [\varepsilon^{-1}(x)\mathbf{D}(x)] = [\varepsilon^{-1}(x)] [\mathbf{D}(x)]$$

or

$$[\mathbf{D}(x)] = [\varepsilon^{-1}(x)]^{-1} [\mathbf{E}(x)]$$

which implies that the averaged permittivity is given by

$$[\varepsilon^{-1}(x)]^{-1}$$

which is more commonly known as the harmonic mean of the permittivity. In the limit of large wavelength the medium is anisotropic with a diagonal permittivity given by

$$\overline{\varepsilon}_{\text{eff}} = \begin{pmatrix} [\varepsilon^{-1}(x)]^{-1} & 0 & 0 \\ 0 & [\varepsilon(x)] & 0 \\ 0 & 0 & [\varepsilon(x)] \end{pmatrix}$$

keeping in mind that this is for a stack oriented in the  $x$  direction in the limit of  $\lambda \gg d$ .

The essential aspect of this derivation is that we have a quantity that is constant or almost constant which is a product of two quantities that may have rapid variations which ‘‘cancel out’’ in some sense. For instance when  $E$  is normal to the layers  $D$  is constant (or varies very slowly) even though it is the product of two quantities,  $\varepsilon$  and  $E$  which vary quickly. In Fourier terminology, we have a quantity that does not contain high frequencies that is the product of two quantities that do. Truncating the spectra

(another way to see the averaging) therefore leaves the first quantity, in this case  $D$ , unmodified, but it modifies both of the other two, with the result that  $[D] \neq [\varepsilon][E]$ , but rather  $[D] = [\varepsilon^{-1}]^{-1}[E]$ .

But homogenization is not the only situation when we are interested in truncating the spectrum of a given quantity. Another very common situation is when attempting to represent a periodic quantity by its Fourier series in view of using it in a numerical algorithm. Periodic quantities must be approximated by a truncated *finite* subset of their infinite set of Fourier coefficients. Cutting off the Fourier series at some point amounts to imposing a low-pass filter on the signal, in a way that is completely analogous to the homogenization approach described above. In some cases the Fourier spectrum of the quantities involved, notably  $\varepsilon$ , does not extend too far into the high frequencies, but there is one case when it does: when the permittivity has discontinuities. In this case truncation throws away some information no matter how high the cutoff. By truncating the Fourier spectrum of a quantity one implicitly modifies the constitutive relations in an anisotropic way.

In direct space the truncation with a very high cutoff essentially comes down to a sliding average using a very small volume for the  $f(x)$  function. Since the volume is much smaller than the distances over which the fields vary it leaves them unmodified *almost* everywhere. The smoothing is noticeable only around discontinuities. The effect of this smoothing is to replace the normal constitutive relations with the following anisotropic constitutive relations in the vicinity of discontinuities

$$\begin{aligned} \mathbf{D}_{\perp}(x) &= [\varepsilon^{-1}(x)]^{-1} \mathbf{E}_{\perp}(x) \\ \mathbf{D}_{\parallel}(x) &= [\varepsilon(x)] \mathbf{E}_{\parallel}(x) \end{aligned} \tag{1.44}$$

This result has also been obtained with exquisite mathematical rigor by Li [27]. Analogous relations hold for the magnetic constitutive relations. The orientations of the normal and parallel fields depend, of course, on the orientation of the discontinuity at any given point. The numerical implementation of these ideas is discussed in the next chapter.

## Chapter 2

# Photonic crystals – the super-prism effect

### 2.1 Introduction

The initial idea of a photonic crystal, that is, of an artificial dielectric medium patterned at the scale of optical wavelengths, was proposed simultaneously at the end of the 1980's by Eli Yablonovitch of Bell Labs [1] and Sajeev John of Princeton University [2]. It was the result of work which, at the time, was trying to improve the performance of lasers by finding ways of controlling their main limiting factor: spontaneous emission. By analogy with the well known electronic band-gaps responsible for the behavior of most dielectrics and semiconductors in the optical range, it was hoped that spontaneous emission could be controlled by strongly limiting the density of states available to it. It thereby seemed possible to improve laser characteristics considerably. Once the idea was put forward it quickly became clear that more efficient lasers is only one of the many possible applications and of the wealth of remarkable optical phenomena which can be realized in photonic crystals.

During the following fifteen years photonic crystal research was dominated by the study and characterization of photonic band gaps (PBGs) and associated phenomena, most especially the confinement and guiding of light on a scale close to the fundamental limit, the wavelength. It was imagined that optical integrated circuits, filters, multiplexers, routers, and of course, lasers could all be realized by using the remarkable properties resulting from the existence of PBGs.

More recently, though, another idea was put forward. Photonic crystals could be used not only as photonic “insulators” but as photonic “conductors”. Aside from the properties related to the PBGs, photonic crystals also exhibit remarkable light *transmission* properties. Surprising phenomena were predicted and subsequently experimentally confirmed, such as negative refraction, the super-prism effect [28, 29, 30], and propagation with very low group velocities (slow light), making possible the amplification of non-linear effects [31, 32]. Whereas the PBG related properties of photonic crystals are due to the absence of propagating modes at a given frequency within the crystal, these latter phenomena are due to the peculiar properties of the propagating modes when they *do exist* in the crystal.

In the following section we present succinctly the methods most widely used for the study of photonic crystals. The rest of the chapter will be concerned with an investigation of the super-prism effect with particular emphasis on crystals with rectangular unit cells.

## 2.2 Numerical Methods

Photonic crystals are structures which function in the regime  $a \ll s \approx \lambda$  where  $a$  represents the atomic size,  $\lambda$  the free space wavelength and  $s$  the scale of the unit cell, or the scale on which the medium is structured. Consequently, as argued in the preface, the notions of index of refraction or impedance are of little use. However, since macroscopic dielectric structures are described accurately by the macroscopic Maxwell's equations, it is possible to solve for the total fields through a computational method.

The macroscopic electric and magnetic fields are continuous quantities in the sense that they have a value at every point in space. Computers, however, can only deal with discrete data sets. Consequently the Maxwell's equations must be discretized. Time and space can be discretized in either direct or reciprocal space, giving four possible combinations:  $(\mathbf{x}, t)$ ,  $(\mathbf{k}, t)$ ,  $(\mathbf{x}, \omega)$ ,  $(\mathbf{k}, \omega)$ . As it turns out only three of them have been applied to any significant extent.

Discretizing in real space is appropriate when the structure is not periodic, which is the case particularly when one is dealing with a finite structure. In this case one takes a scattering approach to the problem and has the option between a time or frequency domain treatment. Of the  $(\mathbf{x}, t)$  methods the most widespread is the Finite Difference Time Domain method, which we discuss briefly below. Frequency domain  $(x, \omega)$  methods are also available. This is the case of harmonic finite element (FEM) solvers such as those provided with commercial packages such as COMSOL MULTIPHYSICS or CST MICROWAVE STUDIO. We will use these in chapter 3.

When the structure under study is periodic then it is natural to discretize in reciprocal space since the Fourier transform of a periodic quantity is discrete. However periodic structures are implicitly infinite, which means that all fields must come from infinity (both in space and in time). The temporal Fourier transform is likewise appropriate. Of the various  $(\mathbf{k}, \omega)$  methods available the most popular is the plane wave method as implemented in the MIT PHOTONIC BANDS package which is described below. No  $(\mathbf{k}, t)$  method has been put forward this far to the author's knowledge but equivalent results can be obtained by using a  $(\mathbf{k}, \omega)$  implementation, running it many times and performing the inverse temporal Fourier transform on the *results*.

The methods mentioned above are usually quite general, since they start directly with the Maxwell's equations and do not rely crucially on any other major assumptions. However, when dealing with structures with certain kinds of symmetry one has also the option of using a method adapted for that particular type of geometry. Examples are transfer/scattering matrix methods, or fictitious sources. We shall see an example of a scattering matrix method in detail in chapter 3 and Appendix A.

We now give a short presentation of the principle of the plane wave method of the MPB package.

### 2.2.1 The plane wave expansion method

In a periodic medium the fields obey Bloch's theorem, and they take the form of Bloch waves. For instance the magnetic field can be written

$$\mathbf{H}(\mathbf{x}) = \mathbf{U}_{\mathbf{k}_B}(\mathbf{x})e^{i\mathbf{k}_B \cdot \mathbf{x}}$$

where  $\mathbf{U}_{\mathbf{k}_B}$  is a function which has the periodicity of the lattice,

$$\mathbf{U}_{\mathbf{k}_B}(\mathbf{x}) = \mathbf{U}_{\mathbf{k}_B}(\mathbf{x} + n_1\mathbf{T}_1 + n_2\mathbf{T}_2 + n_3\mathbf{T}_3), \quad n_1, n_2, n_3 \in \mathbb{Z}.$$

Here  $\mathbf{k}_B$  is the Bloch wavevector, as in the first chapter, and  $\mathbf{T}_{1,2,3}$  are any set of direct lattice vectors. Since the function  $\mathbf{U}_{\mathbf{k}}$  is periodic it can be decomposed in a Fourier series, which in general may be three

dimensional. The magnetic field can be written

$$\mathbf{H}_{\mathbf{k}_B}(\mathbf{x}) = \sum_{\mathbf{G}} \mathbf{H}_{\mathbf{G}} e^{i(\mathbf{k}_B + \mathbf{G}) \cdot \mathbf{x}} \quad (2.1)$$

where  $\mathbf{G}$  are the reciprocal lattice vectors and  $\mathbf{H}_{\mathbf{G}}$  are the Fourier coefficients. Physically this expression can be interpreted as considering the Bloch wave as a superposition of plane waves, each with a weight given by the corresponding Fourier coefficient. The expansion is justified because the plane waves form a complete basis of piecewise continuous periodic functions. A knowledge of the Fourier coefficients is therefore sufficient to completely specify the magnetic field inside the structure. However, the series above is infinite, meaning that the Fourier coefficients form a discrete but infinite set. An exact numerical treatment is therefore not possible and a truncation of the Fourier series is required. The field is therefore approximated by a finite series where the plane waves with very large wavevectors are neglected.

As discussed in section 1.9 the truncation of the Fourier series has the effect of modifying the constitutive relations which connect the different fields and the permittivity and the permeability, especially when these parameters contain discontinuities. Moreover the constitutive relations are modified in a way which *depends on the structure*, particularly on the orientation and position of the discontinuities in the two parameters. There are two options available to deal with this aspect.

First, one may implement the method using *one or the other* of the two following constitutive relations:

$$\begin{aligned} \mathbf{D} &= [\varepsilon] \mathbf{E} \\ \mathbf{D} &= [\varepsilon^{-1}]^{-1} \mathbf{E} \end{aligned}$$

The first implementation will be efficient when the electric field is parallel to discontinuities while the second implementation will be efficient when the field is normal to them. For instance, in 2D calculations when the electric field is chosen to be parallel to the infinite direction ( $\mathbf{E}_{\parallel}$  polarization) the first implementation will be optimal. However, in the  $\mathbf{H}_{\parallel}$  polarization the second implementation will be preferable, though *not optimal*, because the electric field will not be normal to the discontinuities everywhere. There will be regions where  $\mathbf{E}$  has a component that is parallel to the discontinuities, leading to slow convergence. In 3D the convergence is sub-optimal for all polarizations.

The advantage of this approach, however, is that it is rigorous in the sense that by letting the number of plane waves go to infinity (increasing the cutoff of the truncation) the solution will approach the rigorous solution for the Maxwell's equations in the given structure, though non-uniformly. In other words the results may not be very reliable for low cutoffs, but one is assured to simulate the actual structure, including the discontinuities in the permittivity and/or permeability. This approach was taken, for instance, by Cassagne [33].

The second possibility is to implement a single algorithm, using a single constitutive relation, but with a slightly modified permittivity:

$$\mathbf{D} = [\varepsilon_m] \mathbf{E}$$

As we have seen in the previous chapter, truncating the Fourier series of a given quantity is equivalent to taking a sliding average of that quantity. The two constitutive relations of Eq. (1.44) can therefore be interpreted as a kind of anisotropic smoothing of the permittivity. Roughly speaking one takes the arithmetic average for components parallel to the discontinuity and the harmonic average for components normal to the discontinuity. Even if the permittivity is scalar everywhere in the original structure, the new structure will have a permittivity that is smooth but anisotropic, at least in the regions where it changes abruptly.

This approach has the advantage that the same algorithm is optimally efficient for all polarizations (in the sense that it converges uniformly to a solution), which includes 3D calculations, whereas the previous approach was optimal only in some 2D or 1D cases. However, since the permittivity that is entered into the computation has been modified to smooth over the discontinuities, this method converges faster, but *to something different* from the rigorous mathematical solution for the original non smoothed structure. Any features of the solution which depend in any essential way on the discontinuities of the permittivity will have been lost.

Experience has shown that the second of the two approaches is largely preferable, and it is the approach that we use in the rest of this work. It is the method adopted by the MPB package, which we have used to obtain the results of this chapter, and which is described in detail by Johnson [34] and very briefly below.

The equations obeyed by the field in the structure are the macroscopic source free (a.k.a homogeneous) Maxwell's equations:

$$\begin{aligned}\nabla \times \mathbf{H}(\mathbf{x}) + i\omega\mathbf{D}(\mathbf{x}) &= 0 \\ \nabla \cdot \mathbf{H}(\mathbf{x}) &= 0 \\ \nabla \times \mathbf{E}(\mathbf{x}) - i\omega\mathbf{B}(\mathbf{x}) &= 0 \\ \nabla \cdot \mathbf{D}(\mathbf{x}) &= 0\end{aligned}$$

where the medium is nonmagnetic and lossless and the constitutive relations take the form

$$\begin{aligned}\mathbf{B}(\mathbf{x}) &= \mu_0\mathbf{H}(\mathbf{x}) \\ \mathbf{D}(\mathbf{x}) &= \varepsilon_0\varepsilon_m(\mathbf{x})\mathbf{E}(\mathbf{x})\end{aligned}$$

The relative permittivity is position dependent but real and positive. The time dependence of all fields is in  $e^{-i\omega t}$ . Eliminating  $\mathbf{E}$  from the two curl equations we obtain a second order linear equation in  $\mathbf{H}$

$$\nabla \times (\varepsilon_m(\mathbf{x})^{-1}\nabla \times \mathbf{H}(\mathbf{x})) = \frac{\omega^2}{c^2}\mathbf{H}(\mathbf{x}) \quad (2.2)$$

This equation must now be Fourier transformed. In reciprocal space the curl operator becomes

$$\nabla \times \xrightarrow{\mathcal{F}} \mathbf{k} \times$$

while the product with the inverse permittivity becomes a convolution with  $\widetilde{\varepsilon_m^{-1}}(\mathbf{k})$ . The exact expression for  $\varepsilon_m$  in terms of the original, possibly discontinuous,  $\varepsilon$  can be found in Ref. [34]. The Fourier transform of the field  $\mathbf{H}$  can be written

$$\widetilde{\mathbf{H}}(\mathbf{k}) = \sum_n \mathbf{H}_{\mathbf{G}_n} \delta(\mathbf{k} - \mathbf{k}_B - \mathbf{G}_n) \quad (2.3)$$

By introducing this expression in Eq. (2.2) we obtain a standard eigenvalue problem in the unknowns  $\mathbf{H}_{\mathbf{G}_n}$

$$\hat{\Theta}_{\mathbf{k}_B} \mathbf{H}_{\mathbf{G}_n} = \frac{\omega^2}{c^2} \mathbf{H}_{\mathbf{G}_n}$$

where the operator  $\hat{\Theta}$  is given by

$$\hat{\Theta}_{n,l} = (\mathbf{k}_B + \mathbf{G}_l) \times \widetilde{\varepsilon_m^{-1}}_{n,l}(\mathbf{k}_B + \mathbf{G}_n) \times$$

If we truncate the sum in Eq. (2.3) to  $N$  terms, then the above operator will take the form of a  $3N \times 3N$

matrix. The permittivity is represented by a  $3N \times 3N$  matrix which is Toeplitz-like in the sense that it is constructed through repeated cyclic permutations of the same  $N \times 3 \times 3$  matrices containing the Fourier coefficients of the tensor  $\varepsilon_m(\mathbf{x})$ .

This classic eigenvalue problem can be further simplified. The fact that electromagnetic waves are transverse waves as required by the electric and the magnetic Gauss's laws, Eqs. (1.38) and (1.40) means that the problem can be reduced from a  $3N \times 3N$  problem to a  $2N \times 2N$  problem by eliminating the longitudinal degree of freedom of the fields. Also, since we are usually interested only in the first few lowest eigenvalues (or bands, in solid state physics terminology) it is possible to apply efficient iterative methods that do not require the direct diagonalization of the whole  $2N \times 2N$  problem. These aspects are described in detail in Ref. [34] and references therein.

### 2.2.2 The finite difference time domain method

The FDTD method is one of the first numerical methods used to solve Maxwell's equations. It samples the fields in direct space and in the time domain, so no Fourier transforms are taken. The curl equations are

$$\begin{aligned}\nabla \times \mathbf{H}(\mathbf{x}, t) &= \frac{\partial \mathbf{D}(\mathbf{x}, t)}{\partial t} \\ \nabla \times \mathbf{E}(\mathbf{x}, t) &= -\frac{\partial \mathbf{B}(\mathbf{x}, t)}{\partial t}.\end{aligned}$$

Since real media do not respond to applied fields instantaneously the dependence of the displacement field at any given time depends on the electric field at all times prior. Consequently, assuming a local response, we must write

$$\mathbf{D}(\mathbf{x}, t) = \varepsilon_0 \int_{-\infty}^t \varepsilon(\mathbf{x}, t - t') \mathbf{E}(\mathbf{x}, t') dt'.$$

The displacement field is then represented as a time convolution of the electric field with a time dependent permittivity. This must be the starting point of a formulation of the FDTD method that would take into account the frequency dispersion of real materials. For illustration purposes in this section we assume the medium response is instantaneous, which is to say that the permittivity takes the form of a Dirac-delta in time:  $\varepsilon(\mathbf{x}, t) = \varepsilon(\mathbf{x})\delta(t)$ . This is a good approximation when the electric fields in a given simulation cover a sufficiently small frequency domain. A similar argument applies to the magnetic constitutive relation giving the non physical but numerically convenient

$$\begin{aligned}\nabla \times \mathbf{H}(\mathbf{x}, t) &= \varepsilon(\mathbf{x}) \frac{\partial \mathbf{E}(\mathbf{x}, t)}{\partial t} \\ \nabla \times \mathbf{E}(\mathbf{x}, t) &= -\mu(\mathbf{x}) \frac{\partial \mathbf{H}(\mathbf{x}, t)}{\partial t}.\end{aligned}$$

The electric and the magnetic field are sampled at points in two cubic lattices which are offset from each other by a half step in each Cartesian direction, as shown in Fig. (2.1) representing a Yee cell [35]. The time evolution of the electric field depends on the spatial variation of the magnetic field at neighboring points and vice-versa. The meaning of this can be made clearer by writing out one of the six scalar equations that make up the two vectorial curl Maxwell's equations. We consider the  $x$  component of the  $\mathbf{E}$  field. We have

$$\varepsilon(\mathbf{x}) \frac{\partial E_x(\mathbf{x}, t)}{\partial t} = \frac{\partial H_z(\mathbf{x}, t)}{\partial y} - \frac{\partial H_y(\mathbf{x}, t)}{\partial z}$$

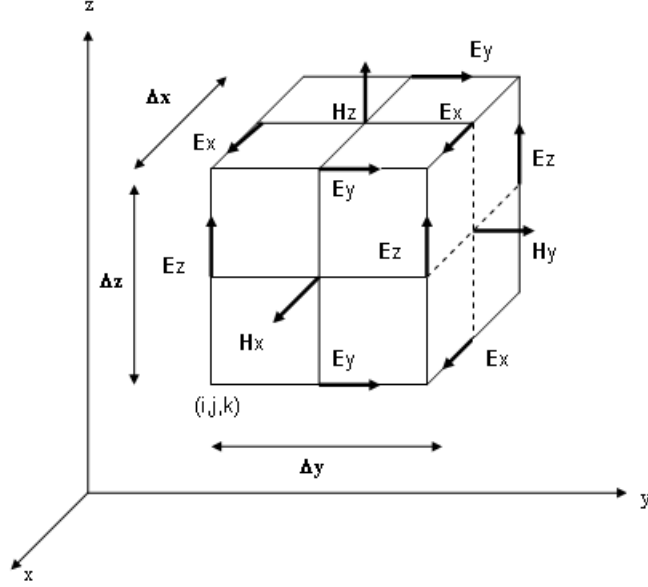


Figure 2.1: The Yee lattice.

This can be rewritten in terms of the fields sampled over the Yee lattice the following way

$$\varepsilon_{x(i,j,k)} \frac{E_{x(i,j,k)}^{n+1} - E_{x(i,j,k)}^n}{\Delta t} = \frac{H_{z(i,j+1,k)}^{n+\frac{1}{2}} - H_{z(i,j,k)}^{n+\frac{1}{2}}}{\Delta y} - \frac{H_{y(i,j,k+1)}^{n+\frac{1}{2}} - H_{y(i,j,k)}^{n+\frac{1}{2}}}{\Delta z}$$

or

$$E_{x(i,j,k)}^{n+1} = E_{x(i,j,k)}^n + \frac{\Delta t}{\varepsilon_{x(i,j,k)} \Delta y} \left( H_{z(i,j+1,k)}^{n+\frac{1}{2}} - H_{z(i,j,k)}^{n+\frac{1}{2}} \right) - \frac{\Delta t}{\varepsilon_{x(i,j,k)} \Delta z} \left( H_{y(i,j,k+1)}^{n+\frac{1}{2}} - H_{y(i,j,k)}^{n+\frac{1}{2}} \right)$$

where at time step  $n$  the unknown quantity  $E_{x(i,j,k)}^{n+1}$  is expressed in terms of the known quantities. Analogous equations can be written for the other five field components and the solution thus evolves from one time step to the next. Remember that the lattices over which the electric and the magnetic field are sampled are shifted, so that for a Yee cell of size  $l$  if  $E_{x(i,j,k)} = E_x(\mathbf{x}_0)$  then  $H_{y(i,j,k)} = H_y(\mathbf{x}_0 - \frac{l}{2}\hat{x} - \frac{l}{2}\hat{y} - \frac{l}{2}\hat{z})$ . A detailed presentation of the FDTD method and its many variations and improvements is available in the authoritative text by Taflové and Hagness [36].

In this work we have used the commercial implementation FULLWAVE/BEAMPROP by RSOFTE.

## 2.3 Iso-frequency curves – the construction line

The superprism effect is the very abrupt change of direction of light at the interface between a homogeneous medium and a photonic crystal. In particular the direction of propagation of light inside the photonic crystal can change dramatically for only a small variation in the angle of incidence on the interface.

In the absence of a well defined index of refraction for the photonic crystal medium the refraction phenomenon must be studied in terms of a more general form of the Ibn Sahl-Snell-Descartes rules<sup>1</sup>. This general rule states that the tangential component of the wavevector is conserved when a plane wave traverses an interface. It is easy to show that when both media are homogeneous this reduces to the

<sup>1</sup>Though rarely acknowledged, the rules of optical refraction at a plane interface were first obtained and used by Ibn Sahl in the year 984 C.E. [37].

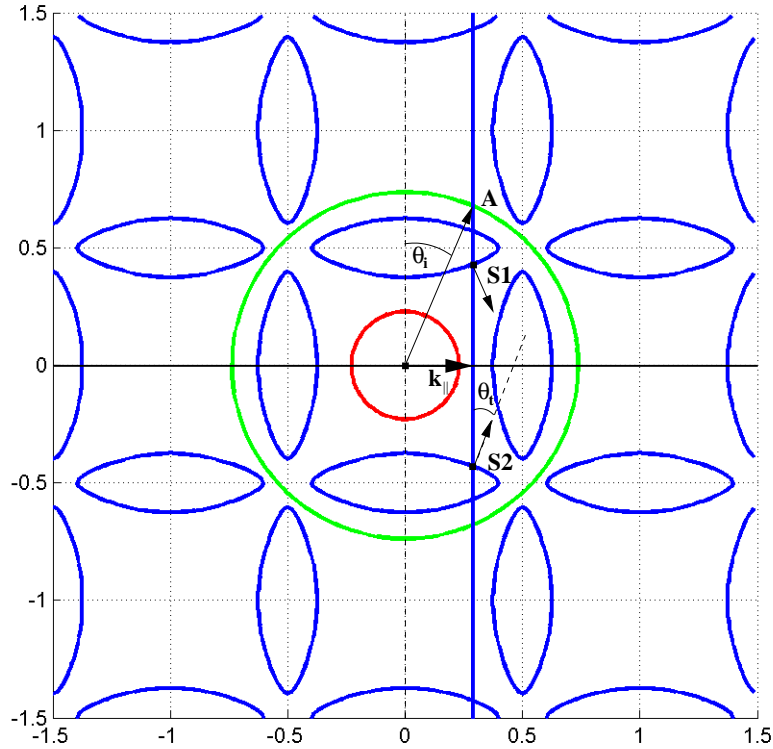


Figure 2.2: Construction line figure.

classic form of the Ibn Sahl-Snell-Descartes rules. However, when one or both of the media are photonic crystals, the more general rule must be used. For a field of given frequency incident on an interface at a given angle, coming from some medium A, the rules select one or more allowed wavevectors inside the medium B which are consistent with the above conservation law. In order to determine these allowed wavevectors it is necessary to have a knowledge of all allowed wavevectors at the given frequency. The diagram of all allowed wavevectors in a photonic crystal at a given frequency is referred to as the iso-frequency diagram, or the iso-frequency curves. The reciprocal plane is infinite but the periodicity of the structure means that all the required information is contained in the first Brillouin zone. In the following we will see that it is advantageous to plot the iso-frequency curves over several neighboring Brillouin zones as well.

Before going any further we should say something about the units. In this chapter we use the fact that Maxwell's equations are scale invariant. Consequently the exact dimensions are not important, only the ratios between them. In the following we will measure lengths in units of the period of the structure under study, which we call  $a$ . Moreover, it is a widely used convention on iso-frequency diagrams to express the frequency in units of  $a/\lambda$ . As concerns the wavevectors we have adopted the convention used by MPB, which expresses wavevectors in units of  $2\pi/a$ .

In order to determine the direction of propagation of the transmitted fields inside a photonic crystal we use the method of the construction line, as illustrated on the iso-frequency diagram of Fig. 2.2. In this figure we see superimposed the isofrequency diagrams of the two media on each side of the interface. The wave in this case is incident from a homogeneous isotropic medium, which therefore has an isotropic isofrequency curve in the form of the green circle. The smaller red circle is the iso-frequency curve of free space. The ratio of the radii of the two circles is the refractive index of the homogeneous medium. Another, more important, reason for plotting the red circle (or the light cone) is discussed in the next section.

The blue curves are the iso-frequency curves of a photonic crystal consisting of a 2D square lattice of circular air holes of radius  $r$  in a homogeneous background with index  $n = 3.21$ . The interface between the media is represented by the thick horizontal black line. At a point of wavevector  $k$  the wave propagates in a direction given by the group velocity of the Bloch wave at that point. It can be shown that at any point on a given iso-frequency curve the group velocity is directed along the line perpendicular to the curve. In isotropic media (e.g. green circle) this direction is collinear with the direction of the wavevector, but in strongly anisotropic media like photonic crystals this is no longer the case, as can be seen by inspecting the blue curves of Fig. 2.2.

We consider a plane wave incident on the interface at the angle  $\theta_i$ . The wavevector of the incident wave is represented by point A in Fig. 2.2 and  $k_{\parallel}$  is the component of A which is parallel to the interface. The vertical line indicates that the wavevectors allowed inside the crystal must have a horizontal component equal to  $k_{\parallel}$ . The points of intersection of this vertical line with the iso-frequency curves of the photonic crystal are the set of allowed wavevectors inside the crystal. Only two of these points are within the first Brillouin zone. The direction of propagation of the waves at the two allowed points are indicated by the vectors  $S1$  and  $S2$ , the group velocities. Since the incident wave is propagating upwards, the transmitted wave must also propagate upwards, and we must choose  $S2$ . The wave is therefore transmitted with an angle  $\theta_t$ . The restriction of choosing only waves propagating away from the interface in the second medium has eliminated one of the two possible solutions. In general, for media with mirror symmetry the number of solutions allowed by the construction line is even, but half of them are eliminated by the above condition.

Since the direction of propagation of the waves inside the crystal is not given by the wavevector but by the group velocity, the super-prism effect is obtained by exploiting a region where the iso-frequency curves exhibit an abrupt change of direction. In this way a small change in incident wavevector can lead to a large variation in the direction of propagation inside the photonic crystal.

### 2.3.1 The effective index approximation

Since in the following we will consider only two dimensional photonic crystals, we must also say a few words on the third direction. In theory we take it to be infinite, but this is clearly not realistic. In fact experiments are carried out on dielectric slabs whose thickness is on the same order of magnitude as the crystal period. We need a way to take into account the effect of the finite height of the photonic crystals. Rigorously speaking this must be done using three dimensional calculations, which are heavy to implement as they require considerable computing power and time. An example of a three dimensional study of photonic crystal slabs is available, for instance in Ref. [38].

However, it is possible to obtain results similar to the rigorous calculations with much less time consuming methods using the effective index approximation. This approximation states that at least for the lower bands the modes of the finite structure can be obtained by a *two-dimensional* calculation by using slightly different value of the refractive index of the dielectric slab in which the photonic crystal (the air holes) is etched. Since in a finite structure the fields are expected to extend to a certain extent into the air half spaces above and below the slab, it seems reasonable to assume that, to first order, this will have the effect that the field will see a somewhat lower refractive index than for the two-dimensional structure. One way to approximate this effective index is to consider a homogeneous dielectric slab of index  $n$  and thickness  $h$  and to calculate the dispersion relation of its guided modes. Any given mode propagating in the slab can be assigned an effective index given by

$$n_e = \frac{\beta}{k}$$

where  $\beta$  is the wavevector of the slab mode and  $k$  is the free space wavevector at the same frequency.

If the waveguide is symmetrical then there is no cutoff for the two lowest modes, usually designated as the  $TE_0$  and the  $TM_0$  modes. There is a frequency region where these two modes are the only two modes confined to the slab. The  $TE_0$  modes have an electric field parallel to the slab, and a magnetic field perpendicular to it, while the  $TM_0$  mode has a magnetic field parallel to the slab and an electric field perpendicular to it. The effective index of the  $TM_0$  mode is generally lower than the  $TE_0$  mode which means that the  $TE_0$  mode is the better confined of the two. Consequently it is the mode that is generally used. In fact, it is convenient to draw an analogy between the  $TM_0$  and the  $TE_0$  modes of the finite slab and the  $\mathbf{E}$  and  $\mathbf{H}$  modes of the two-dimensional photonic crystal structure respectively. The former are the finite height version of the latter. Consequently when calculating the dispersion relation of a photonic crystal in polarization  $\mathbf{H}$  one must use the effective index of the  $TE_0$  mode to account for the finite height of the structure. The index used for the rest of this chapter is that given in Ref. [39] for the In-P slab used in their etching process:  $n_{\text{eff}} = 3.21$ .

There are two aspects that must be kept in mind when modeling finite height structures using 2D simulations.

First, the effective index approximation is good only below the cutoff of the higher order  $TE_1$  and  $TM_1$  slab modes. We only expect the few lowest bands of the photonic crystal dispersion relation to be reproduced accurately using this method. For higher bands full three-dimensional calculations are required.

Second, vertical losses due to poorly confined slab modes are not accounted for. In fact, the red circle mentioned in the previous section delimits the region of reciprocal space where the fields are not confined within the PC slab and the modes are lossy. This is why the region inside the red circle is to be avoided.

## 2.4 The super-prism effect

In order to find a configuration appropriate for observing the super-prism effect several conditions must be satisfied.

- Optimizing vertical confinement requires us to use  $\mathbf{H}$  polarized fields propagating in slabs containing air holes (rather than rods suspended in air or in a lower index medium).
- There must be an abrupt change of group-velocity for a small change of wavevector.
- There must be only one transmitted mode, giving one transmitted beam.
- The propagating mode inside the crystal cannot be a lossy mode.
- Practical fabrication considerations limit us to considering *circular* air holes.

In order to satisfy these requirements we have a series of knobs, or degrees of freedom:

- The shape of the unit cell, given by its aspect ratio,  $x = \frac{|\mathbf{T}_2|}{|\mathbf{T}_1|}$  where  $T_1$  and  $T_2$  are the lattice vectors, which are perpendicular to each other, and  $|\mathbf{T}_2| \geq |\mathbf{T}_1| = a$ .
- The frequency, specified by the normalized frequency parameter  $f = \frac{a}{\lambda}$  where  $\lambda$  is the free space wavelength.
- The hole radius, specified by the normalized parameter  $r = \frac{R}{a}$ .

We have used the length of the shorter lattice vector as the length unit.

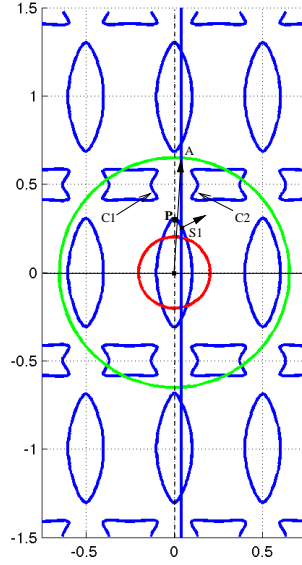


Figure 2.3: Iso-frequency diagram for  $r = 0.25$ ,  $x = 2$ ,  $f = 0.203$ . For an angle of incidence is  $\theta_i = 3.52^\circ$  (point A) the angle of refraction is  $\theta_t = 63.4^\circ$ .

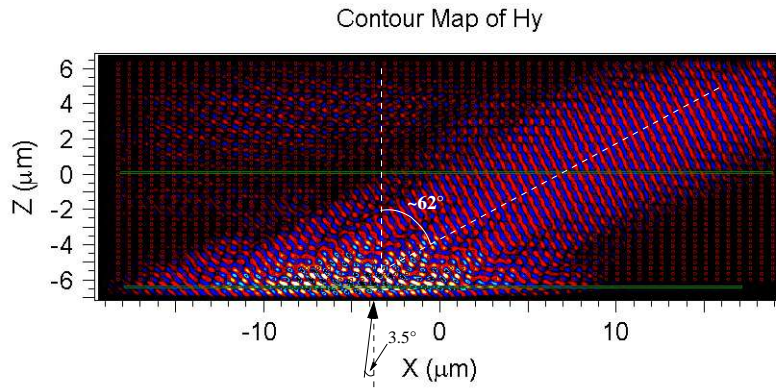


Figure 2.4: FDTD simulation of the configuration of Fig. 2.3. We have a refracted beam at  $62^\circ$  for an incident beam at  $3.5^\circ$  from the normal.

By suitably adjusting these knobs it is possible to find a crystal satisfying all of the above constraints [40]. The isofrequency curves are represented in Fig. 2.3. For small angles of incidence we see that we have a single transmitted mode with group velocity  $S1$ . When changing the angle of incidence several degrees around normal incidence we can obtain a very large variation in the direction of the group velocity. A calculation of the group velocity  $S1$  at this frequency gives  $(0.25, 0.1252)$  in units of  $c$ , giving an angle of refraction  $\theta_t = 63.4^\circ$  for an angle of incidence of  $\theta_i = 3.52^\circ$ . Since the medium is symmetrical with respect to the normal to the interface, by changing the angle of incidence by  $7^\circ$  we can obtain a variation of the angle of refraction of over  $126^\circ$ . The FDTD simulation confirms this result as can be seen in Fig. 2.4. For an angle of incidence of 3.5 degrees we obtain an angle of refraction of approximately 62 degrees. This is in good agreement with the prediction based on the dispersion relation of the infinite crystal.

The FDTD simulation also illustrates another important aspect of the super-prism effect. The main criterion for observing it is that there be a strong variation of the group velocity over a small region in reciprocal space. Consequently, if the incident field contains a range of spatial frequencies that is of a size similar to this region, then the effect will be difficult to see due to the fact the some components

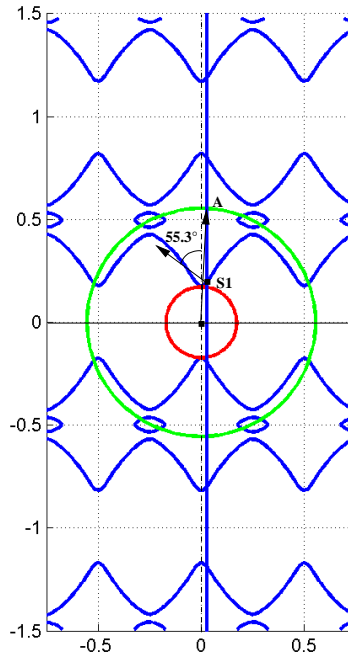


Figure 2.5: An alternative configuration which satisfies the conditions set out above, but at a lower frequency. Here we have  $x = 2$ ,  $r = 0.3$  and  $f = 0.173$ . For an angle of incidence of  $3.5^\circ$  we have an angle of refraction of  $-55.3^\circ$ .

of the incident field will refract differently from other components resulting in a *fanning out* of the refracted beam inside the crystal. Rather than selecting a point on a iso-frequency curve a whole section will be excited, corresponding not to a construction *line* but to a construction *band*. In this case it is no longer possible to define a refraction angle properly speaking. Consequently, in order to observe the super-prism effect one must use very large beams, corresponding to narrow spatial frequency profiles in reciprocal space in order to be able to select a single point rather than a whole section of a given iso-frequency curve. This is why the FDTD simulations in this chapter use very wide beams incident on very wide crystals. In the case of Fig. 2.4 the incident Gaussian beam has width  $W \approx 50a$ .

But the configuration which uses the elliptical curve at the center of Fig. 2.3 is not the only possible configuration. Another possibility exists, which also has the advantage that it illustrates a remarkable refraction phenomenon in photonic crystals: negative refraction. This configuration is represented on the iso-frequency diagram of Fig. 2.5. The incident field at point *A* is going to the right, but the refracted field has a group velocity traveling to the left. The angle of refraction is therefore negative. This is seen also on the FDTD simulation of Fig. 2.6. The incident beams in Figs. 2.4 and 2.6 are identical but in the first case the transmitted beam is going right, while in the second case the transmitted field is going left.

## 2.5 Transmission efficiency

There is one drawback to the iso-frequency diagrams that we use to find useful configurations for the super-prism effect. The diagrams tell us what wavevectors are allowed inside the crystal, consistent with the continuity conditions at the interface, but they do not tell us how efficiently they are transmitted. It is not possible to obtain this information simply from the iso-frequency curves. What is needed is a more detailed knowledge of the underlying Bloch modes and their field profiles.

A situation illustrating this aspect is the crystal with the iso-frequency diagram of Fig. 2.3 illuminated

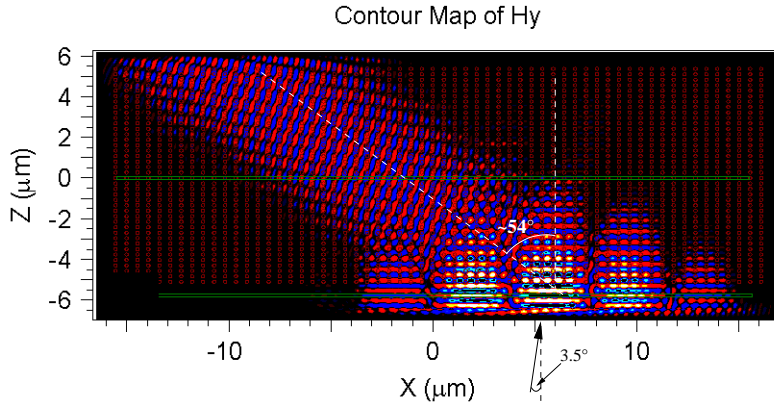


Figure 2.6: FDTD simulation corresponding to the configuration of Fig. 2.5. The refraction angle is approximately  $54^\circ$  which is in excellent agreement with the iso-frequency diagram prediction.

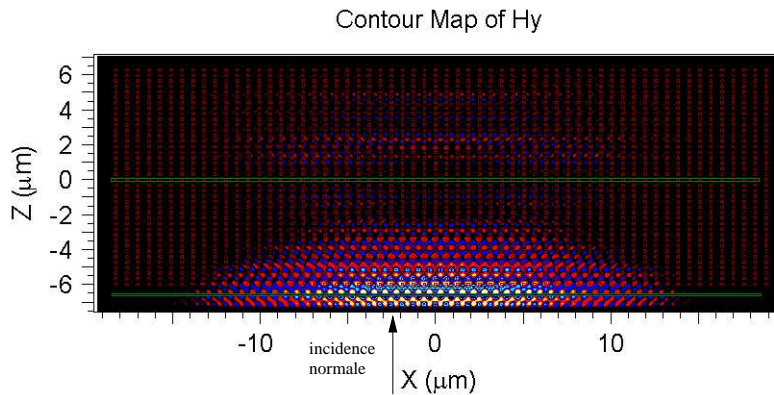


Figure 2.7: The same crystal as in Figs. 2.3 and 2.4 is illuminated in normal incidence. The light is completely reflected.

at normal incidence. The construction line would indicate that the field should couple to the mode at point P and propagate straight up, undeviated. The FDTD simulation of Fig. 2.7 indicates that this is not the case. In fact the field is completely reflected at the interface.

This phenomenon can be understood by looking at the field profile of the mode at point P, see Fig. 2.8. It is clear that since the incident wave is directed along a nodal plane of the Bloch mode, the incident field and the Bloch wave are therefore orthogonal and the wave is completely reflected. This is also the reason why the transmission for small angles of incidence as in Fig. 2.4 is quite small as can be seen in Fig. 2.9.

We must therefore find a way to avoid normal or near-normal incidence but still somehow couple to the Bloch modes near the top of the ellipse shaped iso-frequency curve of Fig. 2.3. If we take a closer look at the field profile of Fig. 2.8 we notice that even though a normally incident field is incompatible with the field profile, a plane wave incident obliquely may have a better chance of coupling efficiently to this mode. This is illustrated in Fig. 2.10. A plane wave lined up with the red and blue rows of the field profile would seem more promising. This optimal angle is estimated by inspection of Fig. 2.10 to be around  $\arctan(3/2) \approx 56^\circ$ . However, such a plane wave would couple to this mode only if the  $k_{\parallel}$  conservation at the interface is respected. As it turns out it is possible to satisfy this condition as illustrated on the iso-frequency diagram of Fig. 2.11. An incidence angle of  $50.12^\circ$  makes it possible to couple to the same Bloch mode as in normal incidence. This is made possible by the anisotropy of the unit cell. The transmitted efficiency for this case is over 40%. For this configuration the angle at which

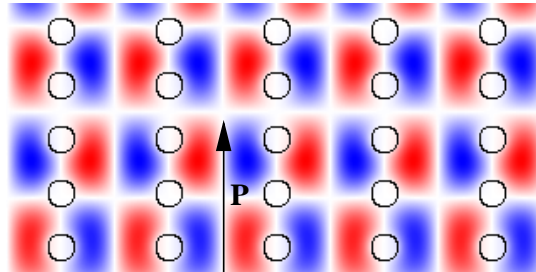


Figure 2.8: The field map of the Bloch mode at point P in Fig. 2.3. For a plane wave incident vertically the coupling is zero due to the symmetry.

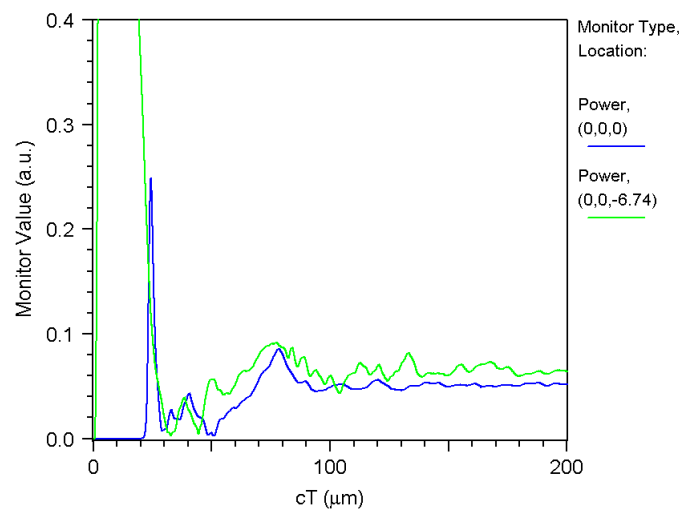


Figure 2.9: Transmitted power detected numerically in the FDTD simulation of Fig. 2.4 just below (green) and within (blue) the photonic crystal as a function of time. The scale is normalized to the power of the incident beam, and the transmission is seen to stabilize at a steady state value around 6%.

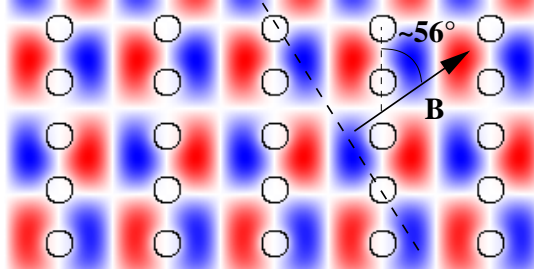


Figure 2.10: An incident beam at a larger angle would be more likely to couple efficiently to the modes close to point P, but of course, only if the conservation of the tangential component of the wavevector allows it. An incidence around  $56^\circ$  seems the most promising.

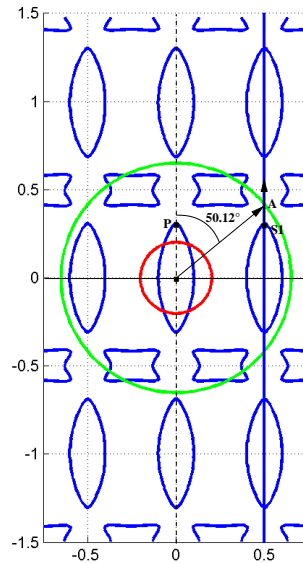


Figure 2.11: Another option for coupling to mode P. The parameters are the same as in Fig. 2.3 but the angle of incidence is  $50.12^\circ$ .

the coupling is allowed by the boundary condition is not exactly equal to  $56^\circ$  but it is close enough to be a real improvement.

As we increase the angle of incidence we observe similar deviation of the transmitted beam as for near normal incidence. For an angle of incidence of  $53.5^\circ$  we observe a refracted angle of around  $54^\circ$  and a transmission efficiency of almost 80%.

We have therefore used the particular features of the isofrequency diagram together with a knowledge of the field profile of the mode to find a near-optimal configuration for observing the super-prism effects. In particular, due to the anisotropy of the unit cell the transmission can be improved from between 0 and 6 percent for near normal incidence to between 40 and 80 percent for incidence around  $\theta_i = 53^\circ$ .

## 2.6 The incident beam width

In the previous section we have discussed the super-prism effect for the case of very large beam widths. A beam that is large in real space is very narrow in reciprocal space meaning that it is well approximated by a single point. The beam can be thought of as a plane wave and the fact that it contains more than just a single spatial frequency can be ignored. In this section we discuss the issues that appear when the size of the beam is decreased.

Let us consider the example of Fig. 2.7 of a beam incident normally on the photonic crystal with

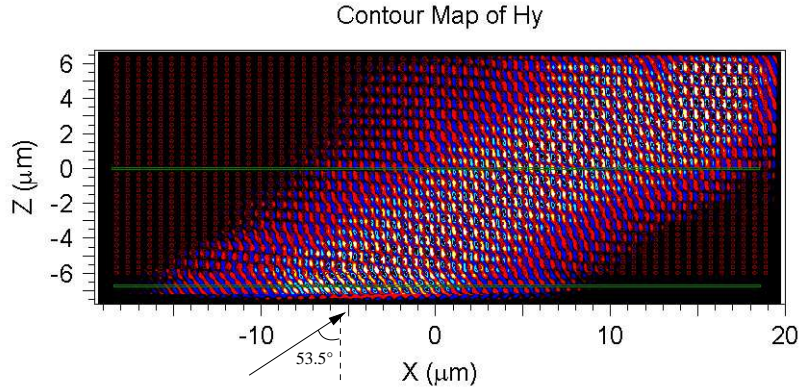


Figure 2.12: For an angle of incidence of  $\theta_i = 53.5^\circ$  the angle of refraction is around  $54^\circ$ . The super-prism effects is only slightly altered, but the transmission is increased to almost 80%.

the iso-frequency diagram of Fig. 2.3. We saw that a broad beam is completely reflected from the interface due to the symmetry mismatch with the field profile of the Bloch wave. If we reduce the size of the beam, this mismatch is no longer perfect, because a smaller beam must contain  $\mathbf{k}$  components that are somewhat off-normal. These components are expected to be transmitted, as is illustrated on the iso-frequency diagram of Fig. 2.13.

The FDTD simulation of Fig. 2.14 confirms that the narrow beam is partially transmitted. The transmitted components to the left and right are weak but noticeable. No energy is transmitted normally due to the symmetry mismatch.

These results show that when dealing with photonic crystals whose dispersion relations are complicated and may contain regions where the behavior changes abruptly in reciprocal space, one must exercise care in approximating incident beams by plane waves. In general, the different  $k$  components making up a beam may be transmitted in different directions and with different efficiencies at a photonic crystal interface. These effects must be taken into account in order to make correct predictions, particularly when dealing with photonic crystals exhibiting the super-prism effect.

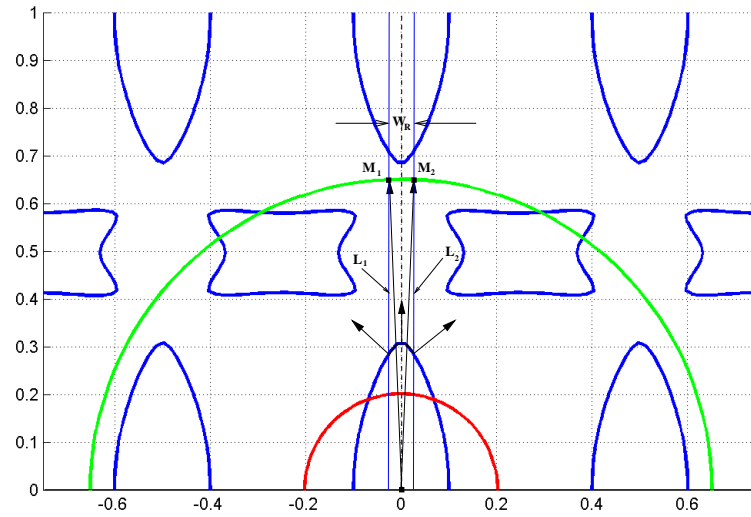


Figure 2.13: The construction line method applied to a narrow incident beam. The width of the beam in direct space is  $W_D \approx 16a$ , the angle of incidence is  $0^\circ$  and the parameters of the photonic crystal are  $r = 0.25$ ,  $x = 2$ ,  $f = 0.203$ . The two vertical construction lines delimit the construction “band” of width  $W_R = 2/W_D$ .

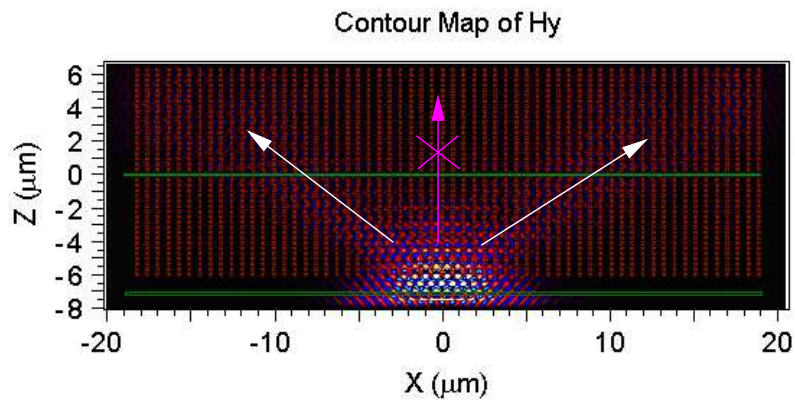


Figure 2.14: FDTD simulation of narrow incident beam with  $W_D \approx 16a$ . The beam is more than three times narrower than the beam used in the FDTD simulation of Fig. 2.7 for which  $W_D \approx 50a$ .

## Chapter 3

# Negative Index Composite Metamaterials

### 3.1 Prior art

After the initial theoretical work of Pendry in which he and coworkers suggested that a negative index medium may be obtained by intermixing magnetic resonators and thin wires [13, 14, 15], the race was on for the experimental confirmation of his theoretical prediction. The race was won by Shelby, Smith and Schultz [41] but it is an open secret today that the results presented in that first paper, while encouraging, are not sufficient to rigorously demonstrate the existence of a negative index of refraction.

In fact the notions of negative index, negative refraction, double negative medium, backward wave and non-diffraction limited focusing (a.k.a. subwavelength resolution) have a history peppered by several confusions. Much effort has gone into clarifying the relationship between them. In this section we attempt to outline the main emerging conclusions as well as some of the remaining issues. We conclude the section by placing our work on the larger map of the research activity in the field.

Chronologically, the first idea that grabbed the spotlight was Pendry's "superlens" [15] which we discuss in detail in the next section. It argued that media with  $n = -1$  (more precisely  $\varepsilon = \mu = -1$ , hence "double negative") exhibit negative refraction which leads to flat (slab-shaped) lens focusing, and that in addition *subwavelength resolution* can be achieved due to the "amplification" of evanescent fields emanating from the object. This article had a major impact on the field of classical electromagnetics but very broadly speaking it resulted in an increase in research activity in two different directions: one focused on the subwavelength imaging applications, and one focused on the physics and realization of double negative materials.

#### 3.1.1 The Imaging Studies

The first course was pursued by rather engineering minded workers, who were interested in obtaining novel imaging devices based on the superlens idea. The essential point that workers in this group focused on was the idea of subwavelength resolution and the possibility of obtaining such resolution with flat lenses. They found that the essential ingredient for this purpose was the existence of resonances in the lens. These resonances could be interface modes or bulk modes, and the problem became that of designing these modes to produce the right amplification of evanescent waves which is responsible for the subwavelength image. Since photonic crystals technology had reached a certain maturity compared to composite metamaterials which were in their infancy, this group of researchers focused mainly on

photonic crystal implementations of subwavelength resolution flat lenses.

This research direction started somewhat on the wrong foot, with two articles that led to some confusion regarding the relationship between the negative index and several other deceptively related notions.

First was the article by Notomi [42] where photonic crystal structures were put forward which, even though *not* homogenizable, behaved in a way like homogeneous media. More precisely, the dispersion relation was such that a Ibn Sahl-Snell-Descartes [37] rule-like dependence could be established between the direction of propagation of an incident beam and that of a transmitted beam at an interface. In this sense, and in this sense only, Notomi showed that it is possible to define a refractive index-like parameter which would correctly give the angle of refraction at an interface between a homogeneous medium and a given photonic crystal. However, the usefulness of the notion of index of refraction as it is most widely defined and used in all electromagnetics texts is that it also describes properties of the medium *other than refraction*. For instance the transmission efficiency and the phase shift of a beam which traverses a slab of material are *also* given by the refractive index as defined in most texts, but NOT by the effective index defined by Notomi in photonic crystals. This was stated by Notomi in the original paper, but it did not prevent numerous *other* workers from overlooking the very important distinction between the real index which must be obtained through some homogenization theory (see the Introduction and Chapter 1) and Notomi's index which only has a meaning as far as the Ibn Sahl-Snell-Descartes rule [37] is concerned, *and no further*. This confusion proved remarkably resilient.

The second article that contributed to the confusion was by Luo et al. [43]. In this work it was correctly pointed out that negative refraction and negative index are two different notions and that it is possible to obtain negative refraction without a true negative index. This cleared up to some extent the confusions which Notomi's work had (inadvertently) generated, but overlooked another detail. In this work Luo et al. claim to have illustrated subwavelength resolution by using the negative refraction obtained using a photonic crystal which, very importantly, had been cut to a *thickness* that would optimize the transmission efficiency. The fact that the interfaces had been modified by cutting the crystal is given no importance. It later turned out, as the same authors admitted the following year [44], that the improvement in the resolution was due not to the negative refraction or to the thickness of the slab but to the modified surfaces. By cutting the photonic crystal they had modified the surfaces and introduced surface states (or resonances) [45] which had allowed higher resolution information to be carried from one interface of the slab to the other. However, even having clarified this subtle point, two unclear aspects of the first 2002 article remained: first, the term "superlens" had been used even though no negative index could properly be defined and the phase of the image had no relation to the phase of the object, and second, an image size of  $0.67\lambda$  was interpreted as proof of superlensing, whereas in fact such image sizes do not demonstrate superlensing as is shown by Li et al. [46].

These two articles led to a considerable literature relying on Notomi's effective index to characterize photonic crystals, neglecting the role of the effective impedance, and claiming superlensing when in fact only negative refraction was present. This continued [47, 48] until the comprehensive clarification of Li et al. [46, 49] in early 2006.

A very telling example of misleading results is the work of Cubukcu et al. in Ref. [50]. In this article the authors claim to demonstrate superlensing experimentally. They use a photonic crystal composed of rods disposed in a square pattern, with a unit cell of  $4.79 \times 4.79$  mm. The photonic crystal slab lens is oriented such that the interfaces are aligned with the 11 lattice vector. Consequently the direction normal to the slab, that in which the light propagates, is also in the 11 direction. The authors then place two sources near the surface of the crystal, at a distance of 6.78mm from one another and 7mm from the crystal. They then measure the field near the output interface and observe two well resolved peaks at about 6.78mm from one another.

However, it is easily seen that the choice of placing the two sources 6.78mm apart could not have been random. In fact 6.78mm just happens to be exactly the surface periodicity, the distance between two adjacent rods in the 11 lattice direction:  $4.79\sqrt{2}\text{mm} \cong 6.78\text{mm}$ . The distance between two successive field peaks at the exit interface is therefore always around 6.78mm and the sources were placed this distance apart specifically in order to be able to neatly claim “superlensing”. In reality the results prove nothing, of course. At best one may say that the photonic crystal behaves as an array of near field probes, and therefore, the resolution is given by the size of the probes, in this case the size of the dielectric rods, which are smaller than the wavelength. There is nothing new or interesting in this “imaging technology”.

### 3.1.2 The Materials Studies

The second major direction stimulated by Pendry’s article was driven by the more physics-minded workers, which were interested rather in the notion of a double negative medium ( $\mu, \varepsilon < 0$ ), and the physics and various exotic phenomena that it leads to. For this group, the essential ingredient seemed to be the cohabitation of negative permittivity and negative permeability within the same material. These researchers generally prefer to speak of the “Veselago lens”, in recognition to the physicist that first put forward the idea of a flat lens based on a negative index medium, Victor Veselago, working in Russia in 1967.

This second group of researchers had to fight simultaneously on two fronts. On one hand, several theoretical aspects related to the homogenization theory (a.k.a. effective medium theory) of composite metamaterials had to be clarified. On the other, design and fabrication techniques for the composites had to be developed in order to reduce their size from the microwave regime, in which the vast majority of experiments were initially conducted, to the optical regime, where most interesting applications awaited.

In an initial phase the experimental work advanced faster than the theoretical work, partly due to the theoretical and numerical difficulties involved in studying an intrinsically three dimensional structure. Since constructive bottom up quantitative homogenization theories were hard to come by at first, the more pragmatic spirits resorted to the holistic phenomenological approach. As already discussed in Section 1.8 the phenomenological approach is the last resort, when more intuitive physical models are unavailable. In this case the holistic model took the shape of the widely used parameter extraction procedure [51]. This procedure considers the medium as a black box. There is an incident electromagnetic wave, a reflected and a transmitted wave. From the magnitudes and the phases of these waves it is possible to estimate the effective refractive index and the effective impedance of the material inside the box. This procedure works best for very thin layers of material, though in the case of a metamaterial, one is limited by the period. The main advantage as well as the main drawback of this procedure is that it always gives a result. No matter what is inside the box, the extraction procedure will give a result. Even if the medium exhibits strong spatial dispersion, and is far from behaving homogeneously, for any given frequency and angle of incidence, an effective index and an effective impedance can be defined. The difficulty resides in estimating how much meaning may be usefully attached to the values given by the extraction procedure. But this is not possible without looking inside the box, or to put it differently, *a constructive approach is needed*. The extraction procedure and its drawbacks are discussed in more detail in Appendix B.

In order to obtain a medium with simultaneously negative permittivity and negative permeability, a double negative medium, the two media (wires and resonators respectively) must be combined in some way. However, not all ways to mix the two types of elements are equal. This was suggested among others by the theoretical work of Pokrovsky and Efros [52] and Marques and Smith [53] as well as the numerical work of Woodley et al. [54].

Pokrovsky and Efros showed that by placing thin metallic wires within a medium with a negative

permeability the negative permittivity of the wires is lost. In other words the interaction between the thin wires and a negative permittivity medium is strong enough as to destroy the negative effective permittivity phenomenon. This produced some consternation until Marques and Smith showed that there are ways to combine negative permittivity and negative permeability such that they do not mutually undermine each other. Their arguments were the first indications that the correct *placement and orientation* of resonators and wires within the unit cell may have a central role to play in the design of double negative composites. They were later confirmed by the numerical results of Woodley et al. which showed that the orientation and placement of the wires and resonators has a strong effect on the homogeneous properties of the composite. The problem was not treated in depth but some suggestions were made. In particular it became clear that the deterioration of the double negative homogeneous properties of the composites was closely linked to the near field inductive coupling between wires and resonators.

These issues are clarified by the simple arguments of Maslovski [55].

In a previous work [56] he and coworkers had shown that the permittivity of a wire medium is related to the effective inductance of the wires. This effective inductance depends in turn in a natural way on the *permeability* of the surrounding medium. Consequently the expression giving the effective permittivity of the wire structure depends on the *permeability* of the medium in which these wires are placed. More precisely, for wires placed in a medium with relative permittivity  $\epsilon_m$  and permeability  $\mu_m$  the effective relative permittivity of the structure is given approximately by

$$\epsilon_{\text{eff}} = \epsilon_m - \frac{2\pi}{\mu_m (kd)^2 \log\left(\frac{d^2}{4r(d-r)}\right)}. \quad (3.1)$$

It is clear that by placing wires in a medium with a positive permittivity and a negative permeability it is impossible to obtain a medium with a negative effective permittivity. This means that, for instance, if the distances between wires are much larger than the distances between resonators, then each wire will behave as if surrounded by a negative permeability medium, and the negative permittivity will be lost. However, this is not the situation under study, due to the need for the wavelength of light to be larger than the element spacing in the metamaterial. Consequently, the distances between adjacent resonators and adjacent wires must be the same. Each unit cell must contain one of each. It is necessary to find a unit cell configuration that reduces inductive coupling. Both Maslovski and Marques and Smith argue that this can be achieved by placing the resonators at nodes of the quasi-static magnetic field of the wire medium. In this way each resonator behaves as if it is surrounded only by resonators, and each wire as if it is surrounded only by wires.

Perhaps one of the most careful and detailed theoretical studies of the composites we are interested in is that of Simovski et al. [57]. In this work the structure studied is similar to ours in all respects with the very important exception that the resonators are placed directly in between nearest neighbor wires. Thus between any given wire and its nearest neighbors one finds a resonator. The theoretical study then reports strong interference between the wires and resonators manifested in a strong modification of the effective permittivity of the composite in the vicinity of the double negative pass band, that is, of the magnetic resonance. The authors of this study then refer to the work of Belov et al. [58] in which it is shown that wire media exhibit strong *spatial dispersion* in conical incidence. We quote the relevant paragraph in full:

*“When the wave propagates strictly in the plane orthogonal to the axis of the wires one can still neglect the spatial dispersion since all parameters are independent of the coordinate [parallel to the wires]. Thus, the problem is 2-D and possible to be homogenized [59]. However, if there is a lattice of scatterers with which the wires interact, the situation becomes quite different (even for propagation orthogonal to the wires). Here the problem is not 2-D and the wire current is influenced by all the SRR particles positioned*

along its infinite length. It results in the abnormal frequency behavior of the effective permittivity of the structure.”

We believe that the strong interference and spatial dispersion reported in this work can be avoided by correctly placing the resonators such that their inductive coupling to the magnetic field of the wires is minimized. Since the effective permittivity is in essence an interference phenomenon (see Section 3.4), it is very susceptible to the presence of any scatterers in the spaces between the wires. The solution is to place resonators at nodes of the magnetic field of the wires, in which case the negative permittivity and negative permeability cohabit harmoniously as shown by the results of Section 3.7.

These ideas reconcile the skepticism of theoretical workers such as Pokrovsky and Efros with the optimism of experimental workers such as Smith and Soukoulis. It turns out that the experiments were built (perhaps due to extensive trial and error) in exactly the right way to avoid the problems discussed by Pokrovsky, Woodley or Simovski. The basic elements, the resonators and the wires must be placed in such a way as to minimize the inductive coupling.

However, even though these aspects have been known for some time, their verification has left a lot to be desired. In particular, numerical studies of composite metamaterials have had two main drawbacks: they make use of the parameter extraction technique, which is a phenomenological black box technique, and also the media are characterized quantitatively only for normal incidence. Eventual spatial dispersion effects have therefore been ignored, since they would only appear when comparing the behavior of the medium for different angles of incidence. Spatial dispersion also remains a major obstacle to the realization of the superlens, which is, after all, one of the main motivations for the research in this field.

In this work we provide a *constructive* physical model for negative index composites, and we provide a metamaterial design that *avoids the deleterious effects of spatial dispersion*. We support our results with comprehensive numerical computations, including sweeping the angle of incidence, a first in the literature.

## 3.2 Negative index of refraction – the superlens

In this section we consider in more detail the notion of index of refraction and we focus particularly on media in which both the permittivity and permeability are negative as compared to media where both are positive. We will see that the medium with  $\mu, \varepsilon = -1$  behaves, in a sense as the optical inverse of free space. The evolution of a wave while propagating in free space can be in all respects undone by having it propagate an equal distance in the  $\mu, \varepsilon = -1$  medium. In the following we will refer to waves or materials where  $\mu, \varepsilon > 0$  as positive index or double positive, and to waves or materials where  $\mu, \varepsilon < 0$  as negative index or double negative. Our development follows the formulation of Ref. [60]

Let us consider the propagation of a scalar wave in the region of space for which  $0 < z < Z$ . We assume this region contains no sources, so that the field satisfies the homogeneous Helmholtz equation. The vector nature of the electromagnetic field does not pose an obstacle in this context. We can for instance consider the fields component-wise, or consider a field polarized parallel to the  $z = \text{const.}$  plane.

We write the total field

$$V(x, y, z, t) = U(x, y, z)e^{-i\omega t}$$

and the spatial part  $U(x, y, z)$  satisfies

$$(\nabla^2 + k^2)U(x, y, z) = 0$$

where we have written  $k^2 = \mu_0\varepsilon_0\mu\varepsilon\omega^2 = \mu\varepsilon\frac{\omega^2}{c^2} = \mu\varepsilon k_0^2$ . The field in a plane  $z = \text{const.}$  can be represented

as a Fourier integral

$$U(x, y, z) = \int \int_{-\infty}^{\infty} \tilde{U}(u, v; z) e^{i(u x + v y)} du dv. \quad (3.2)$$

Replacing this in the Helmholtz equation we obtain

$$\int \int_{-\infty}^{\infty} (\nabla^2 + k^2) \left( \tilde{U}(u, v; z) e^{i(u x + v y)} \right) du dv = 0$$

or

$$\int \int_{-\infty}^{\infty} \left( (k^2 - u^2 - v^2) \tilde{U}(u, v; z) + \frac{\partial^2 \tilde{U}(u, v; z)}{\partial z^2} \right) e^{i(u x + v y)} du dv = 0$$

This is the Fourier development of the null function, so each coefficient must be null independently. We have

$$\frac{\partial^2 \tilde{U}(u, v; z)}{\partial z^2} + (k^2 - u^2 - v^2) \tilde{U}(u, v; z) = 0.$$

If we introduce  $w^2 = k^2 - u^2 - v^2$  then the general solution of this equation takes the form

$$\tilde{U}(u, v; z) = A(u, v) e^{i w z} + B(u, v) e^{-i w z}.$$

If we assume that the sources of the fields are all in the  $z < 0$  half space then all waves must propagate in the positive  $z$  direction and  $B(u, v) = 0$ . We obtain

$$\tilde{U}(u, v; z) = A(u, v) e^{i w z} \quad (3.3)$$

It is therefore clear that in the plane  $z = 0$  the Fourier components of the field distribution are given by  $A(u, v)$ . As the field propagates in the  $z$  direction, the Fourier composition in the plane  $z = Z$  is  $\tilde{U}(u, v; Z) = A(u, v) e^{i w Z}$ . The propagation in an isotropic homogeneous medium therefore has the effect of transforming the Fourier components of the field distribution according to the factor  $e^{i w z}$ . In the following we shall refer to it as the evolution operator, by analogy to the quantum mechanical time evolution operator. The parameter  $w$  therefore seems to be of paramount importance. Recall that it is defined as

$$w^2 = \mu \varepsilon k_0^2 - u^2 - v^2.$$

In this relation all quantities are real except possibly  $\mu$ ,  $\varepsilon$  and  $w$ . As in section 1.6 we introduce the index  $n^2 = (\beta + i\alpha)^2 = \mu \varepsilon$  where  $\beta$  and  $\alpha$  are real,  $\alpha$  is positive, and they are given by Eqs. (1.42) and (1.43). As before the imaginary part  $\alpha$  is only introduced in order to determine the signs of the real parts, and it is then made to tend to zero. We also write  $w = b + ia$ . Since no active media are present all imaginary parts must be positive. We write out the real and imaginary parts of the above relation:

$$\begin{aligned} b^2 - a^2 &= (\beta^2 - \alpha^2) k_0^2 - u^2 - v^2 \\ ab &= \alpha \beta k_0^2 \end{aligned}$$

By letting  $\alpha$  tend to zero in the second equation it results that either  $a$  or  $b$  must also go to zero. It is the first equation that will determine which. For vanishing  $\alpha$  we have  $b^2 - a^2 = \beta^2 k_0^2 - u^2 - v^2$ . When  $u^2 + v^2 < \beta^2 k_0^2$  then the quantity is positive and  $a$  must be the one that vanishes along with  $\alpha$ . If on the other hand  $u^2 + v^2 > \beta^2 k_0^2$  then it must be  $b$  that vanishes. The sign of the remaining quantity is determined by the sign of  $\beta$  from the second relation above. We have already seen in Section 1.6 that when the medium is double positive, then  $\beta$  is positive, but that when it is double negative, then  $\beta$  is negative. The different possibilities are summarized in Table (3.1).

	$u^2 + v^2 < \beta^2 k_0^2$	$u^2 + v^2 > \beta^2 k_0^2$
$\beta > 0$	$w$ real positive $w = \sqrt{\beta^2 k_0^2 - u^2 - v^2}$	$w$ imaginary positive $w = +i\sqrt{u^2 + v^2 - \beta^2 k_0^2}$
$\beta < 0$	$w$ real negative $w = -\sqrt{\beta^2 k_0^2 - u^2 - v^2}$	$w$ imaginary negative $w = -i\sqrt{u^2 + v^2 - \beta^2 k_0^2}$

Table 3.1: Table of the possible behaviors of  $w$ .

From the table it is clear that features corresponding to large spatial frequencies  $u, v$  correspond to evanescent waves, while low frequency features correspond to propagating waves. In the course of propagating between the planes  $z = 0$  and  $z = Z$  the low frequency components have undergone a unitary transformation, or a change of phase. In the case of the high spatial frequencies, however, the phase does not evolve, but the amplitude does. When  $\beta$  is positive this amplitude is attenuated, while when  $\beta$  is negative this amplitude is amplified. If we note  $w_p$  for a double positive medium and  $w_n$  for the corresponding double negative medium, then from the above table we have

$$w_p = -w_n.$$

The evolution of the field in the  $z$  direction in the material with  $\beta = -1$  is exactly the reverse from the evolution of the field when  $\beta = 1$  if losses are ignored. In fact, it is as if the time runs backwards. This is not an accident. If we consider the Maxwell curl equations (1.37) and (1.39) then changing the signs of the constitutive parameters  $\mu$  and  $\varepsilon$  is formally equivalent to taking the inverse of the time dependent term  $e^{-i\omega t}$ .

It is tempting now to consider a system of two such complementary slabs, by simply multiplying the exponential evolution operators. However this is in general not correct. The reason is that above we have assumed the sources of all fields were to the left of the region of interest and that consequently all fields propagate in the same direction, the  $z$  positive direction. But this can only be the case if the medium is homogeneous *and infinite* in the  $z$  direction. If an interface or a scattering element of any kind is present then this is no longer true. In such cases we must consider both left and right going waves, and the way that they couple at interfaces. This is done by calculating transmission and reflection of each wave at each interface, by employing the notion of impedance. Without going into the details, we will only point out that these reflections are absent when the materials are matched. If the two media have the same impedance then our assumption is justified and it is possible to simply multiply the evolution operators.

It is then possible to consider a region of free space of width  $z$  as a filter with a transfer function given by the evolution operator  $e^{iwz}$ , while a similar region filled with  $\mu = \varepsilon = -1$  placed next to the first provides the *inverse* filter,  $e^{-iwz}$ . After propagation through the two layers the field is reproduced **exactly**. It is, however, well known in the theory of linear systems that inverse filtering is sensitive to noise. If the initial filter has reduced the amplitude of some frequency components to values close to the noise amplitude at those frequencies, then when the inverse filter re-amplifies them, it amplifies the noise as well, resulting in a very noisy reconstructed signal. A way to avoid this problem is to avoid small signal amplitudes. Since extinction and amplification are given by a term exponential in the distance  $z$  then it may be advantageous to use many thin alternating regions of double positive and double negative media, rather than two thick ones. As long as they both occupy equal volumes the signal will be reproduced exactly. The noise limitation also places an upper bound on the distance between an object and the

surface of the lens, for any given required resolution. The higher the resolution we seek, the higher the  $k$  components that must be resolved. Higher  $k$  components, in turn, attenuate faster with distance, which means that the object must be placed closer to the lens surface in order for the signal level at the large  $k$  to be larger than the noise. The alternative, of course, is cooling the lens to very low temperatures.

The signal reproducing device described above is known as a super-lens because it is capable of reproducing an image including the high spatial frequencies, which in normal optical systems are inevitably lost. The loss of high frequency components of an image is known as the “diffraction limit” and it is often said in the literature that the super-lens can overcome the diffraction limit. The fact that high frequency components of a signal are carried by waves that are attenuated in space, or evanescent waves, has also led some workers to say that the superlens can “focus” the evanescent waves, in addition to focusing the propagating waves. This is however subject to the noise limitation mentioned above, even when absorption is ignored.

In fact the possibility of having media with negative constitutive parameters was investigated for the first time in a speculative article by Victor Veselago in 1967, translated in English in 1968 [61]. Veselago showed that such media would have many exotic and unexpected properties such as a reversed Doppler shift, reversed Cerenkov radiation and negative refraction. He did not, however, point out that evanescent waves would be amplified in such a medium. This was done over thirty years later by John Pendry in the now famous Physical Review Letters article [15] which can be said to have truly launched the field of negative index metamaterials.

In order to picture these ideas it is convenient to illustrate them by plotting the transmission coefficient of a homogeneous slab as a function of the tangential component of the incident wavevector. It is especially interesting to observe the behavior for tangential components that are larger than the wavevector. The plots below show the magnitude of the transmission coefficient through various slab media as a function of the parameter  $\alpha = \frac{\sqrt{u^2+v^2}}{k_0} = \frac{\sqrt{k^2-w^2}}{k_0}$  for  $H_{\parallel}$  polarization. Though it may sound paradoxical,  $\alpha$  is the tangential component of the incident wavevector in units of the magnitude of the said wavevector. When  $\alpha > 1$  this means that the “incident” wave in question is an evanescent wave (also known as inhomogeneous because equiphase planes and equiamplitude planes do not coincide). Large  $\alpha$  correspond to high spatial harmonics of the object in front of the lens. When  $\alpha > 1$  the wave is evanescent in free space, though it may propagate in media with a high enough index. In other words  $w$  is not necessarily imaginary when  $\alpha > 1$  but only when  $\alpha > |k|/|k_0|$ .

We begin by comparing the transmission of a regular dielectric slab, with both  $\mu$  and  $\varepsilon$  positive with the transmission of a superlens, that is, a slab with  $\mu = \varepsilon = -1$ . On the left side of Fig. 3.1 we have plotted the transmission through a slab of thickness  $d = 1\text{au}$  with  $\mu = 1$ ,  $\varepsilon = 12$  for  $\lambda = 5\text{au}$  while on the right side we have plotted the transmission of free space in blue and of the superlens with  $\mu = \varepsilon = -1$  in green.

There are several important differences between the three situations. Propagating waves, that is, for which  $\alpha < 1$ , are perfectly transmitted by the two media in the right plot, though not quite in the left plot. The transmission is exponentially decreasing with increasing  $\alpha$  for the case of the dielectric slab as well as for free space, though the decay is faster in the dielectric slab. In addition, the transmission of the slab exhibits poles, or values of  $\alpha$  for which the transmission diverges. This happens close to  $\alpha = 1$  and  $\alpha = 2.5$  for our choice of parameters. No such divergences appear in the right plot. These poles correspond to guided modes in the dielectric slab. The absence of poles in the right plot indicates that neither free space nor the superlens support guided modes. The reason for this difference is that in the right plot, neither of the two media exhibits total internal reflection at the interface with free space. Light cannot be guided using these media.

It would be tempting to explain the absence of guided modes by the fact that these media are impedance matched to free space. However, this is not correct. It is possible to have guided modes

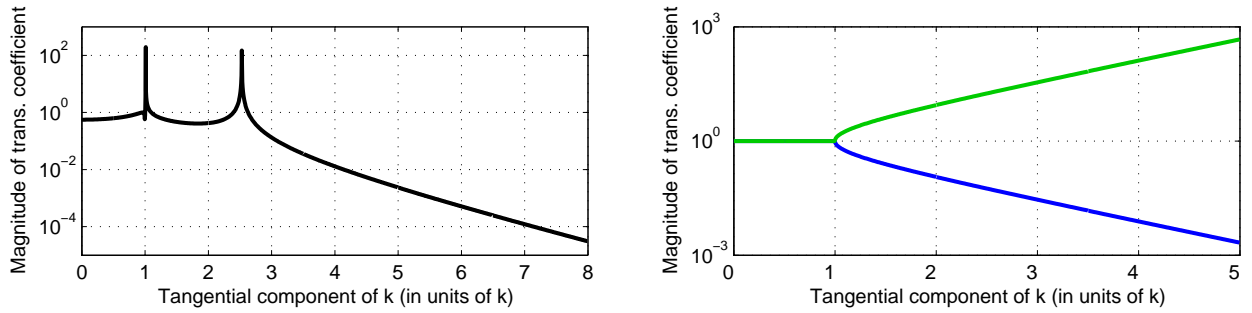


Figure 3.1: Transmission of dielectric slab with thickness  $d = 1\text{au}$ , with  $\mu = 1$ ,  $\varepsilon = 12$  and wavelength  $\lambda = 5\text{au}$  as a function of the tangential component of the wavevector (left). On the right side we compare the transmission of free space slab (blue) and the superlens (green).

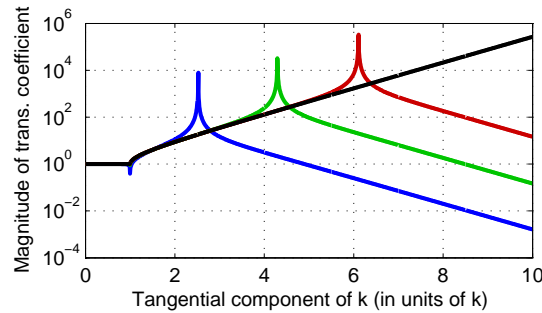


Figure 3.2: As the permittivity and permeability of the slab approach  $-1$  the guided mode is shifted to infinity and the transmission approaches that of the superlens, the black curve, also the green curve in Fig. 3.1.

in structures that are perfectly impedance matched to free space. In order to see this we plot the transmission for three slabs with  $\mu = \varepsilon = \{1.1, 1.01, 1.001\}$ , traced in blue, green and red respectively. The superlens transmission is plotted in black. All three slabs support a single guided mode, though the mode is shifted to higher  $\alpha$  the closer we get to the superlens condition. In fact the superlens can be seen as having a guided mode at infinity on the  $\alpha$  axis. Total internal reflection and partial reflection at transmission through an interface are two physically distinct and unrelated phenomena. Partial reflection is related to the impedance mismatch while total internal reflection is an effect which is related to the translation symmetry of the interface and the conservation of the tangential component of the wavevector which the symmetry requires. Partial reflection is an impedance phenomenon, while total internal reflection is a symmetry and index phenomenon.

One interesting aspect that can also be seen in Fig. 3.2 is that in principle one does not need a perfect superlens in order to observe the amplification of evanescent waves across a slab. If one can obtain a medium with  $\mu = \varepsilon = -1.01$ , corresponding to the green curve, then spatial harmonics up to about  $\alpha = 4$  will be transmitted accurately across the slab. Higher spatial harmonics will still be lost, but speaking very loosely one may say that the traditional diffraction limit has been beaten very roughly by a factor of 4. Of course, this is nothing revolutionary, since resolutions far better than this can be achieved using widespread near field scanning optical probe microscopy. What is novel in this case is the means used: a double negative medium.

Before going on to discuss two examples of the larger family of flat lenses let us emphasize once more that the remarkable properties of the superlens reside essentially in its response to an incident evanescent field, that is, for  $\alpha > 1$  and that negative refraction is a phenomenon which is a behavior which pertains to fields with  $\alpha < 1$ . Consequently while the superlens may exhibit negative refraction, negative refraction does not imply a superlens.

### 3.3 Flat lenses

In the previous section we have seen that the operation of the superlens depends on three crucial factors.

1. The phase evolution of propagating waves in free space is reversed resulting in refocusing.
2. The amplitude evolution of evanescent waves in free space (decay) is reversed resulting in amplification.
3. Interfaces are perfectly matched so that there are no back reflections and no guided modes distorting the transmission of evanescent fields.

When one speaks of a super-lens it is understood that all three of the above conditions are fulfilled. However, since it is clear that such superlenses are ideal situations and since no way was immediately available to design and construct true double negative media, various workers have attempted to emulate superlensing action using other means. This work has given rise to a series of flat lens proposals, some of which are closer kin than others to the true superlens as described in the previous section. In this section we discuss two of these proposals.

But we must say a few words about surface modes.

Let us consider Eq. (3.3). This equation gives the 2D Fourier transform of the field in the  $x-y$  plane at coordinate  $z$ . If we use the definition of Eq. (3.2) then we can write the field in plane  $z$  as

$$U(x, y, z) = \iint_{-\infty}^{\infty} A(u, v) e^{i(ux+vy+wz)} du dv + \iint_{-\infty}^{\infty} B(u, v) e^{i(ux+vy-wz)} du dv.$$

It can be seen immediately that the exponential terms in the integrals are solutions of the Helmholtz equation, because applying the scalar Laplace operator to them is equivalent to multiplication by  $u^2 + v^2 + w^2 = k^2 = k_0^2 n^2$ . Consequently the expression above can be seen as a mode representation of the field in the slab geometry where each mode has a weight given by the coefficients  $A(u, v)$  or  $B(u, v)$ . The space of modes in the slab geometry has two degrees of freedom since it can be parameterized by two parameters,  $u$  and  $v$ . These field distributions are called modes because each of them is independently a solution of the Helmholtz equation in the slab.

The slab modes can be distinguished into two types, according to whether  $w$  is real or imaginary. When  $w$  is imaginary then the fields are either exponentially increasing or exponentially decreasing with  $z$ . This means that if the slab is sufficiently thick (or if  $w$  is sufficiently large), any given mode will be confined to one of the surfaces, either at  $z = 0$  or  $z = Z$ . For this reason, these exponentially increasing or decreasing modes are also known as surface modes. The majority of the energy that they carry is strongly localized in the vicinity of one or the other of the two slab faces.

From this point of view, one can say that the particularity of the superlens as compared to any other slab, is that it allows the electromagnetic energy emitted by an object close to the input surface to couple efficiently to the surface mode *on the output surface*. This peculiar phenomenon makes it seem like the field is “amplified” as it “propagates” across the slab. The net result is that the evanescent field due to a source, by coupling to the interface mode on the far side, gets a new lease on life, so to speak. A detector placed close to the output surface can detect fields associated with  $k$  components that would not have been detectable without the negative index slab, or lens.

The trick is to have the high  $k$  components of the incident field couple efficiently to the surface mode on the exit surface. In the ideal superlens this happens directly, with no intermediary. In other types of flat lenses this process is mediated by the surface mode of the input surface. The energy first couples to the input interface, and then it propagates to the output interface. The translation symmetry of the slab

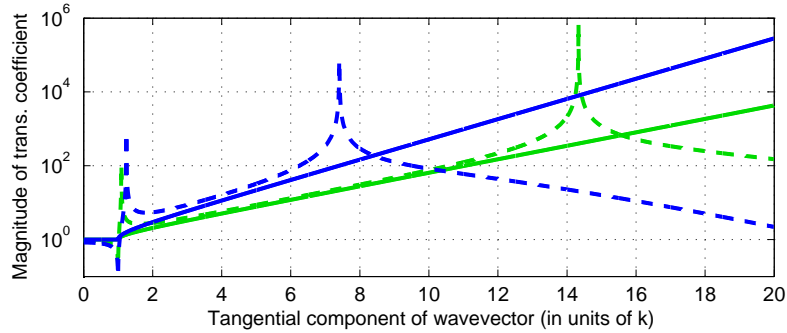


Figure 3.3: We compare the transmission of a superlens ( $\varepsilon = \mu = -1$ , solid curves) with that of an ideal plasma slab ( $\varepsilon = -1$ ,  $\mu = 1$ , dashed curves) for wavelengths of  $\lambda = 10\text{au}$  (blue) and  $\lambda = 15\text{au}$  (green). The slab thickness is  $1\text{au}$ .

insures the conservation of the component of the wavevector in the  $x - y$  plane, thereby transmitting it across the slab.

Two main flat lens ideas have been put forward. The first consists of a thin metal film, while the second consists of a photonic crystal slab. The metal film is conceptually very closely related to the superlens, while the photonic crystal idea is considerably farther removed. We first discuss the metallic film.

Being aware that a true superlens may not be easy to come by experimentally in the short term, in his original article Pendry also put forward a “watered down” simplified version of the superlens, that he hoped may be more easily amenable to experimental test. He claimed that superlensing action may also be observed in a slab of negative permittivity material, provided it was very thin, or equivalently, the wavelength was very large, and *independently of the value of the permeability*. His argument goes like this. In the absence of free charges the behavior is electrostatic and the two curl Maxwell equations can be ignored. The electromagnetic problem is reduced to one of electrostatics, in other words, a problem independent of the permeabilities of the media involved. The very thin negative permittivity lens should therefore behave as a superlens regardless of the value of the permeability. The only requirement is that the permittivity be close to  $-1$ . Pendry therefore claimed that electrostatics reduces the requirement  $\mu = \varepsilon = -1$  to simply  $\varepsilon = -1$  and it is known that there are frequencies where this value is approached by several metals.

It is true that metals do contain free charges, which is expressed in the fact that their permittivity has a non-negligible imaginary part. However, for silver, this imaginary part has been shown to be sufficiently small to allow experimental workers to observe limited superlensing action. This impressive experimental feat is due to Fang and coworkers [62] at UC Berkeley and took almost five years to achieve.

In the case of silver the interface modes required to carry the higher spatial harmonics of the input signal are surface plasmons. Joule absorption notwithstanding, the field at the exit interface is sufficiently strong to be detected. As far as the phase evolution of the electromagnetic field across the slab, in a negative permittivity medium this phase shift is null and would not correspond to a superlens-like behavior, but for the fact that in this case we are dealing with a layer much thinner than the wavelength, so the phase evolution across it would have been close to zero in any case.

Figure 3.3 compares the transmission through a true superlens (solid curves) with the transmission through a slab with  $\mu = 1$ ,  $\varepsilon = -1$  (dashed curves). The thickness is  $1\text{au}$  (arbitrary unit) and the wavelength is  $\lambda = 10\text{au}$  for the blue curves and  $\lambda = 15\text{au}$  for the green curves. It is easy to see that as the wavelength increases the negative permittivity slab becomes a good approximation to the superlens. Thus the theoretical predictions of Pendry were, at least partially, experimentally confirmed.

The second flat lens design that was put forward was the photonic crystal flat lens [43]. In this design one mimics superlensing action without the presence of either a well defined permittivity, permeability or index. The three points at the beginning of this section may potentially be satisfied, at least in theory, but in a way that is far from straightforward. The photonic crystal slab is a generalization of the homogeneous slab due to the fact that the translation symmetry group is discrete, rather than continuous and moreover the optical size of the unit cell may be quite large, depending on the relative permittivity and geometry. The notions of propagating, evanescent and anti-evanescent modes must be generalized. This can and has been done within the monodromy matrix formalism [63].

The monodromy matrix is to a photonic crystal slab what the evolution operator  $e^{i\omega z}$  is to the homogeneous slab as discussed in the previous section. It is an operator that propagates the field from one side of the slab to the other. And just as in the case of the homogeneous slab, the operator has propagating or evanescent components (corresponding to unimodular and nonunimodular eigenvalues respectively), with the added complication that they are always both present to some extent and no straightforward simplifications can be made, even when the incident field is a homogeneous plane wave. In the case of vanishing index contrast this formalism reduces to the angular spectrum representation discussed in the previous section and in Ref. [60] and the monodromy matrix reduces to the evolution operator  $e^{i\omega z}$ .

A further complication which appears in the case of the slab, but which we were able to ignore with the superlens is the impedance mismatch at the input and output interfaces. The superlens was perfectly matched to free space but photonic crystals slabs rarely are, if ever. Consequently the monodromy matrix is not sufficient to characterize the imaging properties of the slab and one must consider it in conjunction with the conventional transmission matrix formalism which takes into account the impedance mismatch at the interfaces. This mismatch further complicates the study of such flat lenses and a systematic study of these phenomena from this fundamental point of view has yet to appear in the literature. The intricate interplay between propagating and evanescent fields in the image formation in PC lenses has been studied numerically and somewhat empirically by several authors, but with ambiguous results [44, 46, 49].

## 3.4 Thin wire medium

In this section we obtain an expression for the effective permittivity of a thin wire medium with a rectangular unit cell. We begin by considering the transfer matrix of a single thin metal-wire grating of period  $d$  in an air slab of thickness  $h$ , and comparing it directly with the transfer matrix of a homogeneous slab with a permittivity  $\varepsilon$  and the same thickness. This holistic phenomenological approach leads to the expressions of Eqs. (3.9) and (3.10). We then show that the permittivity given by these expressions should in fact be seen as the effective permittivity of a rectangular *two-dimensional* wire medium, with periodicities  $d$  and  $h$ . This interpretation is not immediately obvious and is another illustration of the dangers of doing physics phenomenologically, as discussed in Section 1.8 and Appendix B. The expressions derived below have the benefit of applying also to media with rectangular unit cells, by contrast with the best existing estimate, that of Maslovski et al. [56], which applies only to media with a square unit cell.

### 3.4.1 Transfer matrix of thin metal wire grating

This derivation follows closely that of Ref. [64]. The case of thin metal wires is a special case of the theory outlined in Sections A.4 to A.7. More precisely, only the polarization whereby the  $\mathbf{E}$  field is parallel to the wires needs to be considered since thin metal wires are transparent to  $\mathbf{H} \parallel$  polarized waves. In addition, since the wires are much smaller than the period which is itself smaller than the

wavelength, this means that only the central monopolar term needs to be kept, namely  $b^0$  from Eq. (A.3). Consequently the scattering equation, Eq. (A.6), becomes

$$b^0 = \frac{i^0}{(S_0)^{-1} - A_0}$$

where  $i^0$  is given by Eq. (A.2),  $S_0$  is given by Eq. (A.1) and  $A_0$  is given by Eq. (A.9). According to Eq. (A.12) the transmission coefficient becomes

$$\begin{aligned} t &= 1 + b^0 \frac{K}{\pi k \cos \phi} \\ &= 1 + \frac{1}{-\frac{H_0(ka)}{J_0(ka)} - \frac{K}{\pi k \cos \theta} + 1 - i\frac{2}{\pi} (\ln(\frac{K}{k}) + \ln 2 - \gamma)} \frac{K}{\pi k \cos \phi} \\ &= 1 + \frac{1}{-\frac{K}{\pi k \cos \theta} - \frac{2i}{\pi} \ln(Ka)} \frac{K}{\pi k \cos \phi} \\ &= 1 + \frac{1}{-1 + ikL \cos \phi} \end{aligned} \quad (3.4)$$

where the asymptotic forms of  $H_0$  and  $J_0$  for small arguments are found in Ref. [65], and where

$$L = -\frac{2}{K} \ln(Ka). \quad (3.5)$$

The reflection coefficient is therefore

$$r = \frac{1}{-1 + ikL \cos \phi}. \quad (3.6)$$

We should now have enough information to obtain the transfer matrix characterizing the electromagnetic properties of the thin wire grating. We use the unimodular matrix formalism (see Refs. [66, 67]) where transfer matrix  $\mathbf{T}$  for a layer with given reflection and transmission coefficients  $r$  and  $t$  satisfies the relation:

$$\mathbf{T} \cdot \begin{pmatrix} 1 + r \\ ik \cos \phi (1 - r) \end{pmatrix} = \begin{pmatrix} t \\ ikt \cos \phi \end{pmatrix} \quad (3.7)$$

Let us note

$$\mathbf{T} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

We would like this transfer matrix to describe a layer of zero thickness and reflection and transmission coefficients given by Eqs. (3.4) and (3.6), with the additional requirement of field continuity across the layer.

Field continuity immediately requires

$$\begin{aligned} a &= 1 \\ b &= 0 \end{aligned}$$

In addition the unimodular property of the transfer matrix then also requires  $d = 1$ .  $c$  therefore remains the only unknown. By writing out the bottom part of Eq. (3.7) we obtain:

$$c = \frac{2}{L}.$$

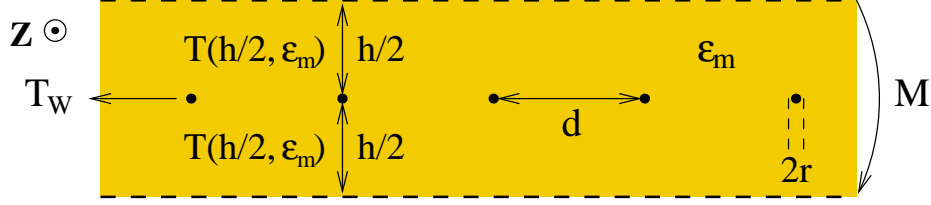


Figure 3.4: We obtain a homogeneous model for the wire grating by adding layers of the surrounding medium on the top and the bottom of the grating. The total transfer matrix given by  $M = T(h/2, \varepsilon_m) \times T_W \times T(h/2, \varepsilon_m)$  is then compared with the transfer matrix of a homogeneous layer of the same total thickness  $h$ ,  $\mathbf{T}(h, \varepsilon_{\text{eff}})$ .

The transfer matrix therefore takes the very simple form [64]

$$\mathbf{T}_W = \begin{pmatrix} 1 & 0 \\ \frac{2}{L} & 1 \end{pmatrix}.$$

### 3.4.2 Effective permittivity - analytical

In order to model the row of wires as a slab of homogeneous material with a finite thickness we must sandwich it between layers of a homogeneous surrounding medium as shown in Fig. 3.4.

To obtain a layer of total thickness  $h$  we must consider the grating itself as a surface of thickness zero characterized by  $\mathbf{T}_W$  between two layers of homogeneous dielectric of thickness  $h/2$  each, characterized by  $\mathbf{T}(h/2, \varepsilon_m)$ . Then the total transfer matrix takes the form:

$$\begin{aligned} \mathbf{M} &= \mathbf{T}(h/2, \varepsilon_m) \times \mathbf{T}_W \times \mathbf{T}(h/2, \varepsilon_m) \\ &= \begin{pmatrix} \cos(\frac{\beta h}{2}) & \frac{1}{\beta} \sin(\frac{\beta h}{2}) \\ -\beta \sin(\frac{\beta h}{2}) & \cos(\frac{\beta h}{2}) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{2}{L} & 1 \end{pmatrix} \begin{pmatrix} \cos(\frac{\beta h}{2}) & \frac{1}{\beta} \sin(\frac{\beta h}{2}) \\ -\beta \sin(\frac{\beta h}{2}) & \cos(\frac{\beta h}{2}) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\beta h) + \frac{1}{\beta L} \sin(\beta h) & \frac{2}{\beta^2 L} \sin^2(\frac{\beta h}{2}) + \frac{1}{\beta} \sin(\beta h) \\ -\beta \sin(\beta h) + \frac{2}{L} \cos^2(\frac{\beta h}{2}) & \cos(\beta h) + \frac{1}{\beta L} \sin(\beta h) \end{pmatrix} \end{aligned}$$

where  $\beta^2 = k^2 \varepsilon_m - k^2 \sin^2 \phi$  and  $h$  is the total thickness. In order to obtain an estimate for an effective permittivity of the structure we compare the transfer matrix above with the transfer matrix of a homogeneous layer of thickness  $h$  and permittivity  $\varepsilon_{\text{eff}}$ :

$$\mathbf{T}(h, \varepsilon_{\text{eff}}) = \begin{pmatrix} \cos(\frac{\beta_{\text{eff}} h}{2}) & \frac{1}{\beta_{\text{eff}}} \sin(\frac{\beta_{\text{eff}} h}{2}) \\ -\beta_{\text{eff}} \sin(\frac{\beta_{\text{eff}} h}{2}) & \cos(\frac{\beta_{\text{eff}} h}{2}) \end{pmatrix}$$

where  $\beta_{\text{eff}}^2 = k^2 \varepsilon_{\text{eff}} - k^2 \sin^2 \phi$ . In fact in order to obtain an estimate for  $\varepsilon_{\text{eff}}$  we need only compare one of the elements, the most convenient one, the  $_{11}$  element:

$$\cos(\beta h) + \frac{1}{\beta L} \sin(\beta h) = \cos(\frac{\beta_{\text{eff}} h}{2}).$$

We obtain:

$$\varepsilon_{\text{eff}} = \sin^2 \phi + \frac{1}{k^2 h^2} \left( \arccos \left( \cos(\beta h) + \frac{1}{\beta L} \sin(\beta h) \right) \right)^2. \quad (3.8)$$

Let us now check that this estimate is physically reasonable. As we will see, the answer is “almost”. One more modification will be necessary in order for the formula to be correct.

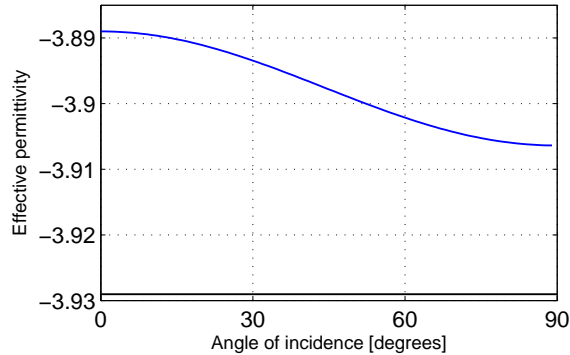


Figure 3.5:  $\phi$ dependence of Eq. (3.8) for a grating of period  $d = 1$ , wire radius  $r = 0.01$  and  $\lambda = 10$ . The horizontal line at the bottom of the figure represents the estimate given by Maslovski's formula [cite]. The  $\phi$  dependence is clearly very weak (less than a half of a percent) while the agreement with Maslovski's estimate is very good (about half of a percent relative difference).

First, we notice that the effective medium seems to depend on the incidence angle  $\phi$ . To have an idea of the strength of this dependence we plot  $\varepsilon_{\text{eff}}$  as a function of  $\phi$  in Fig. 3.5. It is clear that the dependence on the angle of incidence is negligible. Over the whole range between normal and grazing incidence the effective permittivity varies by less than a half of a percent. We can therefore safely fix the angle to  $90^\circ$  with negligible error. We discuss that choice below. At the same time we will see below that the formula gives results which agree very well with the existing estimate by Maslovski et al. [56]. The formula we obtain is therefore:

$$\varepsilon_{\text{eff}} = 1 + \frac{1}{k^2 h^2} \left( \arccos \left( \cos(\beta_g h) + \frac{1}{\beta_g L} \sin(\beta_g h) \right) \right)^2 \quad (3.9)$$

where  $\beta_g = k\sqrt{\varepsilon_m - 1}$  or more usefully to a form resembling Maslovski's expression of Eq. (3.1):

$$\varepsilon_{\text{eff}} = \varepsilon_m + \frac{1}{k^2 h^2} \left( \arccos \left( 1 + \frac{h}{L} \right) \right)^2 \quad (3.10)$$

Obtaining this neat formula is the reason why we chose to fix the angle at  $90^\circ$  in Eq. (3.8) above.

This expression seems, at first sight, satisfactory, and its accuracy is excellent as shown in the next section. However, there is a question as to its interpretation. Can one consistently speak of the effective permittivity of a single wire grating? In other words, is a single wire grating in a way analogous to a monoatomic layer? It is known that it is possible to introduce the concept of the permittivity of a monoatomic adsorbed layer [68], even though it is a derived quantity and less physically meaningful than the permittivity in the bulk.

In fact the more physically meaningful quantities in the context of an atomic monolayer are the tangential and normal susceptibilities per unit area noted by Dignam and Moskovits as  $\gamma_t$  and  $\gamma_n$ . In our case, the wire medium is only of interest in the case of polarization corresponding to the electric field parallel to the wires, such that only  $\gamma_t$  appears. It is then known that in order for the monolayer permittivity to be a useful quantity, one must have

$$\gamma_t = (1 - \varepsilon_t)h$$

be independent of  $h$  where  $h$  is the thickness of the effective medium with which one would like to replace

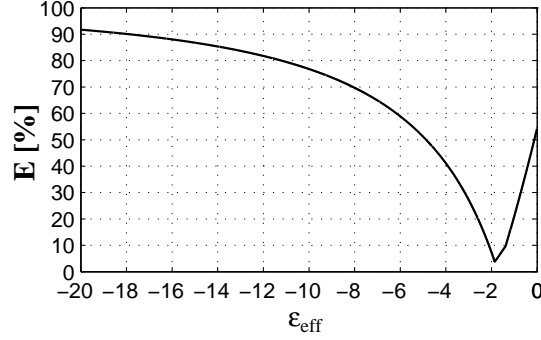


Figure 3.6:  $E$  is plotted as a function of the effective permittivity. It can be seen that a slab with  $\epsilon_{\text{eff}} \approx -1.9$  (henceforth the *numerical estimate*) has a transfer matrix which reproduces to within 5% the transfer matrix  $M$  of the structure under study. The minimum value of  $E$  in percent, which we denote  $E_{\text{min}}$  (in this case 5%), is then a measure of how “homogenizable” the structure is.

the monolayer. For our wire structure this comes down to requiring

$$(\epsilon_{\text{eff}} - 1)h = \text{const.}$$

A quick look at Eq. (3.10) is sufficient to see that this is not the case. It is a good approximation when  $h \ll L$  but since  $L$  is a logarithmic function of the wire radius, it does not increase very quickly, and  $h/L$  cannot be treated as a perturbation in many cases of interest. There seems to be something inconsistent about extracting an effective permittivity from a single thin wire grating.

This puzzle is clarified by a close consideration of the results of Pokrovsky and Efros [52] and the discussions in Refs. [53] and [69]. The results of these workers indicate that the negative permittivity of thin metal wires is due to the effect of the interference of the *multiple* scattered waves from the wires composing the structure, while the negative permeability of resonator media is due to the effect of the resonances in each *individual* metal resonator. In particular, the metal resonators can be interpreted by analogy to atoms characterized by a (magnetic) polarizability, whereas this cannot be said of the thin metal wires. The effective permeability of the resonator medium can be obtained by estimating the polarizability of each resonator and the number of resonators per unit volume, where the polarizability depends, to a very good approximation, *only* on the internal structure of the resonators, and not on their spacing. This is not the case for the wire medium, because if one were to define an effective wire polarizability, it would have to depend not only on the radius of the wires, but also on the distances between them and on their spatial disposition. The negative permittivity of wire media is then an intrinsically *collective* effect. This is why the permittivity of a single wire grating is an ill defined quantity. One can only speak meaningfully of the permittivity of a two dimensional structure.

Consequently the expressions of Eqs. (3.9) and (3.10) are correctly interpreted as the effective permittivity of a rectangular *two-dimensional* wire medium with periods  $d$  and  $h$ .

### 3.4.3 Effective permittivity - numerical

We now numerically check the accuracy of Eq. (3.8) by comparing the transfer matrix of the three layer structure of Fig. 3.4,  $M = T(h/2, \epsilon_m) \times T_W \times T(h/2, \epsilon_m)$ , with the transfer matrix of the equivalent homogeneous slab,  $T(h, \epsilon_{\text{eff}})$ . We define  $E(\epsilon_{\text{eff}}) = \frac{\text{norm}(M - T(h, \epsilon_{\text{eff}}))}{\text{norm}(M)}$  using Matlab’s “infinity” norm and we plot it as a function of  $\epsilon_{\text{eff}}$  in Fig. 3.6. There is a pronounced minimum around  $\epsilon_{\text{eff}} = -1.9$  where the homogeneous slab approximates the wire grating to within better than 5%. This numerical estimate is in very good agreement with the analytical formula of Eq. (3.9) which gives a value of  $\epsilon_{\text{eff}} = 1.85$ .

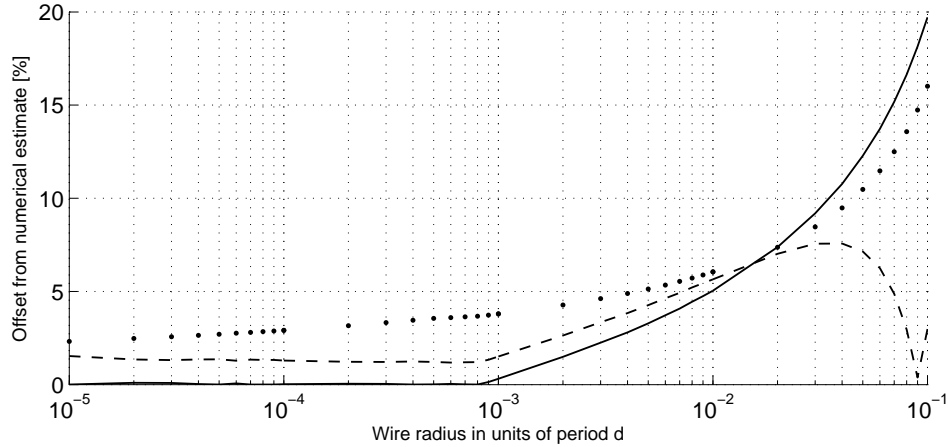


Figure 3.7: Comparison of analytical formula of Eq. (3.9) (solid line) with Maslovski’s formula (dashed line). The figure shows the relative offset of both estimates in percent from the numerical estimate given by the minimum of  $E$  as shown in Fig. 3.6. The dots represent the “homogenizability”  $E_{\min}$  in percent (see caption of Fig. 3.6)

We now compare our estimate with that given by Maslovski [56] in Fig. 3.7:

$$\varepsilon_{\text{eff}}^M = 1 - \frac{\lambda^2}{2\pi d^2 \ln\left(\frac{d^2}{4r(d-r)}\right)}.$$

Eq. (3.9) works better than Maslovski’s formula for wire radii smaller than about  $d/50$ . For wire radii larger than this Maslovski’s formula works better, but at the same time,  $E_{\min}$  becomes larger than 10% meaning that the structure is less and less homogenizable. For wire radii smaller than  $d/1000$  the accuracy of Eq. (3.9) is limited only by the numerical precision.

### 3.5 Dielectric resonators

O’Brien and Pendry were the first to point out that resonances in high dielectric fibers can be used to create artificial magnetism in metamaterials [70]. One starts by observing the field distribution inside and around a dielectric fiber illuminated at its first resonance. For the case of large relative permittivities the images are quite striking. In Fig. 3.8 we plot the magnetic field and the displacement current in a circular fiber of radius 0.5 and  $\varepsilon = 200$ . It is clear that the resonance induces a large magnetic field within the fiber due to the circulation of the displacement current. In fact, the circulating displacement current plays the role of a current loop, which then acts as a source for the magnetic field. The question then naturally arises as to whether by filling space with such dielectric rods one may obtain a metamaterial with an effective magnetic response, which may therefore be described by an effective permeability. Such a possibility has the element of novelty that it allows one to construct an effectively magnetic medium from purely dielectric components (the individual rods). As O’Brien and Pendry point out, at the right wavelength, such a medium behaves as a magnetic conductor, or a magnetic plasma, the Mie resonances of the rods acting as fictitious magnetic poles.

The homogenization theory of high dielectric resonating media was placed on a rigorous mathematical basis by the work of Felbacq and Bouchitté [71, 69, 72]. For instance, in the case of a medium of square rods of period 1 and side  $a$  the effective permeability around the first two  $s$ -type (monopolar, or non-null

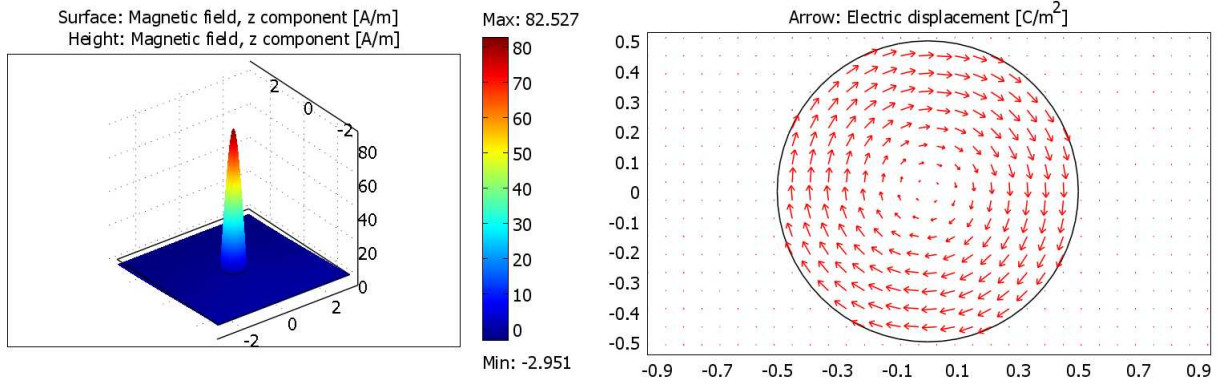


Figure 3.8: Magnetic field (left) and electric displacement field (right) of resonating dielectric rod, with  $\varepsilon = 200 + 5i$ ,  $r = 0.5$  at  $\lambda = 18.4$ . The displacement field outside the rod is too small for the arrows to be visible on the plot, which also explains the weak magnetic field there (left).

mean value) resonances is given by:

$$\mu_h = 1 + \frac{64}{\pi^4} \left( \frac{k^2}{\left(\frac{2\pi^2}{a^2\varepsilon_r} - k^2\right)} + \frac{k^2}{9\left(\frac{10\pi^2}{a^2\varepsilon_r} - k^2\right)} \right). \quad (3.11)$$

It is interesting to point out that even though the homogenization theory referred to above is valid for an infinite 2D medium of dielectric rods, the model continues to be a good approximation even for the case of a single row, a monolayer, see Fig. 3.9. This is due to the fact that the high relative permittivity of the rods strongly localizes the field, leading to very weak evanescent coupling between neighboring rods. Consequently each rod acts as a scattering center in the quasi-static approximation (due to the large wavelength), and the monolayer thereby falls within the domain of applicability of the Mossotti-Clausius model. It is one of the remarkable features of this model and an indication of its robustness that over 92% of the “bulk” local field seen by any particular dipole is due to other dipoles in the same plane. In other words a monolayer approximates the bulk to within better than 8%, and the agreement improves quickly as one adds more layers [68, 73].

### 3.6 Metal resonators

The changing magnetic field at the position of the resonator produces an induced electro-motive force  $\mathcal{E}_H$  in each ring:

$$\begin{aligned} \mathcal{E}_H &= -\mu_0 S \frac{\partial H}{\partial t} \\ &= j\omega\mu_0 S H \end{aligned}$$

where  $S$  is the surface area of the ring,  $S = \pi r^2$  and  $H$  is the amplitude of the magnetic field at the position of the resonator. We assume that the field is approximately homogeneous over distances on the scale of  $r$ . The current induced in each ring is  $I$  and obeys Ohm’s law

$$I \left( \frac{1}{j\omega C} + j\omega(M + L_s + L_i) + R \right) = \mathcal{E}_H$$

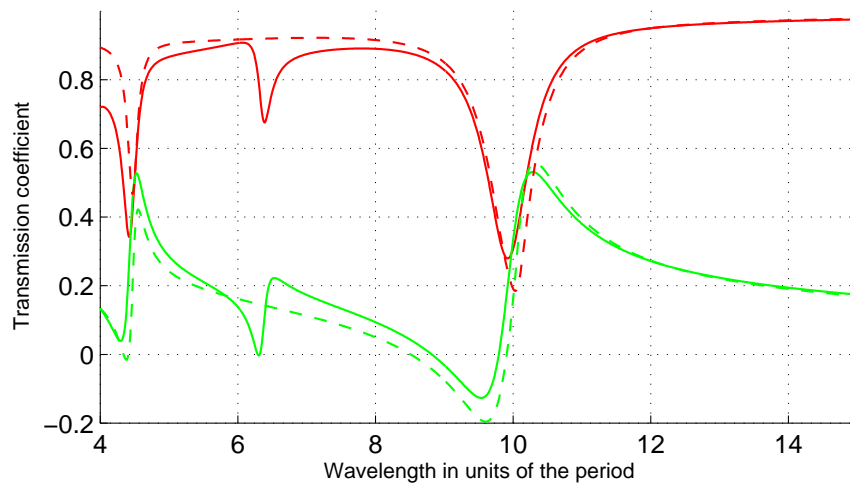


Figure 3.9: Real (red) and imaginary (green) parts of the transmission coefficient for grating (solid) and equivalent homogeneous slab (dashed). The grating has period 1, rods are squares of side 0.5, and  $\varepsilon_r = 200 + 5i$ . The effective permeability of the slab is given by Eq. (3.11) and the effective permittivity is 1.7, see [Felbacq Bouchitté].

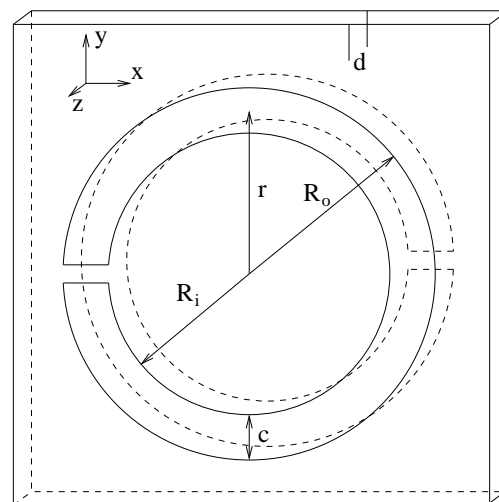


Figure 3.10: Broadside coupled resonator. The thickness of the metallic strips as well as widths of the gaps do not appear in the formulas characterizing the resonance subject to reasonable conditions discussed in the text.

Each ring is modeled as a series RLC circuit. The capacitance  $C$  is the sum of the mutual capacitance of the two rings and the gap capacitance. The inductance is given by the sum of the mutual inductance of the rings,  $M$ , the self-inductance of each ring,  $L_s$ , and the low frequency internal inductance of the conductors,  $L_i$ . In the following we shall note  $L = M + L_s + L_i$ . Since the rings are considered to be made of a perfect conductor, the resistive term is not related to Joule losses but rather to radiative losses. Each ring acts as a small emitting antenna. This term is necessary for maintaining the energy balance, as discussed at the end of Section 1.4.1.

The magnetic moment  $m$  of a small current loop is given by  $m = IS$  where  $I$  is the current, and  $S$  is the surface area. Since there are two loops we have

$$\begin{aligned} m &= 2IS \\ &= \frac{2S\mathcal{E}_{\mathbf{H}}}{\left(\frac{1}{j\omega C} + j\omega L + R\right)} \\ &= \frac{2j\omega\mu_0 S^2}{\left(\frac{1}{j\omega C} + j\omega L + R\right)} H \end{aligned}$$

and we can define the complex magnetic polarizability of the broadside coupled resonator as

$$\gamma_{\text{cplx}}^{\text{m}} = \frac{2j\omega\mu_0 S^2}{\left(\frac{1}{j\omega C} + j\omega L + R\right)}$$

which can be put in the form

$$\gamma_{\text{cplx}}^{\text{m}} = \frac{2\omega^2\mu_0 S^2}{L} \frac{1}{\omega^2 - \omega_1^2 - \frac{j\omega^2 k^3 \mu_0 S^2}{3\pi L}}. \quad (3.12)$$

We have introduced  $\omega_1 = 1/(LC)$  and the radiative losses are accounted for through the  $R$  term with

$$R = \frac{k^3 \omega \mu_0 S^2}{3\pi}$$

where  $k$  is the free space wavevector of the radiation at frequency  $\omega$  [74, 23].

The appearance of this imaginary term may seem surprising because it is immediately seen that by putting expression (3.12) into the Mossotti-Clausius relation we obtain an imaginary permeability even though no absorbing media are present. This aspect was already discussed in Section 1.4.1 with the conclusion that the complex polarizability  $\gamma_{\text{cplx}}^{\text{m}}$  must be replaced, in the Mossotti-Clausius relation by the real polarizability given by

$$\gamma^{\text{m}} = \frac{1}{\text{Real}\left(\frac{1}{\gamma_{\text{cplx}}^{\text{m}}}\right)}. \quad (3.13)$$

$\gamma_{\text{cplx}}^{\text{m}}$  may be seen as the free-space polarizability while  $\gamma^{\text{m}}$  is the polarizability in the bulk.

We now discuss the inductive terms,  $L$  and  $M$ . All inductances are given in SI units of Henry.

As far as the inductive properties of thin metal strips of width  $c$  are concerned, at large wavelengths, they are equivalent to cylindrical wires of radius  $r_w = c/4$ . Consequently we shall treat each ring as a loop of cylindrical wire of radius  $r_w$ . The mutual inductance of two parallel loops of wire of radius  $r$  spaced  $d$  apart is given by [75]

$$M = \mu\mu_0 r \left[ \left( \frac{2}{\xi} - \xi \right) K(\xi) - \frac{2}{\xi} E(\xi) \right]$$

where

$$\xi^2 = \frac{4r^2}{d^2 + 4r^2},$$

$$E(x) = \int_0^{\pi/2} \sqrt{1 - x^2 \sin^2(\phi)} d\phi$$

and

$$K(x) = \int_0^{\pi/2} \frac{d\phi}{\sqrt{1 - x^2 \sin^2(\phi)}}.$$

The self-inductance of each loop is given by

$$L = \mu\mu_0(2r - r_w) \left[ \left(1 - \frac{\xi^2}{2}\right) K(\xi) - E(\xi) \right]$$

where

$$\zeta = \frac{4r(r - r_w)}{(2r - r_w)^2}.$$

The low frequency internal inductance of round wires is given by [75]

$$L_i \approx 2\pi r \frac{\mu\mu_0}{8\pi} = \mu\mu_0 \frac{r}{4}.$$

It is easy to verify that this quantity is about two orders of magnitude smaller than the other two terms making up the total inductance (the mutual- and the self-inductance) and we ignore it in the rest of this work.

The total capacitance of the resonator is given by

$$C_{\text{tot}} = C_m + 2C_s.$$

Since this must be divided between the two loops, the series capacitance  $C$  seen by each loop is given by

$$C = C_s + \frac{C_m}{2}.$$

However, it is easy to see that the mutual capacitance will be far larger than the split capacitance as long as the two loops are very close together and the split is not too narrow. The validity of this assumption is discussed further below.

We now evaluate the mutual capacitance. If the charge distribution over the two loops were uniform, such that one loop would be positively charged and the other negative, with the charge distributed uniformly over their surface, then the mutual capacitance of the two loops would simply be equal to  $C = \pi\epsilon\epsilon_0 rc/d$ . However, the loops must remain neutral, and the charge distribution cannot be uniform. It is shown in the Appendix of Ref. [76] that for the first (lowest frequency) resonance this has the effect of reducing the mutual capacitance by a factor of  $\pi$ . Consequently the capacitance of the loops takes the form

$$C = \epsilon\epsilon_0 \frac{rc}{d}. \quad (3.14)$$

Once we have determined the magnetic polarizability of the resonators, we can insert it into the magnetic Mossotti-Clausius relation to obtain the effective permeability of a medium composed of such resonators. If the wavelength is much larger than the period of the structure, then the permeability is approximately local and we have

$$\mu_z = 1 + \chi^m = 1 + \frac{N\gamma^m}{1 - N\gamma^m/3}, \quad (3.15)$$

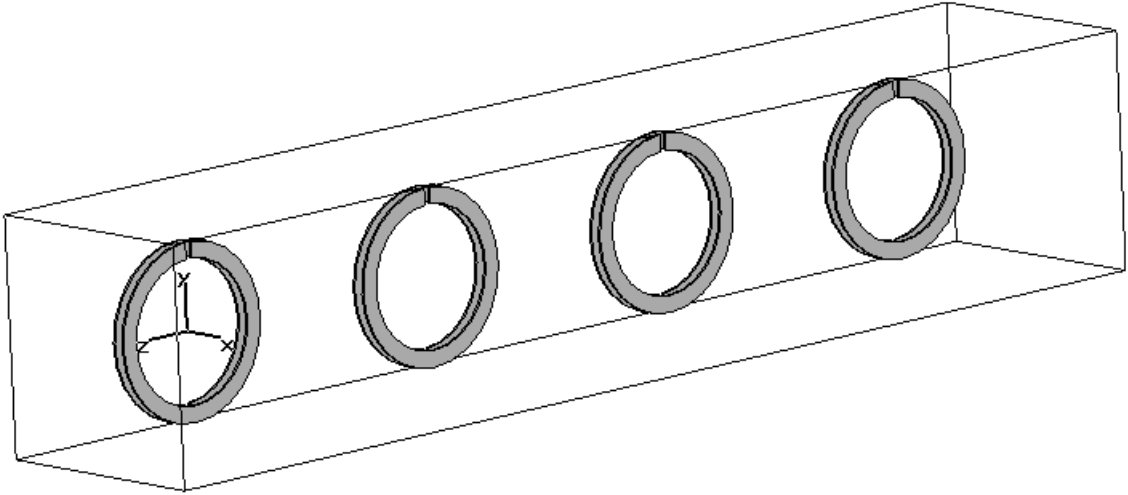


Figure 3.11: The unit cell of the slab of resonators. The structure is repeated in the  $x$  and  $y$  directions with period 1cm. The resonators are spaced by 1.3cm in the  $z$  direction.

where  $N$  is the resonator number density, or  $N = V^{-1}$  where  $V$  is the volume of the unit cell which contains a single resonator.

We have now obtained a model for the *magnetic* behavior of a medium filled with metallic resonators, as long as the axes of all resonators are aligned with the magnetic field of the incident wave, and the resonators are evenly (e.g. periodically) distributed. But the resonators also have an *electric* response in addition to the magnetic response. The electric field of the incident wave produces a polarization of the metal rings, resulting in an induced electric dipole moment. In the polarization under study the magnetic field is parallel to the axis of the loops and the electric field must lie in the  $x - y$  plane (See Fig. 3.10).

It has been pointed out by various authors that when the electric field is oriented in the  $y$  direction, across the ring splits, an electric resonance appears which complicates both theoretical and experimental studies of the magnetic behavior of the metamaterials. Consequently we prefer to orient the electric field in the  $x$  direction, in which case the behavior is independent of the presence of the splits. In this case, the polarizability of a round loop of radius  $r$  of round wire of radius  $r_w$  can be shown to be equal to [77]

$$\gamma^e = \frac{4\pi^2 r_{\text{eff}}^3}{\log\left(\frac{8r_{\text{eff}}}{r_{w,\text{eff}}}\right) - 2}$$

and using the electric Mossotti-Clausius relation we have

$$\varepsilon_x = 1 + \frac{N\gamma^e}{1 - N\gamma^e/3}.$$

It is worth noting that the radius of the loop and of the wire which must be used in the expression for the electric polarizability are not the same as  $r$  and  $r_w = c/4$  as used above, because in that case we were considering the *inductive* properties of the loops, whereas here the *polarization* properties are under study. Consequently,  $r_{\text{eff}}$  is not necessarily equal to  $r$  and  $r_{w,\text{eff}}$  is not necessarily equal to  $c/4$  as above. Numerical simulations show that the effective permittivity of the resonator medium is approximately given by  $\varepsilon_y = 1.41$ . Since we are working in  $E_y$  polarization, this is the only component of the permittivity tensor which appears in the effective medium model.

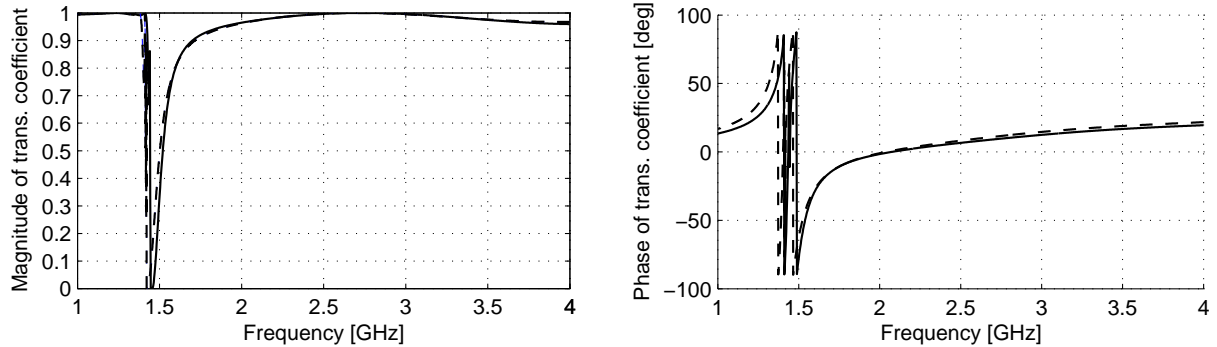


Figure 3.12: Magnitude of the transmission coefficient for normal incidence onto a slab of four layers of resonators with  $r = 3.5\text{mm}$ ,  $c = 1\text{mm}$ ,  $d = 0.05\text{mm}$ . Numerical (solid) and theoretical (dashed).

With this, our model for the effective permeability and permittivity of a metamaterial composed of broadside coupled split ring resonators distributed uniformly in space is complete. Fig. 3.12 shows excellent agreement between our analytical model (dashed) and the numerical calculation (solid), *in both amplitude and phase*. The plot compares the magnitude of the transmission coefficient through a medium composed of four planes of resonators. The unit cell is shown in Fig. 3.11. The numerical computations are done using the finite element method as implemented in the commercial package CST MICROWAVE STUDIO. All metal components are modeled as perfect conductors because they only require surface meshing resulting in considerably faster calculations than for real conductors, which must be meshed throughout their volume.

However, we must point out that the very good agreement between theory and numerical calculations is due to the fact that a configuration was chosen where the simplifying assumptions made along the way hold to a good approximation. In particular, the distance between the two metal loops was chosen extremely small, 200 times smaller than the period. We now discuss why this is necessary.

There are four different effects that have been left out of the analytical calculation. First is the split capacitance  $C_s$ , second is the internal inductance of the strips  $L_i$ , third is the correction to the mutual inductance  $L_m$  due to the metal being thin strips rather than cylindrical wires, and fourth is the edge effect due to the fact that the two metal surfaces are not infinite, and the electric field lines close to the edges bulge out into space or even couple to nearby objects. The first two effects tend to increase the capacitance and the inductance pushing the resonance frequency lower, while the third effect tends to decrease the mutual capacitance  $C_m$  pushing the resonance frequency higher. In the case of Fig. 3.12 the fourth effect is slightly stronger than the first three, resulting in a very slightly higher resonance frequency than that predicted theoretically; the solid curve resonates just above the dashed curve. Overall, however, the agreement is very good; due to the fact that since the two rings are very close together, both the mutual capacitance and the mutual inductance are large and the various corrections are small.

However, as we increase the separation  $d$  we expect our analytical values to *underestimate* the effective capacitance and inductance of the resonators since the split capacitance and the internal inductance are no longer negligible. The increasingly important effects of the edges complicate the treatment further, with the result that for larger separations  $d$  we treat  $L$  and  $C$  as fit parameters which are determined phenomenologically before proceeding any further. They are obtained by fitting  $\gamma^m$  of Eq. (3.13) to the numerical data.

The table below gives the values of effective inductance and effective capacitance for different resonator geometries. The unit cell in all cases is a brick of  $1 \times 1 \times 1.3$  centimeters where the structure is periodic in the  $x$  and  $y$  directions with period of 1cm. The period in the  $z$  direction is therefore 1.3cm. The

$R_o$ [cm]	$R_i$ [cm]	$d$ [cm]	$L^{\text{eff}}$	$C^{\text{eff}}$
0.4	0.3	0.005	21.39nH	0.534pF
0.4	0.3	0.01	20.63nH	0.298pF
0.37	0.3	0.02	15.76nH	0.203pF
0.37	0.3	0.04	16.79nH	0.11pF
0.37	0.3	0.06	16.32nH	0.09pF
0.37	0.3	0.08	14.64nH	0.087pF
0.37	0.3	0.1	13.52nH	0.086pF

Table 3.2: Effective capacitance and inductance of broadside coupled split ring resonators for different geometries.

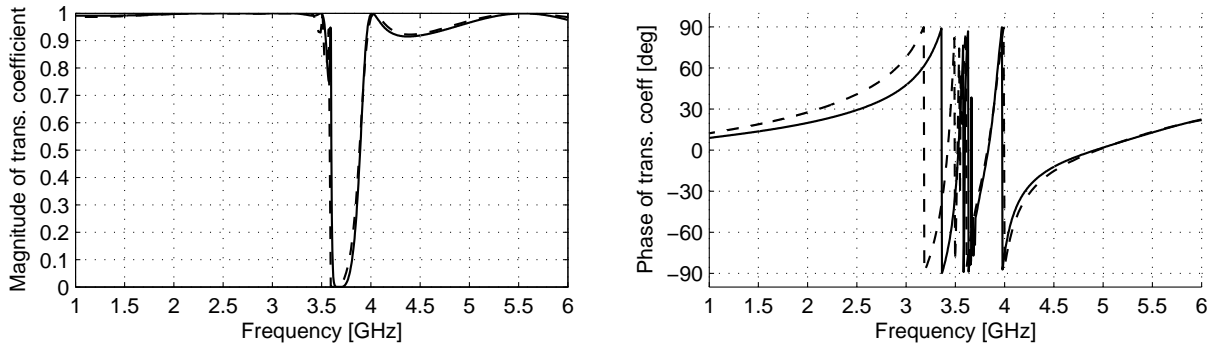


Figure 3.13: Magnitude of the transmission coefficient for normal incidence onto a slab of four layers of resonators with  $R_o = 3.7\text{mm}$ ,  $R_i = 3\text{mm}$ ,  $d = 0.4\text{mm}$ . Numerical (solid) and theoretical (dashed).

outside and inside radii of the resonators are noted  $R_o$  and  $R_i$ , such that in the notation of Fig. 3.10  $c = R_o - R_i$  and  $r = (R_o + R_i)/2$ .

As the distance between the loops increases, both the mutual capacitance and the mutual inductance decrease, leading to larger resonance frequencies. Whereas the resonance frequency for  $d = 0.05\text{mm}$  is slightly below 1.5GHz, for  $d = 0.4\text{mm}$  the resonance is higher than 3.5GHz as can be seen in Fig. 3.13. It can be seen that for larger metal loop separations the fit is not quite as good as above, at least insofar as the phase is concerned. However, in the frequency range of interest, in this case just below 4GHz, the agreement continues to be very good.

It is important to recall, however, that these simulations are done for normal incidence only. Since the most interesting applications of composite metamaterials involve fields incident at angles far from normal incidence we now check the validity of our homogeneous model as a function of angle of incidence. In the plots below, we compare the magnitude and phase of the transmission coefficient obtained from 3D finite element calculations and from transfer matrix homogeneous slab calculations for the structure of Fig. 3.12 for incidence angles varying from  $0^\circ$  to  $80^\circ$  in steps of  $10^\circ$  at a frequency of 1.5GHz. The agreement is excellent.

*The wavelength at this frequency is  $\lambda = 20\text{cm}$  which is over 15 times larger than the period of the structure and 25 times larger than the diameter of the resonators.*

Given the excellent agreement between rigorous simulation and theoretical model, we may confidently claim that the resonator medium behaves as an anisotropic homogeneous medium with an effective permeability tensor with one negative element in a given frequency range <sup>1</sup>. The permeability can be accurately calculated using the model presented in this section.

<sup>1</sup>The permeability tensor has both positive and negative elements making this an indefinite homogeneous effective medium [78].

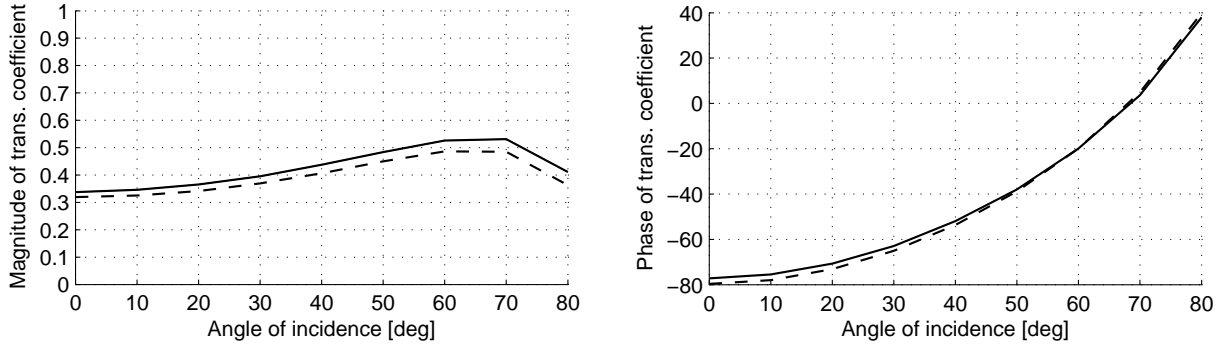


Figure 3.14: Magnitude and phase of the transmission coefficient for a slab of four layers of resonators with  $R_o = 4\text{mm}$ ,  $R_i = 3\text{mm}$ ,  $d = 0.05\text{mm}$  at a frequency of  $1.5\text{GHz}$  and for an incidence angle varying between  $0^\circ$  and  $80^\circ$ . At this frequency we have effective permeabilities  $\mu_x = -0.838$ ,  $\mu_y = 1$  and  $\mu_z = 1$ . The effective permittivity is  $\epsilon_y = 1.41$ . 3D finite element calculation (solid) and effective medium theory (dashed).

### 3.7 Composite metamaterial – the 1D stack approach

In this section we put together the results of Section 1.4 with the results of the previous sections of the current chapter. Since, as emphasized in Section 1.8 and in Appendix B we search for a constructive model of the metamaterials under study, we must work within the Small Atoms configuration. What is required for this, however, is not only a wavelength sufficiently large compared to the period, but also that the atoms be sufficiently far from one another and their fields sufficiently localized. We avoid squeezing together the wires and resonators as done in other studies [79] and moreover we choose a resonator geometry exhibiting fields that are strongly localized within the narrow space between the two metal loops. By, in addition, placing the elements in such a way as to minimize the near field overlaps (see section 3.1.2), we create a situation where the elements composing our material see each other effectively as point *dipoles*, as opposed to point *multipoles*, making the medium amenable to analytical methods.

In the previous sections we have studied the effective media associated with collections of thin wires on one hand, and of split ring resonators on the other. The thin wire medium exhibits an effective negative permittivity for frequencies *below* a certain cutoff which depends on the radius of the wires. For thinner wires the cutoff frequencies are lower. The variation of the cutoff with the wire radius is slow, since the radius appears inside a logarithm function, see Eq. (3.5). The resonators placed in a regular array exhibit a negative effective permeability within a certain small frequency range around the resonance frequency for fields in a single space direction. The resonance frequency depends on the mutual inductance and capacitance of the two rings, and is generally speaking shifted to lower frequencies as the two rings are brought closer together and interact more strongly.

As discussed in section 3.1.2, the mixing of the two types of media has raised questions relating to how the negative permittivity and negative permeability effects influence and perturb each other. The best solution seems to be to place resonators at nodes of the magnetic field of the wires, as illustrated in Fig. 3.15. In this way the interference of the scattered waves of neighboring wires is only minimally perturbed by the presence of the resonators.

This we first illustrate on the structure composed of wires and *closed* resonators. As is well known and discussed in Ref. [80] as well as in section 3.6 above, metal ring resonators have not only a magnetic response but also an electric response. By closing the resonator one eliminates the magnetic response, and is able to study the effect that the electric polarization of the resonators due to the incident field has on the behavior of the wires. For the resonator geometries used here the value of the effective permittivity seems to be close to  $\epsilon_y \approx 1.41$ . That this value gives results in good agreement with 3D

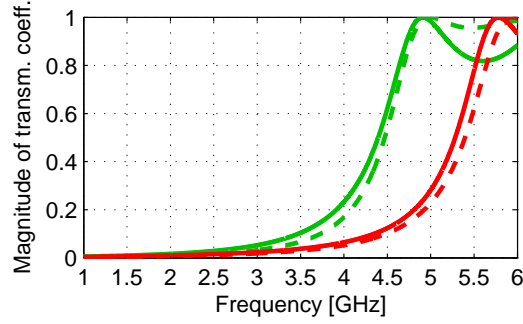


Figure 3.16: Magnitude of the transmission coefficient of wires alone (red) and of wires and closed resonators (green) as a function of frequency. Dashed curves are given by Eq. (3.10) with  $\varepsilon_m = 1$  (red) and  $\varepsilon_m = 1.41$  (green). The wires are of radius 0.05 mm and the structure is composed of 4 rows of period 1 cm and spaced by 1.3 cm from each other.

numerical simulation can be seen in Figs. 3.12 and 3.14.

The effect of the superposition of the electric response of the resonators and of the wires is to lower the plasma frequency of the wires, or, in other words, to increase the effective permittivity of the composite by approximately 0.41. This can be seen immediately from Maslovski's formula giving the effective permittivity of a wire lattice immersed in a medium with  $\varepsilon_m$  and  $\mu_m$ :

$$\varepsilon_{\text{eff}} = \varepsilon_m - \frac{2\pi}{\mu_m(kd)^2 \log\left(\frac{d^2}{4r(d-r)}\right)}.$$

However, we are dealing with a rectangular unit cell, so this formula cannot be applied directly.

Instead we use Eq. (3.10):

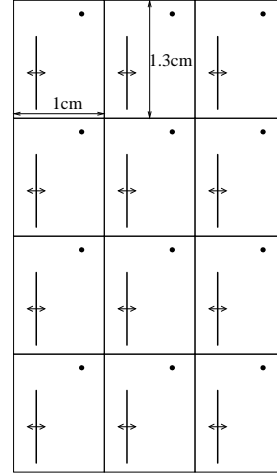


Figure 3.15: Schematic metamaterial geometry.

$$\varepsilon_{\text{eff}} = \varepsilon_m + \frac{1}{k^2 h^2} \left( \arccos\left(1 + \frac{h}{L}\right) \right)^2. \quad (3.16)$$

Thus the electric behavior of the medium is fairly well understood. It remains to combine the permittivity of Eq. (3.10) with the permeability as derived in the previous section. We consider the two resonator geometries that were also studied in the previous section: in the first type of resonators were 1mm wide and  $d = 0.1\text{mm}$  apart, while the second were 0.7mm wide and  $d = 0.4\text{mm}$  apart. For the purpose of this study the most important difference between them is their resonance frequency. In the first case it is close to 2GHz while in the second case it is close to 4GHz. The operating wavelength therefore differs by a factor of two.

The question which we shall attempt to address is that of the homogeneity of the metamaterial, or, in other words, the importance of the spatial dispersion which a medium exhibits. If the spatial dispersion is important then the homogeneous effective medium model is of limited *usefulness*, while if it is negligible, then the homogeneous effective medium constitutes a very useful and simple model, opening the way to the design of materials with remarkable properties.

In general, in order to check that a medium is spatially dispersive, one must perform calculations as a function of two different parameters, the wavelength, and the angle of incidence, and compare the results with the results given by the homogeneous model. The experimental data available is encouraging, but

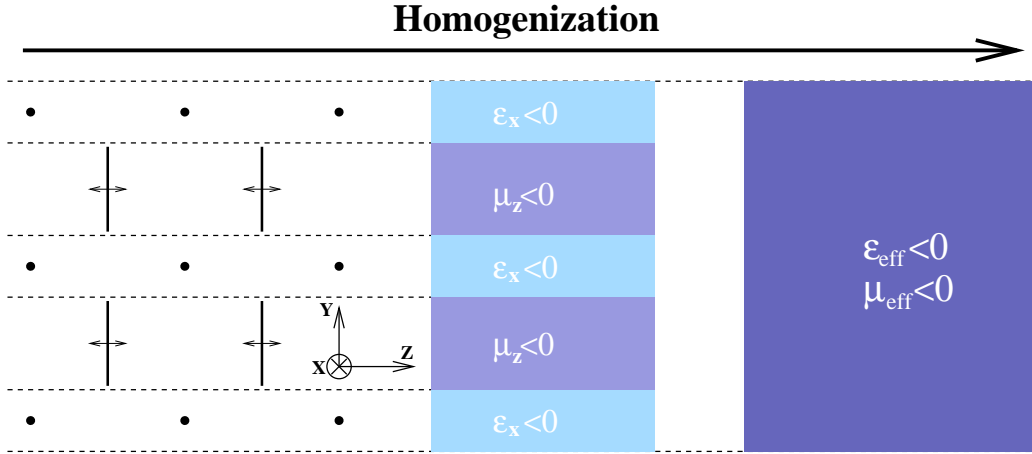


Figure 3.17: Homogenization as a progression.

not conclusive. The existence of spatial dispersion in composite media is rarely acknowledged in the literature [81]. Consequently at the present time, in the view of the author, it is premature to claim that negative index media have been experimentally observed.

In this section I put forward a method whereby a designer may estimate whether a given metamaterial will behave as a homogeneous effective medium, or whether, on the contrary, the spatial dispersion is expected to be strong. The argument below is not meant to be mathematically rigorous, but to provide a certain intuition about the logic behind the method. It follows naturally from the discussion of Section 1.7 and is illustrated visually in Fig. 3.17.

The idea is to homogenize the composite, but only partially, obtaining not a homogeneous effective medium but an effective slab stack. The geometry of this slab stack would mimic the geometry of the composite, as shown in Fig. 3.17. Each wire row would be represented as a negative permeability slab, and each row of resonators as a negative permittivity slab. In this context it seems reasonable to assume that *if the composite behaves as a homogeneous medium, then so does the slab stack*. The homogenization of the metamaterial must “*pass through*” the homogenization of the 1D stack, in some sense. This is what we refer to as the “*squeeze conjecture*”. An equivalent way to state it is that if the partially homogenized medium behaves *inhomogeneously*, then so will the composite medium. We are using an intermediate homogenized model to gauge how far the composite metamaterial is from its homogeneous effective medium.

Let us explain in more detail what we mean by “intermediate model” (see Section 1.7) and to give a more precise meaning to our notion of “homogenization axis”.

We begin by recalling that one of the ways that has been used to obtain homogeneous parameters from composite metamaterials (see Section 1.9 and Ref. [26]) is to define

$$\begin{aligned}\mu_{\text{eff}}^{i,j} &= \frac{\langle B_i \rangle}{\langle H_j \rangle} \\ \epsilon_{\text{eff}}^{i,j} &= \frac{\langle D_i \rangle}{\langle E_j \rangle}\end{aligned}$$

where the indices run over  $x, y, z$  and the brackets indicate spatial averaging. In Chapter 1 it was shown that spatial averaging can be written as a convolution with an averaging function  $f(\mathbf{x})$ , which is equivalent, in reciprocal space, to multiplication by the filter function  $\tilde{f}(\mathbf{k})$ . One may then say that when we are comparing the behavior of the composite with that of its effective medium, we are basically comparing two models of the structure corresponding to two choices of smoothing function  $f$ . The

composite itself corresponds to a choice of  $f$  whose support in reciprocal space *includes all lattice points*:  $f(\mathbf{x}) = \delta(\mathbf{x})$ , corresponding to the identity filter in reciprocal space. The homogeneous model corresponds to a choice of  $f$  whose support in reciprocal space *excludes all lattice points except the  $\Gamma$  point*. In this context then, an intermediate model would correspond to any model obtained by choosing a smoothing function  $f$  between the two extremes. In the case of the 1D stack model, the support of  $\tilde{f}(\mathbf{k})$  is chosen to exclude all lattice points except the  $\Gamma$  point *and also the points corresponding to the periodicity in the  $y$  direction* (Fig. 3.17). In this sense the 1D stack model can be said to be an “intermediate model” because the filter function  $\tilde{f}(\mathbf{k})$  used to obtain it lets through not only the DC component but also one higher harmonic, namely the one associated with the periodicity in the  $y$  direction.

Basically we have ordered the different possible models of the composite according to how many of the spatial harmonics of the 3D periodic structure they keep; the composite itself contains all harmonics; the homogeneous effective medium contains only one: the DC harmonic; the 1D stack contains two: the DC harmonic and the first  $y$  harmonic. This ordering clarifies somewhat what we meant above by “homogenization axis”. It is a graphical representation of a process whereby the spatial harmonics making up the field in a 3D periodic structure are progressively removed to obtain models progressively closer to the homogeneous effective medium.

In this language then, the “squeeze conjecture” above says that if two models obtained from different smoothing functions  $f$  give the same transmission and reflection for all angles of incidence, then for some intermediate smoothing function, the result must be the same. These ideas were also discussed, from a more fundamental point of view, in Chapter 1, Section 1.7.

We now show how to obtain the parameters of the 1D single negative stack. The overall homogeneous medium parameters are given by the equations of previous sections. The unit cell is 13mm deep, and the resonators occupy about 8mm. Consequently we model the structure as a slab-stack with  $h_1 = 8\text{mm}$  and  $h_2 = 5\text{mm}$  with  $\mu_2 = 1$  and where  $\mu_1$ ,  $\varepsilon_1$  and  $\varepsilon_2$  are fixed by requiring

$$\begin{aligned} \frac{h_1 + h_2\varepsilon_2}{h_1 + h_2} &= \varepsilon_{\text{eff}}^{\text{wires alone}} \\ \frac{h_1\mu_1 + h_2}{h_1 + h_2} &= \mu_{\text{eff}}^{\text{res. alone}} \\ \frac{h_1\varepsilon_1 + h_2}{h_1 + h_2} &= \varepsilon_{\text{eff}}^{\text{res. alone}} \end{aligned} \quad (3.17)$$

where  $\varepsilon_{\text{eff}}^{\text{wires alone}}$  is given by Eq. (3.16) and  $\mu_{\text{eff}}^{\text{res. alone}}$  by Eqs. (3.12), (3.13) and (3.15) and  $\varepsilon_{\text{eff}}^{\text{res. alone}} = 1.41$ . The permittivity and permeability are now position dependent, periodic with the period of the lattice, but in such a way that *they average to the fully homogenized effective medium values*,  $\varepsilon_{\text{eff}}$  and  $\mu_{\text{eff}}$ . In a sense, the homogenization of the composite metamaterial must “*pass through*” the homogenization of the single negative stack [82].

Let us now consider a structure analogous to the ones studied experimentally in the literature. We design the resonators such that the resonance frequency be just below the plasma frequency of the wires+closed resonators. Looking at Fig. 3.16 the frequency range between 3.5 and 4 GHz seems promising. The geometry that resonates at this frequency is the second geometry studied in the previous section namely, with  $R_o = 3.7\text{mm}$ ,  $R_i = 3\text{mm}$ ,  $d = 0.4\text{mm}$ . We shall refer to this as the “A resonator”, or the “A type geometry”. The wires are of radius  $r = 0.05\text{mm}$  and the unit cell layout is as in Fig. 3.15.

The frequency range between 3.5 and 4GHz corresponds to a free space wavelength around 6 times larger than the period, and about 10 times larger than the diameter of the resonators. This is similar to the experimental configurations used, among others, by Ref. [80]. The simulation results are presented in Fig. 3.18. The plot on the left side compares the magnitude of the transmission coefficient of the 3D composite (in blue) with the equivalent homogeneous effective medium, with  $\varepsilon_{\text{eff}}$  and  $\mu_{\text{eff}}$  as discussed

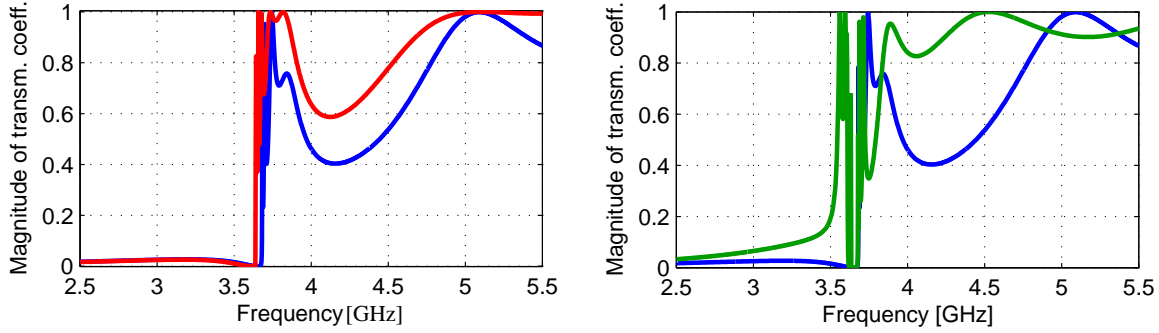


Figure 3.18: Magnitude of the transmission coefficient as a function of frequency; A type resonators with  $R_o = 3.7\text{mm}$ ,  $R_i = 3\text{mm}$ ,  $d = 0.4\text{mm}$ . 3D composite (blue), homogeneous model (red) and single negative stack (green).

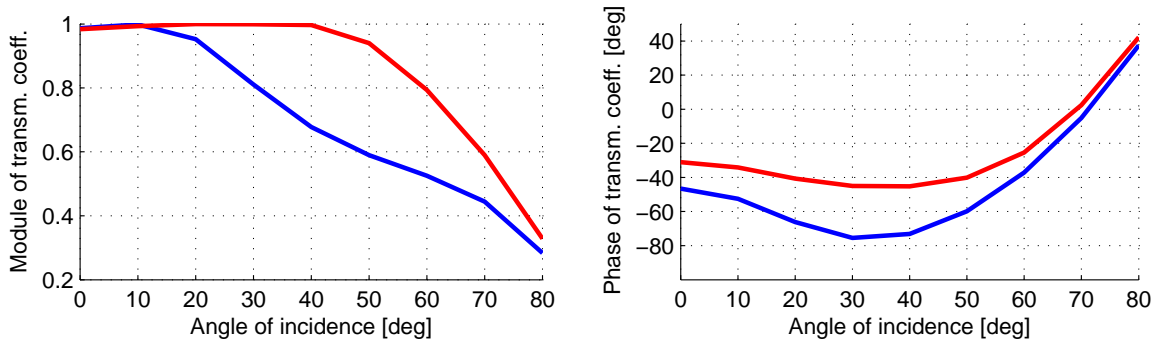


Figure 3.19: Module and phase of transmission coefficient as a function of angle of incidence. 3D composite (blue) and homogeneous model (red). The frequency is  $f = 3.748\text{GHz}$ . The homogeneous parameters values at this frequency are  $\mu_{\text{eff}} = -1.13$  and  $\varepsilon_{\text{eff}} = -0.412$ .

above (in red). We see that there is a transmission peak, indeed at the right frequency, as expected, and as reported experimentally in the literature, but that the agreement with the homogeneous model is not good. Only the general position of the peak is predicted correctly by the homogeneous model, not its precise shape. This peak is analogous to what Katsarakis et al. [80] refer to as their “true left handed” peak. However, on the right side of Fig. 3.18 we also compare the 3D composite with the single negative stack, with parameters that can be obtained from Eqs. (3.17). The agreement is very poor.

This is to us an indication that what we are looking at is not a transmission peak due to homogeneous behavior, but rather an inhomogeneous peak. In order to test this, we study the transmission of the material at a single frequency where the red and the blue curves agree ( $f = 3.748\text{GHz}$ ), by sweeping the angle of incidence. The results are presented in Fig. 3.19. Our impression is confirmed. The homogeneous model is a poor description of the behavior of this material and the agreement for normal incidence at this frequency is accidental. The magnitude as well as the phase of the transmission coefficient off normal incidence disagree for the 3D composite and the homogeneous model. *The most important consequence here is that simply reporting a transmission peak below the plasma frequency of a wire+resonator medium for a single angle of incidence does not “demonstrate” anything [83, 84, 85].*

We believe these results show that when using the A type resonators the wavelength is too small for homogenization theory to be truly useful. We need to work at longer wavelengths. This can be done by shifting the resonance frequency of the resonators to lower frequencies. For this purpose we use resonators with the “type B” geometry, that is with  $R_o = 4\text{mm}$ ,  $R_i = 3\text{mm}$ ,  $d = 0.1\text{mm}$ . In this second geometry we have reduced somewhat the wire radius, to  $r = 0.02\text{mm}$ . This is because at this lower frequency we are further away from the plasma cutoff frequency, which means that the effective

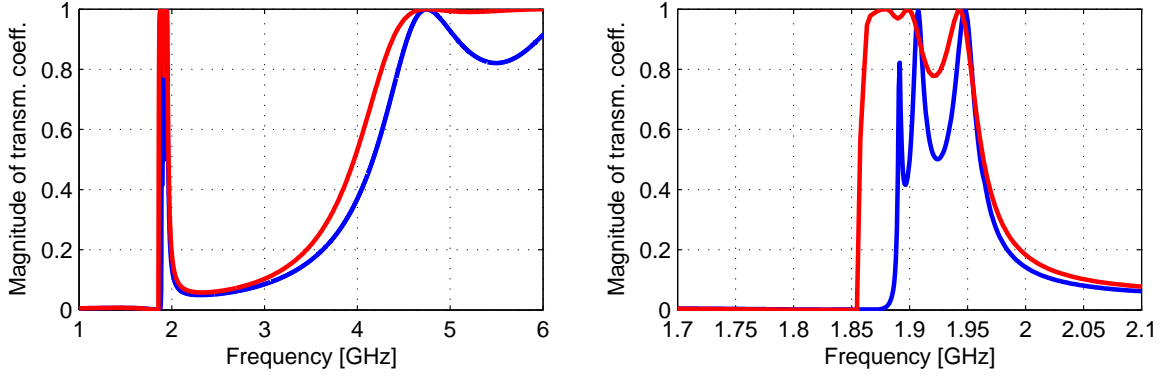


Figure 3.20: Magnitude of the transmission coefficient for normally incident plane wave on 3D structure (blue) and on equivalent homogeneous slab (red). See text for parameters.

permittivity of the wires is larger, more negative. This in turn makes the medium less homogeneous. By reducing the wire radius we lower the plasma frequency and reduce the effective permittivity of the wires making the medium more homogenizable.

The transmission curve looks as in Fig. 3.20. The homogeneous model used to obtain the red curve is very slightly different (several percent) from that of the resonators alone. The effective capacitance  $C$  of the resonators is slightly increased by the presence of the wires ( $C_{\text{eff}}^{\text{res}} = 0.336\text{pF}$ , about 10% higher than in Table 3.2), but this shift is so small that we consider this as confirmation that the placement of the resonators and wires in the configuration of Fig. 3.15 achieves its main aim, that of decoupling the near fields of the wires and resonators. Most importantly *the permittivity of the wires and the effective inductance of the resonators are unaffected*. See the discussion of subsection 3.1.2 above.

The two curves seem to be in good agreement, however, a closer look at the transmission peak will show that the agreement is not uniform. At some frequencies the agreement is better than at others. For instance above a frequency of 1.95 GHz the medium seems to behave homogeneously. According to our hypothesis outlined above, we expect the single negative stack model to agree with the homogeneous model at those frequencies where the homogeneous model agrees with the 3D composite, and to disagree with it where it does not. This hypothesis is confirmed by the results of Fig. 3.21. At those frequencies where the red curve agrees with the blue curve, it also agrees with the green curve. However, for frequencies between 1.85 and 1.95GHz the medium clearly does not behave homogeneously and the red and green curves disagree.

Moreover, and most remarkably, we see that in this case the agreement between the green and the blue curves is almost perfect everywhere, in the vicinity of the transmission peak. *At least in normal incidence the composite seems to behave exactly like a single negative stack*. Once more this confirms that the correct placement of the resonators has decoupled them from the inductive interaction with the wires, and that the part of the double-negative peak that is not homogeneous in nature is not due to the spatial dispersion of the infinite wires, as hypothesized by Simovski et al. [57], but rather to the periodicity and layerwise arrangement of the composite components.

But as mentioned above, plotting the transmission as a function of frequency is not enough to justify claims about the effective permittivity and permeability. We must also study how it varies as a function of the angle of incidence. This is done in Fig. 3.22 for a frequency of 1.958GHz. The agreement is far better than in the first case, where the type A resonators were used. In this second case we have also reduced the wire radius to  $r = 0.02\text{mm}$  in order to bring the effective permittivity closer to -1. At this lower frequency we are operating farther from the plasma frequency of the wires, resulting in a larger negative effective permittivity. A higher permittivity contrast between the wires and the resonators alone

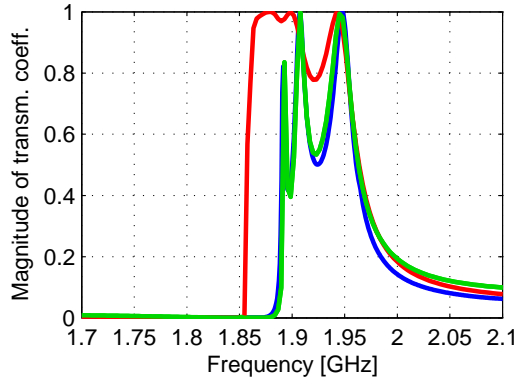


Figure 3.21: Comparing the transmission of the 3D composite (blue), the homogeneous model (red), and the slab stack (green). The agreement between the blue and green curves is striking.

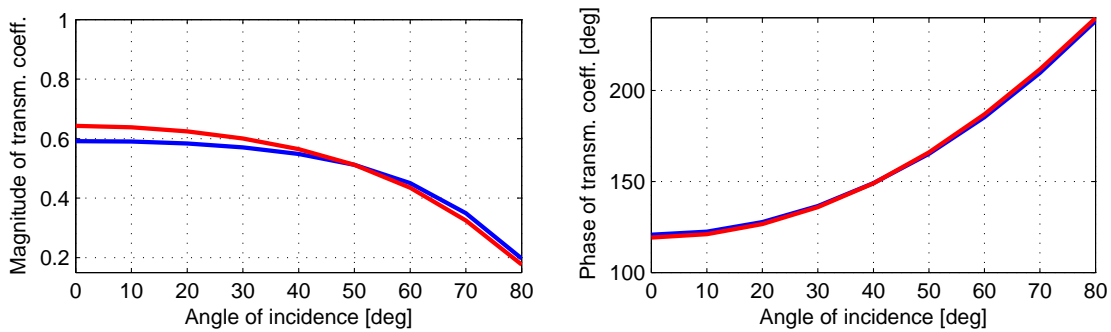


Figure 3.22: Module and phase of transmission coefficient as a function of angle of incidence. 3D composite (blue) and homogeneous model (red). The frequency is  $f = 1.958\text{GHz}$  and the wires are of radius  $0.02\text{mm}$ , a factor of 2.5 thinner than in Fig. 3.19. The homogeneous parameter values at this frequency are  $\mu_{\text{eff}} = -0.295$  and  $\varepsilon_{\text{eff}} = -4.332$ .

(which, recall have  $\varepsilon_{\text{eff}}^{\text{res. alone}} = 1.41$ ), is harder to homogenize.

These results show that it is possible to model composite metamaterials quantitatively, constructively. We have shown that the composite behaves as an anisotropic homogeneous slab at least at some frequencies, for all angles of incidence. In addition the above results show that it is possible to determine the frequency regions where the medium is likely to have a homogeneous behavior from regions where it is not by comparing the predictions of the homogeneous model with the 1D single negative stack model given by Eqs. (3.17). At the frequencies at which the two models disagree, it is highly likely that the spatial dispersion of the composite will be too large for it to be *usefully* modelled as a homogeneous medium. This is the first time that a straightforward method is proposed for gauging the homogeneity of a double negative composite metamaterial.

In order to relate these results to the discussion of Chapter 1 one may say that the 1D model is a way of accounting for the effect of the lattice dispersion in the medium. Recall that the lattice contribution is expected to dominate the spatially dispersive response of not-quite-homogeneous composites, and that it is due to the nonuniformity of the macroscopic field over the volume of the smoothing volume  $f(\mathbf{x})$ . The 1D model is the first and simplest way to account for this field nonuniformity and it can be seen that it works remarkably well, at least in cases where the medium is close to homogeneous, such as in Fig. 3.20. In cases where the medium is farther from homogenization, such as in Fig. 3.18, the spatial dispersion is too strong and a more complex model than the 1D stack is required, one which includes Bloch harmonics in other directions than the  $y$  direction. At such wavelengths the metamaterial becomes a meta-photon crystal, as discussed in Section 1.7.



# Conclusion

This work has been concerned with the study of the interaction of matter with the electromagnetic field in a classical context. In particular, we have been most interested in the study of configurations situated in the intermediate region where matter does not behave quite homogeneously, but where the scale is such that neither ray, beam, nor Fourier optics are very useful. These intermediate configurations are characterized by a wavelength that is on a scale comparable to the structure period.

In such cases there are two approaches to the characterization of the medium.

The first is a method that is a ramification of work done on solid state physics, and it relies on a Bloch wave description of the field propagation. This could be seen as the microscopic approach, since it provides an exact and rigorous description of the field in the structure. However, it has the drawback of being computationally intense, which means that it is intuitively opaque and much design work is reduced to an elevated form of trial and error. This approach is described and illustrated in Chapter 2 where it is used to study the superprism effect in dielectric photonic crystals. It is shown that photonic crystals with anisotropic unit cells are well suited for applications employing the superprism effect, particularly due to the fact that they offer the possibility of optimizing the transmission by changing the angle of incidence.

The second method originates in the field of optics, as it attempts to describe the medium by introducing the effective medium notions of permittivity and permeability (or index and impedance). This approach is macroscopic in character, in that it does not provide a complete description of the electromagnetic field. In the language of linear time invariant systems, the homogeneous description is an *external* description, since it renders inaccessible the internal variables related to the exact field distribution. The external description only includes enough information about the medium and the field to be able to correctly predict its interaction with the outside world in terms of incoming and outgoing fields. This effective medium approach has the advantage of being simple, intuitive, and readily amenable to design, with the drawback that it is an incomplete description. In its quest for simplicity and intuitiveness this description leaves out much information about the structure, and not always in a controlled manner. This is why, unfortunately, effective medium models are often too easily used to “describe” structures where such simplifications are not warranted.

The first Chapter of this work is concerned with a detailed discussion of the effective medium description from first principles, with much attention focused on the assumptions and hypotheses underlying the introduction of effective medium parameters such as the permittivity. The novel concepts of custom-made effective medium models and meta-photonic crystals are discussed, both being based on the use of the spatial smoothing function  $f(\mathbf{x})$  as a dial useful for trading off non-locality against inhomogeneity. The distinction between holistic and constructive approaches is also discussed, particularly in relation to the usefulness of the Mossotti-Clausius relation between the susceptibility and the atomic polarizability. It is shown that from a constructive point of view the Mossotti-Clausius relation is only valid in the case of Small Atoms, where the scatterers composing the medium are sufficiently far away from each other as to be considered oscillating *point dipoles*. In any other case the dependence of the susceptibility on

the polarizability becomes either heavily computational, or, more often, holistic, or phenomenological, thereby losing much of its usefulness.

Moreover, in the derivation of the Mossotti-Clausius relation, one crucial assumption is made, namely, that the same electric field acts on all atoms in the (infinite) lattice. In reality this is clearly not the case, but the approximation is in many cases far better than one would expect due to the fact the the strength of the interaction between lattice sites falls off rapidly with distance. Consequently, it is not necessary for the field to be uniform everywhere, but only over a volume the size of the “electromagnetic neighborhood” of the typical atom. Here, “electromagnetic neighborhood” of a given lattice site is understood as the region including all atoms interacting non-negligibly with the given site. When defining effective medium parameters, the important parameter is therefore not the ratio of the wavelength to the period, but rather the ratio of the wavelength to the size of the EM neighborhood. It is therefore useless to squeeze together atoms since even though this reduces the period, it is unlikely to reduce the size of the EM neighborhood. Placing the atoms too close together is also detrimental to the Small Atoms approximation rendering a constructive description more difficult.

The insights developed in Chapter 1 are applied in Chapter 3 to the study of negative index composite metamaterials. These media are composed of thin wires, exhibiting a strong electric response, and resonators, exhibiting a strong magnetic response. They are first studied separately in order to characterize their electric/magnetic behavior in isolation. In the last section of Chapter 3 they are combined and the behavior of the composite is submitted to detailed numerical scrutiny. In particular the external, homogeneous description is tested not only as a function of frequency, but also as a function of angle of incidence, *a first in the literature*. The agreement between the homogeneous model and the 3D finite element simulations is remarkable, in both magnitude and phase, as can be seen in Figure 3.22 in particular.

A model is proposed whereby the composite structure is modeled as a 1D stack of alternating negative permittivity/negative permeability layers. The predictive power of this model is remarkable, particularly when the medium is not quite homogenized, as can be seen in Fig. 3.21. It is shown that this 1D structure can be used to gauge the homogeneity of the 3D composite. The predictions made in Chapter 1, Section 1.7, are tested and confirmed by the numerical results of this chapter. When the 1D structure behaves inhomogeneously, then so does the 3D composite. This is the first time that a method is proposed that is able to distinguish in a straightforward way frequency regions where an effective medium model of a composite medium can be expected to be useful from regions where this is not the case. Most importantly these results open the way to a new and potentially very fruitful way of looking at composite metamaterials, as discussed in Section 1.7.

# Acknowledgements

Few are the doctorate students who can claim to have spent as much time with their doctorate advisor as I have. By his scientific enthusiasm, his spirit of scientific inquiry and his all-encompassing curiosity, Prof. Felbacq stands apart from a scientific herd guided by fashion, trend, and economic growth as opposed to the quest for understanding. I am profoundly indebted to him for the myriad thought provoking discussions and advice, and above all for his patience and understanding, especially through the rougher patches over the last three and a half years.

The working conditions in the Nanophotonics Group are beyond most doctorate students' wildest dreams. For this I think I speak for everybody when I thank David Cassagne, without whom the group would not have existed. His pragmatic thinking and efficiency have contributed to setting a high standard for the work in the Group. It was mainly through his efforts that our team participated in projects such as Cristel, PHAT and the European Network of Excellence PHOREMOST and was able to derive the corresponding, far from negligible, financial benefits.

I wish to thank Prof. Jean-Paul Albert for his wisdom and his faith in my abilities, faith I must confess I did not always share... I am deeply grateful to him for his encouragement and support, they always came exactly when I needed them.

My teaching advisor, Emmanuel Centeno is a rock of equanimity and Zen calm in the Group. I thank him for the innumerable discussions and for his level-headed, pragmatic approach, from which I have much to learn. Even after four years in the group I am still impressed by his efficiency.

My teaching activity during the last 4 years could not have gone smoother, in no small part due to the affable and efficient Prof. André Raymond, pedagogic director of the Physics Department, to whom I am grateful for his flexibility and patience, always managing to successfully slalom through a mire of constraints and preferences and find the best possible compromise.

I have to send out a big hearty thank you to the GDBM, composed of, in alphabetical order, Mathieu Nespoulous, Catherine Nicolas, Lauréline Roger, Guillaume Toquer and Kevin Vynck, who tolerated unerringly all my erratic philosophical flights of fancy and without whom my years spent in Montpellier would have been much darker and bleaker. I consider myself lucky to have met each one of you.

I wish to thank above all my parents, Cleopatra and Jean Căbuz for their moral and financial support. Nothing would have been possible without their unwavering faith and assistance. My debt to them has no bottom.



# Appendix A

## Multiscattering theory of circular rod gratings

### A.1 Single circular rod

In this appendix we develop a multiscattering theory of gratings composed of circular rods (whether metallic or dielectric). Only dipolar terms are kept, such that the theory only accounts for interactions between the rods that are mediated by dipolar fields. Comparing the results of this theory with those of an exact, rigorous method one can determine frequency domains where the rods behave as Small Atoms in the sense of subsection 1.4.1.

The basic unit of the grating is the individual circular rod. Its optical properties are characterized by what we will call the scattering matrix. This matrix is defined by the relationship between the incident waves and the diffracted waves. In cylindrical coordinates it is most convenient to express these in terms of infinite series of Bessel and Hankel functions respectively,  $J_n$  and  $H_n$ . Hankel functions of the first or second kind must be used as outgoing waves, according to which convention is used for the frequency  $e^{i\omega t}$  or  $e^{-i\omega t}$ .

The scattering matrix is obtained by matching the fields inside and outside the rod using the boundary conditions appropriate to the material of which the rod is made. For the case of cylindrically symmetrical objects the scattering matrix takes a diagonal form. For example for a metallic cylindrical rod surrounded by dielectric with  $\varepsilon = \nu^2$  we have

$$S_n^{metal} = \begin{cases} -\frac{J_n(k\nu R)}{H_n(k\nu R)} & \text{for } \mathbf{E} \parallel \text{polarization} \\ -\frac{J'_n(k\nu R)}{H'_n(k\nu R)} & \text{for } \mathbf{H} \parallel \text{polarization} \end{cases} \quad (\text{A.1})$$

while for a dielectric rod with  $\varepsilon_1 = \nu_1^2$  surrounded by dielectric with  $\varepsilon_2 = \nu_2^2$  we have

$$S_n^{dielectric} = \begin{cases} \frac{\nu_2 J_n(k\nu_1 R) J'_n(k\nu_2 R) - \nu_1 J_n(k\nu_2 R) J'_n(k\nu_1 R)}{\nu_1 H_n(k\nu_2 R) J'_n(k\nu_1 R) - \nu_2 H_n(k\nu_1 R) J'_n(k\nu_2 R)} & \text{for } \mathbf{E} \parallel \text{polarization} \\ \frac{(1/\nu_2) J_n(k\nu_1 R) J'_n(k\nu_2 R) - (1/\nu_1) J_n(k\nu_2 R) J'_n(k\nu_1 R)}{(1/\nu_1) H_n(k\nu_2 R) J'_n(k\nu_1 R) - (1/\nu_2) H_n(k\nu_1 R) J'_n(k\nu_2 R)} & \text{for } \mathbf{H} \parallel \text{polarization} \end{cases}$$

Here  $R$  is the radius,  $k = 2\pi/\lambda_0$  is the free space wavevector and all Hankel functions are of the first kind.

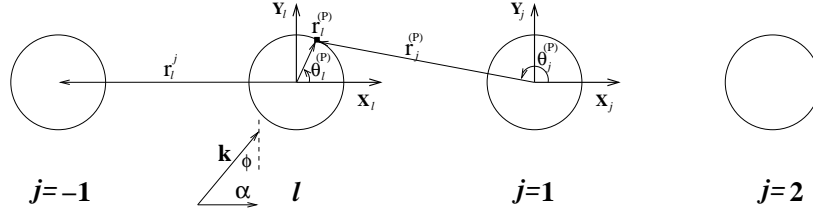


Figure A.1: Infinite grating notation.

## A.2 Infinite grating

Figure (A.1) shows the notation which we will use below. The continuity conditions are written on the boundary of the rods. Thus fields are evaluated at points such as P, of cylindrical coordinates  $r_l^{(P)}$  and  $\theta_l^{(P)}$  in the reference frame of rod  $l$ . The same point has coordinates  $r_j^{(P)}$  and  $\theta_j^{(P)}$  in the reference frame of rod  $l + j$ . Finally, the position of the center of rod  $l + j$  in the reference frame of rod  $l$  is given by  $r_l^j$  and  $\theta_l^j$ . The incident field has wavevector  $\mathbf{k}$  in the surrounding medium (generally assumed to be vacuum) and its tangential component  $\alpha$  gives the quasi-periodicity of the field.

The field at P contains two contributions. The first comes from the incident field, and the second comes from the fields diffracted by all the other rods.

A plane wave incident with wavevector  $\mathbf{k}$  on a grating as depicted in Fig. (A.1) can be written as a sum of Bessel functions as follows:

## A.3 Plane wave - Bessel series representation

We start with the formula

$$e^{ir \sin \theta} = \sum_{n=-\infty}^{\infty} J_n(r) e^{in\theta}$$

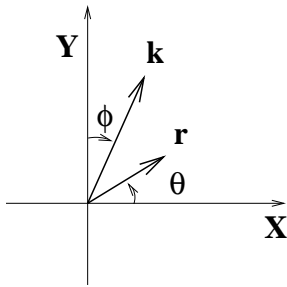


Figure A.2: Angle definitions

In our case we would like to obtain a formula for  $e^{i\mathbf{k}\cdot\mathbf{r}}$  where  $\mathbf{k} = k \begin{pmatrix} \sin \phi \\ \cos \phi \end{pmatrix}$  and  $\mathbf{r} = r \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$ . Here  $\theta$  is referenced to the X axis while  $\phi$  is referenced to the Y axis, as it is the angle of incidence of the plane wave on the horizontal grating. We therefore have

$$\begin{aligned} e^{i\mathbf{k}\cdot\mathbf{r}} &= e^{ikr(\sin \phi \cos \theta + \sin \theta \cos \phi)} \\ &= e^{ikr \sin(\phi + \theta)} \\ &= \sum_{n=-\infty}^{\infty} J_n(kr) e^{in(\phi + \theta)} \\ &= \sum_{n=-\infty}^{\infty} e^{in\phi} J_n(kr) e^{in\theta} \end{aligned}$$

By considering  $J_n(kr)e^{in\theta}$  as the basis of a vector space, and by considering only the three lowest harmonics then a plane wave with an angle of incidence of  $\phi$  can be represented by the column vector

$$Q = \begin{pmatrix} i^{-1} \\ i^0 \\ i^1 \end{pmatrix} = \begin{pmatrix} e^{-i\phi} \\ 1 \\ e^{i\phi} \end{pmatrix}. \quad (\text{A.2})$$

Keeping only the three lowest terms is a good approximation for wavelengths larger than the optical size of the individual dielectric rods or larger than the diameter for the case of metallic rods.

In the notation introduced above we can now estimate the field produced by a plane wave at point P at the surface of rod  $l$  in the reference frame of rod  $l$ . It is given by:

$$e^{i\mathbf{k}\cdot\mathbf{r}_l^{(P)}} \approx \sum_{-1}^1 e^{in\phi} J_n(kr_l^{(P)}) e^{in\theta_l^{(P)}}.$$

## A.4 Multi-scattering contribution to the local field

The field produced at P by rod  $j$  is simply the series of Hankel functions representing the outgoing field corresponding to rod  $j$ , explicitly

$$\sum_{-\infty}^{\infty} b_j^n H_n(kr_j^{(P)}) e^{in\theta_j^{(P)}}$$

or, keeping only the first three terms, as explained above,

$$b_j^{-1} H_{-1}(kr_j^{(P)}) e^{-i\theta_j^{(P)}} + b_j^0 H_0(kr_j^{(P)}) + b_j^1 H_1(kr_j^{(P)}) e^{i\theta_j^{(P)}}. \quad (\text{A.3})$$

We now sum over all the rods except for the center one to obtain the total multi-scattering contribution at point P:

$$\sum_{j \neq 0} b_j^0 H_{-1}(kr_j^{(P)}) e^{-i\theta_j^{(P)}} + \sum_{j \neq 0} b_j^0 H_0(kr_j^{(P)}) + \sum_{j \neq 0} b_j^1 H_1(kr_j^{(P)}) e^{i\theta_j^{(P)}}. \quad (\text{A.4})$$

The coefficients  $b$  are unknown and determining them is the purpose of the multi-scattering theory outlined below. It is important to note that for each value of  $j$  the cylindrical coordinates  $r$  and  $\theta$  correspond to a different system of coordinates, the  $j$ -th. In order to write the continuity conditions at rod  $l$  we must express all geometrical quantities in the coordinate system of rod  $l$ . This is done by using Graf's formula to go from coordinates  $j$  to coordinates  $l$ . In our case, for which  $r_l^{(P)} < r_j^j$ , it is written:

$$H_m(kr_j^{(P)}) e^{im\theta_j^{(P)}} = \sum_{-\infty}^{\infty} e^{i(m-q)\theta_l^j} H_{q-m}(kr_l^j) J_q(kr_l^{(P)}) e^{iq\theta_l^{(P)}},$$

or, by denoting the grating period by  $d$  and the rod radius by  $a$ ,

$$H_m(kr_j^{(P)}) e^{im\theta_j^{(P)}} = \sum_{-\infty}^{\infty} e^{i(m-q)\theta_l^j} H_{q-m}(k|j|d) J_q(ka) e^{iq\theta_l^{(P)}}.$$

Applying this to Eq. (A.4) and keeping only terms with  $q \in \{-1, 0, 1\}$ , we can write the multi-scattering contribution to the field at P in the form:

$$TJ_{-1}(ka) e^{-i\theta_l^{(P)}} + UJ_0(ka) + VJ_1(ka) e^{i\theta_l^{(P)}}$$

where

$$\begin{aligned}
T &= b_0^{-1} \sum_{j \neq 0} e^{ij\alpha d} \mathbf{H}_0(k|j|d) + b_0^0 \sum_{j \neq 0} e^{ij\alpha d} \text{sgn}(j) \mathbf{H}_{-1}(k|j|d) + b_0^1 \sum_{j \neq 0} e^{ij\alpha d} \mathbf{H}_{-2}(k|j|d) \\
U &= b_0^{-1} \sum_{j \neq 0} e^{ij\alpha d} \text{sgn}(j) \mathbf{H}_1(k|j|d) + b_0^0 \sum_{j \neq 0} e^{ij\alpha d} \mathbf{H}_0(k|j|d) + b_0^1 \sum_{j \neq 0} e^{ij\alpha d} \text{sgn}(j) \mathbf{H}_{-1}(k|j|d) \\
V &= b_0^{-1} \sum_{j \neq 0} e^{ij\alpha d} \mathbf{H}_2(k|j|d) + b_0^0 \sum_{j \neq 0} e^{ij\alpha d} \text{sgn}(j) \mathbf{H}_1(k|j|d) + b_0^1 \sum_{j \neq 0} e^{ij\alpha d} \mathbf{H}_0(k|j|d)
\end{aligned}$$

and where we have used the identities  $\mathbf{H}_2 = \mathbf{H}_{-2}$  and  $\mathbf{H}_{-1} = -\mathbf{H}_1$ , we denoted  $e^{i\theta_j} = \text{sgn}(j)$ , and where the field quasi-periodicity tells us that  $b_j = b_0 e^{ij\alpha d}$ . The field diffracted by the  $j$ -th rod is the same as the field diffracted by the central  $l$  rod with a phase shift which depends on  $\alpha$ .

To clarify the notation we introduce the quantities

$$\begin{aligned}
A_0 &= \sum_{j \neq 0} e^{ij\alpha d} \mathbf{H}_0(k|j|d) \\
A_1 &= \sum_{j \neq 0} e^{ij\alpha d} \text{sgn}(j) \mathbf{H}_1(k|j|d) \\
A_2 &= \sum_{j \neq 0} e^{ij\alpha d} \mathbf{H}_2(k|j|d)
\end{aligned} \tag{A.5}$$

which enable us to write the multiscattering equations self-consistently as

$$\begin{aligned}
b^{-1} &= S_{-1}(i_{-1} + b^{-1}A_0 - b^0A_1 + b^1A_2) \\
b^0 &= S_0(i_0 + b^{-1}A_1 + b^0A_0 - b^1A_1) \\
b^1 &= S_1(i_1 + b^{-1}A_2 + b^0A_1 + b^1A_0).
\end{aligned}$$

Rearranging and writing in a matrix form we obtain

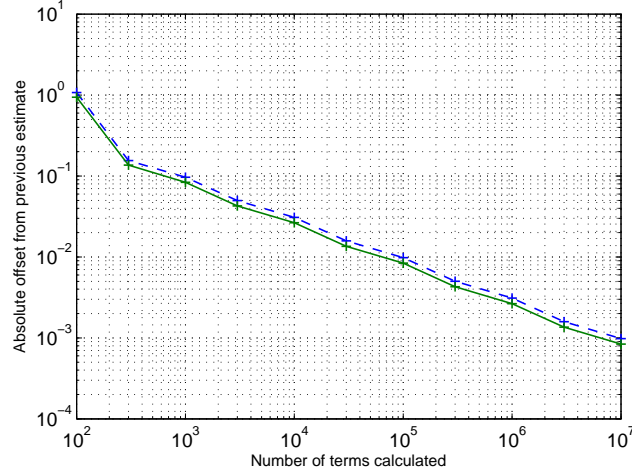
$$\begin{pmatrix} b^{-1} \\ b^0 \\ b^1 \end{pmatrix} = - \begin{pmatrix} A_0 - (S_{-1})^{-1} & -A_1 & A_2 \\ A_1 & A_0 - (S_0)^{-1} & -A_1 \\ A_2 & A_1 & A_0 - (S_1)^{-1} \end{pmatrix}^{-1} \begin{pmatrix} i^{-1} \\ i^0 \\ i^1 \end{pmatrix}. \tag{A.6}$$

We should emphasize that superscripts indicate the harmonic on  $i$  and  $b$  coefficients *only*. On all other quantities a superscript is an exponent.

The multi-scattering problem is therefore solved once the quantities  $A_0$ ,  $A_1$  and  $A_2$  are determined. This is what we shall undertake in the remainder of this Appendix.

## A.5 Computing $A_0$ , $A_1$ and $A_2$ numerically

Hankel series expansions such as that of Eq. (A.9) converge *extremely* slowly. We illustrate by plotting the convergence of  $A_0$  (solid) and  $A_2$  (dashed) for a wavelength equal to ten times the period, and for an angle of incidence of 30 degrees. See Figure A.3. On the horizontal axis is the number of terms calculated and on the vertical axis is the absolute difference between two consecutive estimates. One needs to calculate over ten million terms in order to estimate the sums to better than two decimal places. This is clearly not satisfactory especially since these sums depend on both the wavelength and on the angle of incidence which means they need to be evaluated often. An analytical approach is indispensable.

Figure A.3: Convergence study of Hankel series  $A_0$  (solid) and  $A_2$  (dashed)

## A.6 Estimating $A_0$ , $A_1$ and $A_2$ analytically

The derivation begins by writing the series as a limit:

$$A_0 = \lim_{\mathbf{r} \rightarrow 0} \left( \sum_{j \neq 0} e^{ij\alpha d} \mathbf{H}_0(k \|\mathbf{r} - jd\hat{x}\|) \right)$$

where we assume the limit exists, and where  $\hat{x}$  is the unit vector in the  $x$  direction. If the limit exists, then we can let  $\mathbf{r}$  approach the origin in any way we like. In particular, choosing the  $x$  axis, we can set  $y$  to zero immediately, in which case  $\|\mathbf{r} - jd\hat{x}\| = |x - jd|$ .

We now apply Weyl's formula:

$$\mathbf{H}_0(k \|\mathbf{r}\|) = \frac{1}{\pi} \int \frac{1}{\xi} e^{i\xi|y|} e^{i\sigma x} d\sigma \quad (\text{A.7})$$

where the Hankel function is of the first kind as throughout this Appendix,  $\xi = \frac{1}{\sqrt{k^2 - \sigma^2}}$  and  $\mathbf{r} = \begin{pmatrix} x \\ y \end{pmatrix}$ .

Since we have already set  $y$  to zero, we have

$$\begin{aligned} \sum_j e^{ij\alpha d} \mathbf{H}_0(k |x - jd|) &= \sum_j e^{ij\alpha d} \frac{1}{\pi} \int \frac{1}{\xi} e^{i\sigma(x-jd)} d\sigma \\ &= \frac{1}{\pi} \int \frac{1}{\xi} e^{i\sigma x} \left( \sum_j e^{ijd(\alpha - \sigma)} \right) d\sigma \end{aligned}$$

But the sum of complex exponentials inside the integral can be written as a sum of Dirac delta "functions".

By denoting  $K = \frac{2\pi}{d}$  and  $\alpha_n = \alpha + Kn$  we obtain

$$\begin{aligned} \frac{1}{\pi} \int \frac{1}{\xi} e^{i\sigma x} \left( \sum_j e^{ijd(\alpha - \sigma)} \right) d\sigma &= \frac{1}{\pi} \int \frac{1}{\xi} e^{i\sigma x} K \left( \sum_n \delta(\sigma - (\alpha + Kn)) \right) d\sigma \\ &= \frac{K}{\pi} \sum_n \int \frac{1}{\xi} e^{i\sigma x} \delta(\alpha_n - \sigma) d\sigma \\ &= \frac{K}{\pi} \sum_n \frac{1}{\beta_n} e^{i\alpha_n x} \end{aligned}$$

Summarizing, we now have

$$\sum_j e^{ij\alpha d} \text{H}_0(k|x-jd|) = \frac{K}{\pi} \sum_n \frac{1}{\beta_n} e^{i\alpha_n x} \quad (\text{A.8})$$

It is important to notice that by letting  $x \rightarrow 0$  both sides of this equation diverge, even though on the right side this is due to a divergent series, while on the left side this is due to a single diverging term, the *zeroth* term which contains the troublesome  $\text{H}_0(k|x|)$ . Fortunately, the quantity that we seek to evaluate is the left hand sum *without* the zeroth term (see Eq. (A.9)). In what follows we will show that the right hand sum diverges in such a way that, in the limiting case of very small  $x$ , it exactly cancels the diverging zeroth term of the left side resulting in a meaningful limiting process.

The asymptotic form of  $\text{H}_0$  for small positive arguments can be found in Abramowitz and Stegun [65]:

$$\lim_{x \rightarrow 0} \text{H}_0(kx) = 1 + i \frac{2}{\pi} (\ln(kx) + \gamma - \ln 2)$$

where  $\gamma$  is the Euler-Mascheroni constant. It is clear that it is the natural logarithm which drives the divergence for small  $x$ . By extracting a natural logarithm from the right hand side of Eq. (A.8) we may be able to cancel out the diverging terms.

We rewrite

$$\begin{aligned} \frac{K}{\pi} \sum_n \frac{1}{\beta_n} e^{i\alpha_n x} &= \frac{K}{\pi} \sum_n \frac{e^{i(\alpha+nK)x}}{\sqrt{k^2 - \alpha_n^2}} \\ &= \frac{K}{\pi} e^{i\alpha x} \left( \frac{1}{k \cos \theta} + \sum_{n \neq 0} \frac{e^{inKx}}{\sqrt{k^2 - \alpha_n^2}} \right) \end{aligned}$$

Let us now focus on the last sum.

$$\begin{aligned} \sum_{n \neq 0} \frac{e^{inKx}}{\sqrt{k^2 - \alpha_n^2}} &= \sum_{n \neq 0} \left( \frac{e^{inKx}}{\sqrt{k^2 - \alpha_n^2}} - \frac{e^{inKx}}{iK|n|} + \frac{e^{inKx}}{iK|n|} \right) \\ &= \sum_{n \neq 0} \left( \frac{1}{\sqrt{k^2 - \alpha_n^2}} - \frac{1}{iK|n|} \right) + \sum_{n=1}^{\infty} \frac{e^{inKx} + e^{-inKx}}{iKn} \\ &= \Delta - \frac{2}{iK} \ln(Kx) \end{aligned}$$

Here  $\Delta$  converges quickly to a small value for large  $\lambda$  and we can safely ignore it. We obtain

$$\begin{aligned} \lim_{x \rightarrow 0} \sum_{j \neq 0} e^{ij\alpha d} \text{H}_0(k|x-jd|) &= \lim_{x \rightarrow 0} (-\text{H}_0(k|x|) + \frac{K}{\pi} \sum_n \frac{1}{\beta_n} e^{i\alpha_n x}) \\ &= \lim_{x \rightarrow 0} \left( -1 - i \frac{2}{\pi} (\ln(kx) + \gamma - \ln 2) + \frac{K e^{i\alpha_0 x}}{\pi k \cos \theta} - \frac{K}{\pi} \frac{2}{iK} \ln(Kx) e^{i\alpha_0 x} \right) \\ &= \frac{K}{\pi k \cos \theta} - 1 + i \frac{2}{\pi} \left( \ln\left(\frac{K}{k}\right) + \ln 2 - \gamma \right). \end{aligned}$$

We are now able to evaluate

$$A_0 \approx \frac{K}{\pi k \cos \theta} - 1 + i \frac{2}{\pi} \left( \ln\left(\frac{K}{k}\right) + \ln 2 - \gamma \right). \quad (\text{A.9})$$

Derivations similar to the one above result in convenient approximate forms for  $A_1$  and  $A_2$ :

$$A_1 \approx -\frac{2}{\pi} \sin \phi + i \frac{K}{\pi k} \tan \phi$$

$$A_2 \approx \frac{K \cos(2\phi)}{\pi k \cos \phi} - i \frac{1}{\pi} \left( \frac{K^2}{3k^2} - \cos(2\phi) \right)$$

where  $\phi$  is the angle of incidence, as above.

## A.7 Transmission and reflection coefficients

Let us now assume that we have calculated the scattering coefficients  $b^{-1}$ ,  $b^0$  and  $b^1$  from Eq. (A.6). We now show how the reflection and the transmission coefficients can be obtained.

The field diffracted by the  $l = 0$ -th rod at any point of coordinates  $\mathbf{r}$  and  $\theta$  in the coordinate system of rod  $l = 0$  is given by

$$b_0^{-1} \mathbf{H}_{-1}(k \|\mathbf{r}\|) e^{-i\theta} + b_0^0 \mathbf{H}_0(k \|\mathbf{r}\|) + b_0^1 \mathbf{H}_1(k \|\mathbf{r}\|) e^{i\theta}. \quad (\text{A.10})$$

We must sum these contributions for all the rods to obtain the total field at  $\mathbf{r}$  and by following the derivation leading up to Eq. (A.8) but *without* letting  $y$  go to zero. One can show that

$$b_0^0 \sum_l e^{il\alpha d} \mathbf{H}_0(k \|\mathbf{r} - ld\hat{x}\|) = b_0^0 \frac{K}{\pi} \sum_n \frac{1}{\beta_n} e^{i(\alpha_n x + \beta_n |y|)}. \quad (\text{A.11})$$

It is easy to see that this formula basically translates a field produced by an infinite row of isotropic point scatterers (left side) into a sum of plane waves (right side). For the case of a wavelength larger than the grating period we must also point out that  $\alpha_n$  and  $\beta_n$  are real only for  $n = 0$ . This means that only the first plane wave is propagating, all the other being evanescent in the upper and lower half spaces. Consequently in the far field we can ignore all but the  $n = 0$  term. The amplitude of the plane wave due to the  $b^0$  term is therefore given by

$$b_0^0 \frac{K}{\pi k \cos \phi}. \quad (\text{A.12})$$

In fact, we should point out that there are *two* plane waves produced by this term, one in the upper half plane, for which  $y > 0$  and one in the lower half plane, for which  $y < 0$  in Eq. (A.11). The first will contribute to the transmitted field, while the second will contribute to the reflected field.

We now calculate the contributions of the  $b^{-1}$  and  $b^1$  terms from Eq. (A.10). First we need a way to also translate the series of  $\mathbf{H}_1$  and  $\mathbf{H}_{-1}$  functions into series of plane waves, as in Eq. (A.11). We begin once again with Weyl's formula, Eq. (A.7):

$$\mathbf{H}_0(k \|\mathbf{r}\|) = \mathbf{H}_0(k \sqrt{x^2 + y^2}) = \frac{1}{\pi} \int \frac{1}{\xi} e^{i\xi|y|} e^{i\sigma x} d\sigma$$

We now apply the differential operator  $\frac{\partial}{\partial x} + i \frac{\partial}{\partial y}$  to both sides. On the left side we obtain

$$\begin{aligned} \left( \frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) \mathbf{H}_0(k \sqrt{x^2 + y^2}) &= -\mathbf{H}_1(k \sqrt{x^2 + y^2}) \left( \frac{kx}{\sqrt{x^2 + y^2}} + i \frac{ky}{\sqrt{x^2 + y^2}} \right) \\ &= -k \mathbf{H}_1(k \sqrt{x^2 + y^2}) e^{i\theta} \end{aligned}$$

and therefore

$$\begin{aligned} \mathbf{H}_1(k \|\mathbf{r}\|) e^{i\theta} &= \frac{1}{\pi k} \int \frac{\xi \operatorname{sgn}(y) - i\sigma}{\xi} e^{i(\sigma x + \xi|y|)} \\ \mathbf{H}_{-1}(k \|\mathbf{r}\|) e^{-i\theta} &= \frac{1}{\pi k} \int \frac{\xi \operatorname{sgn}(y) + i\sigma}{\xi} e^{i(\sigma x + \xi|y|)}. \end{aligned}$$

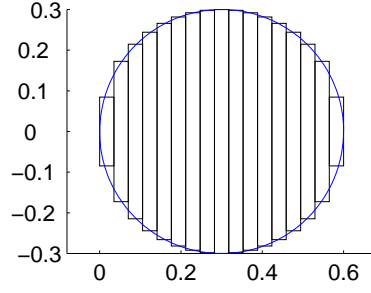


Figure A.4: Dielectric cylinder of radius 0.3 as represented in the RCW grating simulation.

From here one can follow a development similar to the one leading up to (A.8) to obtain the formulas analogous to Eq. (A.11):

$$b_0^1 \sum_l e^{il\alpha d} \mathbf{H}_1(k \|\mathbf{r} - ld\hat{x}\|) e^{i\theta} = b_0^1 \frac{K}{\pi k} \sum_n \frac{\beta_n \operatorname{sgn}(y) - i\alpha_n}{\beta_n} e^{i(\alpha_n x + \beta_n |y|)}$$

$$b_0^{-1} \sum_l e^{il\alpha d} \mathbf{H}_{-1}(k \|\mathbf{r} - ld\hat{x}\|) e^{i\theta} = b_0^{-1} \frac{K}{\pi k} \sum_n \frac{\beta_n \operatorname{sgn}(y) + i\alpha_n}{\beta_n} e^{i(\alpha_n x + \beta_n |y|)}$$

The contributions to the transmitted (upper half plane,  $y > 0$ ) plane wave are therefore:

$$b_0^1 \frac{K}{\pi k \cos \phi} (\cos \phi - i \sin \phi) \quad \text{and} \quad b_0^{-1} \frac{K}{\pi k \cos \phi} (\cos \phi + i \sin \phi)$$

while the contributions to the reflected (lower half plane,  $y < 0$ ) plane wave are

$$b_0^1 \frac{K}{\pi k \cos \phi} (-\cos \phi - i \sin \phi) \quad \text{and} \quad b_0^{-1} \frac{K}{\pi k \cos \phi} (-\cos \phi + i \sin \phi).$$

The reflection and transmission coefficients are therefore given by

$$r = \frac{K}{\pi k \cos \phi} (b_0^0 + b_0^{-1} (-\cos \phi + i \sin \phi) - b_0^1 (\cos \phi + i \sin \phi))$$

$$t = 1 + \frac{K}{\pi k \cos \phi} (b_0^0 + b_0^{-1} (\cos \phi + i \sin \phi) + b_0^1 (\cos \phi - i \sin \phi))$$

## A.8 Domain of validity - comparison with rigorous simulations

In this section we compare the results obtained using the method described above with the Rigorous Coupled Wave (RCW) transfer matrix method as described by Nevière and Popov [86]. In this way we can gain a better understanding of the domain of validity of the approximations made above.

Since in the above calculations we have assumed a single plane wave incident on the grating, and a single plane wave diffracted respectively above and below the grating, our calculations only apply in the single-order-of-diffraction regime, where  $\lambda > d$ ,  $d$  being the grating period. In the figures that follow the horizontal axis will therefore represent the wavelength measured in units of the period and  $\lambda > 1$ .

The RCW method is one in which a given geometry is divided up into thin layers and the transfer matrices of the individual slices are assembled to give the transfer matrix of the whole structure. As a result, when circular rods are considered the actual structure being simulated is something resembling Fig. A.4 rather than a perfect cylinder. However, for the large wavelengths under study here the difference is negligible.

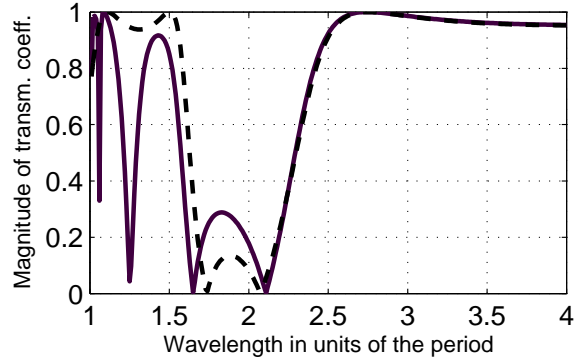


Figure A.5: The absolute value of the transmission coefficient for a grating of dielectric rods with  $\varepsilon = 10$  and  $r = 0.3$  at normal incidence. Comparison between RCW simulation (solid) and the asymptotic analytic method (dashed).

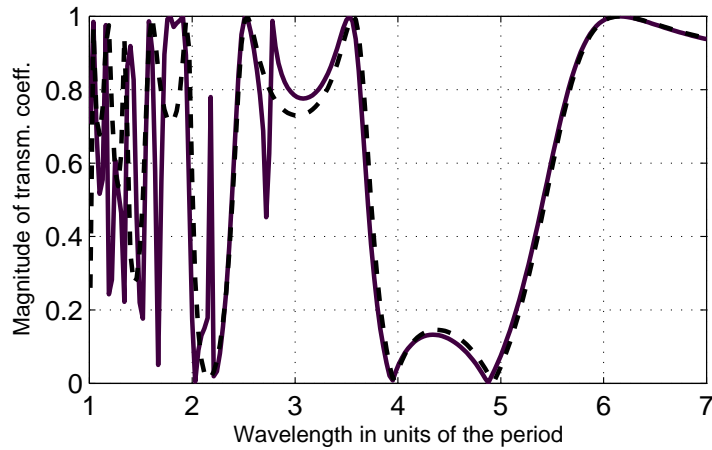


Figure A.6: The absolute value of the transmission coefficient for a grating of dielectric rods with  $\varepsilon = 30$  and  $r = 0.4$  at normal incidence. Comparison between RCW simulation (solid) and the asymptotic analytic method (dashed).

We begin with a typical structure, with a relative permittivity of 10 and rod radius of 0.3, in units of the period. We see in Fig. A.5 that the agreement is excellent down to wavelengths about twice as large as the period. In other words the asymptotic method seems to give good results for  $\lambda \gtrsim 2d$ . This is quite satisfactory for a study of the homogeneous properties of this kind of structure where the wavelengths under study are usually considerably larger than the period. The agreement is quite robust for different values of relative permittivity or rod radius but there are two cases in which it breaks down.

The first one is illustrated in Fig. A.6. In this case we chose a very large relative permittivity,  $\varepsilon = 30$ , with a slightly larger rod radius  $r = 0.4$ . We see that a series of resonances appear which are accounted for reasonably well by the asymptotic theory. However, it is clear that the resonances around  $\lambda = 2.2$  and  $\lambda = 2.8$  are not reproduced. The reason for this is that they are higher order multipolar resonances which cannot be reproduced by our theory which keeps only monopolar ( $b^0$ ) and dipolar ( $b^{-1}, b^1$ ) terms.

The second limitation is, more interestingly, for the case of very low relative permittivities ( $\varepsilon = 3$ ), and for rods that are very close to each other ( $r = 0.49$ ), as illustrated in Fig. A.7. We can see that the two methods agree well only for much larger wavelengths, of  $\lambda \gtrsim 4 - 5$ . The explanation for this lies in the fact that the asymptotic method relies on a tight-binding-like expansion of the fields. The fields are expressed in terms of functions localized around individual fibers. As a consequence the results are good when the grating is in the *tight-binding regime*, in other words when the relative permittivity is

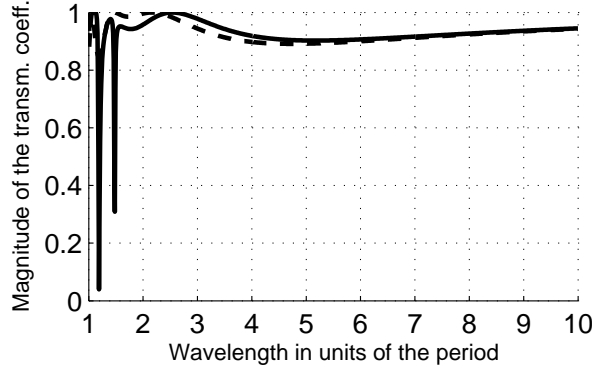


Figure A.7: The absolute value of the transmission coefficient for a grating of dielectric rods with  $\varepsilon = 3$  and  $r = 0.49$  at normal incidence. Comparison between RCW simulation (solid) and the asymptotic analytic method (dashed).

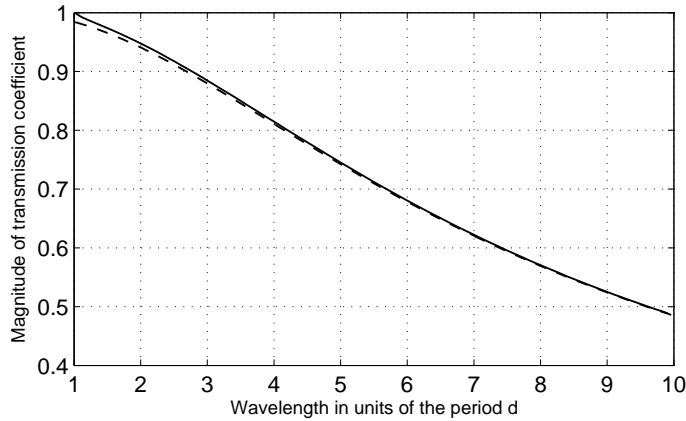


Figure A.8: Magnitude of transmission coefficient for normal incidence obtained through the analytical method above (dashed) and by the finite element method employed by FEMLAB MULTIPHYSICS (solid). The wire radius is  $r = 0.01d$  and the incidence is normal.

large and the fibers are relatively far from each other. In this situation it is a good approximation to consider that the total field is produced by a superposition of the fields generated by individual fibers. When the relative permittivity is small and the fibers are almost touching, then the grating is closer to a *perturbation regime* where the coupling between the fibers is strong, and the grating behaves more similarly to a flat dielectric guide with a periodic perturbation.

A good agreement between the analytic theory and a rigorous simulation is therefore, generally speaking, an indication that the grating is in the tight-binding regime and that the dominant phenomenon is the individual rod resonance. From this perspective it is possible to use this method to test the extent to which the rods are coupled through non-dipolar fields, also known as near fields or evanescent fields. When the code works well, this is an indication that the rods see each other as dipole scatterers. In the context of the development presented in Chapter 1 this is quite interesting, because we now have a criterion for determining whether the all important dipolar approximation is justified or not in a given structure.

For the case of thin metal wires the agreement is even better than for dielectric rods, since the wires act as monopolar scatterers and since even wavelengths close to 1 are much larger than the size of the scatterers. We illustrate by comparing the results of the analytic method above and the finite element method employed by the commercial suite COMSOL MULTIPHYSICS in Fig. A.8.

## Appendix B

# The parameter extraction procedure

As mentioned in Section 3.1, the initial absence of detailed constructive approaches to the homogenization of composite metamaterials led some authors to take a holistic black box approach to modeling metamaterials. This approach took the form of a procedure which would be able to deduce the effective index and the effective impedance of a slab of material from the reflection and transmission coefficients. Both the magnitude and phase of these coefficients are required, as well as the thickness of the structure. Since the phases play an important role in the parameter extraction and since they depend on the conventions and choice of coordinates during the derivation we give the details below.

If we consider a homogeneous slab of thickness  $d$ , index  $n$  and impedance  $z$  surrounded by air, above and below, oriented normally to the  $y$  axis, then the field above the slab takes the form

$$u_a(y) = a_a^- e^{-ik_0 y} + a_a^+ e^{ik_0 y}$$

and the field below the slab takes the form

$$u_b(y) = a_b^- e^{-ik_0 y}.$$

Since we are in normal incidence we have the choice of writing the continuity conditions in  $E_{\parallel}$  or  $H_{\parallel}$ , the result will be the same. We choose  $E_{\parallel}$ . In this case the field  $u$  is continuous at the two interfaces and also the quantity  $\frac{1}{\mu_0 \mu} \frac{\partial u}{\partial y}$  is continuous. Writing out the continuity conditions in matrix form we can define the transfer matrix for any given interface between medium  $j$  and medium  $j+1$  as the matrix

$$\mathcal{T}_j = \begin{pmatrix} s_j & d_j \\ d_j & s_j \end{pmatrix}$$

where

$$s_j = \frac{k_j + \frac{\mu_{j+1}}{\mu_j} k_{j+1}}{2k_j} \quad d_j = \frac{k_j - \frac{\mu_{j+1}}{\mu_j} k_{j+1}}{2k_j}.$$

The propagation of the field within the layer is accounted for by the matrix

$$\mathcal{C} = \begin{pmatrix} e^{-ink_0 d} & 0 \\ 0 & e^{ink_0 d} \end{pmatrix}.$$

The overall transfer matrix through the slab is then given by

$$\mathcal{M} = \mathcal{T}_0 \mathcal{C} \mathcal{T}_1$$

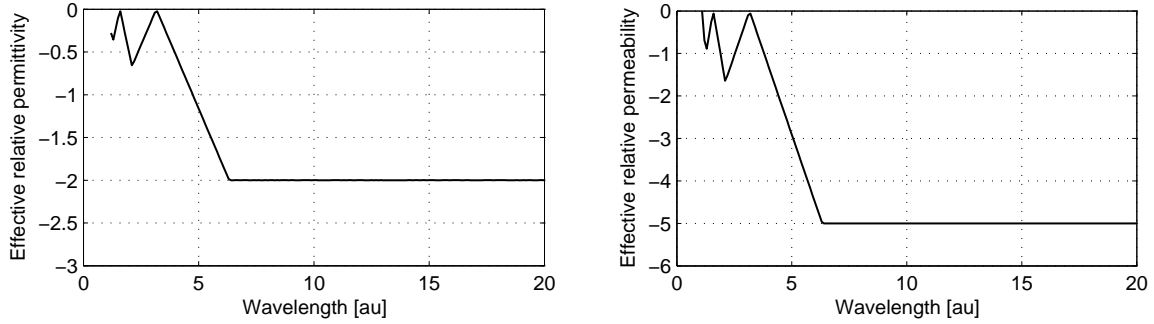


Figure B.1: Effective permittivity and effective permeability extracted from a slab with  $\varepsilon = -2$  and  $\mu = -5$ ,  $d = 1$ [au].

and the reflection and transmission coefficients are given by

$$r = \frac{\mathcal{M}_{2,1}}{\mathcal{M}_{1,1}} \quad t = \frac{1}{\mathcal{M}_{1,1}}.$$

Writing these out in terms of the index and impedance of the slab, we have

$$t^{-1} = \left( \cos(nk_0d) - \frac{i}{2}(z - 1/z) \sin(nk_0d) \right) e^{ik_0d}$$

and

$$r = -t \frac{i}{2}(z - 1/z) \sin(nk_0d) e^{ik_0d}.$$

These relations can be inverted in order to express  $n$  and  $z$  in terms of  $r$  and  $t$ . We have [51]

$$n = \frac{1}{k_0d} \cos^{-1} \left( \frac{1}{2te^{ik_0d}} [1 - (r^2 - t^2 e^{2ik_0d})] \right) + \frac{2\pi m}{kd} \quad (\text{B.1})$$

where  $m$  is an integer, and

$$z = \pm \sqrt{\frac{(1-r)^2 - t^2 e^{2ik_0d}}{(1+r)^2 - t^2 e^{2ik_0d}}}. \quad (\text{B.2})$$

The sign in front of the square root in the expression for the impedance is fixed by the requirement that in passive materials the real part of the impedance is always positive. The imaginary part of the index must likewise be positive in passive media. If the structure in question is lossless, all imaginary parts must be null. The remaining ambiguity is the choice of branch for the inverse cosine, in the expression for the index. In order to fix the value of  $m$  at least two measurements must be performed, for two different thicknesses. The value of  $m$  which gives the same value of  $n$  must then be chosen. However, for samples with an optical thickness that is sufficiently small we have  $m = 0$ . The optical thickness depends on both the physical thickness and the optical index. Therefore, we expect that the expression for  $n$  give good results with  $m = 0$  for situations where the sample is thin, and the index is not much larger than 1.

We illustrate first on a homogeneous slab, and then on a very high contrast grating (see Section 3.5).

The dielectric slab has thickness  $d = 1$ ,  $\mu = -5$  and  $\varepsilon = -2$ . We extract the effective permittivity and permeability for wavelengths ranging from  $\lambda = 1.1$  to  $\lambda = 20$ . Clearly, when the wavelength is shorter than about six times the thickness of the slab, a different branch must be chosen for the inverse cosine of Eq. (B.1), namely that with  $m = 1$ . The extraction seems to work well for this simplest of cases, especially because the medium is not dispersive and we know that the extraction works when we see a straight horizontal line on the extraction plots. The situation is considerably more complicated when we

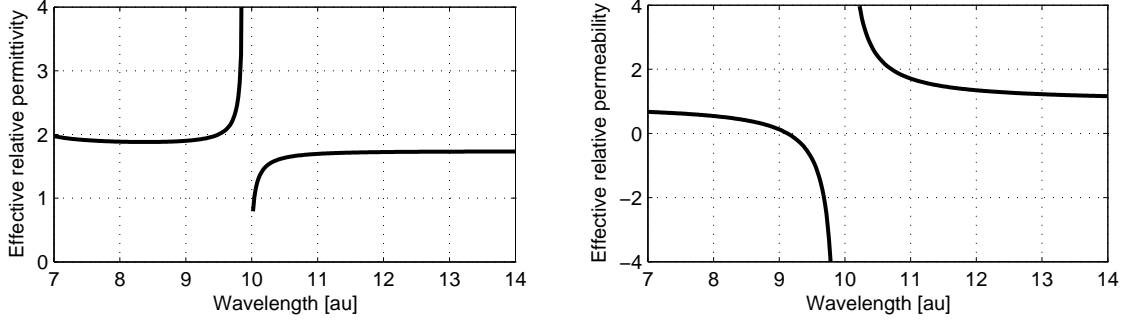


Figure B.2: Effective permittivity and effective permeability extracted from a single grating of square rods or side  $0.5[\text{au}]$  suspended in air. The thickness of the effective homogeneous slab is  $d = 1[\text{au}]$ .

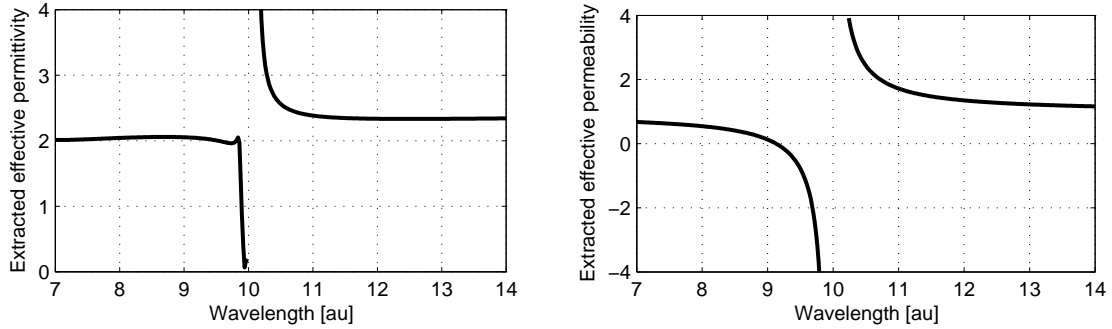


Figure B.3: Effective permittivity and effective permeability extracted from two gratings of square rods of side  $0.5[\text{au}]$  suspended in air. The thickness of the effective homogeneous slab is  $d = 2[\text{au}]$ .

are dealing with a strongly dispersive structure.

We use the example of the grating of Fig. 3.9, which consists of a series of dielectric rods suspended in air, with  $\epsilon = 200$ . The rods are square, of side  $0.5[\text{au}]$  and the period is  $p = 1[\text{au}]$ . We try to replace this structure with a homogeneous slab of thickness  $d = 1[\text{au}]$ . The extraction is presented in Fig.

The permeability is resonant around the first Mie resonance at  $\lambda \approx 10$  as expected, but the permittivity seems to have a resonance also at about the same position. However, it is not clear whether the form of the permittivity is an actual part of the homogeneous description of the two dimensional medium or whether it is an artefact due to the fact that we are extracting parameters from a single monolayer, a single grating. It is to be expected that a monolayer will not behave exactly as the bulk crystal, and it is likely that that is the reason for the appearance of the unexpected resonance of the permittivity. We must therefore attempt to extract parameters from a thicker structure, composed of more stacked gratings.

Unfortunately, however, as we increase the thickness of the structure the extraction procedure ceases to give reliable results. Figures B.2 and B.3 show the results of the extraction for a monolayer and a bilayer respectively. The permeability seems to be consistent, in the two cases, but the results for the effective permittivity are quite different. One cannot say from these results whether the medium truly has a resonant effective dielectric response or if it is simply an artefact of the extraction method due to the small thickness of the structures. It is encouraging to notice, though, that at least around a wavelength of about  $9.5\text{au}$ , wavelength for which the effective permeability is close to  $-1$  and which is of particular interest to us, the effective permittivity given in both cases is close to  $2$ . This is reasonably close to the theoretically predicted value of  $1.7$  of Felbacq and Bouchitté [72].

The best confirmation of the theoretical model, though, is a direct comparison between the result of the simulation of the rod structure and the effective homogeneous structure. The difference between

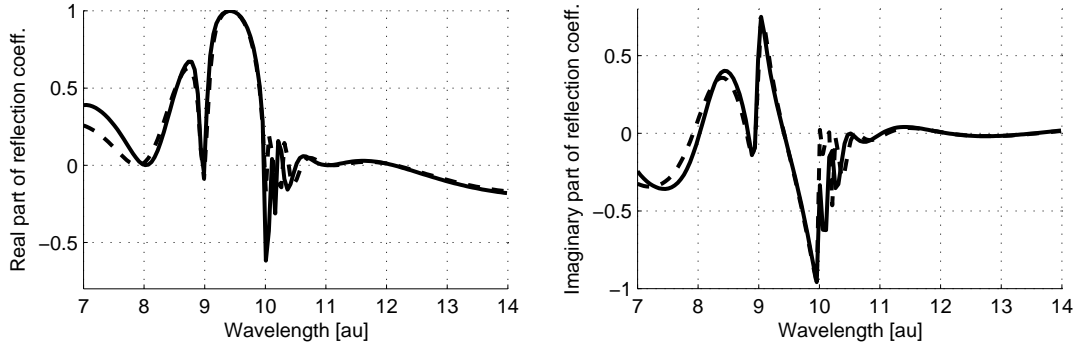


Figure B.4: Comparison of the real and imaginary parts of the reflection coefficient of a structure composed of 8 layers of square dielectric rods of  $\varepsilon = 200$  and the effective homogeneous slab with  $\mu$  given by Eq. (B.3).

the two approaches is important because one is a holistic black box model that may have no physical content, and the other is a constructive model where the physics appears explicitly, taking the form of the closed form expressions:

$$\mu_{\text{eff}} = 1 + \frac{16}{\pi^4} \frac{k^2}{8\pi^2/\varepsilon - k^2} \quad (\text{B.3})$$

$$\varepsilon_{\text{eff}} = 1.7. \quad (\text{B.4})$$

In the first approach, design is purely trial and error, while in the second, design is direct and almost automatic. When taking a constructive approach we can be sure to have captured the physical phenomenon involved. Figure B.4 compares the real and imaginary parts of the reflection coefficient from a structure of 8 stacked gratings, with the same geometry as above, with the reflection from the effective homogeneous slab with parameters given in Eqs. (B.3) and (B.4).

The agreement is excellent almost everywhere, with the exception of a small region just above  $\lambda = 10$ . In particular, the agreement is almost perfect in the region of interest where the effective permeability becomes negative,  $9 < \lambda < 10$ . This validates the constructive analysis of Ref. [72] and constitutes a strong argument in favor of constructive methods, as opposed to holistic, phenomenological methods.

# Bibliography

- [1] Yablonovitch, E. *Physical Review Letters* **58**(20), 2059–2062 (1987).
- [2] John, S. *Physical Review Letters* **58**(23), 2486–2489 (1987).
- [3] Jackson, J. D. *Classical Electrodynamics*. John Wiley and Sons, (1999).
- [4] Maxwell-Garnett, J. *Phil. Trans. R. Soc. Lond. A* **203**, 385 (1904).
- [5] Wiener, O. *Abh. Math.-Phys. Königl. Sächs. Ges.* **32**, 509 (1912).
- [6] Bruggeman, D. *Ann. Phys. Lpz.* **24**, 636 (1935).
- [7] Hashin, Z. and Shtrikman, S. *J. Appl. Phys.* **33**, 3125 (1962).
- [8] Milton, G. W. *Applied Physics Letters* **37**(3), 300–302 (1980).
- [9] Tsang, L. and Kong, J. A. *Radio Science* **16**(3), 303–320 (1981).
- [10] Bergman, D. J. *Physical Review Letters* **44**(19), 1285–1287 (1980).
- [11] Bensoussan, A., Lions, J., and Papanicolaou, G. *Asymptotic Analysis for Periodic Structures*. North-Holland, Amsterdam, (1978).
- [12] Kock, W. E. *Bell System Technical Journal* **27**(1), 58–82 (1948).
- [13] Pendry, J. B., Holden, A. J., Robbins, D. J., and Stewart, W. J. *J. of Phys.-Cond. Matt.* **10**(22), 4785 (1998).
- [14] Pendry, J. B., Holden, A. J., Robbins, D. J., and Stewart, W. J. *IEEE Transactions on Microwave Theory and Techniques* **47**(11), 2075 (1999).
- [15] Pendry, J. B. *Physical Review Letters* **85**(18), 3966–3969 (2000).
- [16] Russakoff, G. *American Journal of Physics* **38**(10), 1188–1195 (1970).
- [17] Robinson, F. N. H. *Macroscopic electromagnetism*. Pergamon Press Ltd., (1973).
- [18] Schram, K. *Physica* **26**, 1080–1090 December (1960).
- [19] de Groot, S. R. *The Maxwell Equations*, volume IV of *Studies in Statistical Mechanics*. North-Holland, Amsterdam, (1969).
- [20] Dendy, R. O. *Plasma Dynamics*. Oxford University Press, (1990).
- [21] Kittel, C. *Introduction to Solid State Physics*. John Wiley and Sons, (1996).
- [22] Ashcroft, N. W. and Mermin, N. D. *Solid State Physics*. Brooks Cole, (1976).

- [23] Yatsenko, V. V., Maslovski, S. I., Tretyakov, S. A., Prosvirnin, S. L., and Zouhdi, S. *IEEE Transactions On Antennas And Propagation* **51**(1), 2–11 January (2003).
- [24] Grigorenko, A. N., Geim, A. K., Gleeson, H. F., Zhang, Y., Firsov, A. A., Khrushchev, I. Y., and Petrovic, J. *Nature* **438**(7066), 335–338 November (2005).
- [25] Zhou, J. F., Zhang, L., Tuttle, G., Koschny, T., and Soukoulis, C. M. *Physical Review B* **73**(4), 041101 January (2006).
- [26] Smith, D. R., Vier, D. C., Kroll, N., and Schultz, S. *Applied Physics Letters* **77**(14), 2246–2248 October (2000).
- [27] Li, L. *Journal of the Optical Society of America A* **13**(9), 1870–1876 (1996).
- [28] Baba, T. and Nakamura, M. *IEEE Journal of Quantum Electronics* **38**, 909–914 (2002).
- [29] Baba, T. and Matsumoto, T. *Applied Physics Letters* **81**, 2325–2327 (2002).
- [30] Kosaka, H., Kawashima, T., Tomita, A., Notomi, M., Tamamura, T., Sato, T., and Kawakami, S. *Journal Of Lightwave Technology* **17**(11), 2032–2038 (1999).
- [31] Sakoda, K. and Ohtaka, K. *Phys. Rev. B* **54**, 5742 (1996).
- [32] Shi, B., Jiang, Z., Zhou, X., and Wang, X. *Journal of Applied Physics* **91**, 6769–6771 (2002).
- [33] Cassagne, D. *Annales de Physique* **23**, 3–91 (1998).
- [34] Johnson, S. G. and Joannopoulos, J. D. *Optics Express* **8**(3), 173–190 (2001).
- [35] Yee, K. *IEEE Transactions Antenna and Propagation* **14**, 302 (1966).
- [36] Taflove, A. and Hagness, S. *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Artech House Publishers, 3rd edition, (2005).
- [37] Rashed, R. *Isis* **81**, 464 (1990).
- [38] d' Yerville, M. L. V. *Modélisation des Cristaux Photoniques Bidimensionels de Hauteur Finie*. PhD thesis, Université Montpellier II Sciences et Techniques de Languedoc, (2002).
- [39] Pommereau, F., Legouezigou, L., Hubert, S., Sainson, S., Chandouineau, J. P., Fabre, S., Duan, G. H., Lombardet, B., Ferrini, R., and Houdre, R. *Journal Of Applied Physics* **95**(5), 2242–2245 March (2004).
- [40] Cabuz, A. I., Centeno, E., and Cassagne, D. *Applied Physics Letters* **84**(12), 2031–2033 (2004).
- [41] Shelby, R. A., Smith, D. R., and Schultz, S. *Science* **292**(5514), 77 (2001).
- [42] Notomi, M. *Physical Review B* **62**(16), 10696–10705 October (2000).
- [43] Luo, C., Johnson, S. G., Joannopoulos, J. D., and Pendry, J. B. *Physical Review B* **65**(20), 201104 (2002).
- [44] Luo, C., Johnson, S. G., Joannopoulos, J. D., and Pendry, J. B. *Physical Review B* **68**(4), 045115 (2003).
- [45] Meade, R. D., Brommer, K. D., Rappe, A. M., and Joannopoulos, J. D. *Phys. Rev. B* **44**(19), 10961–10964 Nov (1991).

- [46] Li, C. Y., Holt, J. M., and Efros, A. L. *Journal Of The Optical Society Of America B-Optical Physics* **23**(5), 963–968 May (2006).
- [47] Wang, X., Ren, Z. F., and Kempa, K. *Optics Express* **12**(13), 2919–2924 June (2004).
- [48] Wang, X., Ren, Z. F., and Kempa, K. *Applied Physics Letters* **86**(6), 061105 February (2005).
- [49] Li, C. Y., Holt, J. M., and Efros, A. L. *Journal Of The Optical Society Of America B-Optical Physics* **23**(3), 490–497 March (2006).
- [50] Cubukcu, E., Aydin, K., Ozbay, E., Foteinopolou, S., and Soukoulis, C. M. *Physical Review Letters* **91**(20), 207401 November (2003).
- [51] Smith, D., Schultz, S., Markos, P., and Soukoulis, C. *Phys. Rev. B* **65**(19), 195104 (2002).
- [52] Pokrovsky, A. L. and Efros, A. L. *Phys. Rev. Lett.* **89**(9), 093901 (2002).
- [53] Marques, R. and Smith, D. R. *Phys. Rev. Lett.* **92**(5), 059401 (2004).
- [54] Woodley, J. F., Wheeler, M. S., and Mojahedi, M. *Phys. Rev. E* **71**, 066605 (2005).
- [55] Maslovski, S. I. *Technical physics letters* **29**(1), 32–34 (2003).
- [56] Maslovski, S. I., Tretyakov, S. A., and Belov, P. A. *Microwave And Optical Technology Letters* **35**(1), 47–51 October (2002).
- [57] Simovski, C. R., Belov, P. A., and He, S. L. *IEEE Transactions On Antennas And Propagation* **51**(10), 2582–2591 October (2003).
- [58] Belov, P. A., Marques, R., Maslovski, S. I., Nefedov, I. S., Silveirinha, M., Simovski, C. R., and Tretyakov, S. A. *Phys. Rev. B* **67**(11), 113103 (2003).
- [59] Brown, J. *Progress in Dielectrics: IEE* **2**, 195–225 (1960).
- [60] Mandel, L. and Wolf, E. *Optical Coherence and Quantum Optics*, chapter 3, 109–127. Cambridge University Press (1995).
- [61] Veselago, V. G. *Soviet Physics USPEKHI-USSR* **10**(4), 509 (1968).
- [62] Fang, N., Lee, H., Sun, C., and Zhang, X. *Science* **308**(5721), 534–537 April (2005).
- [63] Smaâli, R., Felbacq, D., and Granet, G. *Physica E* **18**, 443 (2003).
- [64] Zolla, F., Felbacq, D., and Bouchitte, G. *Physical Review E* **74**(5), 056612 (2006).
- [65] Abramowitz, M. and Stegun, I. A. *Handbook of mathematical functions*. Dover, (1972).
- [66] Lekner, J. *J. Opt. Soc. Am. A* **11**(11), 2892 (1994).
- [67] Felbacq, D., Guizal, B., and Zolla, F. *Optics Communications* **152**(1-3), 119 (1998).
- [68] Dignam, M. J. and Moskovits, M. *Journal Of The Chemical Society-Faraday Transactions II* **69**(1), 56–64 (1973).
- [69] Felbacq, D. and Bouchitte, G. *Opt. Lett.* **30**, 1189 (2005).
- [70] O’Brien, S. and Pendry, J. B. *J. Phys.: Condens. Matter* **14**, 4035 (2002).
- [71] Bouchitte, G. and Felbacq, D. *Comptes Rendus Mathematique* **339**(5), 377–382 September (2004).

- [72] Felbacq, D. and Bouchitte, G. *Phys. Rev. Lett.* **94**(18), 183902 (2005).
- [73] Simovski, C. R., Tretyakov, S. A., Sihvola, A. H., and Popov, M. M. *European Physical Journal-Applied Physics* **9**(3), 195–204 March (2000).
- [74] Yatsenko, V., Maslovski, S., and Tretyakov, S. *Progress in Electromagnetics Research* **25**, 285 (2000).
- [75] Ramo, S., Whinnery, J. R., and Van Duzer, T. *Fields and Waves in Communication Electronics*. John Wiley and Sons, third edition, (1994).
- [76] Sauviac, B., Simovski, C. R., and Tretyakov, S. A. *Electromagnetics* **24**(5), 317–338 July (2004).
- [77] Tretyakov, S. A., Mariotte, F., Simovski, C. R., Kharina, T. G., and Heliot, J. P. *IEEE Transactions On Antennas And Propagation* **44**(7), 1006–1014 July (1996).
- [78] Smith, D. R. and Schurig, D. *Phys. Rev. Lett.* **90**(7), 077405 (2003).
- [79] Markos, P. and Soukoulis, C. M. *Optics Express* **11**(7), 649 (2003).
- [80] Katsarakis, N., Koschny, T., Kafesaki, M., Economou, E. N., Ozbay, E., and Soukoulis, C. M. *Physical Review B (Condensed Matter and Materials Physics)* **70**(20), 201101 (2004).
- [81] Schurig, D. *International Journal Of Numerical Modelling-Electronic Networks Devices And Fields* **19**(2), 215–228 March (2006).
- [82] Cabuz, A. I., Felbacq, D., and Cassagne, D. *Physical Review Letters* **98**(3), 037403 (2007).
- [83] Smith, D. R., Padilla, W. J., Vier, D. C., Nemat-Nasser, S. C., and Schultz, S. *Physical Review Letters* **84**(18), 4184–4187 May (2000).
- [84] Koschny, T., Kafesaki, M., Economou, E. N., and Soukoulis, C. M. *Physical Review Letters* **93**(10), 107402 (2004).
- [85] Zhou, J. F., Koschny, T., Zhang, L., Tuttle, G., and Soukoulis, C. M. *Applied Physics Letters* **88**(22), 221103 May (2006).
- [86] Neviere, M. and Popov, E. *Light Propagation in Periodic Media: Differential Theory and Design*. CRC, (2002).



---

**RESUME en français**

Les métamatériaux composés sont des structures périodiques metallo-dielectriques fonctionnant à des longueurs d'onde plus grandes que leur périodicité. A des longueurs d'onde suffisamment grandes ces structures se comportent comme des milieux homogènes caractérisés par des paramètres tensoriels de permittivité et perméabilité. Ces paramètres peuvent prendre des valeurs en dehors du domaine naturellement disponible; notamment il est possible de concevoir des matériaux avec une permittivité et perméabilité simultanément négative, donnant lieu à un indice de refraction négatif. Cependant, il n'est pas toujours évident que un modèle homogène soit adapté pour décrire le comportement d'une structure donnée à une longueur d'onde donnée. C'est un aspect souvent passé sous silence dans la littérature.

Dans ce mémoire je commence avec une analyse fondamentale des notions de perméabilité et permittivité dans l'absence des charges libres. Les conclusions de cette discussion sont utilisés pour analyser le comportement des métamatériaux composites à indice négatif. Je propose une méthode qui permet de identifier des domaines de fréquence où une structure puisse être décrite quantitativement par des paramètres homogènes de permittivité et perméabilité. Ce travail ouvre la voie vers une compréhension plus détaillée de la transition entre comportement homogène et inhomogène dans les métamatériaux composites, et introduit, notamment, les notions nouvelles de *modèle effectif sur mesure*, et de *meta-cristaux photoniques*.

---

**TITRE en anglais****Electromagnetic Metamaterials – From Photonic Crystals to Negative Index Composites**

---

**RESUME en anglais**

Composite metamaterials are periodic metal-dielectric structures operating at wavelengths larger than the structure period. If properly designed these structures behave as homogeneous media described by effective permittivity and permeability parameters. These effective parameters can be designed to take values in domains that are not available in naturally occurring media; notably it is possible to design composite metamaterials with simultaneously negative permittivity and permeability, or, in other words, with a negative refractive index. However, in many experimental or numerical studies it is far from obvious that the use of a homogeneous model is justified for a given structure at a given wavelength. This issue is often glossed over in the literature.

In this work I take a detailed look at the fundamental assumptions on which effective medium models rely and put forward a method for determining frequency domains where a given structure may or may not be accurately described by homogeneous effective medium parameters. This work opens the door to a more detailed understanding of the transition between homogeneous and inhomogeneous behavior in composite metamaterials, in particular by introducing the novel notions of *custom made effective medium model*, and of *meta-photonic crystal*.

---

**DISCIPLINE**

Physique de la matière condensée

---

**MOTS-CLES**

Effective medium theory, homogenization, negative index, permittivity, permeability, superlens, photonic crystal, magnetic resonator, spatial dispersion, non-local, metamaterial, meta-photonic crystal, inhomogeneous effective medium.

---

**INTITULE ET ADRESSE DE L'U.F.R. OU DU LABORATOIRE :**

Groupe d'Etude des Semiconducteurs, UMR 5650, Bat 21, Place E. Bataillon, 34095 Montpellier