

On electrons, dielectrics and beyond - the classical theory of electromagnetic metamaterials

Alexandru I. Căbuz

*Institut Fresnel, Ave. Escadrille Normandie
Domaine Universitaire de St. Jérôme, 13397 Marseille, France.*

December 10, 2009

Abstract

Electromagnetic metamaterials have attracted much theoretical and experimental attention over the last two decades, due to their rich and even exotic properties. These developments, however, have exposed some gaps in the standard expositions of the classical theory of the dielectric response. For instance, the transition between the short wavelength domain (e.g. in photonic crystals) and the large wavelength domain (i.e. homogeneous effective media) is often treated as a “no man's land”, in spite of the fact that experimental studies of composite effective media often find themselves in exactly that intermediate zone. Also, the limitations and assumptions intrinsic to the widely used Mossotti-Clausius relation are rarely clarified. In the following I review the classical theory of the dielectric response, leading from the microscopic electromagnetic field equations, to the macroscopic homogeneous Maxwell equations. The roles and limitations of spatial averaging and of ensemble averaging are motivated and analyzed, with special attention to the phenomenon of spatial dispersion. The concept of inhomogeneous effective media is introduced and discussed in detail. The last sections discuss several notions which are important building blocks for an objective understanding of current metamaterials literature: one-dimensional stacks, superlenses, and flat lenses.

The main distinguishing feature of this review is its emphasis on conceptual discussion and physical understanding; it is hoped that this approach may complement the more calculational treatments available elsewhere, thereby filling an important gap in the literature.

Contents

Introduction	3
1 Electromagnetic materials and scale	5
2 The Lorentz model and its discontents	9
3 Spatial averaging as truncation	11
4 Ensemble averaging	16
5 Periodic media	21
6 Polarizability, susceptibility – electric	25
6.1 Small atoms – electric	27
6.2 Large atoms – electric	32
7 Polarizability, susceptibility – magnetic	33
7.1 Small atoms – magnetic	34
8 Permittivity and permeability – index and impedance	36
9 Homogeneous models and local models	38
10 Constructive versus holistic	44
11 Truncation revisited - 1D stacks	46
12 Negative index of refraction – the superlens	48
13 Flat lenses	54
Conclusion	58
Acknowledgements	59
References	59

Introduction

The two decades that have passed since the landmark contributions of Eli Yablonovitch [1] and Sajeev John [2] have produced a large amount of theoretical, numerical and experimental work in the field known at first as photonic crystals, and later more generally as electromagnetic metamaterials. Together with the advances in computing power and in microfabrication technologies, these developments have made it possible to conceive and design a rich array of new optical materials and devices.

While it is not possible to do justice to the massive volume of work that has emerged over the last two decades, I try to broadly highlight the main directions in this short introduction.

The original work of E. Yablonovitch and S. John aimed to use light confinement on the wavelength scale in order to gain greater control over light emission and propagation in artificial media, in view of improving the characteristics of devices such as lasers and solar cells. Today, light confinement continues to be a highly active line of research in optical microcavities with an escalating race to ever higher quality factors [3, 4, 5, 6, 7, 8], as well as the study of quantum electrodynamics effects in the coupling of such cavity modes to quantum dots [9, 10, 11, 12].

The idea of confining light in zero dimensional defects (cavities) also led to the study of waveguides (one-dimensional defects) [13, 14], and eventually to the control of light without using the bandgap, but rather the dispersion relation of the crystal itself: auto-collimation, the super-prism effect, and the negative refraction effect [15, 16, 17, 18].

Perhaps the most promising application that has emerged, from an industrial point of view, is the photonic crystal fiber, or microstructured fiber [19]. It was originally developed mainly in the group of P. Russell [20, 21, 22] leading to applications such as fiber lasers [23], supercontinuum generation [24, 25], and other nonlinear optics phenomena [26, 27].

But nonlinear optics is a much broader sub-field within photonic crystals research. It includes, among other things, the study of solitons [28, 29, 30, 31, 32], bistable switching [33, 34, 35, 36], slow light [37, 38, 39, 40], and second harmonic generation [41, 42, 43, 44, 45].

Even though initially dielectric structures received the most attention, the research community has also become interested in metallic structures, particularly after several surprising results came to light, such as the extraordinary transmission through sub-wavelength hole arrays [46, 47, 48], the existence of electromagnetic surfaces with exotic conduction properties [49], the possibility of using surface plasmons to explore the quantum nature of light [50], and J. Pendry's revival of the theoretical possibility of creating artificial materials (metamaterials) with a negative refractive index [51, 52, 53, 54, 55], i.e., materials with simultaneously negative permittivity and negative permeability.

The theoretical work of Pendry, followed by the various experimental tests of his ideas, has set off a flurry of activity in the theory, design and experimental scrutiny of very diverse effective medium structures. By far the two

most glamorous ideas that have come out of this research are the superlens [51, 56, 57, 58, 59, 60] and the electromagnetic cloak [61, 62, 63, 64, 65].

As will be discussed in more detail in the following sections, the latter contributions (starting with Pendry’s negative index materials) are distinguished from all of the rest of the research directions mentioned above by the fact that the phenomena they refer to operate on different *scales*. Whereas in the rest of the research mentioned above the wavelength λ is comparable to the periodicity or scatterer spacing d , the notion of “homogeneous effective medium” makes sense only in the domain where $\lambda \gg d$.

The shortcoming in the existing literature, which this article attempts to address, is the lack of a coherent and unified framework for understanding the *transition* between phenomena functioning on different scales, specifically, between the $d \approx \lambda$ and the $d \ll \lambda$ regimes. This lack has manifested itself, among other ways, through a large amount of experimental as well as numerical work where the effective index concept is invoked and used, with no attempt to show whether, or to what extent, the critical condition: $d \ll \lambda$ is satisfied. More explicitly, questions such as “Is λ large enough?”, “How would we know that?” and “What if it isn’t?” have been systematically swept under the rug in a large portion of the effective medium literature. We believe this is mostly due to the lack of adequate discussion of these issues in the existing standard electromagnetics texts, which were written when such issues would have been so purely academic as to be of no practical interest, and were therefore safely ignored.

Consequently, the aim of this article is to provide a much needed back-to-basics, thoroughly pedagogical review of effective medium theory from the perspective of the modern field of composite metamaterials. Special emphasis is placed on *motivating* each step of the derivations, and on clearly defining the domain of applicability and the *limitations* of the results. We begin, as we must, by emphasizing the concept of scale, in section 1. Then we review the most widespread theoretical model for artificial dielectrics: the Lorentz model, and its shortcomings (Sec. 2). The transition from the microscopic ($d \approx \lambda$) to the macroscopic ($d \ll \lambda$) requires two averaging procedures: ensemble averaging and spatial averaging (as argued persuasively by Russakoff [66] and Robinson [67]). I have chosen to discuss first spatial averaging, in section 3 and then ensemble averaging in section 4. The former is required for defining macroscopic *fields* (\mathbf{D} , \mathbf{H}), while the latter is required for defining macroscopic *parameters* (χ , ε , μ). The main goals of the treatment are to obtain the Mossotti-Clausius relation, in section 6.1, to discuss its meaning and limitations in Sections 6.2 and 10, and to propose a more technologically relevant approach to effective medium theory, in Section 9. Subsequent sections discuss several examples of effective media of current interest in the metamaterials literature: one-dimensional stacks, the superlens and other flat lenses. The Conclusion outlines several directions for future research.

1 Electromagnetic materials and scale

In the discussion below, I will divide phenomena and devices into categories following the relative scale of *four* characteristic lengths. The first, is the wavelength of light, λ . The second is the typical size of atoms, a , where an “atom” is generally and somewhat arbitrarily understood as the basic electromagnetic scattering unit in the structure. This somewhat vague definition covers not only actual atoms but also any electromagnetically convenient grouping thereof, including molecules, unit cells of crystals, macroscopic scatterers like dipole antennas or dielectric rods. 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 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.

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 ε and permeability μ . Thus, at this scale *all* traces of the medium’s true underlying discreteness are lost, and it looks smooth, continuous and homogeneous. However, as we will see, in order to reconcile this fact with the knowledge that the medium is really discrete, it is common to introduce the (purely fictitious but useful) idea of considering each atom in the material as a point scatterer which acquires a *dipole moment* in an external field [68]. 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 ε and permeability μ , if the dipolar approximation is not valid (e.g. when quadrupolar moments are non-negligible), then the *usefulness* of the homogeneous parameters μ, ε is drastically reduced. These parameters are in general dyadic (square matrices) and depend on the frequency.

It is shown below in Sec. 6 and further argued in Sec. 10 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 (difficult to prevent electron clouds from overlapping), but when dealing with metamaterials composed of macroscopic scatterers this is not an obstacle: it is up to the engineer to design them appropriately.

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, coax-

ial, 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 μ and ε 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 a given *field*, or *wave*, while the permittivity and permeability as properties of *matter*, or the *medium*, in which the wave is propagating. They are obviously related. Engineers tend to prefer the first set, while physicists the second. I 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 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 [69] and the Wiener bounds several years later [70], and continued through the work of Bruggeman [71], Hashin and Shtrikman [72], Milton [73] and finally Tsang and Kong [74] (see also Refs. [75, 76]). 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 [77]. 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 [78, 79, 80]. 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. 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 [81] but they generated relatively little interest before the late 1990's, see below.

Broadly speaking this is where electromagnetic meta-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 fraction 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 were lumped together under the name of *spatial dispersion* and they were known to appear when the scale of the electronic wavelength was comparable with the atomic scale: $\lambda_e \approx a$. The usefulness of these features of electronic transport in periodic media is well known, giving rise to the modern "information age".

In this context, the birth of photonic crystals at first, and of effective media later, is due to two ideas. Yablonovitch and John realized that the phenomena associated with spatial dispersion are independent of: A. Whether the periodicity is on the a or the s scale (i.e. large objects/inhomogeneities can play the role of atoms too), and B. The kind of wave that is propagating through the lattice, whether electronic (scalar wavefunction) or electromagnetic (vector \mathbf{E}/\mathbf{B} fields). 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 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. Thus the rich array of phenomena and devices of the previous section (except negative index metamaterials and their offspring, the superlens and the cloak) have in common that they function on a scale where $a \lesssim d \approx \lambda \ll s$ ¹.

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 $d \approx \lambda$: the

¹Since in this work we are concerned with characterizing materials themselves rather than devices using them, we will henceforth assume $s \rightarrow \infty$ and omit it for brevity.

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 Prof. 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 and of John to realize how deep the analogy goes.

Finally, the last stop on our journey through scales, and the configuration that will concern us for the rest of this review, is that where $a \lesssim d \ll \lambda$. Yablonovitch's and John's realization that by designing the medium on the a and d scales we can tailor the (complicated) dispersion relation when $\lambda \approx d$ naturally led to considering the - in principle simpler - situation of designing media with $\lambda \gg d$, as suggested by Kock in 1948 [81]. 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 [53, 54, 51] finally revived the field of effective medium 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 d 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 or double negative metamaterials.

Since all the media we will be interested in are periodic, for the rest of this article we will take d (inter-"atomic" distance) equal to the period. Thus what we mean by "atom" will depend on the context: in natural 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 λ . 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 gray area between heterogeneous ($a \lesssim d \approx \lambda$) and homogeneous ($a \lesssim d \ll \lambda$).* This gray area is characterized by *spatial dispersion*.

One of the main motivations of this review is that the standard electromagnetism texts such as Jackson's "Classical Electrodynamics" include no mention of the new perspective allowed by the last 20 years of work on electromagnetic metamaterials. In particular, the presentation of the all important Mossotti-Clausius relation includes no discussion of its scope and limitations or of the assumptions underlying it. The treatment of the transition from microscopic to macroscopic electromagnetism is minimal, as is the discussion of spatial dispersion, a concept which holds center stage in the modern science of metamaterials. It is my hope that this review will help point the way towards a more unified,

coherent, and modern approach to the theory of electromagnetic media.

The discussion below is mainly interested in phenomena occurring on a time scale sufficiently long that retardation effects are negligible. Consequently, the time dependence of the fields can safely be locked inside an ω dependence, which is assumed throughout the following discussion and which we therefore omit for brevity. We shall be concerned above all with the spatial behavior of the fields, that is, their dependence on \mathbf{x} or \mathbf{k} .

In the following sections I 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, the present approach is a constructive one, with an emphasis on *design*. Section 9 introduces the technologically relevant notions of inhomogeneous effective medium model, and of meta-photon crystal. But first, let us begin by reviewing the alma mater of effective medium theory: the Lorentz model.

2 The Lorentz model and its discontents

The Lorentz model [82] is the first successful approach to a microscopic theory of the dielectric response. Its first detailed English language exposition is in Ref. [83], while a useful historical review of the theory and its extensions is available in Ref. [84]. It is one of the most massively useful [85] and fertile [86, 84] models in the electromagnetic theory of matter, and its success in reproducing basic features of the optical properties of matter at large wavelengths has been amply documented.

From a more general point of view however, one of the most striking features of its success is the fact that it was elaborated long *before* the microscopic quantum nature of matter was understood. The Lorentz theory has this in common with another very successful microscopic theory: the kinetic theory of gases. The first version of the latter was elaborated by D. Bernoulli in the 1730's, long before the quantum nature of atoms could even be imagined.

Whereas the Lorentz model treats atoms as electromagnetic versions of tiny harmonic oscillators: point dipoles, the kinetic theory treats them as tiny billiard balls: hard, elastic spheres². However, quantum mechanics has shown (with spectacular accuracy) that they are emphatically neither “hard”, nor “point” particles, that their motion exhibits diffraction, and that they even lack that most central feature of familiar, macroscopic objects: a well defined position.

These models are so simple and effective that they have become part of the

²In fact, one of the main precursors to Lorentz's theory, elaborated by Mossotti in 1836 [87], also treats atoms as hard (conducting) spheres, no doubt taking inspiration from kinetic theory. Moreover, Mossotti's estimates of the radii of atoms thus conceived are in remarkable agreement with the kinetic theory estimates, see the table on page 5 of Ref. [88].

permanent conceptual furniture of the minds of most scientists, to the point that when thinking of a dielectric or a gas, we immediately summon up images of tiny oscillating arrows in regular arrays, or of hard billiard balls bouncing against each other. It is easy to forget how far from reality these mental images really are.

This situation therefore begs the question: How is it possible to make a model that starts from premises that are so wrong, and yet make predictions that are so accurate? The answer, given colloquially by Robert Laughlin [89], and made explicit in the sections below, is that these models work not because they provide an accurate description of the physics, but because *an accurate description is not required*: the phenomena concerned are largely independent of the details of the underlying mechanics. The fact that the simple (and therefore wrong) models work proves that the macroscopic behavior of a dielectric or a gas depends in only a small measure on the physics of their building blocks: the atoms (See also footnote 8 in Section 4 below). Consequently, there are many possible different but equivalent microscopic models consistent with the macroscopic behavior. The particular ones we most commonly use are special only by virtue of their intuitive simplicity.

This is an immediate indication that the derivation of the macroscopic laws from the microscopic laws must involve an irreversible step whereby information is irretrievably lost. Once this step is taken, the description of the medium is no longer unique. Many scenarios different in the small are no longer distinguishable in the large. The tool used to accomplish this is the spatial average, described in Section 3, the consequences of which are encapsulated in the “master equation” of the effective medium, Eq. (22) of Section 6.

However, it can be argued that when studying composite metamaterials, the very strengths of the Lorentz model become weaknesses. This is because it was designed to apply to situations where one either does not know, or is not interested in knowing, too much about what is happening on the microscopic scale. The Lorentz model works well when the wavelength is much larger (at *least* a couple of orders of magnitude) than the interatomic distance, at which scale questions of the internal functioning of the atoms and their interactions are largely irrelevant. These aspects are locked away in black box phenomenological parameters, and students are taught to never ask exactly where the numbers come from as long as they work ³

In contrast, composite metamaterials often work at wavelengths barely one order of magnitude larger than the interatomic distance ($\lambda/d = 6$ in the landmark Refs. [90, 52] !), and moreover one is *highly* interested in the internal physics of the atoms on the microscopic scale for the obvious reason: *design*. Hiding away the complexity of the microscopic physics inside black-box parameters is exactly what we *don't* want to do: it is this very complexity which provides us with the wealth of possibilities for which metamaterials have captured so much attention in the first place! What is needed is a natural mechanism for studying the transition between the *Lorentz regime* ($\lambda \gg d$) and the *Bloch*

³A status which they share with the “fundamental” numbers: c, \hbar , etc.

regime ($\lambda \approx d$). In the following sections it will be shown that this mechanism is provided by the averaging volume f used in the spatial averaging procedure. This procedure may be seen as defining a horizon of physical law: beyond (or rather below) this horizon physical law becomes unobservable, in a given experiment. The physics of a given experiment, and thereby the best way to interpret its results, depends crucially on where this horizon is positioned⁴. This is the main conceptual message of this review.

3 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. 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 underlines indicate the fact that no ensemble average has been applied to the quantities, a necessary operation as argued in Sec. 4. If one assumes the medium behaves as a linear time invariant system, and the ensemble average has been applied (removing the underlines), one can define permittivity and permeability tensors through the relations

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

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 8. 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:

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

⁴The existence of this horizon in many physical situations is what R.B. Laughlin refers to as the phenomenon of “protection” [89] by which nature suppresses microscopic irregularity to give rise to macroscopic stability, predictability and regularity. An observer at a given scale is thereby “protected” from knowing too much about what is happening at much smaller scales by a barrier of irrelevance, i.e., a horizon.

and with an averaging procedure. Various approaches to the averaging have been put forward: spatial, temporal or ensemble averaging. It was argued by Russakoff [66] 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

$$[\underline{\xi}(\mathbf{x})] = \int d^3x' f(\mathbf{x}') \underline{\xi}(\mathbf{x} - \mathbf{x}') \quad (2)$$

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

$$[\underline{\xi}(\mathbf{x})] = f(\mathbf{x}) \circ \underline{\xi}(\mathbf{x}) \quad (3)$$

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. (2):

$$\begin{aligned} [\underline{\xi}(\mathbf{x})] &= f(\mathbf{x}) \circ \underline{\xi}(\mathbf{x}) \\ &= \mathcal{F}^{-1}(\mathcal{F}(f(\mathbf{x}))\mathcal{F}(\underline{\xi}(\mathbf{x}))) \\ &= \mathcal{F}^{-1}(\tilde{f}(\mathbf{k})\tilde{\underline{\xi}}(\mathbf{k})) \end{aligned} \quad (4)$$

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

$$\mathcal{F}(\underline{\xi}(\mathbf{x})) = \int \underline{\xi}(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} d^3x \quad \mathcal{F}^{-1}(\tilde{\underline{\xi}}(\mathbf{k})) = \frac{1}{(2\pi)^3} \int \tilde{\underline{\xi}}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} d^3k. \quad (5)$$

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 [67]. 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

$$\begin{aligned} \nabla \cdot [\mathbf{b}] &= 0 & \nabla \times [\mathbf{e}] + \frac{\partial [\mathbf{b}]}{\partial t} &= 0 \\ \nabla \cdot [\mathbf{e}] &= [\eta]/\varepsilon_0 & \frac{1}{\mu_0} \nabla \times [\mathbf{b}] - \varepsilon_0 \frac{\partial [\mathbf{e}]}{\partial t} &= [\mathbf{j}] \end{aligned} \quad (6)$$

The macroscopic fields $\underline{\mathbf{E}}$ and $\underline{\mathbf{B}}$ are then defined as $\underline{\mathbf{E}} = [\mathbf{e}]$ and $\underline{\mathbf{B}} = [\mathbf{b}]$ and in order to obtain the macroscopic equations we need to write out the average charge and current densities, $[\eta]$ and $[\mathbf{j}]$. 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, we leave this discussion to the section on "Ensemble averaging" below.

For the time being we simply group charges in the medium arbitrarily in clusters, which we refer to as atoms. The issue of stability does not arise since the time t is fixed. 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. (4) to $[\eta(\mathbf{x})]$:

$$\begin{aligned} [\eta(\mathbf{x})] &= \mathcal{F}^{-1}(\tilde{f}(\mathbf{k})\underline{\tilde{\eta}}(\mathbf{k})) \\ &= \mathcal{F}^{-1}(\tilde{f}(\mathbf{k}) \sum_n \underline{\tilde{\eta}}_n(\mathbf{k})) \\ &= \sum_n \mathcal{F}^{-1}(\tilde{f}(\mathbf{k})\underline{\tilde{\eta}}_n(\mathbf{k})). \end{aligned} \quad (7)$$

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 $\underline{\tilde{\eta}}_n(\mathbf{k})$ as a Taylor series around $\mathbf{k} = \mathbf{0}$ and hope that we may only

need to keep a few terms. We have

$$\begin{aligned}\tilde{\eta}_n(\mathbf{k}) &= \tilde{\eta}_n(\mathbf{k})|_{\mathbf{k}=\mathbf{0}} + \mathbf{k} \cdot \nabla_{\mathbf{k}} \tilde{\eta}_n(\mathbf{k})|_{\mathbf{k}=\mathbf{0}} + \mathbf{k} \cdot \tilde{\mathbf{R}}_n(\mathbf{k}) \\ &= \tilde{\eta}_n(\mathbf{k})|_{\mathbf{k}=\mathbf{0}} + \mathbf{k} \cdot \left(\nabla_{\mathbf{k}} \tilde{\eta}_n(\mathbf{k})|_{\mathbf{k}=\mathbf{0}} + \tilde{\mathbf{R}}_n(\mathbf{k}) \right)\end{aligned}\quad (8)$$

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. (5) we have

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

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 Section 1 was referred to as the *dipolar approximation*. We now define the *generalized electric moment*

$$\tilde{\underline{\mathbf{p}}}_n(\mathbf{k}) = i \nabla_{\mathbf{k}} \tilde{\eta}_n(\mathbf{k})|_{\mathbf{k}=\mathbf{0}} + i \tilde{\mathbf{R}}_n(\mathbf{k}) = \underline{\mathbf{p}}_n + i \tilde{\mathbf{R}}_n(\mathbf{k}) \quad (9)$$

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:

$$\tilde{\eta}_n(\mathbf{k}) = q_n - i\mathbf{k} \cdot \tilde{\underline{\mathbf{p}}}_n(\mathbf{k}) \quad (10)$$

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

$$\begin{aligned}\tilde{f}(\mathbf{k}) &= 1 + \mathbf{k} \cdot \nabla_{\mathbf{k}} \tilde{f}_n(\mathbf{k})|_{\mathbf{k}=\mathbf{0}} + \mathbf{k} \cdot \tilde{\mathbf{R}}_f(\mathbf{k}) \\ &= 1 + \mathbf{k} \cdot \tilde{\mathbf{R}}_f(\mathbf{k})\end{aligned}\quad (11)$$

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. (7). We shall see what this constraint entails in the following sections.

The generic term of Eq. (7) takes the form:

$$\begin{aligned}\mathcal{F}^{-1}(\tilde{f}(\mathbf{k})\tilde{\eta}_n(\mathbf{k})) &= \mathcal{F}^{-1}\left(\underline{q}_n\tilde{f}(\mathbf{k}) - i\mathbf{k} \cdot \tilde{\mathbf{p}}_n(\mathbf{k})\tilde{f}(\mathbf{k})\right) \\ &= \underline{q}_n\delta(\mathbf{x} - \mathbf{x}_n) \circ f(\mathbf{x}) - \delta'(\mathbf{x} - \mathbf{x}_n) \circ \mathcal{F}^{-1}\left(\tilde{f}(\mathbf{k})\tilde{\mathbf{p}}_n(\mathbf{k})\right) \\ &= \underline{q}_nf(\mathbf{x} - \mathbf{x}_n) - \nabla \cdot (f(\mathbf{x}) \circ \underline{\mathbf{p}}_n(\mathbf{x} - \mathbf{x}_n))\end{aligned}\quad (12)$$

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. (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:

$$\begin{aligned}\underline{\mathbf{p}}_n(\mathbf{x} - \mathbf{x}_n) &= \mathcal{F}^{-1}(\tilde{\mathbf{p}}_n(\mathbf{k})) \\ &= \mathcal{F}^{-1}(\underline{\mathbf{p}}_n + i\tilde{\mathbf{R}}_n(\mathbf{k})) \\ &= \underline{\mathbf{p}}_n\delta(\mathbf{x} - \mathbf{x}_n) + i\underline{\mathbf{R}}_n(\mathbf{x} - \mathbf{x}_n)\end{aligned}\quad (13)$$

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

$$[\eta(\mathbf{x})] = -\nabla \cdot \underline{\mathbf{P}}(\mathbf{x}) \quad (14)$$

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

$$\underline{\mathbf{P}}(\mathbf{x}) = \left[\sum_n \underline{\mathbf{p}}_n(\mathbf{x} - \mathbf{x}_n) \right] \quad (15)$$

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

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

Note that among the consequences of the spatial averaging is the fact that a *discrete* microscopic quantity (the atomic polarization \mathbf{p}) has been transformed into a *continuous* macroscopic one (the polarization \mathbf{P}). It has been argued that this is a very general feature of the universe: macroscopic continuity hides microscopic discreteness. Many examples of this phenomenon can be found in solid state physics in particular [89].

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. For instance the interaction of neighboring atoms has not even been mentioned. 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 us to introduce *time* into the treatment, and the way to do this is through 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 [68]. In our case, the $\mathbf{P}(\mathbf{x})$ includes the quadrupolar and higher terms of the multipolar expansion of the microscopic charge distribution η , such that Eq. (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. (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 λ and a*** (the other, the proximity of λ and d , is discussed below in Sections 5 and 6). 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 6 where we introduce the susceptibility.

4 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 $\nabla \cdot \mathbf{e} = \eta/\varepsilon_0$ 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. (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. (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 periodic. 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 model time invariant is required. We must “zoom out” both spatially and temporally, and the previous section only performed the former. Thus the question becomes of what is the best way to perform a time average, i.e. to extract the “typical” response of the medium from the chaotic thermal disorder characterizing its microscopic motions.

This is a subtle and vexing question. To start with, it is far from obvious why such macroscopically “typical” behavior should exist *at all*⁵! It is therefore also

⁵Understanding the mechanisms of the emergence of macroscopic reliability from microscopic fickleness remains a very much open problem in physics [89].

far from clear exactly how this typical behavior is best extracted. Some writers, such as Van Vleck [88], initially used time averages over individual molecules and spatial averages (which Van Vleck refers to as “statistical averages”) over large numbers of molecules. Subsequently, however, the approach that has gained favor is the more mathematically rigorous ensemble averaging of Mazur and Nijboer ([91], see also [92, 84, 67]). They prefer the rigor of mathematical manipulation to the physical intuition leading previous authors to use “too much verbal explanation”.

The idea is to imagine a very large (i.e. infinite) number of systems prepared in exactly the same way as the system under study. This is known as the ensemble. Then one introduces the ensemble distribution function f_{ens} , which depends on all the positions and momenta of all particles within the system, and whose value at a given point in configuration space is given by the fraction of the imaginary systems in the ensemble which have a configuration close to that point. f_{ens} is a probability density function on configuration space. The *observed, or experimental* values of macroscopic quantities are then calculated by averaging their instantaneous values over the ensemble, by means of the distribution function.

This argument is used broadly in statistical physics, though it has been criticized on a conceptual basis (see section 2.3 in [93], reprinted in [94]). In particular, experiments are not performed on an infinity of similar systems, but on a single system; nor do they need to be repeated many times on a given system, a single experiment suffices. Moreover, the time averaging approach is not free from problems either, due to the fact that the length of time required for the averaging to converge on the macroscopic (i.e. “typical”) value of some quantity may be longer than the time that the medium is actually subjected to the probe (i.e. field) in question, in a given experimental test.

This having been said, going into the arcane and subtle aspects of the precise nature of the temporal “zooming-out” procedure required resides firmly outside the scope of this review⁶. All we require for the present purposes is that the averaging have 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

⁶The study of how qualitatively different behavior emerges at large scales from other behavior at small scales has generated a vast, fascinating and increasingly active volume of research, in statistical physics, solid state physics, as well as in information theory, pattern recognition, and artificial intelligence. Refs. [95, 96, 97, 98, 99, 89, 100] are a random sampling of some research directions that have emerged.

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 (classical *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.

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 [101]. 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 unchanged⁷. The equations are not invariant with respect to ensemble averaging. We therefore emphasize that one should not expect the divergence equation $\nabla \cdot \mathbf{e} = \eta/\epsilon_0$ to suffice in describing the interaction between the microscopic field \mathbf{e} and the charge distribution η , 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:

$$\mathbf{p}_n(\mathbf{x} - \mathbf{x}_n) = \epsilon_0 \gamma_n^e(\mathbf{x} - \mathbf{x}_n) \mathbf{e}(\mathbf{x}) \quad (16)$$

⁷This is a fact with very deep implications. It implies that the inhomogeneous Maxwell equations in matter are physically hybrid because they would not exist without quantum mechanics. Moreover, since *they have the exact same form in matter as in vacuum*, this has led some to speculate about space itself as a fundamentally inhomogeneous “material”, and the familiar continuous Maxwell equations emerging in the limit of large wavelengths from some yet unknown underlying physics [102, 89].

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 γ_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 be⁸. 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 γ_n^e such that the total polarization can be written as a sum over the atoms, and Eq. (15) remains valid (though only formally) without the underlines:

$$\mathbf{P}(\mathbf{x}) = \left[\sum_n \mathbf{p}_n(\mathbf{x} - \mathbf{x}_n) \right] = [\mathbf{p}(\mathbf{x})] = [\varepsilon_0 \gamma^e(\mathbf{x}) \mathbf{e}(\mathbf{x})] \quad (17)$$

where

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

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. (16) while the polarization charge density $\eta(\mathbf{x}) = -\nabla \cdot \mathbf{p}(\mathbf{x})$ produces an electric field according to Coulomb's law,

$$\mathbf{e}_{\text{pol}}(\mathbf{x}) = \frac{1}{4\pi\varepsilon_0} \nabla_{\mathbf{x}} \int \frac{\nabla_{\mathbf{x}'} \cdot \mathbf{p}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (18)$$

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 η) is tied directly to the time dependence of the electric field. The two equations $\nabla \cdot \underline{\mathbf{e}}(\mathbf{x}, t) = \underline{\eta}(\mathbf{x}, t)/\varepsilon_0$ and $\mathbf{F}_j(t) = q_j \underline{\mathbf{e}}(\mathbf{x}_j, t)$ have been replaced by equations (18) and (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

$$\mathbf{e}(\mathbf{x}) = \mathbf{E}_{\text{ext}}(\mathbf{x}) + \mathbf{e}_{\text{pol}}(\mathbf{x}) = \mathbf{E}_{\text{ext}}(\mathbf{x}) + \frac{1}{4\pi\varepsilon_0} \nabla_{\mathbf{x}} \int \frac{\nabla_{\mathbf{x}'} \cdot \gamma^e(\mathbf{x}') \mathbf{e}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x'$$

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

⁸ As Robinson [67] observes, “[...] for all we are concerned atoms could equally well be held together with glue and rubber bands”.

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 we will not delve into.

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 retain their form.

We note that Eq. (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 λ 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 λ 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.

5 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 ω

propagating in a periodic lattice has a space dependent part of the form

$$\mathbf{e}(\mathbf{x}) = \mathbf{U}(\mathbf{x})e^{i\mathbf{k}_B\mathbf{x}} \quad (19)$$

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 ω 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. [103, 104]. 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. (19) is very close to Γ 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)

$$\text{real}(\tilde{f}(\mathbf{k})) = \exp\left(-\frac{1}{2\nu}\left(\frac{2k}{K}\right)^{2\nu}\right), \quad (20)$$

see Fig. 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

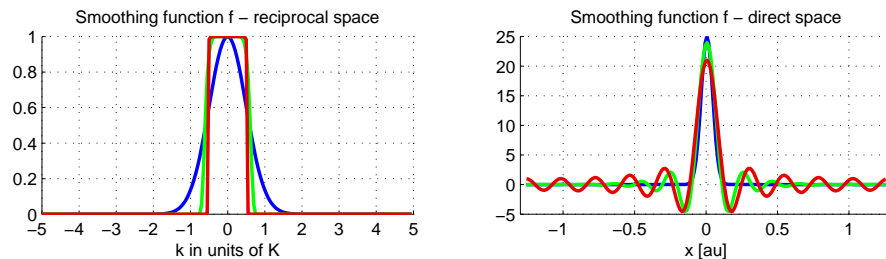


Figure 1: Comparison of smoothing functions f with sharper and sharper cut-offs, in direct and reciprocal space. The blue curves are Gaussians while the green and red curves correspond to ν of 5 and 30 respectively. The sharper cutoff leads to broader smoothing functions in direct space.

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 where we compare the absolute values of f and \tilde{f} for $\nu = 1$ (Gaussian) and ν 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 ξ 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 Γ , 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 ($f(\mathbf{k})$) of Eq. (20) with a very large ν . 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 η_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*.

Note that the model of the medium depends both on its intrinsic features (charge distribution) *and* on choices made by the observer (the volume f which we have total control over). When λ is large, the model we build is mostly dependent on the intrinsic features of the medium, which gives the impression of an independent, classical reality underlying its behavior. However, when λ becomes smaller, the model becomes increasingly complex, and the choice of f increasingly critical. Our model of the medium, i.e., our perception of what it “*is*” starts to depend not only on *it* (i.e. on η) but also on *us* (i.e. on our choice of f) ! We see the first signs of emerging observer participancy. In the words of statistician George E. P. Box: “Essentially, all models are wrong, but some are useful”.

Before moving on we must also 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 gov-

erned 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.

6 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 [105]. 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 [105] of dielectric media, also known as the macroscopic description, therefore must be obtained by 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

$$\mathbf{P}(\mathbf{x}) = \varepsilon_0 \chi^e(\mathbf{x}) \circ \mathbf{E}(\mathbf{x}) = \varepsilon_0 \chi^e(\mathbf{x}) \circ [\mathbf{e}(\mathbf{x})]. \quad (21)$$

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}_{pol} of Eq. (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 in an experiment 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. (21) with Eq. (17) obtaining the master equation of the dielectric medium

$$[\mathbf{p}(\mathbf{x})] = [\varepsilon_0 \gamma^e(\mathbf{x}) \mathbf{e}(\mathbf{x})] = \varepsilon_0 f(\mathbf{x}) \circ (\gamma^e(\mathbf{x}) \mathbf{e}(\mathbf{x})) = \varepsilon_0 \chi^e(\mathbf{x}) \circ [\mathbf{e}(\mathbf{x})] \quad (22)$$

where we have assumed zero intrinsic polarization at zero field (no ferro-electricity). The susceptibility χ^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 χ^e is obviously fundamentally dependent on γ^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-, worlds that are intuitively and physically apart and no straightforward or smooth transition is possible.

Eq. (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 Schrödinger’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 sections, 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. (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 6.2), 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.

6.1 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 of Ref. [106].

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. (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

$$\mathbf{e}_{dipole}(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \nabla \left(\mathbf{p} \cdot \nabla \frac{1}{\|\mathbf{x}\|} \right) = \frac{1}{4\pi\epsilon_0} \frac{(3\hat{\mathbf{x}}(\hat{\mathbf{x}} \cdot \mathbf{p}) - \mathbf{p})}{\|\mathbf{x}\|^3} - \frac{\mathbf{p}}{3\epsilon_0} \delta(\mathbf{x}) \quad (23)$$

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

$$\mathbf{e}_{dipole}^*(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \frac{(3\hat{\mathbf{x}}(\hat{\mathbf{x}} \cdot \mathbf{p}) - \mathbf{p})}{\|\mathbf{x}\|^3} \quad (24)$$

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

$$\mathbf{e}^*(\mathbf{x}) = \mathbf{e}(\mathbf{x}) + \sum_n \frac{\mathbf{p}_n \delta(\mathbf{x} - \mathbf{x}_n)}{3\epsilon_0} \quad (25)$$

Eq. (22) for this case must therefore be written as

$$f(\mathbf{x}) \circ (\gamma^e(\mathbf{x}) \mathbf{e}^*(\mathbf{x})) = \chi^e(\mathbf{x}) \circ [\mathbf{e}(\mathbf{x})] \quad (26)$$

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

$$f(\mathbf{x}) \circ (\gamma^e(\mathbf{x}) \mathbf{e}^*(\mathbf{x})) = f(\mathbf{x}) \circ \left(\sum_n \gamma_n^e(\mathbf{x}) \mathbf{e}^*(\mathbf{x}) \right) \quad (27)$$

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 $\tilde{\gamma}_n^e(\mathbf{k}) \cong \tilde{\gamma}_n^e(\mathbf{0}) \equiv \gamma_s^e$ (all atoms are identical) and we have $\gamma_n^e(\mathbf{x}) \cong \gamma_s^e \delta(\mathbf{x} - \mathbf{x}_n)$ where the s subscript stands for the singular, or DC component of the polarizability. The left side of Eq. (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

$$\tilde{f}(\mathbf{k}) \left(\tilde{\mathbf{e}}^*(\mathbf{k}) \circ \frac{\gamma_s^e}{V_{\text{uc}}} \sum_n \delta(\mathbf{k} - \mathbf{G}_n) \right) \quad (28)$$

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_{uc} is the volume of the unit cell in real space (see Eq. 2-12 of Kittel [103]). 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 (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 [68] 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. (23) and (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

$$N\gamma_s^e\left([\mathbf{e}(\mathbf{x})] + \frac{1}{3}\chi^e(\mathbf{x}) \circ [\mathbf{e}(\mathbf{x})]\right) = \chi^e(\mathbf{x}) \circ [\mathbf{e}(\mathbf{x})] \quad (29)$$

and by grouping the terms containing χ 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:

$$\chi^e(\mathbf{x}) = \frac{N\gamma_s^e}{1 - N\gamma_s^e/3}\delta(\mathbf{x}) \quad (30)$$

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. (29) 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. (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. (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. This energy balance can be accounted for by replacing the free-space imaginary polarizability γ with a real polarizability γ' given by the relation (see Section III of Ref. [107])

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

This relation assumes 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} . More general versions of this relation can be found in Ref. [107]. 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 Γ 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 effect dominates. This explains why it is so useful to treat resonators or other scatterers in composite metamaterials 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.

6.2 Large 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 γ_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. (28) becomes

$$\begin{aligned} \tilde{f}(\mathbf{k}) (\tilde{\gamma}(\mathbf{k}) \circ \tilde{\mathbf{e}}(\mathbf{k})) &= \tilde{f}(\mathbf{k}) \left(\sum_i \gamma_i^e \delta(\mathbf{k} - \mathbf{G}_i) \right) \circ \left(\sum_j \mathbf{E}_j \delta(\mathbf{k} - \mathbf{G}_j - \mathbf{k}_B) \right) \\ &= \tilde{f}(\mathbf{k}) \sum_{i,j} \gamma_i^e \mathbf{E}_j \delta(\mathbf{k} - \mathbf{G}_i - \mathbf{G}_j - \mathbf{k}_B) \end{aligned} \quad (31)$$

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 10. 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 α 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. 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 α into the effective polarizability, which was a phenomenological parameter to start with. We define

$$\gamma_{\text{eff}}^e = \frac{3\gamma_s^e}{3 + N\gamma_s^e(1 - \alpha)} \quad (32)$$

and the Mossotti-Clausius relation remains formally unchanged

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

7 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

$$\mathbf{H}(\mathbf{x}) = \frac{\mathbf{B}(\mathbf{x})}{\mu_0} - \mathbf{M}(\mathbf{x}) \quad (33)$$

the inhomogeneous curl equation from Eq. (6) becomes

$$\nabla \times \mathbf{H}(\mathbf{x}) - \frac{\partial}{\partial t} (\varepsilon_0 \mathbf{E}(\mathbf{x}) + \mathbf{P}(\mathbf{x})) = 0 \quad (34)$$

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. (22) is then obtained by combining the last two equations

$$f(\mathbf{x}) \circ (\gamma^m(\mathbf{x}) \mathbf{b}(\mathbf{x})) = \chi^m(\mathbf{x}) \circ \mathbf{H}(\mathbf{x}) \quad (35)$$

In the following section we consider the small atoms case.

7.1 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. (23) and (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 6.1 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

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

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.

8 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. (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. (34) is then written

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

By rewriting Eq. (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:

$$\nabla \times \mathbf{H}(\mathbf{x}) + i\omega \mathbf{D}(\mathbf{x}) = 0 \tag{36}$$

$$\nabla \cdot \mathbf{B}(\mathbf{x}) = 0 \tag{37}$$

$$\nabla \times \mathbf{E}(\mathbf{x}) - i\omega \mathbf{B}(\mathbf{x}) = 0 \tag{38}$$

$$\nabla \cdot \mathbf{D}(\mathbf{x}) = 0 \tag{39}$$

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. (36) and (38) 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

$$\mathbf{E}(\mathbf{x}) = \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} e^{ikx - i\omega t}. \quad (40)$$

This represents a plane wave of frequency ω propagating in the positive x direction. k is the wavevector and by the Helmholtz equation it is related to ω 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 α, β 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

$$\beta^2 - \alpha^2 = \mu'\varepsilon' - \mu''\varepsilon'' \quad (41)$$

$$2\alpha\beta = \mu'\varepsilon'' + \varepsilon'\mu'' \quad (42)$$

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 β 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. (42) if one keeps in mind that α , ε'' and μ'' must all be positive in a passive medium 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 3x3 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 $n = \mu_y \varepsilon_x$. However note that there is no such thing as the “tensor index” in the anisotropic case. The definition $n^2 = \mu\varepsilon$ is only valid in the scalar case. In the anisotropic case one must begin once more with Eqs. (36) and (38) (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.

9 Homogeneous models and local models

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:

$$f(\mathbf{x}) \circ (\gamma^e(\mathbf{x})\mathbf{e}(\mathbf{x})) = \chi^e(\mathbf{x}) \circ (f(\mathbf{x}) \circ \mathbf{e}(\mathbf{x})) \quad (43)$$

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 review but I must emphasize several aspects related to the role of the function $f(\mathbf{x})$, in particular in the context of periodic media.

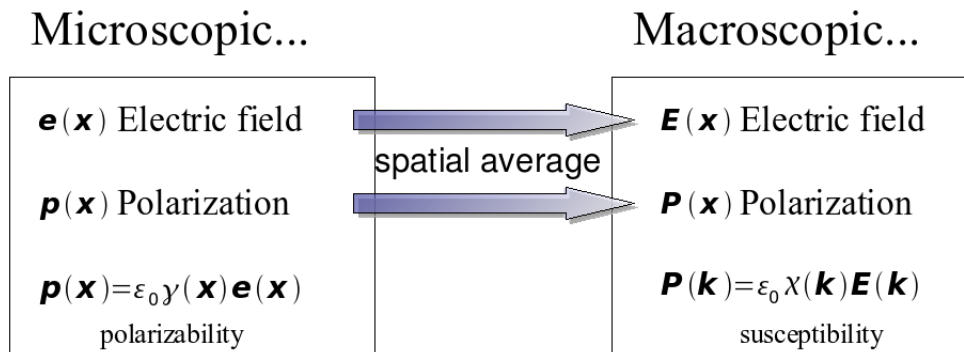


Figure 2: Schematic representation of the homogenization procedure.

In typical formulations of the derivation of the macroscopic Maxwell equations [66, 67, 68], 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 were 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.

The first thing to notice is that it is written in reciprocal space. This means that when we Fourier transform it to real space it takes the form of a convolution integral which we note concisely:

$$\mathbf{P}(\mathbf{x}) = \chi(\mathbf{x}) \circ \mathbf{E}(\mathbf{x}) \quad (44)$$

This implies that the value of the polarization at a given point is related to the values of the macroscopic electric field over a region the size of the susceptibility. It is often said that the medium has a “non-local response”. A medium with a “local response” is then a medium where $\chi(\mathbf{k}) = \text{const.}$, corresponding to a singular real space susceptibility $\chi(\mathbf{x}) = \chi\delta(\mathbf{x})$. I argue that this language is

misleading because what is local or non-local is not the *medium*, but the *model* we have made of it, which depends on the position of the $f(\mathbf{x})$ dial. By tuning $f(\mathbf{x})$ we can always obtain a local model, though in some cases it may be an inhomogeneous model. Non-locality (known as spatial dispersion in reciprocal space) is a direct consequence, not only of the properties of the medium, but also of our choice of ignoring microscopic details of the field and charge distributions.

But why would a local model be preferable to a non-local one? Because in a certain sense which is made more precise in this section and is supported by the results of Ref. [108], a non-local model is an incomplete description of the behavior of the underlying medium. This can be understood by considering a nonlocal susceptibility $\chi(\mathbf{k})$. This quantity describes what will happen when a given macroscopic field interacts with a given medium (distribution of matter). However, the \mathbf{k} dependence indicates that the scientist's knowledge of the medium is incomplete because he/she cannot predict how the medium will react without some additional information about the field, namely the \mathbf{k} vector. Physically speaking the \mathbf{k} dependent susceptibility encapsulates a *partial* description of the distribution of matter in the medium because it requires knowledge of the *field* in order to make any quantitative predictions. A local susceptibility, on the other hand is a complete description because it is independent of the field distribution in the medium (at the given frequency). Thus if a physicist would like to understand a medium at a given frequency, he must favor a complete, and therefore, local, description.

What we mean by an incomplete description is related to the fact that the spatial averaging is designed to eliminate microscopic information. To summarize the above discussion, therefore, as long as the microscopic information which is eliminated by the averaging process is irrelevant, the homogeneous model is complete and thereby local. However, when the microscopic aspects which are averaged over by the smoothing function $f(\mathbf{x})$ have a non-negligible contribution to the behavior of the medium, then the model becomes non-local.

In order to understand where the missing microscopic information resides it is simplest to consider the homogenization of a one-dimensional system. The left side of Fig. 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 0.2au 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. 3) to consist of filtering out all but the lowest harmonic, corresponding to $n = 0$.

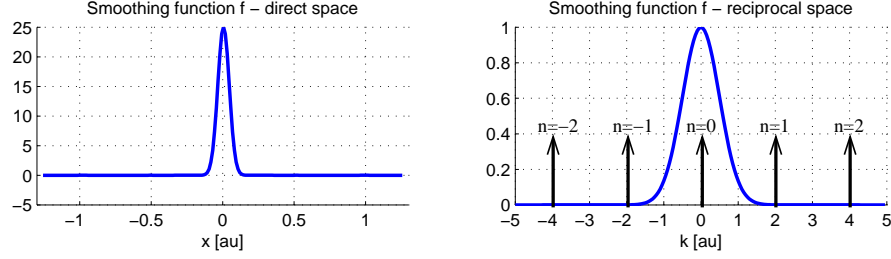


Figure 3: 1D smooth

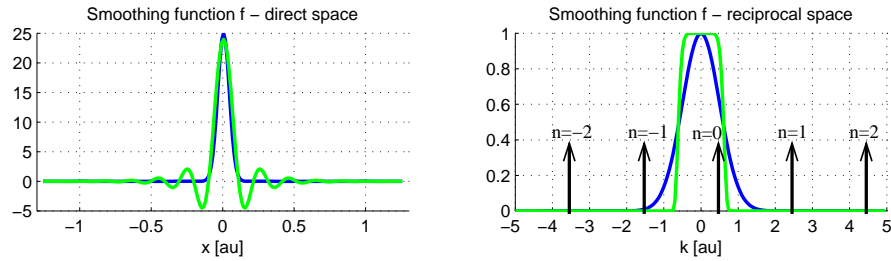


Figure 4: 1D smooth

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. 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 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. This situation is what is known as nonlocality. The homogeneous model has become nonlocal, or spatially dispersive, and as argued above, *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.

The above argument can be extended in a straightforward way to the two-dimensional case. This is illustrated in Fig. 5. From left to right we have a homogeneous model, a 1D model and the identity model ($f(\mathbf{x}) = \delta(\mathbf{x})$ the Dirac delta). 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

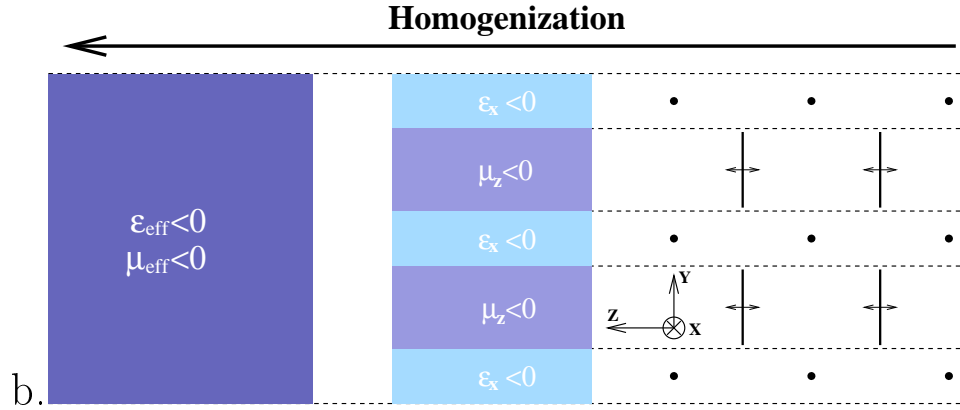
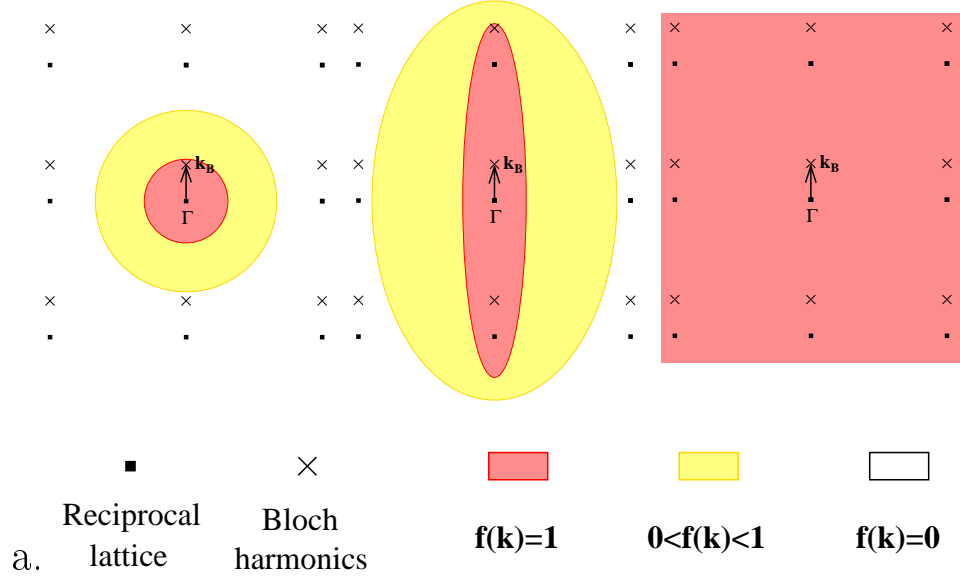


Figure 5: a. Three possible models corresponding to three $f(\mathbf{k})$ functions are illustrated. The first is a homogeneous model, only the lowest harmonic is kept, the rest are averaged over. The model on the right side is the one where all harmonics are kept, corresponding in real space to $f(\mathbf{x}) = \delta(\mathbf{x})$ the Dirac delta. 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 left is homogeneous. But is it also local?

b. Schematic illustrating the result of the intermediate homogenization step in the case of a composite material made of wires (parallel to the X axis) and magnetic resonators (dipole moments parallel to the Z axis). For a more detailed discussion see Ref. [108].

model is related to the importance of the information contained in the higher Bloch harmonics of the field. Our arguments lead to two predictions:

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.

These predictions were numerically tested successfully for the case of the perfectly conducting wires/resonators structure in Fig. 5b, the results are published in Ref. [108]. 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 [53, 54], 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 [109, 110] 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 magnetic resonance 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 the 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 homo-

geneous 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 may 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.

10 Constructive versus holistic

The arguments above have 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*** [95, 89] from the behavior of the parts that no direct, analytic 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 above 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 periodic 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 α .

The non-dipolar coupling parameter α 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 γ_{eff} (Eq. (32)) is therefore conceptually hybrid and in the context of classical treatments such as the one presented above *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 structure 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 the following sections 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 sections below remain within the *constructive* paradigm.

11 Truncation revisited - 1D stacks

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 completely 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 several major advantages. First, 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 inconvenience of the inevitability of quantum mechanics for a complete description. Moreover, working in the macroscopic realm, the ensemble average is no longer necessary, since the media are already time invariant. Only the spatial averaging is required. We are now standing safely on the solid ground of the macroscopic Maxwell's equations, which allows for detailed and rigorous numerical study of complex structures, without the need for ad-hoc, black-box or other phenomenological parameter guesswork.

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. [111]. 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 numerically straightforward, though experimentally cumbersome, since the values of the fields must be known everywhere.

We now obtain a relatively straightforward, but very important and useful result, the homogenization formulas for 1D structures, or stacks. They were first derived by Wiener in 1912 [70], but a more accessible reference is [112].

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

$$\overline{\varepsilon}_{\text{eff}} = \begin{pmatrix} [\varepsilon^{-1}(x)]^{-1} & 0 & 0 \\ 0 & [\varepsilon(x)] & 0 \\ 0 & 0 & [\varepsilon(x)] \end{pmatrix} \quad (45)$$

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 \mathbf{E} is normal to the layers \mathbf{D} is constant (or varies very slowly) even though it is the product of two quantities, ε and \mathbf{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 \mathbf{D} , unmodified, but it modifies both of the other two, with the result that $[\mathbf{D}] \neq [\varepsilon][\mathbf{E}]$, but rather $[\mathbf{D}] = [\varepsilon^{-1}]^{-1}[\mathbf{E}]$.

The seemingly innocuous result of Eq. (45) is important for two reasons. First, the 1D stack is the first and simplest example of an “intermediate”, inhomogeneous effective medium, as described above in Section 9. Second, unlike the arithmetic mean, *the harmonic mean is not bounded!*

It is easy to see that the arithmetic mean of two real numbers must belong to the interval between them: $a \leq \text{mean}_a(a, b) \leq b$. In the case of the harmonic mean an equivalent relation holds *only* when a and b are *both positive*. In the general case however, the harmonic mean is not bounded. In particular, it becomes infinite, when $\frac{a}{b} \rightarrow -1$. Thus interesting and insufficiently explored phenomena are to be expected from 1D stacks composed of layers with both positive and negative permittivities and/or permeabilities [113, 114, 115, 116, 117, 118]. In addition, recent work seems to indicate that the divergence of the

effective parameters in this case is an instance of a more general phenomenon [76].

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 ε , 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

$$\begin{aligned} \mathbf{D}_{\perp}(x) &= [\varepsilon^{-1}(x)]^{-1} \mathbf{E}_{\perp}(x) \\ \mathbf{D}_{\parallel}(x) &= [\varepsilon(x)] \mathbf{E}_{\parallel}(x) \end{aligned} \tag{46}$$

The orientations of the normal and parallel fields depend, of course, on the orientation of the discontinuity at any given point. This result has also been obtained with exquisite mathematical rigor by Li [119] and it has been used by S. Johnson to considerably improve convergence of his MPB implementation of the plane wave method for calculating photonic band structures [120, 121]. Analogous relations hold for the magnetic constitutive relations.

12 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 uses the formalism of Ref. [122]

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

$$U(x, y, z) = \int \int_{-\infty}^{\infty} \tilde{U}(u, v; z) e^{i(ux+vy)} du dv. \quad (47)$$

Replacing this in the Helmholtz equation we obtain

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

or

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

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

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

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

$$\tilde{U}(u, v; z) = A(u, v)e^{iwz} \quad (48)$$

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 $\tilde{U}(u, v; Z) = A(u, v)e^{iwZ}$. The propagation in an isotropic homogeneous medium therefore has the effect of

	$u^2 + v^2 < \beta^2 k_0^2$	$u^2 + v^2 > \beta^2 k_0^2$
$\beta > 0$	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 1: Table of the possible behaviors of w .

transforming the Fourier components of the field distribution according to the factor e^{iwz} . 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 μ , ε and w . As in section 8 we introduce the index $n^2 = (\beta + i\alpha)^2 = \mu\varepsilon$ where β and α are real, α is positive, and they are given by Eqs. (41) and (42). As before the imaginary part α 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 α 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 α 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 α . 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 β from the second relation above. We have already seen in Section 8 that when the medium is double positive, then β is positive, but that when it is double negative, then β is negative. The different possibilities are summarized in Table (1).

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 β is positive this amplitude is attenuated, while when β 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 (36) and (38) then changing the signs of the constitutive parameters μ and ε 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 $e^{i\omega z}$, while a similar region filled with $\mu = \varepsilon = -1$ placed next to the first provides the *inverse* filter, $e^{-i\omega z}$. 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 is cooling the lens to very low temperatures. Note also that noise can be not only of a physical origin but also numerical, due, for instance to the discretization employed in a computer simulation [123].

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.

We must also mention that 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 [124]. 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 also be transmitted by such a medium. This was done over thirty years later by John Pendry in the now famous Physical Review Letters article [51] 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, α 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 α 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 μ and ε positive with the transmission of a superlens, that is, a slab with $\mu = \varepsilon = -1$. On the left side of Fig. 6 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 α 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 α 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

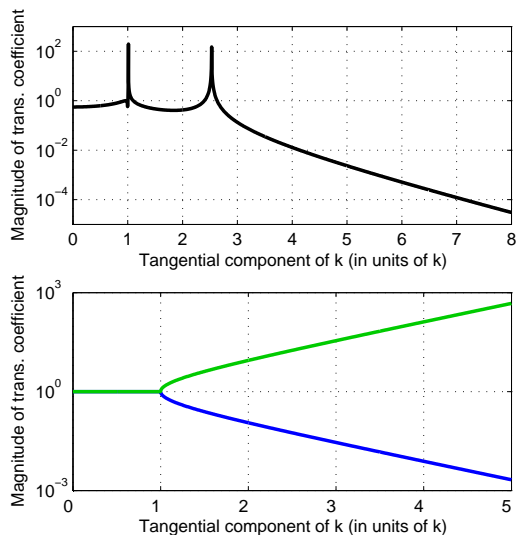


Figure 6: Transmission of dielectric slab with thickness $d = 1\text{au}$, with $\mu = 1$, $\varepsilon = 12$ and wavelength $\lambda = 5\text{au}$ as a function of the tangential component of the wavevector (top). In the lower figure we compare the transmission of the free space slab (blue) and the superlens (green).

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 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 α the closer we get to the superlens condition. In fact the superlens can be seen as having a guided mode at infinity on the α 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. 7 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

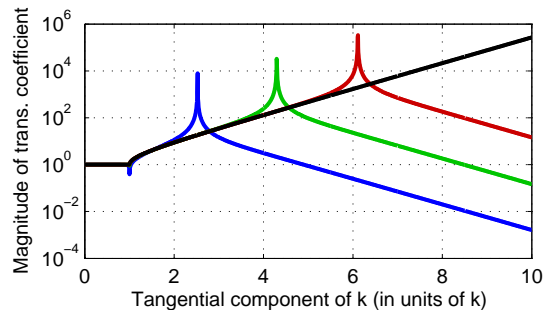


Figure 7: As the permittivity and permeability of the slab approach -1 the guided mode is shifted to infinity and the transmission approaches that of the superlens, the black curve, also the green curve in Fig. 6.

$\alpha = 4$ 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 optical microscopy techniques [125, 126, 127, 128]. 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 pertains to fields with $\alpha < 1$* . Consequently while the superlens may exhibit negative refraction, negative refraction does not imply a superlens.

13 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 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 first, we must say a few words about surface modes.

Let us consider Eq. (48). 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. (47) 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 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 (i.e., “charges free to move over distances comparable to the wavelength”) 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 an infinite metal slab does contain free charges, which is expressed in the fact that its 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 [57] 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 8 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 1au (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 [60]. 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

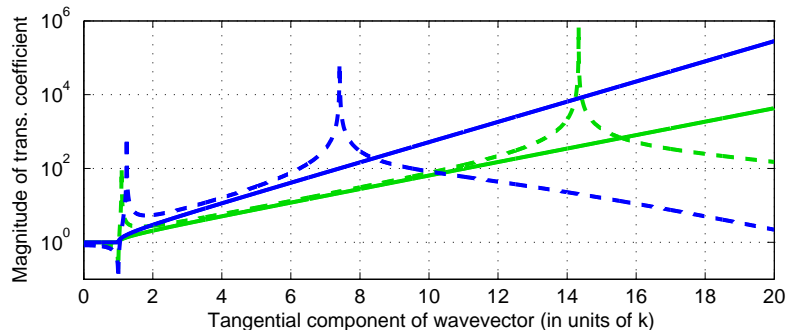


Figure 8: We compare the transmission of a superlens ($\varepsilon = \mu = -1$, solid curves) with that of an ideal plasma slab ($\varepsilon = -1$, $\mu = 1$, dashed curves) for wavelengths of $\lambda = 10\text{au}$ (blue) and $\lambda = 15\text{au}$ (green). The slab thickness is 1au .

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 [129].

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. [122] 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 crystal 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 ([17, 130, 131], see also Chapter 7 of Ref. [132]). The conceptual differences and/or practical advantages (if any) of this kind of flat lens with respect to the various

existing and highly impressive near field optical microscopy techniques [125, 126, 127, 128] are far from having been demonstrated convincingly.

Conclusion

In this review I have provided a detailed treatment of the effective medium theory of neutral media with no free charges, with particular attention to the frequency region intermediate between what I have called the Bloch domain ($d \approx \lambda$) and the Lorentz domain ($d \ll \lambda$), in the special but important case of periodic media. I have emphasized the importance of spatial dispersion throughout, and I have highlighted the central role of the averaging volume f in determining the outcome of the homogenization procedure.

The treatment presented here points toward two promising future research directions, one technological and the other more theoretical. I discuss them in order.

Inhomogeneous effective media. In Sections 6 and 9 the master equation of effective medium modeling was discussed (Eq. (43)), and I emphasized a point of view whereby the averaging volume f is seen not as an academic construct of purely theoretical interest, but as a very useful adjustment knob which can be used to obtain different kinds of effective medium models. In particular, by reducing the size of the averaging volume one may trade non-locality for inhomogeneity, effectively zooming in on the metamaterial, in order to understand the physics within the unit cell. This provides a technique able to explore a frequency domain which until now had remained in limbo: neither truly homogeneous, nor truly in the Bloch regime. This is the concept of the *inhomogeneous effective medium model*.

Moreover, since such inhomogeneous effective media may be much richer than existing photonic crystals (which mostly consist of air holes in classic dielectrics), this also opens the way to the design of *meta-photonic crystals*: metallo-dielectric metamaterials operating in a frequency region where they are *partially* homogenized and may exhibit new properties.

Toward a covariant effective medium theory. The effective medium theory presented above has one important theoretical shortcoming: the derivation of the macroscopic fields is not manifestly covariant, for two reasons. First, the key quantity in the derivation, the averaging volume f , is not coordinate independent. Second, spatial averaging and time averaging are treated in completely different and incompatible ways. For the sake of theoretical coherence and elegance, it would be desirable to obtain a unified, explicitly covariant averaging procedure. A finite approach to electromagnetism [133, 134, 135, 136, 137, 138] may be a suitable starting point in this direction.

It is remarkable that even in a field as mature as classical electromagnetism, no satisfactory derivation exists connecting the microscopic equations (1) to the macroscopic Maxwell equations (36-39). It may seem “excessively pedantic”

[67] to worry about this *derivation* given that the macroscopic equations have already been amply confirmed by experiment. However, there is a larger question lurking beneath the surface. If modern theoretical physics is incapable of convincingly making the transition from the small to the large in this comparatively simplest of situations⁹, how can it expect to make the far more difficult inverse transition from the large to the small in order to make progress in more challenging settings such as quantum field theory or quantum gravity? Such “pedantic” problems are therefore crucial for tackling the *fundamental* problem [89] of how laws at larger scales emerge from laws at smaller scales.

Acknowledgements

I would like to thank my doctoral advisor, Prof. D. Felbacq for his guidance and support, and my thesis reviewers, Prof. J.-M. Raimond, Prof. J.-P. Vignerot, and Prof. A. Lakhtakia for their constructive comments.

This work was funded in part by the French project NALIM of the ACI (Action Concertée Interdisciplinaire) "Nanosciences-nanotechnologies 2003", by an ATER fellowship of the French Ministry of Education, and by the French ANR (Agence Nationale de la Recherche) project POEM.

References

- [1] Eli Yablonovitch. Inhibited spontaneous emission in solid-state physics and electronics. *Physical Review Letters*, 58(20):2059–2062, 1987.
- [2] Sajeev John. Strong localization of photons in certain disordered dielectric superlattices. *Physical Review Letters*, 58(23):2486–2489, 1987.
- [3] H. G. Park, S. H. Kim, S. H. Kwon, Y. G. Ju, J. K. Yang, J. H. Baek, S. B. Kim, and Y. H. Lee. Electrically driven single-cell photonic crystal laser. *Science*, 305(5689):1444–1447, 2004.
- [4] S. Noda, M. Fujita, and T. Asano. Spontaneous-emission control by photonic crystals and nanocavities. *Nature Photonics*, 1(8):449–458, 2007.
- [5] P. Lalanne, C. Sauvan, and J. P. Hugonin. Photon confinement in photonic crystal nanocavities. *Laser & Photonics Reviews*, 2(6):514–526, December 2008.
- [6] S. H. Kwon, T. Sunner, M. Kamp, and A. Forchel. Ultrahigh-q photonic crystal cavity created by modulating air hole radius of a waveguide. *Optics Express*, 16(7):4605–4614, 2008.

⁹In classical electromagnetism *both* the small and the large are accessible to experimental and numerical scrutiny, which is not the case in more difficult settings.

- [7] T. Tanabe, M. Notomi, E. Kuramochi, A. Shinya, and H. Taniyama. Trapping and delaying photons for one nanosecond in an ultrasmall high-q photonic-crystal nanocavity. *Nature Photonics*, 1(1):49–52, 2007.
- [8] M. Notomi, E. Kuramochi, and T. Tanabe. Large-scale arrays of ultrahigh-q coupled nanocavities. *Nature Photonics*, 2(12):741–747, 2008.
- [9] D. Englund, A. Faraon, I. Fushman, N. Stoltz, P. Petroff, and J. Vuckovic. Controlling cavity reflectivity with a single quantum dot. *Nature*, 450:857–861, 2007.
- [10] K. Aoki, D. Guimard, M. Nishioka, M. Nomura, S. Iwamoto, and Y. Arakawa. Coupling of quantum-dot light emission with a three-dimensional photonic-crystal nanocavity. *Nature Photonics*, 2(11):688–692, November 2008.
- [11] M. Winger, A. Badolato, K. J. Hennessy, E. L. Hu, and A. Imamoglu. Quantum dot spectroscopy using cavity quantum electrodynamics. *Physical Review Letters*, 101(22), 2008.
- [12] A. Badolato, M. Winger, K. J. Hennessy, E. L. Hu, and A. Imamoglu. Cavity qed effects with single quantum dots. *Comptes Rendus Physique*, 9(8):850–856, October 2008.
- [13] S. H. Fan, P. R. Villeneuve, and J. D. Joannopoulos. Channel drop tunneling through localized states. *Physical Review Letters*, 80(5):960–963, February 1998.
- [14] J. M. Lourtioz, H. Benisty, A. Chelnokov, S. David, and S. Olivier. Photonic crystals and the real world of optical telecommunications. *Annales Des Telecommunications-Annals Of Telecommunications*, 58(9-10):1197–1237, September 2003.
- [15] Hideo Kosaka, Takayuki Kawashima, Akihisa Tomita, Masaya Notomi, Toshiaki Tamamura, Takashi Sato, and Shojiro Kawakami. Self-collimating phenomena in photonic crystals. *Applied Physics Letters*, 74(9):1212–1214, 1999.
- [16] H. Kosaka, T. Kawashima, A. Tomita, M. Notomi, T. Tamamura, T. Sato, and S. Kawakami. Superprism phenomena in photonic crystals: Toward microscale lightwave circuits. *Journal Of Lightwave Technology*, 17(11):2032–2038, 1999.
- [17] C. Luo, S. G. Johnson, J. D. Joannopoulos, and J. B. Pendry. Subwavelength imaging in photonic crystals. *Physical Review B*, 68(4):045115, 2003.
- [18] J. M. Lourtioz. Photonic crystals and metamaterials. *Comptes Rendus Physique*, 9(1):4–15, 2008.

- [19] Frederic Zolla, Gilles Renversez, Andre Nicolet, Boris Kuhlmeiy, Sebastien Guenneau, and Didier Felbacq. *Foundations of Photonic Crystal Fibres*. Imperial College Press, 2005.
- [20] T. A. Birks, J. C. Knight, B. J. Mangan, and P. S. Russell. Photonic crystal fibres: An endless variety. *Ieice Transactions On Electronics*, E84C(5):585–592, 2001.
- [21] W. J. Wadsworth, A. Ortigosa-Blanch, J. C. Knight, T. A. Birks, T. P. M. Man, and P. S. Russell. Supercontinuum generation in photonic crystal fibers and optical fiber tapers: a novel light source. *Journal Of The Optical Society Of America B-Optical Physics*, 19(9):2148–2155, 2002.
- [22] P. Russell. Photonic crystal fibers. *Science*, 299(5605):358–362, 2003.
- [23] J. Limpert, T. Schreiber, S. Nolte, H. Zellmer, A. Tunnermann, R. Iliev, F. Lederer, J. Broeng, G. Vienne, A. Petersson, and C. Jakobsen. High-power air-clad large-mode-area photonic crystal fiber laser. *Optics Express*, 11(7):818–823, 2003.
- [24] S. Coen, A. H. L. Chan, R. Leonhardt, J. D. Harvey, J. C. Knight, W. J. Wadsworth, and P. S. J. Russell. White-light supercontinuum generation with 60-ps pump pulses in a photonic crystal fiber. *Optics Letters*, 26(17):1356–1358, 2001.
- [25] J. M. Dudley, G. Genty, and S. Coen. Supercontinuum generation in photonic crystal fiber. *Reviews Of Modern Physics*, 78(4):1135–1184, 2006.
- [26] A. V. Husakou and J. Herrmann. Supercontinuum generation, four-wave mixing, and fission of higher-order solitons in photonic-crystal fibers. *Journal Of The Optical Society Of America B-Optical Physics*, 19(9):2171–2182, 2002.
- [27] J. M. Dudley and J. R. Taylor. Ten years of nonlinear optics in photonic crystal fibre. *Nature Photonics*, 3(2):85–90, February 2009.
- [28] P. Tran. Optical switching with a nonlinear photonic crystal: A numerical study. *Optics Letters*, 21(15):1138–1140, 1996.
- [29] A. A. Sukhorukov and Y. S. Kivshar. Spatial optical solitons in nonlinear photonic crystals. *Physical Review E*, 65(3), 2002.
- [30] D. N. Christodoulides and N. K. Efremidis. Discrete temporal solitons along a chain of nonlinear coupled microcavities embedded in photonic crystals. *Optics Letters*, 27(8):568–570, 2002.
- [31] Y. V. Kartashov, V. A. Vysloukh, and L. Torner. Stable ring-profile vortex solitons in besell optical lattices. *Physical Review Letters*, 94(4), 2005.

- [32] D. Trager, R. Fischer, D. N. Neshev, A. A. Sukhorukov, C. Denz, W. Krolikowski, and Y. S. Kivshar. Nonlinear bloch modes in two-dimensional photonic lattices. *Optics Express*, 14(5):1913–1923, 2006.
- [33] S. John and N. Akozbek. Nonlinear-optical solitary waves in a photonic band-gap. *Physical Review Letters*, 71(8):1168–1171, 1993.
- [34] M. Soljacic, C. Luo, J. D. Joannopoulos, and S. H. Fan. Nonlinear photonic crystal microdevices for optical integration. *Optics Letters*, 28:637–639, 2003.
- [35] M. Notomi, A. Shinya, S. Mitsugi, G. Kira, E. Kuramochi, and T. Tanabe. Optical bistable switching action of si high-q photonic-crystal nanocavities. *Optics Express*, 13(7):2678–2687, 2005.
- [36] S. F. Mingaleev, A. E. Miroshnichenko, Y. S. Kivshar, and K. Busch. All-optical switching, bistability, and slow-light transmission in photonic crystal waveguide-resonator structures. *Physical Review E*, 74, 2006.
- [37] M. Soljacic, S. G. Johnson, S. H. Fan, M. Ibanescu, E. Ippen, and J. D. Joannopoulos. Photonic-crystal slow-light enhancement of nonlinear phase sensitivity. *Journal Of The Optical Society Of America B-Optical Physics*, 19:2052–2059, 2002.
- [38] Y. A. Vlasov, M. O’Boyle, H. F. Hamann, and S. J. McNab. Active control of slow light on a chip with photonic crystal waveguides. *Nature*, 438(7064):65–69, 2005.
- [39] H. Gersen, T. J. Karle, R. J. P. Engelen, W. Bogaerts, J. P. Korterik, N. F. van Hulst, T. F. Krauss, and L. Kuipers. Real-space observation of ultraslow light in photonic crystal waveguides. *Physical Review Letters*, 94(7):073903, 2005.
- [40] T. Baba. Slow light in photonic crystals. *Nature Photonics*, 2(8):465–473, 2008.
- [41] K. Sakoda and K. Ohtaka. Sum-frequency generation in a two-dimensional photonic lattice. *Phys. Rev. B*, 54:5742, 1996.
- [42] J. Martorell, R. Vilaseca, and R. Corbalan. Second harmonic generation in a photonic crystal. *Applied Physics Letters*, 70(6):702–704, 1997.
- [43] V. Berger. Nonlinear photonic crystals. *Physical Review Letters*, 81(19):4136–4139, 1998.
- [44] M. Centini, C. Sibilìa, M. Scalora, G. D’Aguanno, M. Bertolotti, M. J. Bloemer, C. M. Bowden, and I. Nefedov. Dispersive properties of finite, one-dimensional photonic band gap structures: Applications to nonlinear quadratic interactions. *Physical Review E*, 60(4):4891–4898, 1999.

- [45] E. Centeno, D. Felbacq, and D. Cassagne. All-angle phase matching condition and backward second-harmonic localization in nonlinear photonic crystals. *Physical Review Letters*, 98(26), 2007.
- [46] T. Thio, H. F. Ghaemi, H. J. Lezec, P. A. Wolff, and T. W. Ebbesen. Surface-plasmon-enhanced transmission through hole arrays in cr films. *Journal Of The Optical Society Of America B-Optical Physics*, 16(10):1743–1748, 1999.
- [47] E. Popov, M. Neviere, S. Enoch, and R. Reinisch. Theory of light transmission through subwavelength periodic hole arrays. *Physical Review B (Condensed Matter and Materials Physics)*, 62(23):16100–16108, 2000.
- [48] H. J. Lezec, A. Degiron, E. Devaux, R. A. Linke, L. Martin-Moreno, F. J. Garcia-Vidal, and T. W. Ebbesen. Beaming light from a subwavelength aperture. *Science*, 297(5582):820–822, 2002.
- [49] D. Sievenpiper, L. J. Zhang, R. F. J. Broas, N. G. Alexopolous, and E. Yablonovitch. High-impedance electromagnetic surfaces with a forbidden frequency band. *Ieee Transactions On Microwave Theory And Techniques*, 47(11):2059–2074, 1999.
- [50] E. Altewischer, M. P. van Exter, and J. P. Woerdman. Plasmon-assisted transmission of entangled photons. *Nature*, 418(6895):304–306, 2002.
- [51] J. B. Pendry. Negative refraction makes a perfect lens. *Physical Review Letters*, 85(18):3966–3969, 2000.
- [52] R. A. Shelby, D. R. Smith, and S. Schultz. Experimental verification of a negative index of refraction. *Science*, 292(5514):77, 2001.
- [53] J. B. Pendry, A. J. Holden, D. J. Robbins, and W. J. Stewart. Low frequency plasmons in thin-wire structures. *J. of Phys.-Cond. Matt.*, 10(22):4785, 1998.
- [54] J. B. Pendry, A. J. Holden, D. J. Robbins, and W. J. Stewart. Magnetism from conductors and enhanced nonlinear phenomena. *IEEE Transactions on Microwave Theory and Techniques*, 47(11):2075, 1999.
- [55] Stephen O’Brien and John B. Pendry. Photonic band-gap effects and magnetic activity in dielectric composites. *J. Phys.: Condens. Matter*, 14:4035, 2002.
- [56] David R. Smith, David Schurig, Marshall Rosenbluth, Sheldon Schultz, S. Anantha Ramakrishna, and John B. Pendry. Limitations on sub-diffraction imaging with a negative refractive index slab. *Applied Physics Letters*, 82(10):1506–1508, 2003.
- [57] N. Fang, H. Lee, C. Sun, and X. Zhang. Sub-diffraction-limited optical imaging with a silver superlens. *Science*, 308(5721):534–537, April 2005.

- [58] X. S. Rao and C. K. Ong. Amplification of evanescent waves in a lossy left-handed material slab. *Physical Review B (Condensed Matter and Materials Physics)*, 68(11):113103, 2003.
- [59] V. A. Podolskiy and E. E. Narimanov. Near-sighted superlens. *Optics Letters*, 30(1):75–77, 2005.
- [60] Chiyan Luo, Steven G. Johnson, J. D. Joannopoulos, and J. B. Pendry. All-angle negative refraction without negative effective index. *Physical Review B*, 65(20):201104, 2002.
- [61] D. Schurig, J. J. Mock, B. J. Justice, S. A. Cummer, J. B. Pendry, A. F. Starr, and D. R. Smith. Metamaterial electromagnetic cloak at microwave frequencies. *Science*, 314(5801):977–980, November 2006.
- [62] U. Leonhardt. Optical conformal mapping. *Science*, 312(5781):1777–1780, 2006.
- [63] J. B. Pendry, D. Schurig, and D. R. Smith. Controlling electromagnetic fields. *Science*, 312(5781):1780–1782, June 2006.
- [64] U. Leonhardt and T. G. Philbin. General relativity in electrical engineering. *New Journal Of Physics*, 8, 2006.
- [65] F. L. Teixeira. Differential form approach to the analysis of electromagnetic cloaking and masking. *Microwave And Optical Technology Letters*, 49(8):2051–2053, August 2007.
- [66] G. Russakoff. A derivation of the macroscopic maxwell equations. *American Journal of Physics*, 38(10):1188–1195, 1970.
- [67] F. N. H. Robinson. *Macroscopic electromagnetism*. Pergamon Press Ltd., 1973.
- [68] John D. Jackson. *Classical Electrodynamics*. John Wiley and Sons, 1999.
- [69] J.C. Maxwell-Garnett. *Phil. Trans. R. Soc. Lond. A*, 203:385, 1904.
- [70] O. Wiener. *Abh. Math.-Phys. Konigl. Sachs. Ges.*, 32:509, 1912.
- [71] D.A.G. Bruggeman. *Ann. Phys. Lpz.*, 24:636, 1935.
- [72] Z. Hashin and S. Shtrikman. *J. Appl. Phys.*, 33:3125, 1962.
- [73] G. W. Milton. Bounds on the complex dielectric constant of a composite material. *Applied Physics Letters*, 37(3):300–302, 1980.
- [74] L. Tsang and J. A. Kong. Scattering of electromagnetic-waves from random-media with strong permittivity fluctuations. *Radio Science*, 16(3):303–320, 1981.

- [75] A. Sihvola. *Electromagnetic mixing formulas and applications*. IEE Press, 1999.
- [76] A. Sihvola. Six-dimensional view of dielectric mixtures as metamaterials. *The European Physical Journal Applied Physics*, 46(3):-, June 2009.
- [77] D. J. Bergman. Exactly solvable microscopic geometries and rigorous bounds for the complex dielectric-constant of a 2-component composite-material. *Physical Review Letters*, 44(19):1285–1287, 1980.
- [78] A. Bensoussan, J.L. Lions, and G.C. Papanicolaou. *Asymptotic Analysis for Periodic Structures*. North-Holland, Amsterdam, 1978.
- [79] D. Felbacq and G. Bouchitté. Homogenization of a set of parallel fibres. *Waves in Random Media*, 7:245, 1997.
- [80] G. Bouchitté and D. Felbacq. Homogenization of a wire photonic crystal: The case of small volume fraction. *SIAM Journal On Applied Mathematics*, 66(6):2061–2084, 2006.
- [81] W. E. Kock. Metallic delay lenses. *Bell System Technical Journal*, 27(1):58–82, 1948.
- [82] H.A. Lorentz. *Proc. Roy. Acad., Amsterdam*, 254, 1902.
- [83] H.A. Lorentz. *The theory of electrons and its applications to the phenomena of light and radiant heat*. G.E. Stechert and Co., 1916.
- [84] S. R. de Groot. *The Maxwell Equations*, volume IV of *Studies in Statistical Mechanics*. North-Holland, Amsterdam, 1969.
- [85] Robert E. Collin. *Field theory of guided waves*. IEEE Press, 1991.
- [86] J. C. Slater. The electron theory of solids. *American Journal Of Physics*, 19(6):368–374, 1951.
- [87] O.F. Mossotti. *Sur les forces qui régissent la constitution intime des corps*. Torino, 1836.
- [88] J.H. Van Vleck. *The theory of electric and magnetic susceptibilities*. Clarendon Press, 1932. URL: <http://www.archive.org/details/theoryofelectric031070mbp>.
- [89] Robert B. Laughlin. *A different universe - reinventing physics from the bottom down*. Basic Books, 2005.
- [90] R. A. Shelby, D. R. Smith, S. C. Nemat-Nasser, and S. Schultz. Microwave transmission through a two-dimensional, isotropic, left-handed metamaterial. *Applied Physics Letters*, 78(4):489–491, 2001.

- [91] P. Mazur and B. R. A. Nijboer. On the statistical mechanics of matter in an electromagnetic field .1. derivation of the maxwell equations from electron theory. *Physica*, 19(8):971–986, 1953.
- [92] K. Schram. Quantum statistical derivation of the macroscopic maxwell equations. *Physica*, 26:1080–1090, December 1960.
- [93] E. T. Jaynes. Foundations of probability theory and statistical mechanics. In M. Bunge, editor, *Delaware Seminar in the Foundations of Physics*, page 77, 1967.
- [94] E.T. Jaynes. *Foundations of probability theory and statistical mechanics*, chapter 6. D. Reidel Publishing Co., Dordrecht, Holland, 1983.
- [95] P. W. Anderson. More is different. *Science*, 177(4047):393–396, 1972.
- [96] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings Of The National Academy Of Sciences Of The United States Of America-Biological Sciences*, 79(8):2554–2558, 1982.
- [97] S.A. Kauffman. *Origins of Order - Self-Organization and Selection in Evolution*. Oxford University Press, 1993.
- [98] Stephan Mertens. Computational complexity for physicists. *Computing in Science and Engg.*, 4(3):31–47, 2002.
- [99] Allon Percus, Gabriel Istrate, and Christopher Moore, editors. *Computational complexity and statistical physics*. Oxford University Press, 2006.
- [100] James P. Sethna. *Statistical mechanics - entropy, order parameters and complexity*. Clarendon Press, 2008.
- [101] R. O. Dendy. *Plasma Dynamics*. Oxford University Press, 1990.
- [102] Carlo Rovelli. *Quantum Gravity*. Cambridge University Press, 2004.
- [103] Charles Kittel. *Introduction to Solid State Physics*. John Wiley and Sons, 1996.
- [104] Neil W. Ashcroft and N. David Mermin. *Solid State Physics*. Brooks Cole, 1976.
- [105] Panos J. Antsaklis and Anthony N. Michel. *Linear systems*. Springer, 2006.
- [106] A. I. Căbuz. *Electromagnetic metamaterials - From photonic crystals to negative index composites*. PhD thesis, University of Montpellier II, 2007.

- [107] V. V. Yatsenko, S. I. Maslovski, S. A. Tretyakov, S. L. Prosvirnin, and S. Zouhdi. Plane-wave reflection from double arrays of small magneto-electric scatterers. *IEEE Transactions On Antennas And Propagation*, 51(1):2–11, January 2003.
- [108] A. I. Căbuz, D. Felbacq, and D. Cassagne. Spatial dispersion in negative-index composite metamaterials. *Phys. Rev. A*, 77(1):013807, 2008.
- [109] A. N. Grigorenko, A. K. Geim, H. F. Gleeson, Y. Zhang, A. A. Firsov, I. Y. Khrushchev, and J. Petrovic. Nanofabricated media with negative permeability at visible frequencies. *Nature*, 438(7066):335–338, November 2005.
- [110] J. F. Zhou, L. Zhang, G. Tuttle, T. Koschny, and C. M. Soukoulis. Negative index materials using simple short wire pairs. *Physical Review B*, 73(4):041101, January 2006.
- [111] D. R. Smith, D. C. Vier, N. Kroll, and S. Schultz. Direct calculation of permeability and permittivity for a left-handed metamaterial. *Applied Physics Letters*, 77(14):2246–2248, October 2000.
- [112] D. E. Aspnes. Local-field effects and effective-medium theory: A microscopic perspective. *American Journal of Physics*, 50(8):704–709, 1982.
- [113] A. Alu and N. Engheta. Pairing an epsilon-negative slab with a mu-negative slab: Resonance, tunneling and transparency. *IEEE Transactions on antennas and propagation*, 51(10):2558–2571, 2003.
- [114] Tom G. Mackay and Akhlesh Lakhtakia. A limitation of the bruggeman formalism for homogenization. *Optics Communications*, 234(1-6):35–42, April 2004.
- [115] D. Schurig and D. R. Smith. Sub-diffraction imaging with compensating bilayers. *New Journal of Physics*, 7:162, 2005.
- [116] Juan D. Baena, Lukas Jelinek, and Ricardo Marques. Reducing losses and dispersion effects in multilayer metamaterial tunneling devices. *New Journal of Physics*, 7:166, 2005.
- [117] P. A. Belov and Y. Hao. Subwavelength imaging at optical frequencies using a transmission device formed by a periodic layered metal-dielectric structure operating in the canalization regime. *Physical Review B*, 73(11):113110, 2006.
- [118] Jinlong Zhang, Haitao Jiang, Boris Gralak, Stefan Enoch, Gérard Tayeb, and Michel Lequime. Towards -1 effective index with one-dimensional metal-dielectric metamaterial: a quantitative analysis of the role of absorption losses. *Opt. Express*, 15(12):7720–7729, 2007.

- [119] L. Li. Use of fourier series in the analysis of discontinuous periodic structures. *Journal of the Optical Society of America A*, 13(9):1870–1876, 1996.
- [120] R.D. Meade, A.M. Rappe, K.D. Brommer, J.D. Joannopoulos, and O.L. Alerhand. Accurate theoretical analysis of photonic band-gap materials. *Physical Review B*, 48:8434, 1993.
- [121] S. G. Johnson and J. D. Joannopoulos. Block-iterative frequency-domain methods for maxwell’s equations in a planewave basis. *Optics Express*, 8(3):173–190, 2001.
- [122] L. Mandel and E. Wolf. *Optical Coherence and Quantum Optics*, chapter 3, pages 109–127. Cambridge University Press, 1995.
- [123] X. S. Rao and C. K. Ong. Subwavelength imaging by a left-handed material superlens. *Physical Review E*, 68(6), 2003.
- [124] V. G. Veselago. Electrodynamics of substances with simultaneously negative values of sigma and mu. *Soviet Physics USPEKHI-USSR*, 10(4):509, 1968.
- [125] B. Hecht, B. Sick, U. P. Wild, V. Deckert, R. Zenobi, O. J. F. Martin, and D. W. Pohl. Scanning near-field optical microscopy with aperture probes: Fundamentals and applications. *Journal Of Chemical Physics*, 112(18):7761–7774, May 2000.
- [126] F. Kulzer and M. Orrit. Single-molecule optics. *Annual Review Of Physical Chemistry*, 55:585–611, 2004.
- [127] C. Girard. Near fields in nanostructures. *Reports On Progress In Physics*, 68(8):1883–1933, August 2005.
- [128] L. Novotny and S. J. Stranick. Near-field optical microscopy and spectroscopy with pointed probes. *Annual Review Of Physical Chemistry*, 57:303–331, 2006.
- [129] R. Smaâli, D. Felbacq, and G. Granet. Bloch waves and non-propagating modes in photonic crystals. *Physica E*, 18:443, 2003.
- [130] C. Y. Li, J. M. Holt, and A. L. Efros. Imaging by the veselago lens based upon a two-dimensional photonic crystal with a triangular lattice. *Journal Of The Optical Society Of America B-Optical Physics*, 23(5):963–968, May 2006.
- [131] C. Y. Li, J. M. Holt, and A. L. Efros. Far-field imaging by the veselago lens made of a photonic crystal. *Journal Of The Optical Society Of America B-Optical Physics*, 23(3):490–497, March 2006.
- [132] G.V. Eleftheriades and K.G. Balmain, editors. *Negative refraction metamaterials*. IEEE Press, 2005.

- [133] Rafael Sorkin. The electromagnetic field on a simplicial net. *Journal of Mathematical Physics*, 16(12):2432–2440, 1975.
- [134] E. Tonti. On the geometrical structure of electromagnetism. In G. Ferrarese, editor, *Gravitation, electromagnetism and geometrical structures*, volume 281-308. Pitagora Editrice Bologna, 1995.
- [135] E. Tonti. Finite formulation of the electromagnetic field. In *PIER 32*, pages 1–44, 2001.
- [136] A. Bossavit. *Differential geometry for the student of numerical methods in Electromagnetism*. 1991. URL: <http://www.em.tut.fi/~bossavit/Books/DGSNME/DGSNME.pdf>.
- [137] Alain Bossavit. *Computational electromagnetism - variational formulations, complementarity and edge elements*. Academic Press, 1998.
- [138] Anil N. Hirani. *Discrete Exterior Calculus*. PhD thesis, California Institute of Technology, 2003.