The Anatomy of Negative Refraction

Abstract — The mechanism of negative refraction (electromagnetic waves bending the ‘wrong’ way at material interfaces) and the closely related phenomenon of backward waves (phase velocity opposite to group velocity) are examined. It is shown that in periodic dielectric structures, such as metamaterials and photonic crystals, backward waves disappear in the homogenization limit, when the lattice cell size becomes negligible relative to the vacuum wavelength. The paper establishes a lower bound for the cell size of metamaterials capable of supporting backward waves. Introductory material and historical notes are included.

I INTRODUCTION

Negative refraction (electromagnetic waves bending the ‘wrong’ way at material interfaces) and the closely related phenomenon of backward waves (phase velocity opposite to group velocity) have become one of the most intriguing areas of research in nanophotonics this century, with thousands of research papers published and a number of books and review papers readily available: P.W. Milonni [20], G.V. Eleftheriades & K.G. Balmain (eds.) [7], S.A. Ramakrishna [33], J.B. Pendry & D.R. Smith [31], and others. The annual numbers of ISI publications (Fig. 1) illustrate the rapidly rising interest in this subject. Negative refraction being an odd phenomenon, it is perhaps fitting that the production of papers is above the average trend in odd years.

The objective of this paper is three-fold: (i) to clarify the mechanism of negative refraction and backward waves in periodic structures; (ii) to examine whether the ‘metamaterial’ and ‘photonic crystal’ species of negative refraction are fundamentally different; (iii) to show that backward waves disappear in the homogenization limit and to establish a lower bound for the lattice cell size of materials capable of supporting backward waves. Much of the introductory and background material in this paper is drawn from the upcoming book [49]; however, the main results related to the minimum lattice cell size are new.

Over 60 years ago, L.I. Mandelshtam in his lecture notes [17, 18] pointed out that waves would refract negatively at an interface boundary between a regular and a backward-wave medium. In his 1967 paper [50],1 V.G. Veselago showed that waves in materials with simultaneously negative dielectric permittivity $\varepsilon$ and magnetic permeability $\mu$ would exhibit quite unusual behavior of propagation and refraction. More specifically:

- Vectors $\mathbf{E}$, $\mathbf{H}$ and $\mathbf{k}$, in that order, form a left-handed system. Consequently, the Poynting vector $\mathbf{E} \times \mathbf{H}$ and the wave vector $\mathbf{k}$ have opposite directions.

- The Doppler and Vavilov-Cerenkov effects are ‘reversed’. The sign of the Doppler shift in frequency is opposite to what it would be in a regular material. The Poynting vector of the Cerenkov radiation forms an obtuse angle with the direction of motion of a superluminal particle in a medium, while the wave vector of the radiation is directed toward the trajectory of the particle.

... at the interface boundary the tangential components of the fields ... must be continuous. It is easy 2

- Light propagating from a regular medium into a double-negative material bends “the wrong way” (Fig. 2). In Snell’s law, this corresponds to a negative index of refraction.

- A slab with $\varepsilon = -1$, $\mu = -1$ in air acts as an unusual lens (Fig. 3).

Subjects closely related to Veselago’s work had in fact been discussed in the literature well before his seminal publication – as early as in 1904. S.A. Tretyakov [47], C.L. Holloway et al. [9] and A. Moroz3 provide the following references:


- A. Schuster’s monograph [35], pp. 313–318; a 1905 paper by H.C. Pocklington [32].

- L.I. Mandelshtam’s short paper [16] and, even more importantly, his lectures on negative refraction and backward waves [17, 18] (more than two decades prior to Veselago’s work, as already noted in the introductory paragraphs. Mandelshtam’s 1945 paper, but not the lecture notes, is cited by Veselago.)

- A number of papers in Russian technical journals from the 1940s to the 1970s: by D.V. Sivukhin (1957) [40], V.E. Pafomov (1959) [26] and R.A. Silin (1959, 1978) [38, 39]. Silin’s earlier review paper (1972) [39], where he focuses on wave propagation in artificial periodic structures.

In one of his lectures cited above, Mandelshtam writes, in reference to a figure similar to Fig. 2 ([18], pp. 464–465):4

1Published in 1967 in Russian. In the English translation that appeared in 1968, the original Russian paper in mistakenly dated as 1964.

2V.G. Veselago remarks that this is not a lens ‘in the usual sense of the word’ because it does not focus a parallel beam to a point.

3http://www.wave-scattering.com/negative.html

4My translation from the Russian. A similar quote is given by S.A. Tretyakov in [47].
to show that these conditions cannot be satisfied with a reflected wave (or a refracted wave) alone. But with both waves present, the conditions can always be satisfied. From that, by the way, it does not at all follow that there must only be three waves and not more: the boundary conditions do allow one more wave, the fourth one, traveling at the angle \( \pi - \phi_1 \) in the second medium. Usually it is tacitly assumed that this fourth wave does not exist, i.e. it is postulated that only one wave propagates in the second medium.

...[the law of refraction] is satisfied at angle \( \phi_1 \) as well as at \( \pi - \phi_1 \). The wave... corresponding to \( \phi_1 \) moves away from the interface boundary. ... The wave corresponding to \( \pi - \phi_1 \) moves toward the interface boundary. It is considered self-evident that the second wave cannot exist, as light impinges from the first medium onto the second one, and hence in the second medium energy must flow away from the interface boundary. But what does energy have to do with this? The direction of wave propagation is in fact determined by its phase velocity, whereas energy moves with group velocity. Here therefore there is a logical leap that remains unnoticed only because we are accustomed to the coinciding directions of propagation of energy and phase. If these directions do coincide, i.e. if group velocity is positive, then everything comes out correctly. If, however, we are dealing with the case of negative group velocity – quite a realistic case, as I already said, – then everything changes. Requiring as before that energy in the second medium flow away from the interface boundary, we arrive at the conclusion that phase must run toward this boundary and, therefore, the direction of propagation of the refracted wave will be at the \( \pi - \phi_1 \) angle to the normal. However unusual this setup may be, there is, of course, nothing surprising about it, for phase velocity does not tell us anything about the direction of energy flow."

"Figure 10 shows refraction phenomena in a periodic waveguide whose effective index... in the modulated region is... higher than... in the unmodulated region. The grating lines, however, are not normal to the boundaries. As a consequence of the boundary conditions... two Floquet-Bloch waves corresponding to the upper and lower branches of the dispersion contour... are excited simultaneously... resulting generally in two rays propagating in different directions. This ray refraction can be described by two effective ray indices: one for ordinary refraction... and the other... with a negative refraction angle..."

Similar observations are made in Silin's 1972 paper [39]; see also a quote in [49].

A turning point in the research on double-negative materials came in 1999–2000, when J.B. Pendry et al. [30] showed theoretically, and D.R. Smith et al. [45] confirmed experimentally, negative refraction in an artificial material with split-ring resonators. A further discovery was made by Pendry in his 'perfect lens' paper in 2000 [29]. While Veselago’s description of his ‘lens’ relied purely on geometric optics, Pendry’s electromagnetic analysis showed that the evanescent part of light emitted by the source is amplified by the slab, with the ultimate result of perfect transmission and focusing of both propagating and evanescent components of the wave.

The first publication on what today would be called a (quasi)-perfect cylindrical lens was a 1994 paper by N.A. Nicorovici et al. [24] (now more detailed follow-up papers by G.W. Milton et al. [22, 21] are available). These authors considered a coated dielectric cylinder, with the core of radius \( r_{core} \) and permittivity \( \varepsilon_{core} \), the shell (coating) with the outer radius \( r_{shell} \) and permittivity \( \varepsilon_{shell} \) – embedded in a background medium with permittivity \( \varepsilon_{bg} \). It turns out, first, that such a coated cylinder is completely transparent to the outside \( H \)-mode field (the \( H \)-field along the axis of the cylinder) under the quasistatic

\[ \varepsilon = 1 \quad \mu = 1 \]

\[ \varepsilon = -1 \quad \mu = -1 \]

Figure 3: The Veselago slab of a double-negative material acts as an unusual lens. Due to the negative refraction at both surfaces of the slab, a point source \( S \) located at a distance \( a < d \) has a virtual image \( S' \) inside the slab and a real image \( I \) outside. The arrows indicate the direction of the Poynting vector, not the wave vector.

Of the earlier contributions to the subject, a notable one was made by R. Zengerle in his PhD thesis on singly and doubly periodic waveguides in the late 1970s. His journal publication of 1987 [54] contains, among other things, a subsection entitled “Simultaneous positive and negative ray refraction”. Quote:
approximation if \( \epsilon_{\text{core}} = \epsilon_{\text{bg}} = 1, \epsilon_{\text{shell}} \to -1 \). (The limiting case \( \epsilon_{\text{shell}} \to -1 \) should be interpreted as the imaginary part of the permittivity tending to zero, while the real part is fixed at \( -1: \epsilon_{\text{shell}} = -1 - i\epsilon_{\text{shell}}', \epsilon_{\text{shell}}'' \to 0 \).) Second, under these conditions for the dielectric constants, many unusual imaging properties of coated cylinders are observed. For example, a line source placed outside the coated cylinder at a radius \( r_{\text{src}} < \frac{r_{\text{shell}}}{r_{\text{core}}} \) would have an image outside the cylinder, at \( r_{\text{image}} = \frac{r_{\text{shell}}}{r_{\text{core}}}(r_{\text{src}} - r_{\text{core}}) \).

Truly homogeneous materials, in the Veselago sense, are not currently known and could be found in the future only if some new molecular-scale magnetic phenomena are discovered. Consequently, much effort has been devoted to the development of artificial metamaterials capable of supporting backward waves and producing negative refraction. Selected developments are as follows (all numbers approximate): D.R. Smith et al. [45] (copper split-ring resonators (SRR) and wires, 4.85 GHz, 2000); R.A. Shelby et al. [37] (copper SRR and strips, 10 GHz, 2001); C.G. Parazzoli et al. [27] (a stack of SRRs with metal strips, 12.6 GHz, 2003); A.A. Houck et al. [10] (composite wire and SRR prisms, 10 GHz, 2003); D.R. Smith & D.C. Vier [43] (copper SRR and strips, 11 GHz, 2004); V.M. Shalaev et al. [36] (pairs of nanorods, 200 THz, 2006); S. Zhang et al. [55] (nano-fishnet with elliptical voids, 170 THz, 2006); G. Dolling et al. [6, 51] (nano-fishnet with rectangular voids, 210 THz, 380 THz, 2006–07).

In these designs, the cell size as a fraction of the vacuum wavelength varies between \( \sim 0.11–0.42 \). Could further improvements in nanofabrication and design bring the cell size down to a small fraction of the wavelength, thereby approaching the Veselago case of a homogeneous material? This question will become central in Section IX.

Separately from the progress in metamaterials, negative refraction was observed in analyzed in singly and doubly periodic waveguides (R. Zengerle [54], late 1970s – 1980s) and in photonic crystals (M. Notomi [25] in 2000). Since 2000, there have been a number of publications on negative refraction and the associated lensing effects in photonic crystals: C. Luo et al. [15], “all-angle negative refraction” in a bcc lattice of air cubes in a dielectric; E. Cubukcu et al. [3], experimental and theoretical demonstration of negative refraction and superlensing in a 2D photonic crystal in the microwave range; R. Moussa et al. [23], experimental and theoretical study of negative refraction and superlensing in a triangular array of rectangular dielectric bars; V. Yannopapas & A. Moroz [52] and M.S. Wheeler et al. [51], negative refraction in a lattice of polaritonic spheres S. Foteinopoulou & C.M. Soukoulis, a general analysis of negative refraction at the air-crystal interfaces, with Notomi’s example as a specific case (a 2D hexagonal lattice of rods with permittivity 12.96 and the radius of 0.35 lattice size). See also P.V. Parimi et al. [28] (left-handed behavior of the waves in a triangular lattice of cylindrical copper rods). An example due to R. Gajic, R. Meisels et al. [8, 19] is discussed in subsequent Sections.

All these intriguing findings have led to the presumption that there are two species of negative refraction, one occurring in photonic crystals and another one in metamaterials. The latter are viewed as conceptually close to the ideal homogeneous “Veselago medium”. There are, indeed, obvious differences between these two cases in terms of the underlying structure, composition and fabrication (e.g. lossless dielectric spheres vs. lossy metallic inclusions such as split ring resonators, pairs of nanorods, fishnets, etc.) On a more fundamental level, however, all cases can be characterized by a periodically varying complex dielectric function \( \epsilon(r) \), and it is legitimate to ask if there might be any principal difference between metamaterials and crystals.

Importantly, can metamaterials, as a matter of principle, be (arbitrarily) close to an ideal Veselago medium? A key point of this paper is that periodic dielectric structures that support backward waves cannot have arbitrarily small lattice cells; there is a fundamental lower bound on the cell size as a fraction of the free-space wavelength.

The remainder of the paper is organized as follows. First, following V.G. Veselago, Section II examines wave propagation in a (hypothetical) medium with simultaneously negative \( \epsilon \) and \( \mu \). Sections III–V provide background information on the governing equations, as well as on Bloch waves and their properties. The Fourier-space formulation of the electromagnetic problem is given in Section VI. Mandelshtam’s interesting examples of backward waves – one mechanical and one electromagnetic – are presented in Sections VII and VIII. Finally, Section IX poses and answers a new question: are there two separate species of negative refraction?

II FORWARD AND BACKWARD PLANE WAVES IN A HOMOGENEOUS ISOTROPIC MEDIUM

In backward waves, energy and phase propagate in opposite directions. Let us first examine this counterintuitive phenomenon in a hypothetical homogeneous isotropic medium with unusual material parameters (the “Veselago medium”). In subsequent Sections, we shall turn to the analysis of forward and backward Bloch waves in periodic dielectric structures.

Consider the behavior of plane waves in a homogeneous isotropic medium with arbitrary constant complex parameters \( \epsilon \) and \( \mu \) at a given frequency. The only stipulation is that the medium be passive (no generation of energy), which under the \( \exp (+i\omega t) \) phasor convention implies negative imaginary parts \( \epsilon'' \) and \( \mu'' \) of \( \epsilon \) and \( \mu \). It will be helpful to assume that these imaginary parts are strictly negative and to view lossless materials as a limiting case of small losses; \( \epsilon'' \to 0, \mu'' \to 0 \). The goal is to establish conditions for the plane wave to be forward or backward. In the latter case, one has a “Veselago medium”.

Let the plane wave propagate along the \( x \) axis, with \( E = E_y \) and \( H = H_z \). Then we have

\[
E = E_y = E_0 \exp(-ikx) \tag{1}
\]

\[
H = H_z = H_0 \exp(-ikx) \tag{2}
\]

where \( E_0, H_0 \) are some complex amplitudes. It immediately follows from Maxwell’s equations that

\[
H_0 = \frac{k}{\omega \mu} E_0 \tag{3}
\]

\[
k = \frac{\omega \sqrt{\mu \epsilon}}{2} \tag{4}
\]

Which branch of the square root “should” be implied in the formula for the wavenumber? In an unbounded medium, there is complete symmetry between the \( +x \) and \( -x \) directions, and waves corresponding to both branches of the root are equally valid. It is clear, however, that each of the waves is unbounded in one of the directions, which is not physical.

For a more physical picture, it is tacitly assumed that the unbounded growth is truncated: e.g. the medium and the wave occupy only half of the space, where the wave decays. With this

---

6The \( \exp (+i\omega t) \) convention is used for complex phasors.
in mind, let us focus on one of the two waves — say, the one with
a negative imaginary part of \(k\):

\[
k'' < 0 \quad (5)
\]

(The analysis for the other wave is completely analogous.) Splitting
up the real and imaginary exponentials

\[
\exp(-ikx) = \exp(-i(k' + ik'')x) = \exp(k''x) \exp(-ik'x)
\]

we observe that this wave decays in the \(+x\) direction. On physical
grounds, one can argue that energy in this wave must flow in the
\(+x\) direction as well. This can be verified by computing the
time-averaged Poynting vector

\[
P = P_x = \frac{1}{2} \Re E_0 H_0^* = \frac{1}{2} \Re \frac{k}{\omega \mu} |E_0|^2 \quad (6)
\]

To express \(P\) via material parameters, let

\[
\epsilon = |\epsilon| \exp(-i\phi_\epsilon); \quad \mu = |\mu| \exp(-i\phi_\mu); \quad 0 < \phi_\epsilon, \phi_\mu < \pi
\]

Then the square root with a negative imaginary part, consistent
with the wave (5) under consideration, gives

\[
k = \omega \sqrt{\mu/\epsilon} \exp\left(-i \frac{\phi_\epsilon + \phi_\mu}{2}\right) \quad (7)
\]

Ignoring all positive real factors irrelevant to the sign of \(P\) in (6), we
get

\[
\text{sign } P = \text{sign } \Re \frac{k}{\mu} = \text{sign } \cos \frac{\phi_\epsilon - \phi_\mu}{2}
\]

The cosine, however, is always positive, as \(0 < \phi_\epsilon, \phi_\mu < \pi\).
Thus, as expected, \(P_x\) is positive, indicating that energy flows in the
\(+x\) direction indeed.

The type of the wave (forward vs. backward) therefore depends
on the sign of phase velocity \(\omega/k'\) — that is, on the sign of \(k'\). As follows from (7),

\[
\text{sign } k' = \text{sign } \cos \frac{\phi_\epsilon + \phi_\mu}{2}
\]

and the wave is backward if and only if the cosine is negative, or

\[
\phi_\epsilon + \phi_\mu > \pi \quad (8)
\]

An algebraically equivalent criterion can be derived by noting
that the cosine function is monotonically decreasing on \([0, \pi]\) and hence \(\phi_\epsilon > \pi - \phi_\mu\) is equivalent to

\[
\cos \phi_\epsilon < \cos(\pi - \phi_\mu)
\]

or

\[
\cos \phi_\epsilon + \cos \phi_\mu < 0
\]

This coincides with the Depine-Lakhtakia condition [4] for backward waves:

\[
\frac{\epsilon'}{|\epsilon|} + \frac{\mu'}{|\mu|} < 0 \quad (9)
\]

This last expression is invariant with respect to complex con-
jugation and is therefore valid for both phasor conventions
\(\exp(\pm \omega t)\).

Note that the analysis above relies only on Maxwell’s equations
and the definitions of the Poynting vector and phase velocity.
No considerations of causality, so common in the literature
on negative refraction, were needed to establish the backward-
wave conditions (8), (9).

### III Field Equations and Bloch Waves

We shall consider the usual 1D, 2D and 3D renditions of time-
harmonic Maxwell’s equations. At optical frequencies, the
intrinsic permeability of all media can be set to \(\mu_0\) (L.D. Landau
and E.M. Lifshitz [12], §60).\(^7\)

In 1D, the equation for the \(E\) field is

\[
E''(x) + k^2 E(x) = 0, \quad \text{with } k^2 = \omega^2 \mu_0 \epsilon(x) \quad (10)
\]

In periodic structures, with \(\epsilon\) a periodic function of \(x\), a funda-
mental solution of the above equation is the Bloch-Floquet wave of the
form

\[
E(x) = E_{\text{PER}}(x) \exp(-iK_B x); \quad (11)
\]

where \(K_B\) is the Bloch wavenumber and subscript ‘PER’ marks periodic functions with a given spatial period \(a\) (the lattice cell
size).

In 2D, the \(E\)-mode (one-component field \(E = E_x\)) is de-
scribed by the equation

\[
\nabla^2 E + k^2 E = 0, \quad \text{with } k^2 = \omega^2 \mu_0 \epsilon(x,y) \quad (12)
\]

Again, if \(\epsilon\) is a periodic function of coordinates, the fundamental
solutions of the field equation are known to be Bloch waves with a
(yet undetermined) wave vector \(K_B = (K_x, K_y)\):

\[
E(r) = E_{\text{PER}}(r) \exp(-iK_B \cdot r); \quad r \equiv (x,y), \quad (13)
\]

Assuming for simplicity a square lattice cell of size \(a\), sub-
script ‘PER’ implies periodicity with respect to any lattice vector
\((n_xa, n_ya)\) with integer \(n_x, n_y\).

The governing equation for the \(H\)-mode (one-component \(H\)
field) is

\[
\nabla \cdot \epsilon^{-1} \nabla H + \omega^2 \mu_0 H = 0 \quad (14)
\]

and the expression for the Bloch \(H\)-wave is completely analog-
ous to (13).

Finally, the 3D \(E\)-field equation and the corresponding Bloch
wave can be written as

\[
\nabla \times \nabla \times E - k^2 E = 0, \quad (15)
\]

\[
E(r) = E_{\text{PER}}(r) \exp(-iK_B \cdot r); \quad (16)
\]

\(r \equiv (x,y,z), \quad K_B = (K_x, K_y, K_z)\)

The governing equation for the \(H\) field is

\[
\nabla \times \epsilon^{-1} \nabla \times H - \omega^2 \mu_0 H = 0 \quad (17)
\]

### IV Fourier Harmonics of Bloch Waves

In 1D, the periodic factor \(E_{\text{PER}}(x)\) (11) can be expanded into a
Fourier series with coefficients \(e_m\) \((m = 0, \pm 1, \pm 2, \ldots)\)

\[
E(x) = \sum_{m=-\infty}^{\infty} e_m \exp(i\kappa_0 a x) \exp(-iK_B x), \quad \kappa_0 = 2\pi a^{-1} \quad (18)
\]

The Fourier coefficients \(e_m\) are given by the usual integral ex-
pressions

\[
e_m = a^{-1} \int_{a} E_{\text{PER}}(x) \exp(-im\kappa_0 x) \, dx \quad (19)
\]

\(^7\)Artificial magnetism can be created in periodic dielectric structures at optical
frequencies (W. Cai et al. [12], S. Linden et al. [13]). The equivalent ‘meso-
scopic’ permeability may then be different from \(\mu_0\), but the intrinsic micro-
scopic permeability of the materials involved is still \(\mu_0\).
where the integration is over any period of length \( a \).

Similarly, in 2D the Fourier series for \( E_{\text{PER}}(x, y) \) has the form

\[
E_{\text{PER}}(r) = \sum_{m \in \mathbb{Z}^2} \tilde{e}_m \exp(ik_m \cdot r),
\]

\[
k_m = 2\pi a^{-1} m \equiv 2\pi a^{-1} (m_x, m_y)
\]

with integers \( m_x, m_y \). A square lattice is assumed for simplicity; \( \tilde{e}_m \) are the Fourier coefficients. The field is treated as vectorial for generality, even if it happens to have only one Cartesian component. The Bloch wave is obtained by multiplying \( E_{\text{PER}} \) with the Bloch exponential:

\[
E(r) = \sum_{m \in \mathbb{Z}^2} E_m \equiv \sum_{m \in \mathbb{Z}^2} e_m \exp(ik_m \cdot r) \exp(-iK_B \cdot r)
\]

This representation of the Bloch wave as a suite of plane waves \( E_m \) is essential for the analysis and physical interpretation of energy flow, phase velocity and other properties of this wave (B. Lombardet et al. [14]).

For \( \mu = \mu_0 \), the above expression for \( E(r) \) leads, via the Maxwell \( \nabla \times E \) equation, to a similar decomposition of the magnetic field:

\[
H(r) = \sum_{m \in \mathbb{Z}^2} H_m \equiv \sum_{m \in \mathbb{Z}^2} h_m \exp(ik_m \cdot r) \exp(-iK_B \cdot r),
\]

\[
h_m = \frac{k_m}{\omega \mu_0} e_m
\]

It is important to note from the outset, as Lombardet et al. do in [14], that the individual plane-wave components of the Bloch wave do not satisfy Maxwell’s equations in the periodic medium and therefore do not represent physical fields. Only taken together do these Fourier harmonics form a valid field.

It is straightforward to verify that the plane waves in the decomposition are orthogonal functions (in the sense of standard vector \( L_2 \) inner product) over the lattice cell. Hence, by Parseval’s theorem, the time- and cell-averaged Poynting vector < \( P > = < \text{Re} \{ E \times H^* \} > / 2 \) can be represented as a sum of the Poynting vectors for the individual plane waves (B. Lombardet et al. [14]):

\[
< P > = \sum_{m \in \mathbb{Z}^2} P_m; \quad P_m = \frac{k_m}{2\omega \mu_0} |e_m|^2
\]

Next, group velocity \( \partial \omega / \partial k \) is clearly the same for all plane wave components, and hence group velocity for the whole Bloch wave can be defined as

\[
v_g = \frac{\partial \omega}{\partial K_B}
\]

In cases of weak dispersion, it can be shown [53, 49] that this velocity, indeed, approximately represents signal velocity in the periodic medium.

V INTRINSIC AND EXTRINSIC PROPERTIES

For the analysis of anomalous wave propagation and refraction, it is important to distinguish intrinsic and extrinsic characteristics of the wave. Intrinsic properties of the wave refer to its characterization as either forward or backward; that is, whether the Poynting vector and phase velocity (if it can be properly defined) are in the same or opposite directions. (Or, more generally, at an acute or obtuse angle.) Extrinsic properties refer to conditions at the interface of the periodic structure with air or another homogeneous medium. A key point is that refraction at the interface depends not only on the intrinsic characteristics of the wave in the bulk, but also on the way the Bloch wave is excited [14].

This can be explained as follows. Let the \( x \) axis run along the interface boundary between air and a material with an \( \alpha \)-periodic permittivity \( \epsilon(x) \). For simplicity, we assume that \( \epsilon \) does not vary along the normal coordinate \( y \). Such a periodic medium can support Bloch \( E \)-modes of the form

\[
E(r) = \sum_{m=-\infty}^{\infty} e_m \exp(im\kappa_0 x) \exp(-iK_B x) \exp(-iK_y y)
\]

Let the first-Brillouin-zone harmonic \( (m = 0) \) have an appreciable magnitude \( e_0 \), thereby defining phase velocity \( \omega/K_B \) in the \( x \)-direction. For \( K_B > 0 \), this velocity is positive.

But any plane-wave component of the Bloch wave can serve as an ’excitation channel’\(^8\) for this wave, provided that for a certain angle of incidence – it matches the \( x \)-component of the incident wave in the air:

\[
K_B x - \kappa_0 m = k_{x,{\text{air}}}
\]

This will correspond to a propagating, rather than evanescent, wave in the air if \( |K_B x - \kappa_0 m| < k_{x,{\text{air}}} \).

First, suppose that the ‘main’ channel \( (m = 0) \) is used, so that \( K_B x = k_{x,{\text{air}}} \). If the Bloch wave in the material is a forward one, then the \( y \)-components of the Poynting vector \( P_y \) and the wave vector \( K_y \) are both directed away from the interface, and the usual positive refraction occurs. If, however, the wave is backward, then \( K_y \) is directed toward the surface (against the Poynting vector) and it can easily be seen that refraction is negative. This is completely consistent with Mandelshtam’s original explanation.

Exactly the opposite occurs if the Bloch wave is excited through an excitation channel where \( K_B x - \kappa_0 m \) is negative (say, for \( m = 1 \)). The matching condition at the interface then implies that the \( x \)-component of the wave vector in the air is negative in this case. Repeating the argument of the previous paragraph, one discovers that for a forward Bloch wave refraction is now negative, while for a backward wave it is positive.

In summary, refraction properties at the interface are a function of the intrinsic characteristics of the wave in the bulk as well as the excitation channel, with four substantially different combinations possible. This conclusion summarizes the results already available but dispersed in the literature [1, 14, 8].

VI FOURIER-SPACE EQUATIONS

Fourier analysis (Plane Wave Expansion, PWE) is the most common way to analyze and compute the band structure in any number of dimensions. The method is very well known and is described here briefly to make the paper self-contained.

Let us start with the 1D case. For simplicity of exposition, let us assume a lossless nonmagnetic periodic medium, where the electric field \( E = E_g(x) \) is governed by the wave equation (10); \( \epsilon \) is assumed to be an \( \alpha \)-periodic function. We are looking for a solution in the form of the Bloch-Floquet wave (11), with \( E_{\text{PER}}(x) \) and \( K_B \) to be determined.

To Fourier-transform the wave equation (10), one expresses the dielectric permittivity via a Fourier series with coefficients \( \tilde{\epsilon}_m \):

\[
\epsilon(x) = \sum_{m=-\infty}^{\infty} \tilde{\epsilon}_m \exp(im\kappa_0 x)
\]

\(^8\)A lucid term due to B. Lombardet et al. [14].
In Fourier space, multiplication $\epsilon(x)E(x)$ (i.e. multiplication of the Fourier series (18) and (25)) turns into convolution and the problem becomes

$$K^2 \xi = \omega^2 \mu_0 \xi$$

(26)

Here $\xi = (\ldots, e_{-2}, e_{-1}, e_0, e_1, e_2, \ldots)^T$ is the (infinite) column vector of Fourier coefficients of the field; $K$ is an infinite diagonal matrix with the entries $k_m = K_B - \kappa_0 m$, or equivalently

$$K = K_B I - \kappa_0 N,$$

(27)

where $I$ is the identity matrix and

$$N = \begin{pmatrix} \ldots & 2 & \ldots & \ldots & \ldots & \ldots \cr \ldots & -1 & \ldots & \ldots & \ldots & \ldots \cr \ldots & 0 & \ldots & \ldots & \ldots & \ldots \cr \ldots & 1 & \ldots & \ldots & \ldots & \ldots \cr \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \cr \ldots & \ldots & \ldots & 2 & \ldots & \ldots \end{pmatrix}$$

(28)

Finally, matrix $\Xi$ in (26) is composed of the Fourier coefficients of $\epsilon$:

$$\Xi_{ml} = \epsilon_{m-l}$$

(29)

for any row $m$ and column $l (-\infty < m, l < \infty)$.

The infinite-dimensional eigenproblem (26) must in practice be truncated to a finite number of harmonics. The computational trade-off is clear: as the number of harmonics grows, both computational complexity and accuracy increase. This paper does not focus on the computational aspects of the problem but may have implications for the simulation and design of photonic structures.

The 2D case is handled in a similar fashion. The dielectric permittivity $\epsilon(x,y)$ is a periodic function of coordinates and can be expanded into its Fourier series

$$\epsilon = \sum_{m \in \mathbb{Z}^2} \tilde{\epsilon}_m \exp(i k_m \cdot r)$$

(30)

For the Bloch-Floquet $E$-field, the negative of the Laplace operator turns, in the Fourier domain, into multiplication with $|K_B - k_m|^2$. Further, the product $\epsilon E$ of the wave equation turns into convolution, and thus the eigenvalue problem in Fourier space reads:

$$|k_m|^2 \tilde{\epsilon}_m = \omega^2 \mu_0 \sum_{m \in \mathbb{Z}^2} \tilde{\epsilon}_{m-m} \epsilon_m$$

(31)

This is an infinite set of equations for the eigenfrequencies and eigenmodes. As in 1D, for computational purposes, the system would need to be truncated to a finite size, but this matter is only tangential in the present paper. Fourier analysis of two- or three-component vector fields in periodic 2D or 3D structures is conceptually similar but technically more involved and is not included in this paper.

VII EXAMPLE: BACKWARD WAVES IN MANDELSHTAM’S CHAIN OF OSCILLATORS

This example, concerned with mechanical rather than electromagnetic waves, is interesting, intuitive and has historical significance. Mandelshtam’s four-page paper [16] published by his coworkers in 1945 after his death is very succinct, so a more detailed exposition below will hopefully prove useful. An electromagnetic analogy of this mechanical example (an optical grating) is the subject of the following Section.

Consider an infinite 1D chain of masses, with the nearest neighbors separated by an equilibrium distance $d$ and connected by springs with a spring constant $f$. Newton’s equation of motion for the displacement $\xi_n$ of the $n$th mass $m_n$ is

$$\ddot{\xi}(n) = \omega_n^2 \left[ (\xi(n-1) - 2\xi(n) + \xi(n+1)) \right], \quad \omega_n^2 = \frac{f}{m_n}$$

(32)

For brevity, dependence of $\xi$ on time is not explicitly indicated. For waves at a given frequency $\omega$, switching to complex phasors yields

$$\omega^2 (\xi(n) + \omega_n^2 [\xi(n-1) - 2\xi(n) + \xi(n+1)]) = 0$$

(33)

Mandelshtam considers periodic chains of masses, focusing on the case with just two alternating masses, $m_1$ and $m_2$. The discrete analog of the Bloch wave has the form

$$\xi(n) = \xi_{\text{PER}}(n) \exp(-iK_B nd)$$

(34)

$\xi_{\text{PER}}$ is a periodic function of $n$ with the period of two and can hence be represented by a Euclidean vector $\xi \equiv (a, b) \in \mathbb{R}^2$, where $a$ and $b$ are the values of $\xi_{\text{PER}}(n)$ for odd and even $n$, respectively.10

Substituting this discrete Bloch-type wave into the difference equation (33), we obtain

$$\begin{pmatrix} \omega_1^2 \lambda^2 + 1 \\
\lambda \omega_2^2 - 2 \omega_1 \omega_2 \end{pmatrix} \begin{pmatrix} a \\
b \end{pmatrix} = 0,$$

(35)

$$\lambda \equiv \exp(-iK_B d)$$

Hence $(a, b)$ is the null vector of the $2 \times 2$ matrix in the left hand side of (35). Equating the determinant to zero yields two eigenfrequencies $\omega_{B_1,B_2}$ of the Bloch wave

$$\omega_{B_1,B_2} = \omega_1^2 + \omega_2^2 \pm \lambda^{-1} \sqrt{(\omega_1^2 \lambda^2 + \omega_2^2)(\omega_1^2 \lambda^2 + \omega_2^2)}$$

To analyze group velocity of Bloch waves, compute the Taylor expansion of these eigenfrequencies around $K_B = 0$ (keeping in mind that $\lambda = \exp(-iK_B d)$):

$$\omega_{B_1} = 2 \frac{\partial \omega_{B_1}^2}{\partial K_B} K_B^2$$

and

$$\omega_{B_2} = 2(\omega_1^2 + \omega_2^2) - 2 \frac{\partial \omega_{B_1}^2}{\partial K_B} K_B^2$$

which coincides with Mandelshtam’s formulas at the bottom of p. 476 of his paper. Group velocity $v_g = \partial \omega / \partial K_B$ of long-wavelength Bloch waves is positive for the ‘acoustic’ branch $\omega_{B_1}$ but negative for the ‘optical’ branch $\omega_{B_2}$.

For $K_B = 0$ (i.e. $\lambda = 1$), simple algebra shows that the components of the second null vector $(a_{B_2}, b_{B_2})$ of (35) are proportional to the two particle masses:

$$\frac{a_{B_2}}{b_{B_2}} = -\frac{m_2}{m_1}$$

(36)

$\xi_{\text{PER}}(n) = \xi(0) + (0)n = \xi(0) + (-1)^n \tilde{\xi}(1)$

where $\tilde{\xi}(1) = \frac{1}{2} (\xi(0) + (1))$: $\tilde{\xi}(1) = \frac{1}{2} (\xi(0) - \xi(1))$

10Alternatively and equally well, $\xi_{\text{PER}}$ can be represented via its two-term Fourier sum, familiar from discrete-time signal analysis:

$$\xi_{\text{PER}}(n) = \tilde{\xi}(0) + \tilde{\xi}(1) \exp(in\pi) = \tilde{\xi}(0) + (-1)^n \tilde{\xi}(1)$$

11On the acoustic branch, by definition, $\omega \to 0$ as $K_B \to 0$; on optical branches, $\omega \to 0$. 

---

9The paper is also reprinted in Mandelshtam’s lecture course [17].
The null vector $a_{B1} = b_{B1}$ corresponding to the zero eigenfrequency for zero $K_B$ represents just a translation of the chain as a whole and is uninteresting.)

Next, consider energy transfer along the chain. The force that mass $n - 1$ exerts upon mass $n$ is

$$F_{n-1,n} = \left[\xi(n-1) - \xi(n)\right]f$$

The mechanical ‘Poynting vector’ is the power generated by this force:

$$P_{n-1,n}(t) = F_{n-1,n}(t)\xi(n,t)$$

the time average of which, via complex phasors, is

$$\langle P_{n-1,n} \rangle = \frac{1}{2} \text{Re}\{F_{n-1,n}\omega\xi(n)\}$$

For the ‘optical’ mode, i.e. the second eigenfrequency of oscillations, direct computation leads to Mandelshtam’s expression

$$\langle P \rangle = \frac{1}{2} \int \omega \text{absin}(K_Bd)$$

The subscripts for $\langle P \rangle$ have been dropped because the result is independent of $n$, as should be expected from physical considerations: no continuous energy accumulation occurs in any part of the chain.

We have now arrived at the principal point in this example. For small positive $K_B$ ($K_Bd \ll 1$), the Bloch wave has a long-wavelength component $\exp(-iK_Bnd)$. Phase velocity $\omega/K_B$ of the Bloch wave — in the sense discussed in more detail below — is positive. At the same time, the Poynting vector, and hence the group velocity, are negative because $a_{B12}$ and $b_{B12}$ have opposite signs in accordance with (36). Thus mechanical oscillations of the chain in this case propagate as a backward wave. An electromagnetic analogy of such a wave is mentioned very briefly in Mandelshtam’s paper and is the subject of the following Section.

VIII: EXAMPLE: BACKWARD WAVES IN MANDELSHTAM’S GRATING

Let us assume a lossless nonmagnetic medium with 1D-periodic permittivity $\epsilon(x)$. The electric field $E = E_0(x)$ is governed by the wave equation (10). We are looking for a solution in the form of the Bloch-Floquet wave (11). Both $E_{\text{PER}}(x)$ and $K_B$ are a priori unknown and need to be determined. In Fourier space, $E_{\text{PER}}(x)$ is given by its Fourier series (18) and $\epsilon$ is expressed via the Fourier series (25).

Mandelshtam’s 1D volume grating mentioned in his paper [16] can, similarly to the chain of mechanical oscillators, support backward waves. For definiteness, let us assume a unit lattice cell and a sinusoidal variation of the permittivity:

$$\epsilon(x) = 2 + \cos 2\pi x = 2 + \frac{1}{2} \exp(i2\pi x) + \frac{1}{2} \exp(-i2\pi x)$$

Thus $\epsilon$ has only three nonzero Fourier coefficients: $\epsilon_{\pm 1} = 1/2$, $\epsilon_{0} = 0$.

The eigenvalue problem (26), with the magnetic permeability normalized to unity for simplicity, is

$$K^2\xi = \omega^2\Xi\xi$$

The diagonal matrix $K^2$ has entries

$$K^2_m = (K_B - 2\pi m)^2, \quad m = 0, \pm 1, \pm 2, \ldots$$

and matrix $\Xi$ is tridiagonal, with the entries in the $m$th row equal to

$$\Xi_{m,m} = \epsilon_0 = 2; \quad \Xi_{m\pm 1,m} = \epsilon_{\pm 1} = \frac{1}{2}$$

For any given value of the Bloch parameter $K_B$, numerical solution can be obtained by truncating the infinite system to the algebraic eigenvalue problem with $2M + 1$ equations ($m = -M, -M + 1, \ldots, M - 1, M$).

The first four dispersion curves $\omega(K_B)$ are shown in Fig. 4; there are two frequency bandgaps in the figure, approximately [1.98, 2.55] and [4.40, 4.68], and infinitely many more gaps beyond the range of the chart. The numerical results are plotted for 41 equally spaced values of the normalized Bloch number $K_B/\pi$ in $[-1,1]$. There is no appreciable difference between the numerical results for $M = 5$ (11 equations) and $M = 20$ (41 equations). The high accuracy of the eigenfrequencies for a small number of plane waves in the expansion is due to the smooth variation of the permittivity. Discontinuities in $\epsilon$ would require a much higher number of harmonics.

In addition to the eigenvalues $\omega^2$ of (37), the eigenvectors $\xi$ are also of interest. As an example, let us set $K_B = \pi/10$. Stem plots of the four eigenvectors corresponding to the four smallest eigenvalues $\omega^2 \approx 0.049, 18.29, 23.12$ and 77.83, are shown in Fig. 5. The first Bloch wave is almost a plane wave; the amplitudes of all harmonics other than $\epsilon_0$ are very small (but not zero, as it might appear from the figure); for example, $e_{-1} \approx 0.00057, e_1 \approx 0.00069$.

It is interesting to note that dispersion curves with positive and negative slopes $\partial\omega/\partial K_B$ (i.e. positive and negative group velocity) alternate in the diagram. Group velocity is positive for the lowest-frequency curve $\omega_1(K_B)$, negative for $\omega_2(K_B)$, positive again for $\omega_3(K_B)$, etc.

We now discuss the splitting of the Poynting vector into the individual ‘Poynting components’ $P_m = k_m|\xi_m|^2/(2\omega^2)$ (23); this splitting has implications for the nature of the wave. The distribution of $P_m$ for the first four Bloch modes in the grating is displayed in Fig. 6. The first mode shown in the upper left corner of the figure is almost a pure plane wave ($P_{\pm 1}$ are on the order of $10^{-5}$; $P_{\pm 2}$ are on the order of $10^{-13}$, and so on) and does not exhibit any unusual behavior.

Let us therefore focus on mode #2 (upper right corner of the figure). There are four non-negligible harmonics altogether. The stems to the right of the origin ($K > 0$) correspond to plane wave

\[ P_m = \frac{k_m |\xi_m|^2}{2\omega^2} \]
phase velocity is worth considering. However, two aspects of the notion of phase velocity are worth considering.

First, boundary conditions at the interface between two homogeneous media are intimately connected with the values of phase velocities and indexes of refraction (defined for homogeneous materials in the usual unambiguous sense). Fundamentally, however, it is the wave vectors in both media that govern wave propagation, and it is the continuity of its tangential component that constrains the fields. Phase velocity plays a role only due to its direct connection with the wavenumber. For periodic structures, there is not one but a whole ‘comb’ of wavenumbers that all need to be matched at the interface. We shall return to this subject later.

Second, in many practical cases phase velocity can be easily and clearly visualized. As an example, Fig. 7 shows two snapshots, at \( t = 0 \) and \( t = 0.5 \), of the second Bloch mode described above. For the visual clarity of this figure, low-pass filtering has been applied – without that filtering, the rightward motion of the wave is obvious in the animation but is difficult to present in static pictures. The Bloch wavenumber in the first Brillouin zone in this example is \( K_B = 0.1 \pi \) and the corresponding second eigenfrequency is \( \omega \approx 4.276 \). The phase velocity – if defined via the first Brillouin zone wavenumber – is \( v_{ph} = \omega / K_B \approx 4.276 / 0.1 \pi \approx 13.61 \). Over the time interval \( t = 0.5 \) between the snapshots, the displacement of the wave consistent with this phase velocity is \( 13.61 \cdot 0.5 \approx 6.8 \). This corresponds quite accurately to the actual displacement in Fig. 7, proving that the first Brillouin zone wavenumber is indeed relevant to the perceived visual motion of the Bloch wave.

So, what is one to make of all this? The complete representation of a Bloch wave is given by a comb of wavenumbers \( K_B - m \kappa_0 \) and the respective amplitudes \( c_m \) of the Fourier harmonics. Naturally, one is inclined to distill this theoretically infinite set of data to just a few parameters that include the Poynting vector, phase and group velocities. While the Poynting vector and group velocity for the wave are rigorously and unambiguously defined, the same is in general not true for phase velocity.

However, there are practical cases where phase velocity is meaningful. The situation is most clear-cut when the Bloch wave has a strongly dominant long-wavelength component. Then the Bloch wave is, in a sense, close to a pure plane wave, but non-trivial effects may still arise. Even though the amplitudes of the components propagating to the right, i.e. in the +\( x \)-direction. Stems to the left of the origin correspond to plane waves propagating to the left, and hence their Poynting values are negative. It is obvious from the figure that the negative components dominate and as a result power flows in the negative-\( x \) direction. The numerical values of the Poynting components and of the amplitudes of the plane wave harmonics are summarized in Table 1.

Now, the characterization of this wave as forward or backward hinges on the definition and sign of phase velocity. The smallest absolute value of the wavenumber in the Bloch ‘comb’ \( K_B = 0.1 \pi \) determines the plane wave component with the longest wavelength (bold numbers in Table 1). If one defines phase velocity \( v_{ph} = \omega / K_B \) based on \( K_B = 0.1 \pi \), then phase velocity is positive and, since the Poynting vector was found to be negative, one has a backward wave.

However, the amplitude of the \( K_B = 0.1 \pi \) harmonic (\( \kappa_0 \approx 0.174 \)) is much smaller than that of the \( K_B - \kappa_0 = \approx -1.9 \pi \) wave (sans-serif typeface in the Table). A common convention (P. Yeh [53], B. Lombardet et al. [14]) is to use this highest-amplitude component as a basis for defining phase velocity. If this convention is accepted in our present example, then phase velocity becomes negative and the wave is a forward one (since the Poynting vector is also negative).

One may then wonder what the value of phase velocity “really” is. This question is not a mathematically sound one, as one cannot truly argue about mathematical definitions. From the physical viewpoint, however, two aspects of the notion of phase velocity are worth considering.

![Figure 5: The amplitudes of the plane wave components of the first four Bloch waves for the volume grating with \( \epsilon(x) = 2 + \cos 2 \pi x \). Solution with 41 plane waves. \( K_B a = \pi/10 \).](image)

<table>
<thead>
<tr>
<th>( K/\pi )</th>
<th>( c_m )</th>
<th>( P_m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-5.9</td>
<td>-0.0023</td>
<td>-1.79 \cdot 10^{-5}</td>
</tr>
<tr>
<td>-3.9</td>
<td>-0.0765</td>
<td>-0.013</td>
</tr>
<tr>
<td>-1.9</td>
<td>-0.948</td>
<td>-0.997</td>
</tr>
<tr>
<td>0.1</td>
<td>0.174</td>
<td>0.00177</td>
</tr>
<tr>
<td>2.1</td>
<td>0.253</td>
<td>0.0783</td>
</tr>
<tr>
<td>4.1</td>
<td>0.0179</td>
<td>0.000767</td>
</tr>
<tr>
<td>6.1</td>
<td>0.000495</td>
<td>8.73 \cdot 10^{-7}</td>
</tr>
</tbody>
</table>

Table 1: The principal components of the second Bloch mode in the grating.

![Figure 6: The Poynting components \( P_m \) of the first four Bloch waves for the volume grating with \( \epsilon(x) = 2 + \cos 2 \pi x \). Solution with 41 plane waves. \( K_B a = \pi/10 \).](image)
individual higher-order harmonics may be small, it is possible for their collective effect to be significant. In particular, as the example in this Section has shown, the higher harmonics taken together may carry more energy than the dominant component, and in the opposite direction. In this case one has a backward wave, where phase velocity is defined by the dominant long-wavelength harmonic, while the Poynting vector is due to a collective contribution of all harmonics.

Here is a modest proposal for defining phase velocity in a more general situation. The rationale is that the motion of the wave as a whole is governed primarily by long-wavelength components; short-wavelength harmonics appear as high-frequency noise on top of a waveform with low spatial frequency. It seems appropriate, therefore, to define the ‘weighted phase velocity’ as

$$v_{\text{ph, weighted}} = \frac{\omega}{K_B} \sum_{m=-\infty}^{\infty} (K_B - m\kappa_0)^{-1} |e_m|^2, \quad \|e\|_2 = 1$$

where the $m = 0$ term is omitted if $K_B = 0$. The weighting factors $(K_B - m\kappa_0)^{-1}$ correspond to integral averaging of the wave over the spatial period. This definition may have mathematical advantages as well, as it can be written in terms of matrix-vector operations involving the diagonal matrix $K$ of (27).

An alternative generalization of phase velocity in 1D is the velocity $v_{\text{field}}$ of points with a fixed magnitude of the $E$ field. From the zero differential

$$dE = \frac{\partial E}{\partial x} dx + \frac{\partial E}{\partial t} dt = 0$$

one obtains

$$v_{\text{field}} = \frac{dx}{dt} = -\frac{\partial E}{\partial x} / \frac{\partial E}{\partial t}$$

Unfortunately, this definition does not generalize easily to 2D and 3D, where an analogous velocity would be a tensor quantity (a separate velocity vector for each Cartesian component of the field).

For illustration and further investigation, it is convenient to have a specific example in mind (however, the analysis and conclusions will be general). As such an example, consider the structure proposed by R. Gajic, R. Meisels et al. [8, 19]. Their photonic crystal is a 2D square lattice of alumina rods ($\epsilon_{\text{rod}} = 9.6$) in air. The radius of the rod is $r_{\text{rod}} = 0.61$ mm, the lattice constant $a = 1.86$ mm, so that $r/a \approx 0.33$. The length of the rods is 50 mm. Gajic, Meisels et al. study various cases of wave propagation and refraction. In the context of our analysis, of most interest is negative refraction for small Bloch numbers in the second band of the $H$-mode (TE-mode). The band diagram, computed using the plane wave method with 441 waves for the $H$-mode, appears in Fig. 8 and, apart from the scaling factors, is very close to the one provided by Gajic et al.

The TE2 dispersion curve is mildly convex around the $\Gamma$ point ($K_B = 0, \omega a/(2\pi c) \approx 0.427$), indicating a negative second derivative $\partial^2 \omega / \partial K_B^2$ and hence a negative group velocity for small positive $K_B$ and a possible backward wave. As we are now aware, an additional condition for a backward wave must also be satisfied: the plane-wave component corresponding to the small positive Bloch number must be appreciable (or better yet, dominant). Let us therefore consider the plane wave composition of the Bloch wave.

The amplitudes of the plane-wave harmonics for the Gajic et al. crystal are shown in Fig. 9. For $K_B = 0$ (i.e. at $\Gamma$) the spectrum is symmetric and characteristic of a standing wave. As $K_B$ becomes positive and increases, the spectrum gets skewed, with the backward components ($K < 0$) increasing and the forward ones decreasing. The amplitudes of the spatial harmonics of this Bloch wave in the first Brillouin zone are quite small. It is therefore debatable whether a valid phase velocity can be attributed to this wave.

The distribution of Poynting components of the same wave and for the same set of values of the Bloch wavenumber is shown in Fig. 10. It is clear from the figure that the negative components outweigh the positive ones, so power flows in the negative direction.

**IX ARE THERE TWO SPECIES OF NEGATIVE REFRACTION?**

The following argument indicates that negative refraction disappears in the homogenization limit when the size of the lattice cells tends to zero, provided that other physical parameters, including frequency, are fixed.

It is natural to normalize the coordinates as $\tilde{x} = x/a, \tilde{y} = y/a, \tilde{z} = z/a$, so that in the tilde-coordinates the 2D / 3D problem is set up in the unit square / cube and the governing equations...
become, in 2D and 3D, respectively,

$$\hat{\nabla}^2 \tilde{E} + \tilde{\omega}^2 \tilde{\mu}_r \epsilon_r \tilde{E} = 0, \quad (39)$$

$$\tilde{\omega} = \frac{\omega_0}{c} = k_0 a = 2\pi \frac{a}{\lambda_0}$$

$$\hat{\nabla} \times \hat{\nabla} \times \tilde{E} = \tilde{\omega}^2 \tilde{\mu}_0 \epsilon_r \tilde{E} \quad (40)$$

$$\hat{\nabla} \times \gamma(r) \hat{\nabla} \times \tilde{H} = \tilde{\omega}^2 \tilde{\mu}_0 \tilde{H} \quad (41)$$

where $c$, $k_0$, $\lambda_0$ are the speed of light, the wavenumber and the wavelength in a vacuum. The relative permeability and permittivity $\mu_r$ and $\epsilon_r$ are periodic functions of coordinates. The homogenization limit is obtained by considering the small cell size – long wavelength limit $a \to 0$, $\tilde{K}_B = K_B a \to 0$. As these limits are taken, the problem and the dispersion curves in the normalized coordinates remain unchanged, and the reference point $\left(\tilde{\omega}, \tilde{K}\right)$ approaches the origin along a fixed dispersion curve – the acoustic branch. In the example of Fig. 8, the acoustic branch is clearly identifiable as two approximately straight lines originating from the $\Gamma$ point.

Around the $\Gamma$ point, phase velocity in any given direction $\tilde{l}$, $\omega/K_1 = \tilde{\omega}/K_1$ is well defined and equal to group velocity $\partial \omega/\partial K_1$ simply by definition of the derivative. No backward waves can be supported in this regime. This argument suggests that the relative cell size $a/\lambda_0$ of a periodic structure with backward waves must lie above a certain threshold. To the best of my knowledge, this question is currently open and has not even been posed explicitly in the literature.

The existence of a lower bound on the cell size for photonic crystals is unsurprising, as they are expected to have features not much smaller than the wavelength in any event. For metamaterials, the situation is more interesting, and the cell size limit has a physical explanation. Indeed, the principal component of artificial metamaterials with negative refraction is some resonating element [45, 43, 33, 36], and the resonance frequency is, roughly, inverse proportional to size (S. Linden et al. [13], S. Tretyakov [48]). As the size of the lattice cell diminishes, the operating frequency increases, so that it is not the absolute frequency $\omega$ but the normalized quantity $\tilde{\omega}$ that remains (approximately) constant. If the absolute frequency were to be fixed as the cell size decreases, the resonance would peter out and the backward waves would disappear, in accord with the general argument above.

The first step toward a more rigorous analysis is a working definition of negative refraction in a metamaterial, as opposed to a photonic crystal. The rationale for the following set of conditions is to make the Bloch wave in a metamaterial resemble, to the extent possible, a plane wave in an ideal “Veselago medium”:

1. The Bloch wave in the metamaterial is, in some sense, close to a long-wavelength plane wave.

2. In particular, the plane-wave component in the first Brillouin zone has an appreciable magnitude. It is this component that defines the phase velocity of the Bloch wave.

3. Other plane-wave components collectively produce energy flow at an obtuse angle with phase velocity.

4. At the air-material interface, it is the long-wavelength, first-Brillouin-zone, component that serves as the excitation channel for the Bloch wave.

By this definition, the metamaterial regime corresponds to small values of $\tilde{K}_B$ ($\tilde{K}_B = K_B a \ll \pi$). Larger values of $\tilde{K}_B$ indicate a photonic crystal or grating regime, where the lattice size is comparable with the wavelength in a vacuum. Our mathematical analysis is therefore focused on the limiting case $\tilde{K}_B \to 0$. Further, only the case of small losses at the operating frequency is of practical interest; large losses would quench all propagating modes.\(^{13}\)

In the absence of losses, $\epsilon$ is real and may be positive or negative. The estimates below depend on the average value of $\epsilon$ over the lattice cell, or equivalently on the Fourier coefficient $\tilde{\epsilon}_0$. It will be assumed that this average value is not close to zero;

\(^{13}\)M. Stockman [46] has recently shown that for negative-index media large losses close to the operating frequency are unavoidable as a matter of principle; this follows from causality relations. Stockman’s analysis is valid for homogeneous media, where the effective $\epsilon$ is rigorously defined, and does not apply to photonic crystals.
The eigenvalue or in matrix-vector form Schur-complement matrix problem algebraically eliminate \( \tilde{\omega} \) for

where symbol ‘\( \sim \)’ indicates that the zeroth entry has been removed; e.g. \( \mathbf{e}_{\sim0} \equiv (\ldots, -2, -1, 1, 2, \ldots, \) etc. The entries of the Schur-modified matrix \( \tilde{\mathbf{E}} \) are

The generalized eigenvalue problem \( \tilde{\mathbf{E}} \mathbf{e} = \eta \mathbf{K} \mathbf{e} \) is equivalent to the regular one with the matrix \( \tilde{\mathbf{E}} = \mathbf{K}^{-1} \mathbf{E} \mathbf{K}^{-1} \); i.e.,

In the ‘metamaterial regime’ \( K_B \ll 1 \),

The eigenvalue \( \eta \) is bounded by the Frobenius norm of \( \tilde{\mathbf{E}} \):
The focus again is on the ‘metamaterial regime’ with a small $K_B$. For $K_B = 0$, $\tilde{\omega}^2 = 0$ is an eigenvalue, with the corresponding constant-field eigenmode. As in 1D, for nonzero $\tilde{\omega}$, the equation for $n = 0$ imposes an algebraic constraint on $\epsilon_m$ ($m \neq 0$). The Schur-complement matrix problem has the same generic form as before:

$$\tilde{\mathcal{E}}_{\omega} = \eta K_{\omega}^2 \tilde{\mathcal{E}}_{\omega}$$  \hfill (51)

where the entries of $\tilde{\mathcal{E}}$ are

$$\tilde{\epsilon}_{nm} = \epsilon_{nm} - \tilde{\epsilon}_0^{-1} \epsilon_{n0} \epsilon_{m0} = \epsilon_{nm} - \tilde{\epsilon}_0^{-1} \epsilon_{n-m}$$

Here some sequential numbering scheme for the integer vectors $\epsilon$ where

$$\tilde{\epsilon}_{\omega}$$

where the entries of $\tilde{\epsilon}_r$ of cylindrical rods of radius $r$ correspond to the row / column of the matrix corresponding to $n, m$.

As in 1D, the generalized eigenvalue problem (51) is equivalent to the regular one with the matrix $\tilde{\mathcal{E}} = K_{\omega}^2 \tilde{\mathcal{E}}_{\omega} K_{\omega}^{-1}$. For $K_B = 0$,

$$\tilde{\mathcal{E}}_{\omega} = \epsilon_{nm} - \tilde{\epsilon}_0^{-1} \epsilon_{n0} \epsilon_{m0} = \frac{1}{(4\pi^2 n \omega)^2}$$

where $n = |n| \neq 0, m = |m| \neq 0$ (by definition, the Schur-modified matrix does not contain the ‘zeroth’ row and column). The Frobenius norm of $\tilde{\mathcal{E}}$ again provides a bound for the eigenvalue $\eta$:

$$|\eta|^2 \leq \|\tilde{\mathcal{E}}\|_F^2 = \sum_{m, n \in \mathbb{Z}^2} |\tilde{\epsilon}_{nm} - \tilde{\epsilon}_0^{-1} \epsilon_{n-m}|^2$$  \hfill (52)

The asterisk indicates that the terms with a zero denominator are omitted. To evaluate the double sum, one needs a more accurate estimate of the Fourier coefficients than in 1D. Let

$$|\bar{\mathcal{E}}_{\omega}| = |\int_{[0,1] \times [0,1]} \epsilon(\tilde{r}) \exp(i2\pi n \cdot \tilde{r}) d\Omega| \leq \frac{g}{n}$$  \hfill (53)

where $g$ is independent of $n$ but does depend on the distribution of $\epsilon$ in the lattice cell. For example, integration by parts yields

$$g \leq \frac{\|\tilde{\mathcal{E}}\|_{L1}}{2\pi}, \quad \text{where} \quad \|\tilde{\mathcal{E}}\|_{L1} \equiv \int_{[0,1] \times [0,1]} |\tilde{\mathcal{E}}| d\Omega$$

Here $\tilde{\mathcal{E}}$ is understood in the sense of distributions and includes surface delta-functions corresponding to the jumps of $\epsilon$ (if any).

Substituting the upper bound (53) for the Fourier coefficients into the Frobenius norm estimate (52) yields, after some algebraic manipulation,

$$|\eta|^2 \leq \|\tilde{\mathcal{E}}\|_F^2 \leq \frac{1}{(4\pi^2)^2} \left[ g^2 \sum_{m = 0}^{n \omega} + \tilde{\epsilon}_0^{-1} g^2 \sum_{m = 0}^{1/\omega} + \tilde{\epsilon}_0^{-1} g^2 \sum_{m = 0}^{1/\omega} \right]$$  \hfill (54)

where $s_{1-3}$ are the following sums:

$$s_1 = \sum_{m, n \in \mathbb{Z}^2} \frac{1}{n^2 |m|^2} \approx 34.8$$

$$s_2 = \sum_{m, n \in \mathbb{Z}^2} \frac{1}{n^4 |m|^2} = \frac{1}{s_1} \approx 27.2$$

$$s_3 = \sum_{n \in \mathbb{Z}^2} \frac{1}{n} \approx 5.2$$

As before, the normalized lattice cell size $a/\lambda_0 = \tilde{\omega}/2\pi$, with $\tilde{\omega} = 1/\eta$.

As an illustrative example, let us apply this result to a lattice of cylindrical rods of radius $r_{rod}$ and permittivity $\epsilon_{rod}$. A simple explicit expression for the Fourier amplitudes $\tilde{\epsilon}$ in this case is available (see e.g. K. Sakoda [34], p. 25), and so parameter $g$ can also be evaluated explicitly:

$$\bar{\epsilon}_n = 2v(\epsilon_{rod} - 1) \frac{J_1(nr_{rod})}{nr_{rod}}; \quad v = \frac{\pi r_{rod}^2}{a^2}, \quad \tilde{r}_{rod} = \frac{2\pi r_{rod}}{a}$$

where $J_1$ is the Bessel function, $v$ is the fraction of the volume occupied by the cylinder, and $\tilde{r}_{rod}$ is its normalized radius. The Fourier coefficients decay even faster than our minimum requirement of $\sim 1/n$ because $J_1$ decays, asymptotically in inverse proportion with the square root of $n$. Estimating $J_1$ very roughly just by its maximum value of $\sim 0.6$, we get

$$g \lesssim 2v(\epsilon_{rod} - 1) \frac{0.6}{\tilde{r}_{rod}} = \frac{0.6\epsilon_{rod}(\epsilon_{rod} - 1)}{a} \leq 0.6|\epsilon_{rod} - 1|$$  \hfill (55)

For a numerical illustration, let us evaluate the lower bound for the lattice cell in the Gajic et al. crystal presented earlier. For $\epsilon_{rod} = 9.6$, equation (55) gives $g \approx 5.16$. Since the volume fraction is $v \approx 0.338$, the average value of $\epsilon$ is $\epsilon_0 \approx 3.91$, and from (54) $\eta \approx 1.71$. Then $\tilde{\omega}^2 = 1/\eta \geq 0.87$ and the normalized cell size $a/\lambda_0 = \tilde{\omega}/2\pi \geq 0.14$. In reality, negative refraction in this crystal is observed above ~68 GHz, which corresponds to $\lambda_0 \leq 4.4$ mm and $a/\lambda_0 \geq 0.42$, comfortably above the theoretical lower bound of 0.14.

If rods were made of an (ideal) plasmonic material with $\epsilon = -2$ (losses neglected), the same calculation would produce a smaller bound on the cell size, $a/\lambda_0 \gtrsim 0.06$. Needless to say, (54) is only a necessary condition for negative refraction and is very far from being sufficient. This condition makes it possible for the higher-order Fourier harmonics of the Bloch wave to outweigh the first-Brillouin-zone harmonic, but does not guarantee that they will do so and that they will have the desirable sign.

X SUMMARY

1. The history of backward waves and negative-index materials can be traced back to the beginning of the 20th century (H. Lamb, A. Schuster, H.C. Pocklington) and to the work of Russian physicists (L.I. Mandelshtam, D.V. Sivukhin, V.E. Pafomov, R.A. Silin, V.G. Veselago) in the 1940s–1970s. Since the experimental demonstration by D.R. Smith and the discovery of the ‘perfect lens’ phenomenon by J.B. Pendry in 2000, there has been an explosion of interest in negative refraction and backward waves in periodic structures.

2. The Fourier-Bloch decomposition of the field into plane-wave components clarifies the nature of backward waves. Phase velocity is governed primarily by the first-Brillouin-zone component of the Bloch wave – assuming that this component has an appreciable magnitude. In contrast, energy flow is a collective effect of all plane-wave harmonics. Backward waves occur when higher-order harmonics, taken together, outweigh the first one and have the opposite sign.

3. Intrinsic and extrinsic characteristics should be distinguished. The forward or backward character of a wave is its intrinsic characteristic in the bulk of the material. At interface boundaries, the Bloch wave can in principle be excited via any of its plane-wave components. The type of refraction depends on both factors: the intrinsic character of the wave and the ‘excitation channel’.

4. Negative-index media are usually subdivided into two kinds: ‘metamaterials,’ typically containing relatively small
resonating elements within a lattice cell, and ‘photonic crystals’ – periodic dielectric structures such as arrays of dielectric cylinders, spheres, etc. Consequently, negative refraction is usually classified as two species: one occurring in metamaterials, another one in photonic crystals.

5. Metamaterials might seem to approximate, in the limiting case of a small cell, an ideal double-negative ‘Veselago medium’ supporting backward waves. However, this paper argues that there is a fundamental lower limit on the lattice cell size, relative to the vacuum wavelength, for a periodic dielectric medium capable of supporting backward waves. As the cell size tends to zero, the operating point on the normalized band diagram of the structure falls on the acoustic branch, thereby eliminating the difference between phase and group velocities. The paper provides explicit lower bounds for the cell size.

6. In short, there are fundamental limitations on how small the lattice size can be for negative index materials. The “Veselago medium” may not in fact be realizable as a limiting case of spatially periodic dielectric structures with a small cell size.

ACKNOWLEDGMENT

I thank Frantisek Čajko for his perceptive comments that helped to debug the manuscript.

XI REFERENCES


