



**HAL**  
open science

# Applicability of effective medium description to photonic crystals in higher bands: Theory and numerical analysis

Vadim A. Markel, I. Tsukerman

## ► To cite this version:

Vadim A. Markel, I. Tsukerman. Applicability of effective medium description to photonic crystals in higher bands: Theory and numerical analysis. *Physical Review B: Condensed Matter and Materials Physics* (1998-2015), 2016, 93 (22), pp.224202 10.1103/PhysRevB.93.224202 . hal-01282090v1

**HAL Id: hal-01282090**

**<https://hal.science/hal-01282090v1>**

Submitted on 3 Mar 2016 (v1), last revised 3 Jun 2016 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# Can photonic crystals be homogenized in higher bands?

Vadim A. Markel\*

*Aix-Marseille Universit, CNRS, Centrale Marseille,  
Institut Fresnel UMR 7249, 13013 Marseille, France*

Igor Tsukerman

*Department of Electrical and Computer Engineering,  
The University of Akron, OH 44325-3904, USA<sup>†</sup>*

(Dated: March 3, 2016)

We consider the conditions under which photonic crystals (PCs) can be viewed as electromagnetically homogeneous at frequencies in the higher photonic bands and, in particular, near the higher-order  $\Gamma$ -points. At these frequencies, the real isofrequency lines of a PC can be very close to a mathematical circle, just like in an isotropic homogeneous material. We however show that the above observation is insufficient for establishing the PC homogenizability. Complex dispersion points must be included into consideration even in the case of strictly non-absorbing materials. By analyzing the complex dispersion relations and the corresponding isofrequency surfaces, we have found that two-dimensional PCs with  $C_4$  and  $C_6$  symmetries are not electromagnetically homogeneous in the higher photonic bands.

## I. INTRODUCTION

The theory of homogenization of periodic electromagnetic media continues to attract significant attention<sup>1-5</sup> due to the proposed important applications such as sub-wavelength optical imaging<sup>6</sup>. However, some fundamental questions remain open in this field of research. Perhaps, the most important of these questions is the following: to what extent can the “exotic” effective parameters obtained via one of the several recently proposed homogenization theories be used without restriction, like the constitutive parameters of homogeneous natural media are conventionally used? Indeed, while it is generally understood that homogenization is an approximate procedure, the accuracy and the applicability range of a given theory is very difficult to ascertain quantitatively.

In this paper, we investigate whether a photonic crystal (PC) can be homogenized at frequencies above the first bandgap and, particularly, near the higher  $\Gamma$ -points. Answering this question is important for the following reason. It is well known that PCs can be characterized by negative dispersion in the higher photonic bands even if the material from which the PC is made has no absorption. On the other hand, in purely dielectric homogeneous media, negative dispersion is obtained only within the absorption bands. A homogeneous material can be simultaneously transparent and characterized by negative dispersion only if it has a nontrivial magnetic permeability<sup>7,8</sup>. Therefore, to obtain this result by homogenizing an intrinsically non-magnetic PC, one has to consider sufficiently large frequencies where the dispersion is negative<sup>9</sup>. Typically, these frequencies are above the first bandgap of the crystal. We note that some fairly intuitive arguments exist suggesting that a PC can be homogenized at frequencies sufficiently close to any of its  $\Gamma$ -points. However, a careful consideration reveals that the first  $\Gamma$ -point  $\omega_1 = 0$  is fundamentally different from the higher ones  $\omega_n > 0$  ( $n = 2, 3, \dots$ ). While homoge-

nization can be arbitrarily accurate near  $\omega_1$ , the same is not true near any  $\omega_n$  with  $n > 1$ .

The above conclusion is consistent with some of the previous numerical investigations of the homogenization problem<sup>10-12</sup>. In fact, the work reported here is conceptually close to Ref. 10 and one of the main ideas on which the present paper is based has been stated in that reference. Namely, it was noticed that the right-hand side of the dispersion equation  $\omega = f(\mathbf{q})$  can be expanded in powers of the Cartesian components of  $\mathbf{q}$  if  $\omega$  is close to one of the  $\Gamma$ -points  $\omega_n$ . Here  $\mathbf{q}$  is the Bloch wave vector. Further, for s-polarized waves in two-dimensional PCs with a center of symmetry, this expansion is of the form  $f(\mathbf{q}) = \omega_n + \beta_x q_x^2 + \beta_y q_y^2 \dots$ , or in the cases of  $C_4$  and  $C_6$  symmetries that are considered in this paper,  $f(\mathbf{q}) = \omega_n + \beta q^2 \dots$ . Here we have written explicitly the first two non-vanishing terms of the expansion and  $\omega$  is assumed to be in the pass-band. By truncating the latter expansion at the second order, we obtain the isotropic isofrequency line  $q^2 = (\omega - \omega_n)/\beta$  so that the wave number  $q$  does not depend on the direction of propagation. Another observation made in Ref. 10 was that, for incident waves with a sufficiently large projection of the wave vector onto the interface (which includes but is not limited to the evanescent waves), higher-order terms must be retained in the expansion of  $f(\mathbf{q})$  no matter how close the frequency is to the  $\Gamma$ -point, and that the resultant law of dispersion is no longer isotropic.

Here we develop this basic idea theoretically and illustrate it with numerical examples of a different kind than what was used in Ref. 10. While the latter reference is focused on the transmission and imaging properties of a PC slab, here we consider in detail the dispersion relations and isofrequency surfaces. We work in a 2D geometry and in s-polarization (one-component electric field) so that Maxwell’s equations are reduced to a scalar wave equation. Under these conditions, only scalar effective parameters can be introduced. The corresponding

effective medium is isotropic and so are its isofrequency surfaces. However, we will show that in the actual PC the isotropy is lost in the higher photonic bands once we include waves with sufficiently large projections of the wave vector onto a given axis. In this case, the medium cannot be described by local and isotropic (scalar) effective parameters. Magnetic anisotropy can be introduced as a phenomenological adjustable parameter, but then one would obtain either a non-symmetric magnetic tensor and hence non-reciprocity, which is not physically present in the system, or otherwise a tensor whose principal axes do not coincide with the crystallographic axes (e.g., for the triangular lattice), or else a tensor that is trivially proportional to the identity (e.g., for the square lattice).

Thus, we will view PCs whose isofrequency surfaces are not isotropic as not electromagnetically homogeneous. We note that the lack of isotropy can be interpreted as an effect of nonlocality (spatial dispersion)<sup>13–15</sup>. This effect can occur in homogeneous materials as well, although in the latter case it is usually very weak. We should keep in mind that nonlocality of medium parameters and anisotropy of *local* parameters can result in somewhat similar phenomena but are not the same effect and should be distinguished. To complicate things further, it is frequently stated that there exists a complete physical equivalence between two alternative descriptions of the electromagnetic properties of continuous media<sup>16–19</sup>. In one description the medium is assigned a nonlocal permittivity  $\epsilon(\omega, \mathbf{q})$  and a trivial permeability  $\mu = 1$  (the so-called Landau-Lifshitz approach<sup>19</sup>). In the other description, the medium is assigned two local parameters  $\epsilon(\omega)$  and  $\mu(\omega)$ . We have previously argued that the two descriptions are not physically equivalent in general<sup>20</sup>, but it is true that in the so-called weak nonlocality regime (usually understood as the regime in which Taylor expansion of  $\epsilon(\omega, \mathbf{q})$  to second order in  $\mathbf{q}$  is an accurate approximation) such an equivalence exists for the refractive index of the medium (but not for the impedance).

For the purpose of this paper, it is unimportant whether the equivalence mentioned above exists or not. If it does exist, then a homogenizable medium should be completely characterized by some local parameters  $\epsilon(\omega)$  and  $\mu(\omega)$ . Whether these local parameters correspond to some nonlocal  $\epsilon(\omega, \mathbf{q})$  and  $\mu = 1$  in the alternative description is irrelevant. On the other hand, if such local parameters do not exist, then the medium cannot be reasonably homogenized and the introduction of the nonlocal permittivity does not solve the problem at all because the knowledge of  $\epsilon(\omega, \mathbf{q})$  for all values of its arguments is not sufficient to solve any boundary value problems in a finite sample<sup>20</sup>, unlike the knowledge of the local parameters  $\epsilon(\omega)$  and  $\mu(\omega)$ . Besides, the typical applications discussed in the literature such as subwavelength imaging require local  $\epsilon(\omega)$  and  $\mu(\omega)$ . Therefore, we say that, in order for a PC to be homogenizable, its dispersion relation must be (at least, approximately) the same as in a truly homogeneous medium with some local parameters  $\epsilon(\omega)$

and  $\mu(\omega)$ . We emphasize that the above condition is necessary but, in general, not sufficient because it does not include the impedance. But we will show that even this necessary condition of homogenizability does not hold in PCs above the first bandgap.

We illustrate the theoretical arguments of this paper with numerical examples using a rather simple but physically relevant model. As was mentioned above, we consider two-dimensional PCs with a one-component electric field polarized perpendicularly to the plane of periodicity (s-polarization). Such PCs have been previously considered in the literature as homogenizable in the higher photonic bands<sup>21–23</sup>. Similarly to these works, we neglect frequency dispersion and absorption in the material that makes up the PC. This is done not for computational convenience (our codes can handle the more general case with equal efficiency) but in order to analyze the exact cases that were previously considered in the literature. Besides, if we include absorption and frequency dispersion into the model, the higher  $\Gamma$ -points do not really exist and the case for homogenizability is even harder to make. Therefore, we have made the approximations that are favorable rather than detrimental to the homogenizability.

We finally note that our results can be understood in a more general framework of the *uncertainty principle* of homogenization<sup>24</sup>. According to this principle, the larger the deviation of the effective magnetic permeability from unity (according to a given theory), the less accurate this theory is in predicting physical observables such as the transmission and reflection coefficients of a composite slab.

The remainder of this paper is organized as follows. We start with some general theoretical considerations relevant to the problem at hand in Sec. II where we explain why the circularity (or sphericity) of a real isofrequency line is not a sufficient condition for homogenizability. Additional mathematical details pertinent to the case of 2-dimensional PCs are given in Sec. III. Extensive numerical examples for 2D square and triangular PCs characterized by  $C_4$  and  $C_6$  symmetries, respectively, are adduced in Sec. IV. Finally, Sec. V contains a discussion of the obtained results.

## II. GENERAL CONSIDERATIONS

For Bloch wave with a nonzero fundamental harmonic  $\mathbf{E}_0$  (as defined in Appendix A), the dispersion equation for a three-dimensional, intrinsically non-magnetic PC can be written in the following general form<sup>20,25</sup>

$$\det [(\mathbf{q} \times \mathbf{q} \times) + k^2 \Sigma(\omega, \mathbf{q})] = 0. \quad (1)$$

Here  $k = \omega/c$  is the free-space wave number and  $\Sigma(\omega, \mathbf{q})$  is a 3D tensor, which is completely determined by the PC geometry and composition and by its two arguments  $\omega$  and  $\mathbf{q}$ . The latter can be considered as mathematically-independent variables and take arbitrary complex values.

We will say that a complex vector  $\mathbf{q}$  is the Bloch wave vector of a PC at some frequency  $\omega$  if the pair  $(\omega, \mathbf{q})$  satisfies (1).

Below we work in the frequency domain, so that  $\omega$  is the true mathematically-independent parameter of the theory, although we do restrict the frequency to be real and positive. The set of all  $\mathbf{q}$ 's that satisfy (1) for a given  $\omega > 0$  forms an isofrequency surface in 3D or a line in 2D. However, the words ‘‘surface’’ and ‘‘line’’ should not be understood literally because  $\mathbf{q}$  is in general complex.

The function  $\Sigma(\omega, \mathbf{q})$  arises in various physical contexts<sup>13,15,26</sup> and is sometimes interpreted as the nonlocal permittivity tensor of the medium due to its explicit dependence on  $\mathbf{q}$ . We have shown<sup>20</sup> that the knowledge of  $\Sigma(\omega, \mathbf{q})$  for all values of its arguments is insufficient for solving boundary value problems in finite samples. However, in this paper we restrict attention to dispersion relations, and to this end the knowledge of  $\Sigma(\omega, \mathbf{q})$  is sufficient. We emphasize that introduction of the external ‘‘impressed’’ current that overlaps with the medium<sup>13,15,26</sup> is not necessary to define  $\Sigma(\omega, \mathbf{q})$  mathematically. Thus, equation (1) has been derived and the function  $\Sigma(\omega, \mathbf{q})$  defined algebraically for a three-dimensional, two-component PC in Refs. 20,25 without considering any external currents. These two references utilize a plane-wave (Fourier) basis for all lattice-periodic functions. In Appendix A, we present a basis-free derivation of (1) for a 2-dimensional PC with an arbitrary lattice-periodic permittivity. In Sec. III below and in Appendix B, we develop a perturbation expansion of  $\Sigma(\omega, \mathbf{q})$  in powers of the Cartesian components of  $\mathbf{q}$  for two-dimensional, two-component PCs. To this end, we use the plane-wave basis, which allows one to obtain expressions that are directly amenable to numerical computation.

We can use (1) to formally define the  $\Gamma$ -points. As one could expect, the definition is confounded by the vector nature of the electromagnetic fields. In particular, the  $\Gamma$ -points can be polarization-dependent. However, the simulations discussed below have been performed for the special case of transverse Bloch waves, which satisfy the condition  $\mathbf{q} \cdot \mathbf{E}_0 = 0$ . Here  $\mathbf{E}_0$  is the amplitude of the fundamental harmonic of the Bloch wave for a given linear polarization state. In this case, (1) simplifies to

$$q^2 = k^2 \Sigma(\omega, \mathbf{q}), \quad (2)$$

where  $\mathbf{q}$  is now a two-dimensional vector orthogonal to  $\mathbf{E}_0$  and  $\Sigma(\omega, \mathbf{q})$  is a scalar (a principal value of the tensor  $\Sigma$  that corresponds to the direction of  $\mathbf{E}_0$ ). Note that  $\mathbf{E}_0$  can be an eigenvector of  $\Sigma(\omega, \mathbf{q})$  for all  $\omega$  and  $\mathbf{q}$ , typically, as a consequence of the PC symmetry. Also, we use the same notation for the tensor and for its principal value but this should not cause confusion since only the latter interpretation will be used below.

In what follows, we will consider only the scalar equation (2) and assume that the two-dimensional Bloch vector  $\mathbf{q} = (q_x, q_y)$  lies in the  $XY$  plane of a rectangular frame while  $\mathbf{E}_0 = E_0 \hat{\mathbf{z}}$  is collinear with the  $Z$ -axis. By

focusing on the special case of transverse waves, we do not disregard any important effects but rather focus on the essential features of the theory. It is important to keep in mind that the above assumptions can hold in both two-dimensional and three-dimensional PCs. However, in the 2D geometry that we use consider in detail in this paper, even stronger conditions are satisfied: (i) not only  $\mathbf{E}_0$  but the total electric field  $\mathbf{E}(x, y) = \hat{\mathbf{z}}E(x, y)$  is collinear with the  $Z$ -axis and (ii) the electric field  $E(x, y)$  is independent of  $z$ . The latter two properties do not generally hold in three-dimensional PCs.

We now proceed with the analysis of Eq. (2). We will say that a frequency  $\omega_n$  or the corresponding free-space wave number  $k_n$  are the  $\Gamma$ -points of a PC if (2) is satisfied for  $\omega = \omega_n$  and  $\mathbf{q} = 0$ . That is, the equation for the  $\Gamma$ -points reads

$$\omega^2 \Sigma(\omega, 0) = 0. \quad (3)$$

The first trivial solution to this equation is  $\omega_1 = 0$ , unless  $\Sigma(\omega, 0)$  diverges as  $1/\omega^2$  or faster. This possibility can be safely ignored and the first (fundamental)  $\Gamma$ -point exists in all PCs, even if they are made of conducting constituents<sup>27</sup>. All higher  $\Gamma$ -points have  $\omega_n \neq 0$  and can be determined from the equation

$$\Sigma(\omega, 0) = 0, \quad \omega > 0. \quad (4)$$

It therefore can be seen that the nature of the first and the higher  $\Gamma$ -points is quite different. At the first  $\Gamma$ -point,  $\Sigma$  does not turn to zero; in fact, it can be readily seen that  $\Sigma(0, 0) = n_{\text{eff}}^2$ , where  $n_{\text{eff}}$  is the effective refractive index of the medium in the homogenization limit. But higher  $\Gamma$ -points all require that  $\Sigma(\omega_j, 0) = 0$ . This equality cannot hold in any homogeneous medium (see Appendix A for more details).

At this point, a few remarks are in order. First, purely real  $\Gamma$ -points of order higher than 1 do not generally exist in PCs with non-negligible absorption. This is so because (4) is in this case a complex equation; both real and imaginary parts of (4) must be satisfied simultaneously. This is unlikely to happen for a purely real  $\omega$ . However, if, as is frequently done in the PC literature, we assume that the dielectric permittivity of the PC constituents is purely real, then (4) can be expected to have real roots. This will be discussed analytically in more detail in Sec. III and illustrated numerically in Sec. IV.

Second, derivation of equations (1) or (2) does not require that  $\text{Re}(\mathbf{q})$  be in the first Brillouin zone (FBZ) of the lattice. However, if  $\mathbf{q}$  is a solution to one of these equations, then  $\mathbf{q} + \mathbf{g}$  is also a solution, where  $\mathbf{g}$  is any reciprocal lattice vector. Therefore, it is sufficient to consider only the solutions with  $\text{Re}(\mathbf{q}) \in \text{FBZ}$ .

Third, and related to the above, homogeneous media do not possess higher-order  $\Gamma$ -points in the sense of the definition (4). It is true that the ‘‘artificially folded’’ dispersion curve of a homogeneous medium, such as the one shown in Fig. 2 of Ref. 28, crosses the vertical axis at some frequencies  $\omega_n$ , and one can conclude that at

$\omega = \omega_n$  we also have  $\mathbf{q} = 0$ . This point is illustrated in Fig. 1 where we plot the purely real isofrequency lines for a homogeneous medium that is artificially discretized on two-dimensional square and triangular lattices. It can be seen that the isofrequency lines are very far from circular and can become quasi-chaotic for large band indexes  $n$ . This behavior distinguishes these pseudo- $\Gamma$ -points from the true  $\Gamma$ -points of a PC with strong interaction. Strictly speaking, the fundamental harmonic of the Bloch wave  $\mathbf{E}_0$  is zero at the frequencies corresponding to these spurious  $\Gamma$ -points, and equation (1) is not applicable in this case. For a more detailed analysis, see Appendix A.

Returning to the question of homogenizability, we can now see why one might expect the medium to be homogenizable in the vicinity of a true  $\Gamma$ -point. Let  $\omega$  be in the transmission band of a PC but close to a  $\Gamma$ -point  $\omega_n$  ( $n > 1$ ), say, slightly below  $\omega_n$ . Let us fix the frequency and expand  $\Sigma(\omega, \mathbf{q})$  in powers of  $\mathbf{q}$ . If, as is often the case, the PC has a center of symmetry, the expansion is to lowest order in  $\mathbf{q}$  of the form

$$\Sigma(\omega, \mathbf{q}) = \alpha + \beta_x \frac{q_x^2}{k^2} + \beta_y \frac{q_y^2}{k^2} + \dots, \quad \alpha \equiv \Sigma(\omega, 0) \ll 1. \quad (5)$$

In the above equation,  $\omega$  is a fixed parameter and therefore the explicit dependence of the expansion coefficients  $\alpha, \beta$  on  $\omega$  is suppressed. Also, it is easy to see that in PCs obeying  $C_4$  or  $C_6$  symmetry,  $\beta_x = \beta_y = \beta$ . If we now treat the first two terms in the expansion (5) as an approximation, write  $\Sigma(\omega, \mathbf{q}) \approx \alpha + \beta q^2/k^2$  and substitute the result into the dispersion equation (2), we will obtain an approximate solution to the dispersion equation, viz,

$$q^2 = k^2 \frac{\alpha}{1 - \beta}. \quad (6)$$

We remark briefly that it is rather typical to interpret  $\alpha$  as the effective permittivity of the medium,  $\epsilon_{\text{eff}}$  and  $1/(1 - \beta)$  as the effective permeability,  $\mu_{\text{eff}}$ . This approach is discussed in detail in Ref. 20. However, since in this paper we only consider infinite media, the breakdown of the squared refractive index  $n_{\text{eff}}^2$  into the product of  $\epsilon_{\text{eff}}$  and  $\mu_{\text{eff}}$  is irrelevant. What is important for our purposes is that (6) describes a perfectly circular isofrequency line. Of course, circular isofrequency lines are also characteristic of electromagnetically-homogeneous materials. The conclusion is then drawn that sufficiently close to a  $\Gamma$ -point, a PC is indistinguishable from a homogeneous medium.

However the above line of arguments has the following deficiency. It is not really true that all Cartesian components of  $\mathbf{q}$  are small near a  $\Gamma$ -point. A more precise statement is that the scalar  $q^2 = \mathbf{q} \cdot \mathbf{q}$  is small. However, the Cartesian components of  $\mathbf{q}$  can still be arbitrarily large. An obvious example is the vector  $\mathbf{q} = (p, ip)$  where  $p$  is a real number.

More specifically, we will consider the following problem. Fix the frequency and assume that the Bloch wave vector has a known and purely real projection onto the real axis,  $q_x$ . Then compute the corresponding values of

$q_y$  that satisfy the dispersion equation. Here  $q_y$  can be real, imaginary or complex, even if the PC is made of purely transparent components. The set of purely real solutions ( $q_x, q_y$ ) would form the traditional isofrequency line. However, the dispersion equation has a solution (in fact, infinitely many solutions) for any  $q_x$ . Some of these solutions will be complex. When  $q_y$  has a nonzero imaginary part, the corresponding Bloch wave is evanescent. Evanescent Bloch waves can be excited in the PC by an incident plane wave that is either propagating or evanescent in vacuum. It is important however, that the projection of the wave vector of the incident wave onto any flat interface is equal to  $q_x$ . Thus, to make sure that a homogenization theory is applicable to a sufficiently large class of incident waves (which must necessarily include evanescent waves), we must consider  $q_x$  in the interval of at least  $0 < q_x \lesssim 2k$ , where  $k = \omega/c$  is the free space wave number.

### III. DISPERSION EQUATION IN A TWO-DIMENSIONAL PC

Consider a two-dimensional PC whose exact permittivity  $\tilde{\epsilon}(\mathbf{r})$  satisfies the periodicity relation

$$\tilde{\epsilon}(\mathbf{r} + n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2) = \tilde{\epsilon}(\mathbf{r}), \quad (7)$$

where  $\mathbf{a}_1$  and  $\mathbf{a}_2$  are two primitive lattice vectors and  $n_1, n_2$  are integers. Here and below the overhead tilde will be used to denote function obeying the lattice periodicity (7). We will work in an orthogonal reference frame whose  $Z$ -axis is perpendicular to both  $\mathbf{a}_1$  and  $\mathbf{a}_2$ . Then these two vectors lie in the  $XY$  plane. We will further consider electromagnetic waves with one-component electric field  $\hat{\mathbf{E}} = (0, 0, E)$  and two-component magnetic field  $\mathbf{H} = (H_x, H_y, 0)$ . In this case, Maxwell's equations are reduced to the scalar wave equation for the electric field

$$[\nabla^2 + k^2 \tilde{\epsilon}(\mathbf{r})] E(\mathbf{r}) = 0, \quad (8)$$

where  $\mathbf{r} = (x, y)$ .

Let  $\mathbf{b}_1$  and  $\mathbf{b}_2$  be the primitive vectors of the reciprocal lattice such that

$$\mathbf{a}_j \cdot \mathbf{b}_k = 2\pi \delta_{jk}. \quad (9)$$

A generic reciprocal lattice vector  $\mathbf{g}$  can be written as

$$\mathbf{g} = \mathbf{b}_1 n_1 + \mathbf{b}_2 n_2, \quad (10)$$

where  $n_1, n_2$  are integers. We can view  $\mathbf{g}$  as a discrete composite index, which maps one-to-one to the pair  $(n_1, n_2)$ . In particular, the exact permittivity of the medium is expandable as

$$\tilde{\epsilon}(\mathbf{r}) = \sum_{\mathbf{g}} \epsilon_{\mathbf{g}} e^{i\mathbf{g} \cdot \mathbf{r}}, \quad \epsilon_{\mathbf{g}} = \frac{1}{S[\mathbf{C}]} \int_{\mathbf{C}} \tilde{\epsilon}(\mathbf{r}) e^{-i\mathbf{g} \cdot \mathbf{r}} d^2r, \quad (11)$$

where  $\mathbf{C}$  is an elementary cell of the medium and  $S[\mathbf{C}]$  is it's area, i.e.,  $a^2$  for a square lattice or  $|(\mathbf{a}_1 \times \mathbf{a}_2) \cdot \hat{\mathbf{z}}|$  in

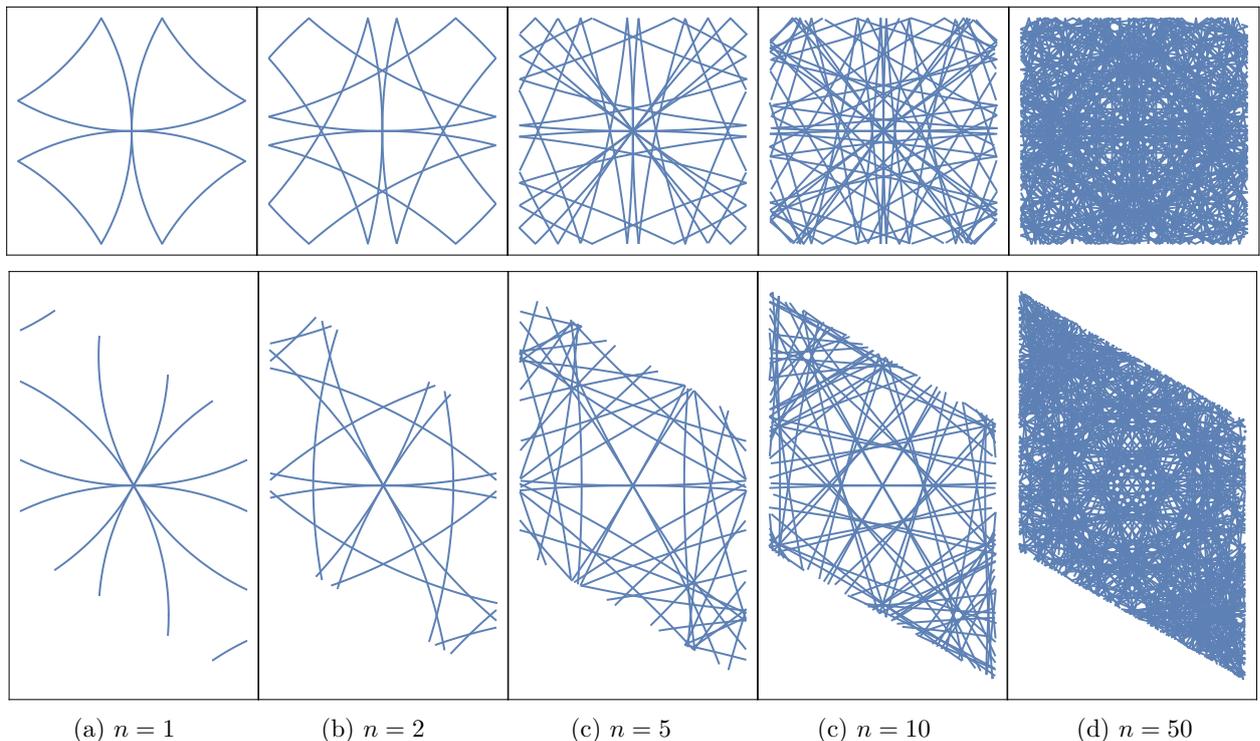


FIG. 1: (color online) Isofrequency lines for a two-dimensional homogeneous medium artificially discretized on a square (top) and triangular (bottom) lattices. The band index is labeled by  $n$ . The plots depict the isofrequency lines contained in the FBZ of the lattice and can be periodically replicated in all directions. A mathematical definition of the folding operation is given in Eq. (34).

the more general case. For a two-component medium we can write

$$\tilde{\epsilon}(\mathbf{r}) = \epsilon_h + (\epsilon_i - \epsilon_h)\tilde{\Theta}(\mathbf{r}), \quad (12)$$

where  $\epsilon_h$  and  $\epsilon_i$  are the permittivities of the host and the inclusions and  $\tilde{\Theta}(\mathbf{r})$  is the lattice-periodic shape function. Let the region of the inclusion be denoted by  $\Omega \in \mathbb{C}$ . Then  $\tilde{\Theta}(\mathbf{r}) = 1$  if  $\mathbf{r} \in \Omega$  and  $\tilde{\Theta}(\mathbf{r}) = 0$  otherwise. Upon Fourier transformation, we find that

$$\epsilon_{\mathbf{g}} = \epsilon_h \delta_{\mathbf{g}0} + \rho \chi M(\mathbf{g}), \quad (13)$$

where

$$M(\mathbf{g}) = \frac{1}{S[\Omega]} \int_{\Omega} e^{-i\mathbf{g}\cdot\mathbf{r}} d^2r, \quad (14a)$$

$$\rho = S[\Omega]/S[\mathbb{C}], \quad (14b)$$

$$\chi = \epsilon_i - \epsilon_h. \quad (14c)$$

Note that in the above equations  $\rho$  is the area fraction of the inclusions and  $\chi$  is the contrast. The function  $M(\mathbf{g})$  contains information about the inclusion geometry but is independent of  $\chi$ .

Bloch-periodic functions such as the electric field  $E(\mathbf{r})$  and displacement  $D(\mathbf{r})$  can be expanded as

$$E(\mathbf{r}) = \sum_{\mathbf{g}} E_{\mathbf{g}} e^{i(\mathbf{q}+\mathbf{g})\cdot\mathbf{r}}, \quad (15a)$$

$$D(\mathbf{r}) = \sum_{\mathbf{g}} D_{\mathbf{g}} e^{i(\mathbf{q}+\mathbf{g})\cdot\mathbf{r}} \quad (15b)$$

where  $\mathbf{q}$  is the Bloch wave vector, which must be determined by substituting (15) into (8). The equation  $D(\mathbf{r}) = \tilde{\epsilon}(\mathbf{r})E(\mathbf{r})$  takes the form

$$D_{\mathbf{g}} = \sum_{\mathbf{p}} \epsilon_{\mathbf{g}-\mathbf{p}} E_{\mathbf{p}} \quad (16)$$

while (8) is reduced in this case to

$$(\mathbf{g} + \mathbf{q})^2 E_{\mathbf{g}} = k^2 D_{\mathbf{g}}. \quad (17)$$

Combining the wave equation and constitutive relation together we obtain the eigenproblem

$$(\mathbf{g} + \mathbf{q})^2 E_{\mathbf{g}} = k^2 \sum_{\mathbf{p}} \epsilon_{\mathbf{g}-\mathbf{p}} E_{\mathbf{p}}, \quad (18)$$

which determines the allowable values of the vector  $\mathbf{q}$  for each frequency  $\omega$  (or for the corresponding wavenumber  $k = \omega/c$ ).

Equation (18) is well known. The family of vectors  $\mathbf{q}$  that satisfy (18) for some fixed  $\omega > 0$  is said to form an isofrequency line. This is however not quite right because the Cartesian components of  $\mathbf{q}$  can be complex and not of the same phase. The set of  $\mathbf{q}$  can be depicted by lines in

a plane only if we restrict consideration to real vectors  $\mathbf{q}$  (or assume that  $q_x$  and  $q_y$  have the same complex phase). This assumption is not always justified, so that we can talk about isofrequency lines only in a generalized sense.

As the first and rather trivial step, let us use the particular expression for  $\epsilon_{\mathbf{g}}$  (13), which is specific to two-component PCs. We will also assume that  $E_0 \neq 0$  (we are restricting ourselves to the solutions of the first kind according to the terminology of Appendix A) and scale all coefficients  $E_{\mathbf{g}}$  by  $E_0$ . Let  $e_{\mathbf{g}} = E_{\mathbf{g}}/E_0$ , so that  $e_0 = 1$ . Upon substitution of (13) into (18) and using the above scaling, we obtain the infinite set of equations

$$[(\mathbf{g} + \mathbf{q})^2 - k^2 \epsilon_h] e_{\mathbf{g}} = \rho \chi k^2 \sum_{\mathbf{p}} M(\mathbf{g} - \mathbf{p}) e_{\mathbf{p}}. \quad (19)$$

The next step is slightly less trivial. As was done by us previously<sup>20,25</sup>, we consider the equation (19) for the cases  $\mathbf{g} = 0$  and  $\mathbf{g} \neq 0$  separately. First, in the  $\mathbf{g} = 0$  case, we obtain the equation (2) in which

$$\Sigma(\omega, \mathbf{q}) = \epsilon_0 + \rho \chi \sum_{\mathbf{g} \neq 0} M(-\mathbf{g}) e_{\mathbf{g}}, \quad (20)$$

where  $\epsilon_0 = \rho \epsilon_i + (1 - \rho) \epsilon_h$  is the cell-averaged permittivity of the PC.

Note that  $\Sigma$  is defined by the relative amplitudes  $e_{\mathbf{g}} = E_{\mathbf{g}}/E_0$  with  $\mathbf{g} \neq 0$ . Therefore, to compute  $\Sigma$  algebraically, we must consider (19) for  $\mathbf{g} \neq 0$ . This yields the following set of equations

$$[(\mathbf{q} + \mathbf{g})^2 - k^2 \epsilon_h] e_{\mathbf{g}} = \rho \chi k^2 \times \left[ M(\mathbf{g}) + \sum_{\mathbf{p} \neq 0} M(\mathbf{g} - \mathbf{p}) e_{\mathbf{p}} \right], \quad \mathbf{g} \neq 0. \quad (21)$$

It is important to note that (21) is a closed (albeit an infinite) set of equations with a nonzero free term  $\rho \chi k^2 M(\mathbf{g})$ . Consequently, (21) is not an eigenproblem but rather a linear set of equations that can be, in principle, solved by matrix inversion.

We are interested in building a perturbation theory in  $\mathbf{q}$  for  $\omega$  in the vicinity of one of the  $\Gamma$ -points  $\omega_n$ . Note that, under this condition, we can expect that  $\Sigma(\omega, 0)$  is small but nonzero and positive (otherwise the frequency is inside a bandgap). We can proceed as follows. Define the following matrices and vectors:

$$D_{\mathbf{g}\mathbf{p}} = (g^2 - k^2 \epsilon_h) \delta_{\mathbf{g}\mathbf{p}}, \quad (22a)$$

$$Q_{\mathbf{g}\mathbf{p}} = (q^2 + 2\mathbf{q} \cdot \mathbf{g}) \delta_{\mathbf{g}\mathbf{p}}, \quad (22b)$$

$$M_{\mathbf{g}\mathbf{p}} = M(\mathbf{g} - \mathbf{p}), \quad (22c)$$

$$b_{\mathbf{g}} = M(\mathbf{g}). \quad (22d)$$

Here only  $Q$  depends on  $\mathbf{q}$ . Also note that the normalization rule  $\langle b|b \rangle = 1/\rho - 1$ . We can now re-write (21) as follows:

$$(Q + D - \rho \chi k^2 M)|e\rangle = \rho \chi k^2 |b\rangle. \quad (23)$$

We can define the  $T$ -matrix of the problem as follows:

$$T = (D - \rho \chi k^2 M)^{-1}, \quad (24)$$

Note that  $T$  is the matrix that we need in order to compute  $\Sigma(\omega, 0)$ . Indeed, we have for  $\mathbf{q} = 0$

$$|e\rangle = \rho \chi k^2 T |b\rangle, \quad \text{for } \mathbf{q} = 0. \quad (25)$$

and

$$\begin{aligned} \Sigma(\omega, 0) &= \epsilon_0 + (\rho \chi k)^2 \langle b|T|b \rangle \\ &= \epsilon_0 + (\rho \chi k)^2 \sum_{\mathbf{g}_1, \mathbf{g}_2 \neq 0} M(-\mathbf{g}_1) T_{\mathbf{g}_1 \mathbf{g}_2} M(\mathbf{g}_2). \end{aligned} \quad (26)$$

We thus assume that  $T = T(\omega)$  has been computed at the working frequency. Computationally, this requires truncation of the basis and one matrix inversion operation. Now the equation we intend to iterate is of the form

$$|e\rangle = \rho \chi k^2 T |b\rangle - T Q |e\rangle, \quad (27)$$

which follows directly from (23) and (24). The formal power series solution is of the form

$$|e\rangle = \rho \chi k^2 \sum_{n=0}^{\infty} (-T Q)^n T |b\rangle \quad (28)$$

and for  $\Sigma$ ,

$$\Sigma = \epsilon_0 + (\rho \chi k)^2 \sum_{n=0}^{\infty} \langle b|(-T Q)^n T|b \rangle \quad (29a)$$

$$= \epsilon_0 + (\rho \chi k)^2 \sum_{n=0}^{\infty} \sigma_n, \quad (29b)$$

where

$$\sigma_n = (-1)^n \langle b|(T Q)^n T|b \rangle. \quad (30)$$

The coefficients  $\sigma_n$  are computed in terms of the matrix  $T$  and Cartesian components of  $\mathbf{q}$  in Appendix B up to fourth order. These results show that the functional form of the expansion (29) (up to the same order  $n = 4$ ) is

$$\Sigma = \beta_0 + \beta_2 q^2 + \beta_4 q^4 + \beta_6 q^6 + \beta_8 q^8 + \delta_4 q_x^2 q_y^2. \quad (31)$$

Here the subscripts of the coefficients  $\beta_n, \delta_n$  are not related directly to the perturbation order. For example, the term  $\beta_8 q^8$  is obtained from the fourth-order term  $\sigma_4$ . The anisotropic term

$$q_x^2 q_y^2. \quad (32)$$

is also contained in  $\sigma_4$ . We can say that, starting from fourth order in  $Q$ , the function  $\Sigma(\omega, \mathbf{q})$  starts to bear the traces of the underlying lattice, which is not circularly-symmetric.

However, in the triangular lattice that obeys the  $C_6$  symmetry, the coefficient  $\delta_4$  is identically zero. This is

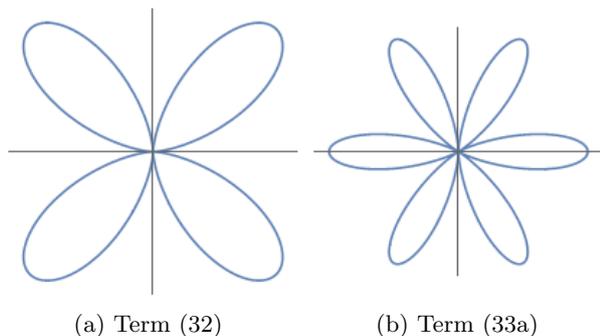


FIG. 2: Graphical illustration of the terms (32) and (33a). The parametric plots show the dependence of the magnitude of each term on the direction of the (purely real) vector  $\mathbf{q}$ . Note that the lines shown in plots (a) and (b) are similar (but not completely identical) to the isofrequency lines shown in Fig. 1 for  $n = 1$ . In the latter case, the lines are deformed (reflected) by the boundaries of the FBZ. However, in the vicinity of the origin, the lines are identical. This is a general manifestation of the applicable rotational symmetry group [ $C_4$  in (a) and  $C_6$  in (b)]. Note however, that the lines shown in this figure are not the isofrequency lines of the PC.

so because the expression (32) is not invariant with respect to rotation by  $\pi/3$ . The simplest anisotropic terms that are invariants of  $C_6$  arise to sixth order in  $Q$  (are contained in  $\sigma_6$ ) and are of the form

$$q_x^2 (q_x^4 - 6q_x^2 q_y^2 + 9q_y^4) , \quad (33a)$$

$$q_y^2 (q_y^4 - 6q_x^2 q_y^2 + 9q_x^4) . \quad (33b)$$

The coefficients in front of these terms are rather complicated and we have not computed them explicitly. The terms (32) and (33a) are graphically illustrated in Fig. 2.

We thus see why the triangular lattice is more amenable to homogenization: the anisotropic terms of the form (33) appear in this case only to sixth order in  $Q$ . However, as soon as these terms yield a noticeable contribution to  $\Sigma$ , isotropy is lost very fast. This observation was made in Ref. 10 and illustrated by considering the transmission coefficient of a PC slab. Below we will illustrate this observation by plotting the isofrequency lines for an extended range of  $q_x$ , which includes not only propagating incident waves (either in vacuum or in PC or in both), but also evanescent incident waves.

#### IV. NUMERICAL EXAMPLES

In the simulations, we did not use any perturbative techniques but rather solved (18) by mapping it onto a linear eigenproblem. This approach does not involve any approximations. However, numerical application of the method requires some truncation of the two-dimensional grid of the reciprocal lattice vectors. Convergence of results with the size of truncated grid was checked by (i) consecutively doubling the size of this grid and (ii)

by comparison with a high-order finite difference method (FLAME)<sup>29</sup>. We have obtained both an excellent convergence with respect to the grid size and an excellent agreement with the finite-difference method, which allows us to conclude that the numerical results shown below are accurate. By this we mean that the results obtained by different methods or by further doubling the size of the grid could not be visually distinguished from the results shown below when plotted.

The solutions obtained numerically in the actual PC will be compared to similar solutions in a truly homogeneous medium that is artificially discretized on the same lattice as that of the PC and has the refractive index determined from the approximate radius of the corresponding quasi-circular isofrequency line. To this end, we need to introduce the operation of “folding” of a general vector  $\mathbf{q}$  into the FBZ of the lattice. This operation can be formally defined as

$$[\mathbf{q}]_{\text{FBZ}} = \mathbf{q} - (n_1 \mathbf{b}_1 + n_2 \mathbf{b}_2) , \quad (34a)$$

$$n_i = \text{Nint} \left( \frac{\mathbf{q} \cdot \mathbf{a}_i}{2\pi} \right) . \quad (34b)$$

In the above equations,  $\text{Nint}(z)$  is the nearest integer to the complex number  $z$ . The vector  $\mathbf{q}$  in the right-hand side of (34a) is not restricted to the FBZ and in a homogeneous isotropic medium it satisfies the dispersion equation  $q_x^2 + q_y^2 = n^2 k^2$  where  $n$  is the index of refraction. We can view  $q_x$  as a mathematically-independent and real-valued variable, compute  $q_y$  as  $q_y = \pm \sqrt{n^2 k^2 - q_x^2}$ , substitute the pair  $(q_x, q_y)$  obtained in the manner in the right-hand side of (34a), and this will yield the dispersion equation of the artificially discretized homogeneous medium. Examples of such folding will be shown in the figures below, and the plots in Fig. 2 were obtained by plotting all real-valued points  $[\mathbf{q}]_{\text{FBZ}}$  in the plane  $(q_x, q_y)$  at the particular frequencies for which these lines cross the origin.

##### A. Square lattice

Consider a two-dimensional square lattice of infinite hollow cylinders embedded in a high-index host. The cylinder axes are aligned with the  $Z$ -axis of a rectangular frame and we assume that the propagation is in the  $XY$  plane so that  $\mathbf{q} = (q_x, q_y)$ . The lattice period is denoted by  $a$  and the radius of the cylinders is  $R = 0.33a$ . The host permittivity is  $\epsilon_h = 9.61$  and the inclusions (cylinders) are assumed to be a vacuum with  $\epsilon_i = 1$ . We thus disregard frequency dispersion either in host or in the inclusions. This is an “inversion” of the model used in Ref. 30 where a 2D square lattice of aluminum oxide rods in air was considered.

We start with the purely real dispersion diagram. The latter is obtained by setting  $q_x = 0$ , computing  $q_y$  for a range of electromagnetic frequencies, and by keeping only real solutions  $q_y$  to the dispersion equation. That is, we will disregard for the moment all complex and

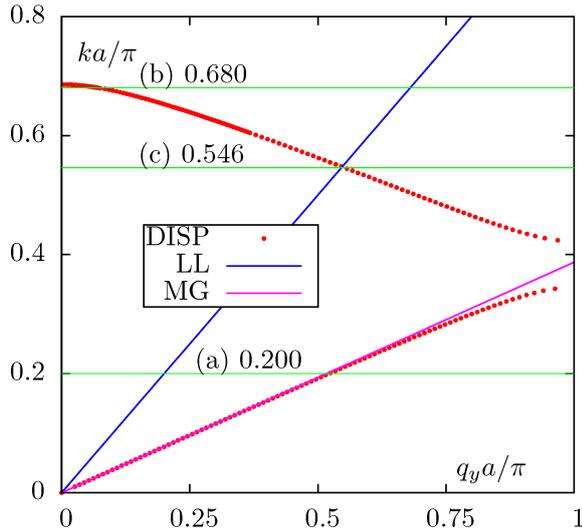


FIG. 3: (color online) Dispersion diagram for the square lattice of hollow cylinders. Frequency is scanned from zero to slightly above the second  $\Gamma$ -point. The variable  $ka/\pi = \omega a/\pi c = 2a/\lambda$  is the dimensionless frequency. The line labeled LL is the light line and the line labeled MG gives the Maxwell-Garnett approximation to the dispersion curve. Thin horizontal lines mark the frequencies at which further simulations have been performed.

imaginary wave numbers  $q_y$  (these solutions will be considered later). This approach is conventional for purely real permittivities of the constituent material. Note that we compute the dispersion diagram for a given propagation direction (along the  $Y$ -axis). Dispersion curves for different propagation directions are, generally, different. However, if a higher  $\Gamma$ -point exists according to definition (4) and the function  $\Sigma(\omega, \mathbf{q})$  is smooth in the vicinity of this  $\Gamma$ -point (see Appendix A), then the frequency of this  $\Gamma$ -point is independent of the propagation direction.

The dispersion diagram for the square lattice of hollow cylinders is shown in Fig. 3. The horizontal lines in this figure mark the dimensionless frequencies at which further simulations have been performed. The frequency  $ka/\pi = 0.200$  [Case (a)] is in the first photonic band. The squared effective refractive index at this frequency is  $n_{\text{eff}}^2 = 6.81$  as determined from the formula  $n_{\text{eff}}^2 = q_y/k$ . Note that the Maxwell-Garnett homogenization result for Case (a) is slightly different:  $n_{\text{eff}}^2 = 6.66$ . The frequency  $ka/\pi = 0.680$  [Case (b)] is in the second transmission band, very close to and slightly below the second  $\Gamma$ -point. This is the main case we are interested in. Finally, the frequency  $ka/\pi = 0.546$  [Case (c)] is at the intersection of the second dispersion branch and the light line. One can expect  $n_{\text{eff}}^2 \approx 1$  and  $n_{\text{eff}} \approx -1$  at this frequency, since the dispersion in the second photonic band is negative. However, we will see that the medium is very far from being electromagnetically-homogeneous in cases (b) and (c) and cannot be characterized by local effective param-

eters.

The isofrequency lines for the three frequencies noted above are shown in Figs. 4-6. As was explained in Sec. II, we fix the frequency and view  $q_x$  as a mathematically-independent and purely real variable, which is scanned in some interval. For each  $q_x$  considered, we find all values of  $q_y$  whose imaginary parts are restricted to some sufficiently large range.

In Fig. 4, we plot the real and imaginary parts of  $q_y$  as functions of  $q_x$ . For each  $q_x$ , there exist infinitely many solutions to the dispersion equation with  $\text{Re}(q_y) = 0$  and  $\text{Im}(q_y) \neq 0$  and we cannot display all such points in the plots. However, the number of solutions with  $\text{Re}(q_y) \neq 0$  is finite, and all such data points are shown in the figure. Note that for each solution  $(q_x, q_y)$  in which both  $q_x$  and  $q_y$  are real, there is also a solution  $(q_y, q_x)$ . However, this symmetry is broken if  $q_y$  has an imaginary part. For this reason, the upper plots are not completely symmetric with respect to the line  $q_y = q_x$  (the diagonal of the square that runs from the bottom-left to the upper-right corner). However, the lobes in the lower-left and upper-right corners consist of purely real solutions and these lobes are symmetric. The line that connects these two lobes consists of complex solutions  $q_y$  with both real and imaginary parts different from zero. Therefore, these connecting lines are not symmetric. We note that such complex solutions cannot be obtained in a truly homogeneous medium and are therefore a manifestation of the complex structure of the PC.

Referring to the data of Fig. 4, we can conclude that at the frequency  $ka/\pi = 0.200$  the dispersion relation in the PC is isotropic and almost indistinguishable from the dispersion relation in a homogeneous medium with  $n^2 = 6.81$ . The correspondence holds well into the evanescent waves. Indeed, the range of  $q_x$  shown in the figure for this frequency is  $0 \leq q_x \leq 5k$  and we can expect that the PC is electromagnetically homogeneous for  $-5k \leq q_x \leq 5k$  (we have extended the interval to negative values due to the obvious reflection symmetry). In fact, the PC is electromagnetically homogeneous in an even wider interval of  $q_x$ . Indeed, the correspondence still holds in the “reflected” segments of the curves in the bottom plot. In homogeneous media, these reflected segments correspond to the values of  $|q_x| > 5k$ , which were folded into the “artificial” FBZ of the lattice according to (34). The only slight discrepancy between the PC and the homogeneous medium can be observed near the edge of the FBZ ( $q_x \approx 5k$ ). Overall, at this frequency, the PC mimics a homogeneous medium with very high precision and in a very wide range of illumination conditions.

Now let us move to Case (b),  $ka/\pi = 0.680$ . This frequency is slightly below the second  $\Gamma$ -point. As one could expect from the theoretical arguments of Sec. III, the upper plot has a nearly circular, purely real lobe in the lower-left corner of the frame. If only this lobe is considered, one might erroneously conclude that the dispersion is isotropic in the PC at the frequency  $ka/\pi = 0.680$  as well and, therefore, the law of dispersion is almost

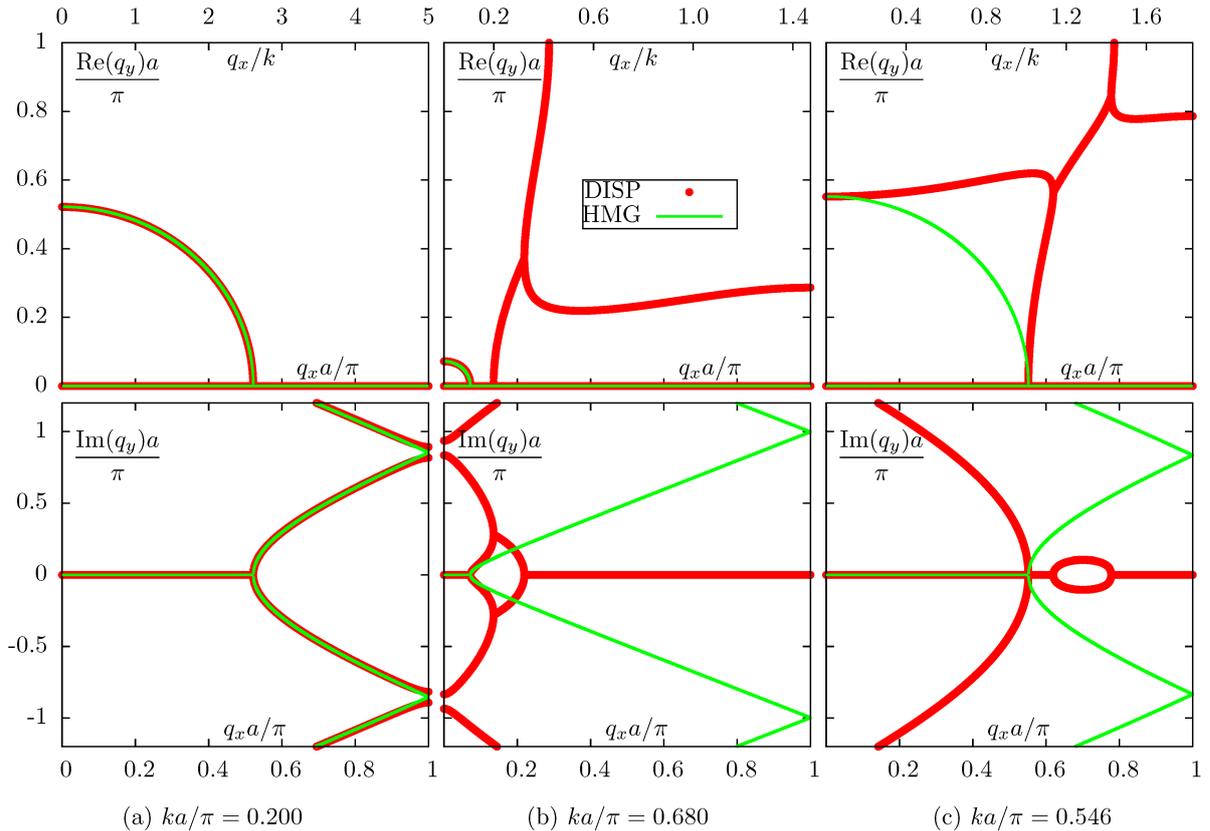


FIG. 4: (color online) Real (upper row of plots) and imaginary (bottom row) parts of  $q_y$  as functions of  $q_x$  (real by definition) for the square lattice of hollow cylinders at various values of the dimensionless frequency  $ka/\pi$ . Only one quadrant of the FBZ is shown in the upper row of plots; this quadrant can be replicated by using the  $C_4$  rotational symmetry of the problem to cover the whole FBZ. The scale of the lower horizontal axes is  $\pi/a$  and the scale of the upper horizontal axes is  $1/k$ , as labeled. All data points with  $\text{Re}q_y = 0$  (shown in the upper row) have nonzero imaginary parts except for the data point  $q_y = 0$  at the intersection of the quasi-circular lobe and the horizontal axis. The corresponding values of  $q_y$  are purely imaginary. The imaginary parts of some of these data points are outside of the plotting range in the lower row of plots. These data points are shown only in the upper row. Data points are labeled 'DISP' for the actual PC and 'HMG' for an electromagnetically homogeneous medium that is artificially discretized on the same lattice. The HMG isofrequency lines were computed according to (34) with  $n^2 = 6.81$  (a),  $n^2 = 0.011$  (b) and  $n^2 = 0.91$  (c). In the case (c), the choice of  $n^2$  guarantees the correct wave number for propagation along  $X$ - or  $Y$ -axes but not for the intermediate directions.

the same as in a homogeneous medium. But a quick purview of the scale of the upper horizontal axis reveals that this isotropic behavior holds only in a very narrow range  $-0.15 \lesssim q_x/k \lesssim 0.15$ . In problems of transmission through a flat interface, this corresponds to incident angles of less than about  $6^\circ$  with respect to the normal. Outside of the above range of  $q_x$ , the isotropy is lost as can be readily seen in the bottom plot. We emphasize that the isotropy is lost in this case for  $|q_x| < k$ . These values of  $q_x$  correspond to propagating waves in a vacuum. It is true that these waves are evanescent in the PC but, upon transmission through a finite slab, incoming propagating waves are always transformed into outgoing propagating waves. Therefore, the transmission properties of the PC at this frequency are very different from those of any homogeneous medium. Note that, at sufficiently large values of  $q_x$ , the waves in PC switch

from being evanescent cito being propagating again (in the upper-right corner lobe). This behavior is not characteristic of any homogeneous medium. Moreover, it can be seen that this PC is characterized by birefringence in some range of  $q_x$  even though no anisotropy is involved - again, an effect not observed in homogeneous media. By birefringence we mean the effect when a single value of  $q_x$  (corresponding to a fixed angle of incidence) gives rise to transmitted waves with two real but different values of  $q_y$  (two different angles of refraction). A narrow ray incident at the interface at this angle will split into two rays refracted at different angles even though both rays have the same linear polarization. This phenomenon cannot occur in any local homogeneous medium.

In the Case (c),  $ka/\pi = 0.546$ , there is obviously no hope to approximate the the law of dispersion of the PC by the law of dispersion of any homogeneous medium.

The quasi-circular isotropic lobe of the isofrequency line is severely distorted and birefringence is quite prominent and occurs in a wide range of incident angles. As expected, the distortion is consistent with the  $C_4$  symmetry of the problem; compare the distortion of the quasi-circular lobe on the upper plot with the shape shown in Fig. 2(a). Note, however, that this type of distortion is suppressed and much less pronounced in the triangular lattice as will be shown below.

Next, we show the purely real quasi-circular lobes of the isofrequency lines in more detail in Fig. 5. It can be seen that the lobe is almost indistinguishable from a mathematical circle at the frequency  $ka/\pi = 0.200$ . A distortion consistent with  $C_4$  rotational symmetry is clearly visible at the frequency  $ka/\pi = 0.680$ . At  $ka/\pi = 0.546$  (far from the  $\Gamma$ -point), this distortion is quite severe.

But perhaps the most clear illustration of the departure from the homogeneous behavior in the cases (b) and (c) is shown in Fig. 6, where we plot  $\text{Re}(q_y^2)$  as a function of  $q_x^2$ . In a homogeneous medium, this function is simply a straight line. This line is folded into the FBZ of the lattice as is shown in Fig. 6(a) if the medium is artificially discretized (we note that for the particular parameters and the lattice type considered at the moment, the folding results only in linear segments; in a more general case the isofrequency line can acquire curvature due to the folding and an example will be shown below). This behavior is reproduced with very good accuracy at the frequency  $ka/\pi = 0.200$  [Case (a)]. However, in Cases (b) and (c) the departure from the linear behavior is obvious and dramatic. We will observe a similar behavior in a triangular lattice as well.

We finally note that the complex branches of the isofrequency lines that connect the purely imaginary and purely real segments of the data point sets is a peculiar feature of the PC which cannot be reproduced in any homogeneous medium with real refractive index. In the latter case  $q_y$  is either purely real or purely imaginary and  $q_y^2$  is always real.

## B. Triangular lattice

We now turn to the case of the triangular lattice. The primitive vectors for the real-space and reciprocal lattices are

$$\begin{aligned} \mathbf{a}_1 &= a(1, 0) , \quad \mathbf{a}_2 = a(1/2, \sqrt{3}/2) , \\ \mathbf{b}_1 &= \frac{2\pi}{a}(1, -1/\sqrt{3}) , \quad \mathbf{b}_2 = \frac{2\pi}{a}(0, 2/\sqrt{3}) . \end{aligned}$$

At the center of each elementary cell, which is now rhombic, we place a hollow cylinder of the radius  $R = 0.42a$ . These inclusions do not cross the boundaries of the elementary cell. It can be seen that the inclusions form a perfect triangular lattice. Since the inclusion obeys the same symmetry as the lattice, the whole structure is  $C_6$ -

symmetric. The permittivity of the host matrix is taken to be  $\epsilon_h = 12.25$ . This model was used in Ref. 31.

The FBZ of the lattice described above is a rhombus and any of the two triangles created by drawing a rhombus diagonal are equivalent. One can construct a hexagon of two complete rhombuses touching at one corner and two such triangles. All six triangles forming this hexagon are equivalent and, moreover, the dispersion points can be replicated periodically on the hexagonal lattice. It is a common practice to plot the isofrequency lines of such lattices inside a hexagonal region of the reciprocal lattice. However, we do not use this approach in the paper and limit attention to the actual FBZ of the lattice, which is shown in all figures below.

The dispersion diagram for the direction of propagation along the  $Y$ -axis ( $q_x = 0$ ) is shown in Fig. 7. As previously, three special frequencies are marked by the horizontal lines in Fig. 7. These frequencies are  $ka/\pi = 0.200$  [Case (a)],  $ka/\pi = 0.720$  [Case (b)] and  $ka/\pi = 0.650$  [Case (c)]. Just as was the case for the square lattice, the frequency (a) is in the first photonic band, the frequency (b) is in the second band slightly below the  $\Gamma$ -point and the frequency (c) is close to the intersection of the light line and the second branch of the dispersion curve.

The complex isofrequency lines for the three frequencies mentioned above are shown in Fig. 8. This figure is analogous to Fig. 4 and the same quantities are plotted using the same scales of the axes. However, the FBZ of a triangular lattice is more complex geometrically. Therefore, in the top row of plots of Fig. 8, we have shown the complete FBZ of the lattice (the black rhombus) while in Fig. 4, only one quarter of the FBZ was shown.

We now analyze Fig. 8 in more detail. First, focus on the real parts of  $q_y$  (the upper row of plots). The central quasi-circular lobe and the horizontal line  $\text{Re}(q_y) = 0$  are analogous to the similar features of the isofrequency lines for a square lattice. The quasi-circular lobes consist of purely real solutions while the line  $q_y = 0$  corresponds to purely imaginary solutions. The upper and lower horizontal lines in the upper row of plots correspond to complex solutions that are specific to the triangular lattice. Unlike in the case of a square lattice, these complex solutions do not connect purely imaginary and purely real segments of the complex isofrequency line (within the FBZ). Another distinction is that the complex solutions in a triangular lattice will appear due to folding of the dispersion equation of an artificially discretized homogeneous medium. For a square lattice, the appearance of complex solutions is the result of interaction that cannot be obtained by artificial folding.

We can understand the appearance of the complex solutions shown in Fig. 8 by replicating the central rhombus of the top row of plots in all directions and noting that the horizontal line  $q_y = 0$  in the central rhombus will connect to the upper or lower horizontal lines in the replicated rhombuses. We can also understand these solutions qualitatively by considering a homogeneous medium artificially discretized on a triangular lat-

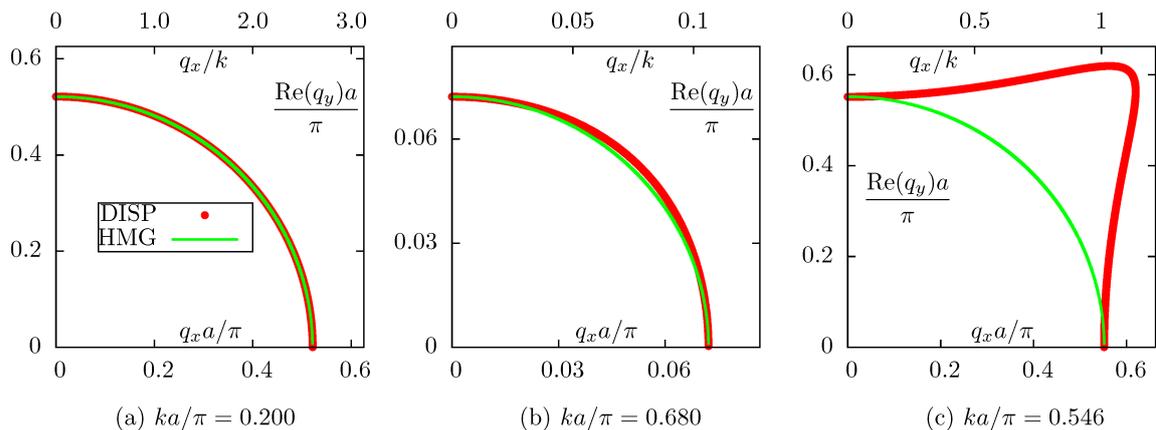


FIG. 5: (color online) Purely real parts of the isofrequency contours shown in Fig. 4. The lines marked 'HMG' are ideal circles.

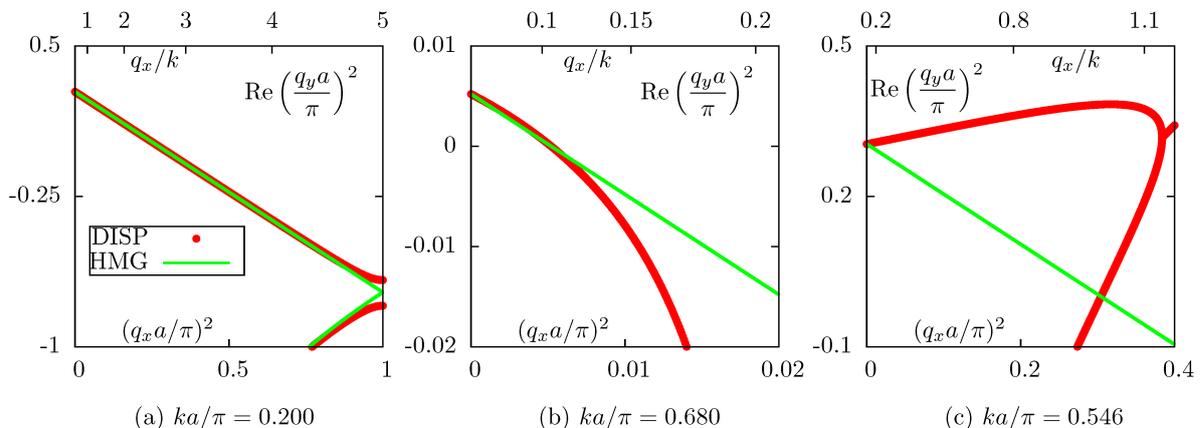


FIG. 6: (color online) Isofrequency contours for the squares of the Cartesian components of Bloch wave vector  $\mathbf{q}$ . More precisely,  $\text{Re}(q_y^2)$  is plotted as a function of  $q_x^2$  for the same set of fixed frequencies that were considered in Figs. 4 and 5. In a homogeneous medium artificially discretized on a square lattice, this dependence is piece-wise linear as is shown in panel (a). For the triangular lattice, the corresponding curves consist of one liner and multiple curved segments (see an illustration below).

tice. In the infinite (non-discretized) medium, the law of dispersion allows for purely imaginary solutions of the form  $q_y = \pm i\sqrt{q_x^2 - n^2 k^2}$  for  $|q_x| > nk$ ,  $n > 0$  being the index of refraction. Now let us fold these solutions to the FBZ of a triangular lattice according to (34). We can ignore the imaginary part of  $\mathbf{q}$  for the purpose of computing the integers  $n_1, n_2$  that are used in this equation. The result of the folding is

$$[q_x]_{\text{FBZ}} = q_x - \frac{2\pi}{a} n_1, \quad (35a)$$

$$[q_y]_{\text{FBZ}} = \pm i\sqrt{q_x^2 - n^2 k^2} - \frac{2\pi}{a\sqrt{3}} (2n_2 - n_1), \quad (35b)$$

$$n_1 = \text{Nint}\left(\frac{q_x a}{2\pi}\right), \quad n_2 = \text{Nint}\left(\frac{q_x a}{4\pi}\right). \quad (35c)$$

These equations are valid for  $|q_x| > nk$ . We note that the integer index  $m = 2n_2 - n_1$  can take only three values:  $0, \pm 1$ . Correspondingly, for the real and imaginary parts

of  $[q_y]_{\text{FBZ}}$ , we have the following results:

$$[\text{Re}(q_y)]_{\text{FBZ}} = m \frac{2\pi}{a\sqrt{3}}, \quad m = 0, \pm 1, \quad (36a)$$

$$[\text{Im}(q_y)]_{\text{FBZ}} = \pm \sqrt{q_x^2 - n^2 k^2}. \quad (36b)$$

We will obtain a dispersion equation containing only the quantities  $[q_x]_{\text{FBZ}}$  and  $[q_y]_{\text{FBZ}}$  if we substitute  $q_x$  in (36b) from (35a), i.e., use the relation

$$q_x^2 = \left\{ [q_x]_{\text{FBZ}} + \frac{2\pi}{a} n_1 \right\}^2 \quad (37)$$

to obtain the following closed-form equation:

$$[\text{Im}(q_y)]_{\text{FBZ}} = \pm \sqrt{\left\{ [q_x]_{\text{FBZ}} + \frac{2\pi}{a} n_1 \right\}^2 - n^2 k^2} \quad (38)$$

$$\text{for } \left| [q_x]_{\text{FBZ}} + \frac{2\pi}{a} n_1 \right| \geq nk.$$

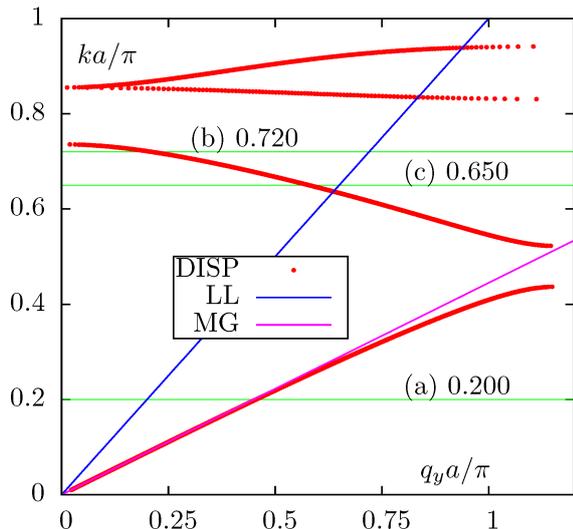


FIG. 7: (color online) Same as in Fig. 3 but for the triangular lattice of hollow cylinders of the radius  $R = 0.42a$  in the host medium of the permittivity  $\epsilon = 12.25$ .

This equation describes a family of hyperbolas parameterized by  $n_1$ . The first pair of these hyperbolas (corresponding to  $n_1 = 0$ ) are labeled as FA and FB in Fig. 8. The second pair of hyperbolas (including two branches with  $n_1 = \pm 1$ ) are labeled as SA and SB. Infinitely many similar curves can be generated by translations along the  $q_x$ -axis, which correspond to arbitrary integer values of  $n_1$ . Note that the two integers  $n_1$  and  $n_2$  that label different “reflected” branches of the solution to the dispersion equation are subject to the selection rule  $m = n_1 - 2n_2 = 0, \pm 1$ . Therefore, for each  $n_1$ , the set of allowable values of  $n_2$  is restricted.

Thus, the FBZ folding in a triangular lattice transforms purely imaginary solutions  $q_y$  into complex solutions  $[q_y]_{\text{FBZ}}$ , which explains the appearance of the upper and lower horizontal lines in the top row of plots in Fig. 8.

Of course, the analytical folding described above is valid only for a homogeneous medium. However, at the frequency  $ka/\pi = 0.200$ , the actual dispersion relation in the PC mimics the dispersion relation in a homogeneous medium with  $n^2 = 5.20$  very closely. This conclusion can be drawn from the data shown in the lower plot for the Case (a) in Fig. 8 and is further confirmed and illustrated in Fig. 9. In the latter figure, we plot  $[\text{Im}(q_y)]_{\text{FBZ}}$  as a function of  $[q_x]_{\text{FBZ}}$  (the symbol  $[\dots]_{\text{FBZ}}$  that signifies Bloch-periodicity of all data points is omitted in this and other figures for brevity). The comparison is made between the respective quantities obtained for the actual PC and a homogeneous medium artificially discretized on the same lattice. The individual hyperbolas described by (38) are shown with green lines and the actual dispersion solutions in the PC are shown by the red points. Both

solutions are only valid within the FBZ (between the two vertical lines) and beyond these two lines they must be periodically replicated. The green lines outside of the FBZ are shown in the figure only to guide the eye (to help visually identify individual hyperbolas).

It can be seen that, at the dimensionless frequency  $ka/\pi = 0.200$ , the law of dispersion in the PC is almost indistinguishable from the law of dispersion in a homogeneous medium, at least up to  $|\text{Im}(q_y)a/\pi| \leq 5$ , which corresponds, approximately, to  $|q_x/k| \lesssim 25$ , that is, very far into the evanescent spectrum. Therefore, we can claim that, at this particular frequency, the PC can be homogenized for many practical purposes.

The situation is quite different at the other frequencies considered. In Case (b) ( $ka/\pi = 0.720$ , just below the second  $\Gamma$ -point), the dependence of  $\text{Re}(q_y)$  on  $q_x$  still looks very “homogeneous”. However, as soon as we look at  $\text{Im}(q_y)$ , it becomes obvious that the law of dispersion departs from that of a homogeneous medium quite dramatically as soon as  $q_x$  approaches the region of evanescent waves ( $|q_x| \gtrsim k$ ). Definitely, homogenization is not possible for incident evanescent waves, and it is inaccurate for propagating waves with large angles of incidence, e.g., for  $\theta_{\text{inc}} \gtrsim 70^\circ$ . For the frequency (c) [ $ka/\pi = 0.650$ ], the quasi-circular lobe is much larger and appears to not be distorted. However, the lower plot again clearly indicates a lack of correspondence between the law of dispersion of a homogeneous medium and the law of dispersion in the PC.

We therefore conclude that the PC is not electromagnetically homogeneous at the frequencies (b) and (c). Assigning the medium some effective parameters at these frequencies can potentially be a valid approximation only for a limited range of incident angles that, at the very least, does not include evanescent waves.

Next, as was done above for the square lattice, we show in Fig. 10 the quasi-circular lobes of the top row of Fig. 8 in more detail. In order to make the small deviations from circularity more visible, we plot in this figure only one quarter of each quasi-circular lobe. The lobes appear to be indistinguishable from mathematical circles in Cases (a) and (b) but some small distortions consistent with the  $C_6$  symmetry are visible in Case (c). The high quality of the quasi-circular lobes can be explained by the fact that the terms of the form (32) in the expansion of  $\Sigma(\omega, \mathbf{q})$  cancel because they are not compatible with the  $C_6$  symmetry of the problem. Therefore, the distortions result from six-order terms in the perturbation expansion of  $\Sigma$  while, for a square lattice, the distortions appear in the fourth order. However, as soon as the six-order terms of the form (33) become non-negligible, the isotropy of  $\Sigma(\omega, \mathbf{q})$  is very quickly lost. Therefore, the high quality of the quasi-circular lobes shown in Fig. 10 is not a sufficient condition for homogenizability.

But the most clear demonstration of non-homogenizability of the PC at the frequencies (b) and (c) can be obtained by considering the squares of the Cartesian components of  $\mathbf{q}$ . In Fig. 11, we plot

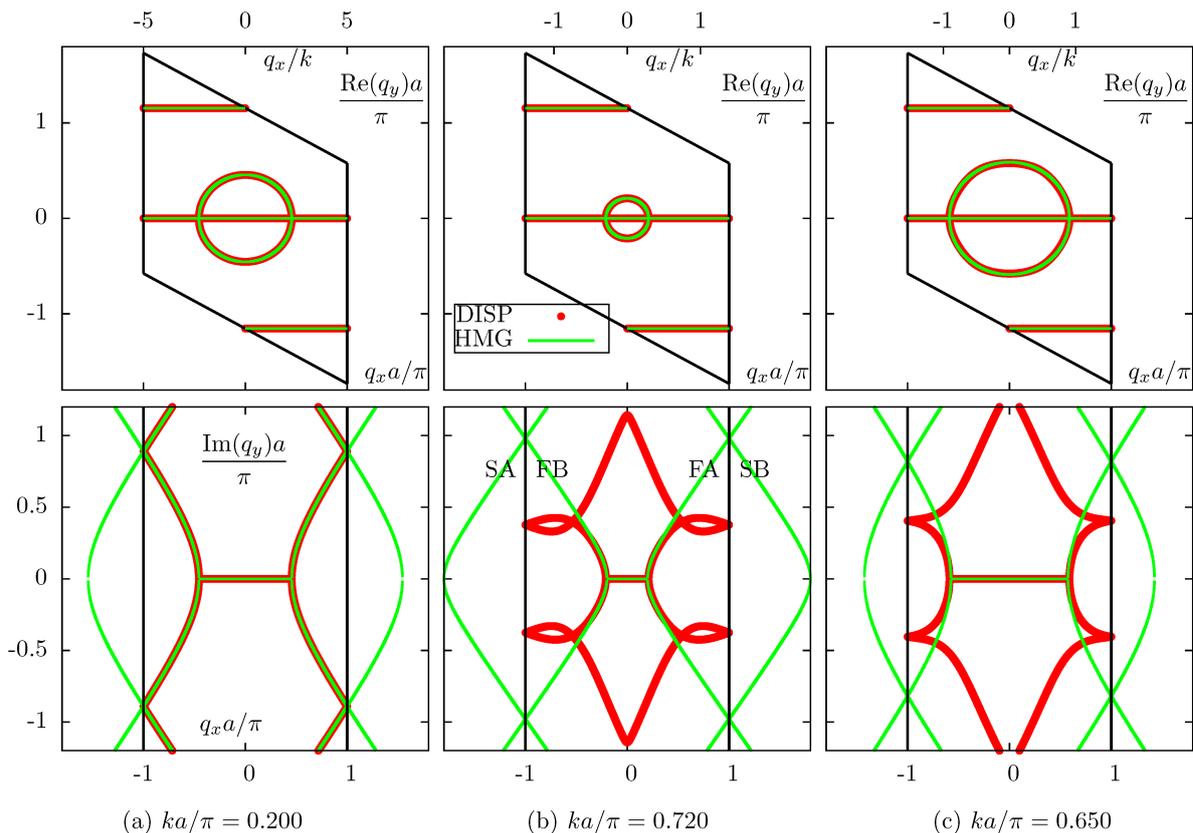


FIG. 8: (color online) Same as in Fig. 4 but for the triangular lattice of hollow cylinders. The 'HMG' isofrequency contours were computed according to (34) with  $n^2 = 5.20$  (a),  $n^2 = 0.085$  (b) and  $n^2 = 0.82$  (c). In the top row of plots, the FBZ of the lattice is shown by the black rhombus. In the lower plot (b), FA and FB mark the first pair of hyperbolas defined by Eq. 38, and SA, SB mark the second pair of hyperbolas.

$\text{Re}(q_y^2)$  vs  $q_x^2$ . In a homogeneous medium artificially discretized on a triangular lattice, this plot consists of one linear and several curved segments. The linear segment corresponds to the eigenvalues that are either purely real or purely imaginary. The curved segments correspond to the peculiar complex eigenvalues that are specific to the triangular lattice and described mathematically by the equations (35) through (38). One such curved segment is shown in every panel of Fig. 11; other curved segments lie outside of the plot frames. We now observe that at the frequency (a) there is a complete correspondence between the artificially discretized homogeneous medium and the PC. This, of course, could have been expected from the data shown in Figs. 8,9,10. At the frequencies (b) and (c), the correspondence is broken in the linear segment. In addition the curved segments in the homogeneous medium and in the PC are completely different. The last point is significant and deserves an additional discussion.

Let us assume that a slab of a homogeneous material contained between two planes  $y = 0$  and  $y = L$  is illuminated by a plane wave with the incidence angle such that  $q_x < \pi/a$ , where  $a$  is the period of artificial discretization

in the  $x$ -direction. We can describe the medium as homogeneous (the traditional approach) or as a PC (by using artificial discretization). Both approaches are mathematically equivalent and will yield the same results for all observables. Assume that we have decided to describe the medium as a PC. In this case, for a given purely real  $q_x$ , there will be infinitely many eigenvalues  $q_y$ . However, the incident radiation will excite only one mode, namely, the mode with  $q_y$  that corresponds to the dispersion relation of the homogeneous material. The modes with other values of  $q_y$  can be excited if we take  $q_x$  to be outside of the FBZ of the lattice. In any case, for a given  $q_x$ , only one mode is excited in the material.

In the case of an actual PC whose law of dispersion closely mimics that of a homogeneous medium, as was the case at the frequency (a), we can expect that the same selection rules will work: at any given  $q_x$ , only one mode will be excited in the PC. Then the transmission and reflection coefficients can be expected to be the same in the PC and the homogeneous medium. This is indeed the case in the homogenization limit<sup>25,32</sup>  $a \rightarrow 0$ .

If we consider the PC with a constant nonzero  $a$  at sufficiently high excitation frequencies (such as the fre-

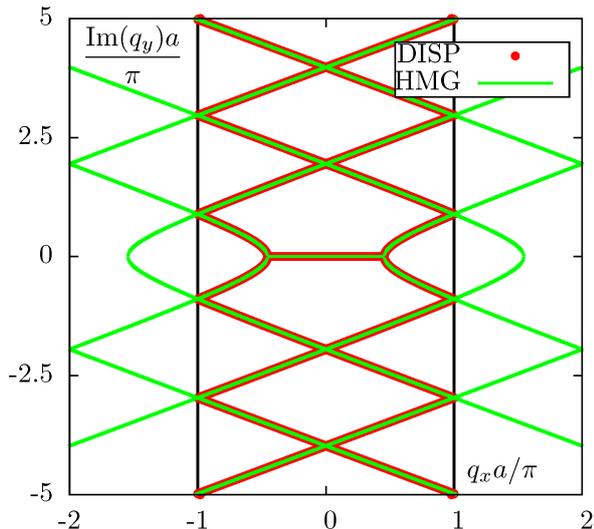


FIG. 9: (color online) An expanded view of  $\text{Im}(q_y)$  as a function of  $q_x$  for  $ka/\pi = 0.200$  [Case (a)]. Here several 'reflected' segments of the isofrequency line are shown. These segments are obtained by folding the corresponding isofrequency line for an infinite homogeneous medium into the FBZ of the triangular lattice according to (38). The analytical lines outside the FBZ (shown by the two vertical lines) are not the actual solutions and are shown only to guide the eye.

quencies (b) and (c) in the above examples), there is no reason to believe that the selection rules will work the same way as in a homogeneous medium. In other words, at a given  $q_x$ , modes with several different values of  $q_y$  can be excited in the PC. Granted, these additional modes have  $q_y$  with nonzero imaginary parts and are therefore exponentially decaying inside the medium. However, it is a mistake to neglect these modes completely. Indeed, even the fields associated with these modes are exponentially decaying with  $y$ , they are not negligibly small at the interface  $y = 0$ . Moreover, these modes do not generally average to zero over the surface of the elementary cell. In these respects, they are different from the surface waves discussed by us previously<sup>25,33</sup>. Therefore, excitation of these additional modes will have an adverse effect on homogenizability.

In Fig. 11, two different effects are illustrated. The first effect is the deviation of the law of dispersion in the PC from that in a homogeneous medium for the "fundamental mode" (for which  $q_y^2$  is always purely real). This effect is manifest at sufficiently large incident angles and, in particular, for incident evanescent waves. The second, more subtle effect is manifest even at small incident angles. Namely, it can be seen that, at  $q_x \sim 0$ , the PC has additional modes (allowable values of  $q_y$ ) that are dramatically different from the respective values in an artificially discretized homogeneous medium. Appearance of these additional solutions can be expected to influence the medium impedance in an angle-dependent manner

and have an additional degrading effect on the PC homogenizability.

## V. DISCUSSION

The main message of this paper is that, in many applications of practical interest, it is insufficient to consider only the purely real segments of the isofrequency lines to decide whether a given photonic crystal (PC) is electromagnetically similar to an isotropic homogeneous medium. The purely real segments can be circular with reasonable precision in the higher photonic bands. In this case, all propagation directions appear to be equivalent and the physical effects of the medium discreteness appear to be minimized or absent. The medium can also be characterized by negative dispersion, which means that the real part of the Bloch wave number tends to decrease with frequency while the imaginary part is negligibly small. Nevertheless, the PC is not electromagnetically homogeneous in this case and cannot be characterized by angle-independent, purely local effective parameters  $\epsilon$  and  $\mu$ . This conclusion can be drawn by considering the angular dependence of the effective parameters and by including evanescent waves into consideration.

A difficulty one faces when restricting consideration to the dispersion relation is that infinite media have no interfaces and therefore it is easy to overlook the important mathematical features of the obtained solutions. This is exactly what happens when one restricts consideration to purely real isofrequency lines. In this paper, we have generalized this approach by considering two orthogonal directions in space,  $X$  and  $Y$ , and assuming that the projection of the Bloch wave vector  $\mathbf{q}$  on one of these axes ( $X$  in our case) is a mathematically-independent variable, which is preserved by the process of reflection and refraction at any planar interface  $x = \text{const}$ .

The results obtained above are consistent with an earlier prediction made by one of the authors regarding the impossibility of negative refraction<sup>34</sup>. The essential assumption of the above reference was that the medium in question is electromagnetically homogeneous. However, this requirement was not clearly defined. On the other hand, it is a common knowledge that some discrete systems such as photonic crystals or chains of interacting particles or resonators can be characterized by *negative dispersion*. This is true, in particular, for PCs in higher photonic bands. This may seem to contradict the conclusions of Ref. 34. In the present work, we show that there is no contradiction since PCs are not homogenizable in the higher bands. We also define more clearly what we mean by the requirement of homogenizability. We note that this requirement has been stated in more mathematical detail in Ref. 32, and the impossibility of negative refraction (in the regime when the medium is homogenizable) was demonstrated in Ref. 35 for the special case of 1D periodic media. The latter case is however compli-

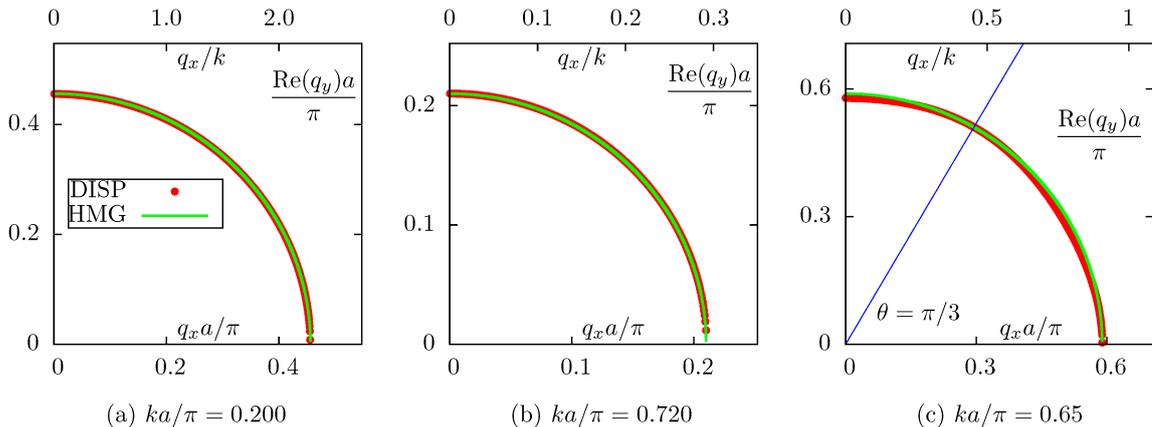


FIG. 10: (color online) Same as in Fig. 5 but for the triangular lattice. Distortions of the quasi-circular lobes are consistent with  $C_6$  symmetry. The straight line in panel (c) makes the  $\pi/3$  angle with the horizontal axis.

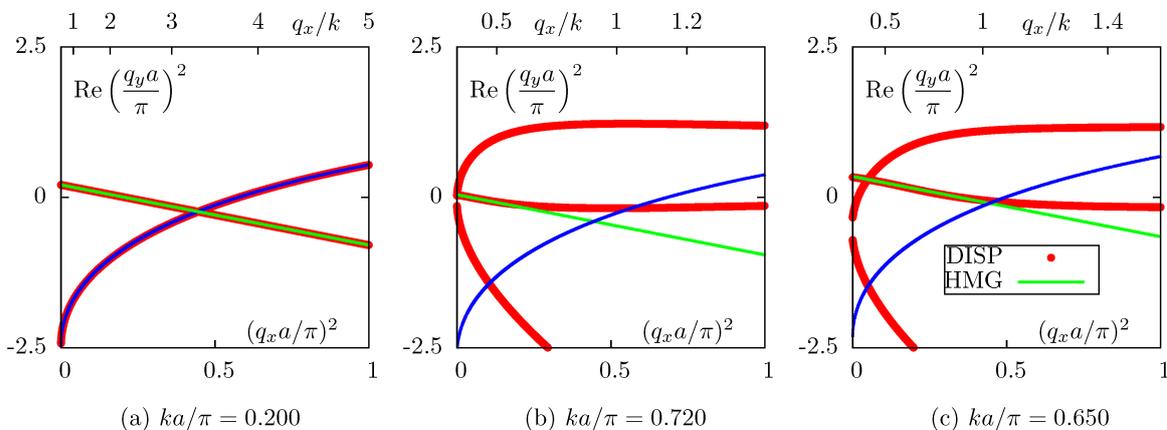


FIG. 11: (color online) Same as in Fig. 6 but for the triangular lattice that is considered in this subsection.

cated by the fact that a 1D layered medium is necessarily anisotropic. Certain types of indefinite anisotropic media are capable of refracting a narrow incident beam on the same side of the normal. This phenomenon can easily be confused with negative refraction. In this paper, we consider a system without anisotropy (either magnetic or electric).

It should be noted that the condition of homogenizability based on the correspondence of the dispersion relation (between a hypothetical homogeneous medium and the actual PC) is a necessary but not the sufficient condition. The sufficient condition must include the medium impedance. This cannot be done without introducing the medium boundary into the problem. This fact was emphasized by us earlier in Ref. 20 using the exactly solvable model of 1D periodic medium as an example.

A more general result closely related to the main con-

clusion of this paper can be stated in the form of an *uncertainty principle* of homogenization, namely: the more an effective magnetic permeability deviates from unity, the less accurate the corresponding homogenization result is<sup>24</sup>.

We finally note that this work is based upon some important theoretical observations made by Li, Holt and Efron<sup>10</sup>. We have developed these observations further with a specific focus on dispersion relations in 2D PCs with square and triangular lattices.

This work has been carried out thanks to the support of the A\*MIDEX project (No. ANR-11-IDEX-0001-02) funded by the “Investissements d’Avenir” French Government program, managed by the French National Research Agency (ANR) and was also supported by the US National Science Foundation under Grant DMS1216970.

\* On leave from the Department of Radiology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA;

Electronic address: [vmarkel@fresnel.fr](mailto:vmarkel@fresnel.fr), [vmarkel@mail.](mailto:vmarkel@mail.)

med.upenn.edu

<sup>†</sup> Electronic address: igor@uakron.edu

- <sup>1</sup> U. C. Hasar, G. Buldu, M. Bute, J. J. Barroso, T. Karacali, and M. Ertugrul, *AIP Advances* **4**, 107116 (2014).
- <sup>2</sup> T. D. Karamanos, S. D. Assimonis, A. I. Dimitriadis, and N. V. Kantartzis, *Photonics and Nanostructures* **12**, 291 (2014).
- <sup>3</sup> N. C. J. Clausen, S. Arslanagic, and O. Breinbjerg, *Photonics and Nanostructures* **12**, 419 (2014).
- <sup>4</sup> A. Ciattoni and C. Rizza, *Phys. Rev. B* **91**, 184207 (2015).
- <sup>5</sup> V. Sozio, A. Vallechi, M. Albani, and F. Capolino, *Phys. Rev. B* **91**, 205127 (2015).
- <sup>6</sup> J. B. Pendry, *Phys. Rev. Lett.* **85**, 3966 (2000).
- <sup>7</sup> A. L. Pokrovsky and A. L. Efros, *Solid State Comm.* **124**, 283 (2002).
- <sup>8</sup> S. G. Rautian, *Phys. Usp.* **51**, 981 (2008).
- <sup>9</sup> There can be no electric anisotropy because the electric field in the considered geometry is always aligned with the same axis, say,  $Z$ . One can attempt to introduce magnetic anisotropy as a phenomenological adjustable parameter, but then one would obtain either a non-symmetric magnetic tensor and hence non-reciprocity, which is not physically present in the system, or otherwise a tensor whose principal axes do not coincide with the crystallographic axes (e.g., for the triangular lattice), or else a tensor that is trivially proportional to the identity (e.g., for the square lattice). Therefore, anisotropic effective magnetic tensor is also not a possibility.
- <sup>10</sup> C. Y. Li, J. M. Holt, and A. L. Efros, *J. Opt. Soc. Am. B* **23**, 963 (2006).
- <sup>11</sup> C. Menzel, C. Rockstuhl, T. Paul, F. Lederer, and T. Pertsch, *Phys. Rev. B* **77**, 195328 (2008).
- <sup>12</sup> C. Menzel, C. Rockstuhl, R. Iliew, F. Lederer, A. Andryeuskii, R. Malureanu, and A. V. Lavrinenko, *Phys. Rev. B* **81**, 195123 (2010).
- <sup>13</sup> M. G. Silveirinha, *Phys. Rev. B* **75**, 115104 (2007).
- <sup>14</sup> C. R. Simovski, *Opt. Spectrosc.* **107**, 766 (2009).
- <sup>15</sup> A. Alu, *Phys. Rev. B* **84**, 075153 (2011).
- <sup>16</sup> V. Agranovich and V. Ginzburg, *Spatial Dispersion in Crystal Optics and the Theory of Excitons* (Wiley-Interscience, New York, 1966).
- <sup>17</sup> L. D. Landau and L. P. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon Press, Oxford, 1984).
- <sup>18</sup> V. M. Agranovich and Y. N. Gartstein, *Phys. Usp.* **49**, 1029 (2006).
- <sup>19</sup> V. M. Agranovich and Y. N. Gartstein, *Metamaterials* **3**, 1 (2009).
- <sup>20</sup> V. A. Markel and I. Tsukerman, *Phys. Rev. B* **88**, 125131 (2013).
- <sup>21</sup> C. Luo, S. G. Johnson, J. D. Joannopoulos, and J. B. Pendry, *Phys. Rev. B* **65**, 201104 (2002).
- <sup>22</sup> C. Luo, S. G. Johnson, and J. D. Joannopoulos, *Appl. Phys. Lett.* **81**, 2352 (2002).
- <sup>23</sup> R. V. Craster, J. Kaplunov, E. Nolde, and S. Guenneau, *J. Opt. Soc. Am. A* **28**, 1032 (2011).
- <sup>24</sup> I. Tsukerman and V. A. Markel, *Phys. Rev. B* (2015), submitted.
- <sup>25</sup> V. A. Markel and J. C. Schotland, *Phys. Rev. E* **85**, 066603 (2012).
- <sup>26</sup> C. Fietz and G. Shvets, *Physica B* **405**, 2930 (2010).
- <sup>27</sup> Usually understood as the regime in which Taylor expansion of  $\epsilon(\omega, \mathbf{q})$  to second order in  $\mathbf{q}$  is an accurate approximation.
- <sup>28</sup> J. D. Joannopoulos, R. D. Meade, and J. N. Winn, *Photonic Crystals: Molding the Flow of Light* (Princeton University Press, Princeton, N.J., 2008).
- <sup>29</sup> I. Tsukerman and F. Čajko, *IEEE Trans. Magnetics* **44**, 1382 (2008).
- <sup>30</sup> R. Gajic, R. Meisels, F. Kuchar, and K. Hingerl, *Opt. Expr.* **13**, 8596 (2005).
- <sup>31</sup> T.-H. Pei and Y.-T. Huang, *J. Opt. Soc. Am. B* **29**, 2334 (2012).
- <sup>32</sup> I. Tsukerman and V. A. Markel, *Proc. R. Soc. A* **470**, 0245 (2014).
- <sup>33</sup> X. Y. Z. Xiong, L. J. Jiang, V. A. Markel, and I. Tsukerman, *Opt. Expr.* **21**, 10412 (2013).
- <sup>34</sup> V. A. Markel, *Opt. Expr.* **16**, 19152 (2008).
- <sup>35</sup> V. A. Markel and J. C. Schotland, *J. Opt.* **12**, 015104 (2010).
- <sup>36</sup> I. Tsukerman, *J. Opt. Soc. Am. B* **25**, 927 (2008).

## APPENDIX A: DISPERSION RELATION IN A BASIS-INDEPENDENT FORM

Let the lattice-periodic permittivity of the medium  $\tilde{\epsilon}(\mathbf{r})$  be a bounded, lattice periodic function, possibly a constant. A Bloch mode solution to the scalar wave equation (8) can be written in the form

$$E(\mathbf{r}) = \tilde{E}(\mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r}), \quad (\text{A1})$$

where  $\mathbf{q}$  is the Bloch wave vector and  $\tilde{E}(\mathbf{r}) \neq 0$  is a nonzero, twice-differentiable, lattice-periodic function, which satisfies the equation

$$\mathcal{L}(\omega, \mathbf{q}) \tilde{E}(\mathbf{r}) = 0 \quad (\text{A2})$$

with the differential operator  $\mathcal{L}(\omega, \mathbf{q})$  given by

$$\mathcal{L}(\omega, \mathbf{q}) = (\nabla + i\mathbf{q})^2 + k^2 \tilde{\epsilon}(\mathbf{r}). \quad (\text{A3})$$

Recall that  $k = \omega/c$ ; hence the dependence of  $\mathcal{L}$  on  $\omega$ . The other reason for this dependence is frequency dispersion (dependence of  $\tilde{\epsilon}$  on  $\omega$ ), which is not indicated here explicitly but is taken into account. The requirement that (A2) has a nontrivial solution  $\tilde{E}(\mathbf{r}) \neq 0$  defines the dispersion equation  $f(\omega, \mathbf{q}) = 0$ . In what follows, we will assume without proof that, if for some pair  $(\omega, \mathbf{q})$  (A2) has a nontrivial solution, then this solution is unique up to multiplication by a constant. This is not generally true in the 3D case where different polarization states can correspond to the same pair  $(\omega, \mathbf{q})$ .

We will refer to the pairs  $(\omega, \mathbf{q})$  for which (A2) has a nontrivial solution as to the solutions to the dispersion equation. Here the variables  $\omega$  and  $\mathbf{q}$  are not restricted and can take general complex values. However, when studying stationary processes, one is typically interested only in solutions with purely real frequencies  $\omega$ . We also note that, if  $(\omega, \mathbf{q})$  is a solution, then  $(\omega, \mathbf{q} + \mathbf{g})$  is also a solution, where  $\mathbf{g}$  is any reciprocal lattice vector. This mathematical property of Bloch waves gives rise to “folding” of the solutions to the dispersion equation. It is a rather nontrivial mathematical question how to distinguish between the solutions that occur due to folding

from those that occur due to multiple scattering (interaction). In particular, it is quite plausible that solutions  $(\omega, \mathbf{q})$  and  $(\omega, \mathbf{q}')$  of different physical origin can co-exist at the same frequency  $\omega$  (or at two very close frequencies). In this Appendix, we present an approach to mathematical classification of these two physically-different solutions.

As was done in Ref. 36, we decompose  $\tilde{E}(\mathbf{r})$  and  $\tilde{\epsilon}(\mathbf{r})$  into the constant and zero-mean components according to

$$\tilde{E}(\mathbf{r}) = E_0 + F(\mathbf{r}) , \quad \langle F(\mathbf{r}) \rangle_{\mathbb{C}} = 0 , \quad (\text{A4a})$$

$$\tilde{\epsilon}(\mathbf{r}) = \epsilon_0 + \eta(\mathbf{r}) , \quad \langle \eta(\mathbf{r}) \rangle_{\mathbb{C}} = 0 . \quad (\text{A4b})$$

where  $\langle \dots \rangle_{\mathbb{C}}$  indicates averaging over the elementary cell  $\mathbb{C}$ . We note that  $E_0$  is the amplitude of the fundamental harmonic of the Bloch wave. Upon substitution of the above decomposition into (A2), we obtain

$$[k^2(\epsilon_0 + \eta(\mathbf{r})) - q^2] E_0 + \mathcal{L}(\omega, \mathbf{q})F(\mathbf{r}) = 0 . \quad (\text{A5})$$

We now introduce the averaging operator  $\mathcal{O}$  and the operator  $\mathcal{P} = 1 - \mathcal{O}$  as projections onto the complementary subspaces of constant and zero-mean functions so that  $\mathcal{O}\tilde{E}(\mathbf{r}) = E_0$ ,  $\mathcal{P}\tilde{E}(\mathbf{r}) = F(\mathbf{r})$  and  $\mathcal{O}\tilde{\epsilon}(\mathbf{r}) = \epsilon_0$ ,  $\mathcal{P}\tilde{\epsilon}(\mathbf{r}) = \eta(\mathbf{r})$ . By acting on (A5) with  $\mathcal{O}$  and  $\mathcal{P}$  from the left, we obtain the following two equations

$$(k^2\epsilon_0 - q^2) E_0 + k^2 \langle \eta(\mathbf{r})F(\mathbf{r}) \rangle_{\mathbb{C}} = 0 , \quad (\text{A6a})$$

$$\eta(\mathbf{r})E_0 + \mathcal{W}(\omega, \mathbf{q})F(\mathbf{r}) = 0 . \quad (\text{A6b})$$

In the first equation above we have used the equality  $\mathcal{O}\mathcal{L}F(\mathbf{r}) = k^2\mathcal{O}\eta(\mathbf{r})F(\mathbf{r}) = k^2 \langle \eta(\mathbf{r})F(\mathbf{r}) \rangle_{\mathbb{C}}$ , which can be proved by using cell-periodicity of  $F(\mathbf{r})$  and integral theorems. In the second equation, we have used  $(\mathcal{P}\mathcal{L})F(\mathbf{r}) = (\mathcal{P}\mathcal{L}\mathcal{P})F(\mathbf{r})$  and defined the operator

$$\mathcal{W}(\omega, \mathbf{q}) = k^{-2}\mathcal{P}\mathcal{L}(\omega, \mathbf{q})\mathcal{P} . \quad (\text{A7})$$

We can now consider the following two different kinds of solutions to (A6):

1. Solutions with  $E_0 \neq 0$ . Bloch waves  $E(\mathbf{r})$  with nonzero fundamental harmonic  $E_0 \neq 0$  can satisfy wave equation (8) for a given pair  $(\omega, \mathbf{q})$  only if the equation  $\mathcal{W}(\omega, \mathbf{q})\phi = \eta$  has a nonzero solution. According to the assumption of uniqueness of  $\tilde{E}$  stated above, this solution is unique if it exists. We will say that  $\phi$  is given in this case by the inverse of  $\mathcal{W}$ , viz,  $\phi = \mathcal{W}^{-1}\eta$ . Then it follows from (A6b) that  $F = -E_0\mathcal{W}^{-1}\eta$ . Substituting this expression into (A6a), we arrive at the dispersion relation (2) in which  $\Sigma(\omega, \mathbf{q})$  is given by the following basis-independent expression:

$$\Sigma(\omega, \mathbf{q}) = \epsilon_0 - \langle \eta(\mathbf{r})\mathcal{W}^{-1}(\omega, \mathbf{q})\eta(\mathbf{r}) \rangle_{\mathbb{C}} . \quad (\text{A8})$$

2. Solutions with  $E_0 = 0$ . There can exist another class of solutions, that is, Bloch waves with zero

fundamental harmonic. For such solutions to exist, the equation  $\mathcal{W}(\omega, \mathbf{q})\phi = 0$  must have a non-trivial solution such that  $\langle \eta(\mathbf{r})\phi(\mathbf{r}) \rangle_{\mathbb{C}} = 0$ . In this case,  $\mathcal{W}(\omega, \mathbf{q})$  is singular. It is easy to see that solutions of this kind exist in the case of a homogeneous medium that is artificially discretized on an arbitrary lattice (see below). However, solutions with  $E_0 = 0$  (or, in practice, with  $E_0$  in some sense very small) can also exist in inhomogeneous media with  $\eta \neq 0$ .

We will say that solutions to the dispersion equation of the first kind generate the ‘‘true’’  $\Gamma$ -points, that is, the  $\Gamma$ -points that result from multiple scattering in inhomogeneous photonic crystals. Solutions of the second kind are due to purely geometrical folding and generate spurious  $\Gamma$ -points.

This classification can be illustrated by considering the special case of a homogeneous medium that is artificially discretized on an arbitrary lattice of a finite pitch. In this case, it is easy to see that  $\mathcal{W}(\omega, \mathbf{q})$  is singular if  $(\mathbf{q} + \mathbf{g})^2 = k^2\epsilon_0$  where  $\mathbf{g} \neq 0$  is any *nonzero* reciprocal lattice vector, and invertible otherwise. Therefore, the dispersion equation for the solutions of the first kind is of the form

$$q^2 = k^2\epsilon_0 , \quad \Sigma(\omega, \mathbf{q}) = \epsilon_0 \quad (\text{first kind}) . \quad (\text{A9})$$

As could be expected, the corresponding lattice-periodic function  $\tilde{E}(\mathbf{r}) = E_0$  is just a constant. Thus, the fundamental harmonic is dominating in this solution.

Solutions of the second kind are of the form

$$(\mathbf{q} + \mathbf{g})^2 = k^2\epsilon_0 \quad \text{for } \mathbf{g} \neq 0 \quad (\text{second kind}) . \quad (\text{A10})$$

For any pair  $(\omega, \mathbf{q})$  satisfying the above condition,  $\mathcal{W}(\omega, \mathbf{q})$  is singular and we have  $\langle \tilde{E} \rangle_{\mathbb{C}} = 0$ . This lattice-periodic function is dominated by higher-order harmonics.

We can investigate the spurious  $\Gamma$ -points generated by the solution of the second kind as follows. If  $\text{Im}(\epsilon_0) = 0$ , equation (A10) is satisfied by a pair  $(\omega_{\mathbf{p}}, 0)$  where  $\omega_{\mathbf{p}}^2 = (c\mathbf{p})^2/\epsilon_0$  and  $\mathbf{p} \neq 0$  is any nonzero reciprocal lattice vector. The isofrequency line at  $\omega = \omega_{\mathbf{p}}$  is then obtained from the equation

$$(\mathbf{q} + \mathbf{g})^2 = p^2 > 0 \quad \text{for any } \mathbf{g} \neq 0 . \quad (\text{A11})$$

Consider for simplicity a square lattice and let  $\mathbf{p} = (2\pi/a)(0, 1)$ . The corresponding frequency  $\omega_{\mathbf{p}} = 2\pi c/a\sqrt{\epsilon_0}$  is the frequency of the first spurious  $\Gamma$ -point out of the infinite sequence. In principle, equation (A11) defines an infinite number of curves but at the first  $\Gamma$ -point most of these curves coincide. To obtain the complete solution, we can start by taking  $\mathbf{g} = \mathbf{p}$ . It can be seen that (A11) defines in this case a circular arc in the  $(q_x, q_y)$  plane, and that this arc crosses the origin. On a square lattice, there are four reciprocal vectors of the same length  $2\pi/a$  and pointing in the directions  $(0, \pm 1)$

and  $(\pm 1, 0)$ . As a result, the isofrequency line contains four circular arcs intersecting at the origin (six arcs in the case of the triangular lattice) as is shown in Fig. 1(a). The arcs are truncated and reflected at the edge of the FBZ. These additional “reflected” segments of the isofrequency line can be obtained by considering additional vectors  $\mathbf{p}$  in (A11).

The above result is quite trivial and we could have obtained it without using the mathematical formalism of this Appendix. However, the derivation is useful to show that the formalism developed herein is consistent with the limit of zero contrast. More importantly, it provides us with a clue how to treat the case when  $E_0$  is in some sense small but nonzero. Let  $\gamma = \|F/E_0\|_2$  where  $\|\cdot\|_2$  is the  $L_2$  norm. We believe that it is physically meaningful to consider the phase and group velocities of a Bloch wave only if  $\gamma$  is sufficiently small so that the fundamental harmonic of the Bloch wave is in some sense dominant. This condition holds if the smallest singular value of  $\mathcal{W}$  is sufficiently far from zero. If this is not so, then  $\gamma$  can be very large or even diverge. In this case, introducing the characteristics of a plane wave such as the phase and group velocities is devoid of physical meaning, even if this can be done formally by considering the dependence  $q(\omega)$ . We therefore conjecture that there are two fundamentally different regimes of propagation,  $\gamma \ll 1$  and  $\gamma \gg 1$ , and in the second regime the group velocity computed as, say,  $\nabla_{\mathbf{q}}\omega$  does not correspond to any physically measurable quantity and should not be invested with any particular interpretation. It is not clear though how to treat the borderline case  $\gamma \sim 1$ ; perhaps, it can only be investigated numerically.

It should be noted that the above discussion is applicable to the case of s-polarization when the scalar electric field is smooth. For the p-polarization, the electric field can jump at the discontinuities of  $\tilde{\epsilon}(\mathbf{r})$ , and a more careful analysis is required. We also note that the condition on  $\gamma$  stated above is closely related to the concept of the smooth field introduced in Ref. 25.

## APPENDIX B: COMPUTATION OF THE EXPANSION COEFFICIENTS $\sigma_n$

In this Appendix, we compute the expansion coefficients  $\sigma_n$  according to (30) up to fourth order in  $Q$ . The derivations are valid for PCs with  $C_4$  and  $C_6$  symmetries (with some notable differences explained below). However, the derivations are not valid for a more general angle between  $\mathbf{a}_1$  and  $\mathbf{a}_2$  and we have not performed a perturbative analysis of what happens when the angle slightly deviates from  $\pi/4$  or  $\pi/3$ .

The results are written in terms of the Cartesian components of  $\mathbf{q}$  and some  $\mathbf{q}$ -independent coefficients. These coefficients can be computed if  $T$  is known. The matrix  $T$  is defined in (24) and can be computed algebraically by one operation of matrix inversion. Therefore, all expressions given below are directly amenable to numerical

computation.

Note that the order in  $Q$  is not the same as the order in  $\mathbf{q}$ . Also note that terms of the form

$$(-1)^j q^{2j} \langle b|T^{j+1}|b\rangle \quad (\text{B1})$$

are generated in all orders of the expansion. These terms are, of course, perfectly isotropic. However, starting from fourth order, more complicated terms appear in the expansion.

In the formulas of this Appendix, we will use the following notations:

(i) The symbol  $\circ$  denotes direct (Hadamard) product of two matrices, e.g.,  $(A \circ B)_{ij} = A_{ij}B_{ij}$ .

(ii) The matrices  $G$ ,  $G_{xx}$ , etc., are defined in terms of the reciprocal lattice vectors as follows:

$$(G_{\alpha\beta})_{\mathbf{g}\mathbf{p}} = g_{\alpha}p_{\beta}, \quad \alpha, \beta = x, y; \quad (\text{B2a})$$

$$G_{\mathbf{g}\mathbf{p}} = \sum_{\alpha} G_{\alpha\alpha} = \mathbf{g} \cdot \mathbf{p}. \quad (\text{B2b})$$

(iii) In all expressions shown below we imply that, for example,  $q^6 = (q_x^2 + q_y^2)^3$ , etc.

a. *Zeroth order.* We start with zero order ( $n = 0$ ), which is rather trivial:

$$\sigma_0 = \langle b|T|b\rangle = \sum_{\mathbf{g}_1, \mathbf{g}_2} M(-\mathbf{g}_1)T_{\mathbf{g}_1\mathbf{g}_2}M(\mathbf{g}_2). \quad (\text{B3})$$

Thus,  $\sigma_0$  is a  $\mathbf{q}$ -independent constant. Obviously,  $\Sigma(\omega, 0) = \epsilon_0 + (\rho\chi k)^2\sigma_0(k)$ , where we have indicated the dependence of  $\sigma_0$  on  $k$  explicitly.

b. *First order.* Here we have

$$\sigma_1 = - \sum_{\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}} M(-\mathbf{g}_1)T_{\mathbf{g}_1\mathbf{g}}(q^2 + 2\mathbf{q} \cdot \mathbf{g}) \times T_{\mathbf{g}\mathbf{g}_2}M(\mathbf{g}_2). \quad (\text{B4})$$

The contribution of the second term in the brackets is zero due to the symmetry  $[\sigma_n(-\mathbf{q}) = \sigma_n(\mathbf{q})]$ . Therefore,

$$\begin{aligned} \sigma_1 &= -q^2 \sum_{\mathbf{g}_1, \mathbf{g}_2} M(-\mathbf{g}_1)T_{\mathbf{g}_1\mathbf{g}_2}^2 M(-\mathbf{g}_2) \\ &= -q^2 \langle b|T^2|b\rangle. \end{aligned} \quad (\text{B5})$$

c. *Second order.* Here we have

$$\begin{aligned} \sigma_2 &= \sum_{\{\mathbf{g}\}} M(-\mathbf{g}_1)T_{\mathbf{g}_1\mathbf{g}_2}T_{\mathbf{g}_2\mathbf{g}_3}T_{\mathbf{g}_3\mathbf{g}_4}M(\mathbf{g}_4) \\ &\quad \times [q^2 + 2(\mathbf{q} \cdot \mathbf{g}_2)] [q^2 + 2(\mathbf{q} \cdot \mathbf{g}_3)], \end{aligned} \quad (\text{B6})$$

where  $\sum_{\{\mathbf{g}\}}$  denotes summation over all relevant indexes. We now expand the product of the square brackets and notice that the terms linear in  $\mathbf{q}$  sum to zero for the reason of inversion symmetry. The two terms that produce nonzero result upon summation are

$$q^4 \quad \text{and} \quad 4(\mathbf{q} \cdot \mathbf{g}_2)(\mathbf{q} \cdot \mathbf{g}_3). \quad (\text{B7})$$

The first of these generates the result of the form (B1) and the second term requires some additional consideration. We can write

$$\begin{aligned} (\mathbf{q} \cdot \mathbf{g}_2)(\mathbf{q} \cdot \mathbf{g}_3) &= (q_x g_{2x} + q_y g_{2y})(q_x g_{3x} + q_y g_{3y}) \\ &= q_x^2 g_{2x} g_{3x} + q_y^2 g_{2y} g_{3y} + q_x q_y (g_{2x} g_{3y} + g_{2y} g_{3x}). \end{aligned}$$

The term proportional to  $q_x q_y$  sums to zero due the inversion symmetry. Moreover, from the symmetry properties of both triangular and square lattices, we find that summation of the coefficients in front of  $q_x^2$  and  $q_y^2$  must yield the same result. Indeed, if this were not so, we would have received a term of the form  $\beta_x q_x^2 + \beta_y q_y^2$  with  $\beta_x \neq \beta_y$ , describing an ellipse of unequal semi-axes, which is inconsistent with both  $C_4$  and  $C_6$  symmetries. Therefore, we can replace the above expression (inside the summation) by

$$\frac{1}{2}(q_x^2 + q_y^2)(g_{2x} g_{3x} + g_{2y} g_{3y}) = \frac{1}{2} q^2 (\mathbf{g}_2 \cdot \mathbf{g}_3).$$

Collecting everything together, we find that

$$\sigma_2 = q^4 \langle b | T^3 | b \rangle + 2q^2 \langle b | T(T \circ G)T | b \rangle. \quad (\text{B8})$$

We thus see that all terms generated in the second order are still circularly-symmetric.

*d. Third order.* Here we have

$$\begin{aligned} \sigma_3 &= - \sum_{\{\mathbf{g}\}} M(-\mathbf{g}_1) T_{\mathbf{g}_1 \mathbf{g}_2} T_{\mathbf{g}_2 \mathbf{g}_3} T_{\mathbf{g}_3 \mathbf{g}_4} T_{\mathbf{g}_4 \mathbf{g}_5} M(\mathbf{g}_5) \\ &\times [q^2 + 2(\mathbf{q} \cdot \mathbf{g}_2)] [q^2 + 2(\mathbf{q} \cdot \mathbf{g}_3)] [q^2 + 2(\mathbf{q} \cdot \mathbf{g}_4)]. \end{aligned} \quad (\text{B9})$$

After expanding the brackets and keeping only the terms that do not sum to zero, we find that  $\sigma_3 = \sigma_3^{(a)} + \sigma_3^{(b)}$  where

$$\sigma_3^{(a)} = -q^6 \langle b | T^4 | b \rangle \quad (\text{B10})$$

is of the form (B1) and

$$\begin{aligned} \sigma_3^{(b)} &= -4q^2 \sum_{\{\mathbf{g}\}} M(-\mathbf{g}_1) T_{\mathbf{g}_1 \mathbf{g}_2} T_{\mathbf{g}_2 \mathbf{g}_3} T_{\mathbf{g}_3 \mathbf{g}_4} T_{\mathbf{g}_4 \mathbf{g}_5} M(\mathbf{g}_5) \\ &\times [(\mathbf{q} \cdot \mathbf{g}_2)(\mathbf{q} \cdot \mathbf{g}_3) + (\mathbf{q} \cdot \mathbf{g}_2)(\mathbf{q} \cdot \mathbf{g}_4) + (\mathbf{q} \cdot \mathbf{g}_3)(\mathbf{q} \cdot \mathbf{g}_4)]. \end{aligned} \quad (\text{B11})$$

We can use the same transformation as was used in the second order to transform the terms of the form  $(\mathbf{q} \cdot \mathbf{g}_2)(\mathbf{q} \cdot \mathbf{g}_3)$  to the form  $(1/2)q^2(\mathbf{g}_2 \cdot \mathbf{g}_3)$ . We thus obtain

$$\begin{aligned} \sigma_3^{(b)} &= 2(qh)^4 \left[ \langle b | T^2(T \circ G)T | b \rangle \right. \\ &\left. + \langle b | T(T^2 \circ G)T | b \rangle + \langle b | T(T \circ G)T^2 | b \rangle \right]. \end{aligned} \quad (\text{B12})$$

The first and last terms in this expression are in fact equal but are written separately for symmetry of expression. Still, all expressions arising to third order in  $Q$  are circularly-symmetric.

*e. Fourth order.* In the fourth order, we can write

$$\sigma_4 = \sigma_4^{(a)} + \sigma_4^{(b)} + \sigma_4^{(c)}, \quad (\text{B13})$$

where the expressions for  $\sigma_4^{(a)}$  and  $\sigma_4^{(b)}$  are obtained in the manner very similar to what was done above. Omitting the intermediate steps, we write the final result for these two terms, viz,

$$\sigma_4^{(a)} = q^8 \langle b | T^5 | b \rangle \quad (\text{B14})$$

$$\begin{aligned} \sigma_4^{(b)} &= 2q^6 \\ &[\langle b | T(T \circ G)T^3 | b \rangle + \langle b | T^3(T \circ G)T | b \rangle \\ &+ \langle b | T(T^2 \circ G)T^2 | b \rangle + \langle b | T^2(T^2 \circ G)T | b \rangle \\ &+ \langle b | T^3(T \circ G)T | b \rangle + \langle b | T^2(T \circ G)T^2 | b \rangle]. \end{aligned} \quad (\text{B15})$$

Note that the terms appearing on each line of the above expression are pair-wise equal. The expressions (B14) and (B15) could, in fact, be anticipated and are directly analogous to the expressions (B10) and (B12). A simple diagrammatic technique can be devised to generate similar expressions that appear in the higher orders of the perturbation theory.

However,  $\sigma_4^{(c)}$  is a term of a different kind and it is the first term we encounter that does not obey the circular symmetry and is consistent with  $C_4$  (but not  $C_6$ ) symmetry. The term is

$$\begin{aligned} \sigma_4^{(c)} &= 16 \sum_{\{\mathbf{g}\}} M(-\mathbf{g}_1) T_{\mathbf{g}_1 \mathbf{g}_2} \dots T_{\mathbf{g}_5 \mathbf{g}_6} M(\mathbf{g}_6) \\ &\times (\mathbf{q} \cdot \mathbf{g}_2)(\mathbf{q} \cdot \mathbf{g}_3)(\mathbf{q} \cdot \mathbf{g}_4)(\mathbf{q} \cdot \mathbf{g}_5). \end{aligned} \quad (\text{B16})$$

We cannot use the same trick as was used above to transform the factor

$$P = (\mathbf{q} \cdot \mathbf{g}_2)(\mathbf{q} \cdot \mathbf{g}_3)(\mathbf{q} \cdot \mathbf{g}_4)(\mathbf{q} \cdot \mathbf{g}_5) \quad (\text{B17})$$

to the form  $\beta q^4$  where  $\beta$  is a scalar expressible in terms of the dot products  $\mathbf{g}_i \cdot \mathbf{g}_j$ . Indeed, let us write the factor  $P$  in terms of Cartesian components of all vectors involved:

$$\begin{aligned} P &= (q_x g_{2x} + q_y g_{2y})(q_x g_{3x} + q_y g_{3y}) \\ &\times (q_x g_{4x} + q_y g_{4y})(q_x g_{5x} + q_y g_{5y}) \\ &= P' + q_x^4 g_{2x} g_{3x} g_{4x} g_{5x} + q_y^4 g_{2y} g_{3y} g_{4y} g_{5y} \\ &+ q_x^2 q_y^2 [g_{2x} g_{3x} g_{4y} g_{5y} + g_{2y} g_{3y} g_{4x} g_{5x} \\ &+ g_{2x} g_{3y} g_{4x} g_{5y} + g_{2y} g_{3x} g_{4y} g_{5x} \\ &+ g_{2x} g_{3y} g_{4y} g_{5x} + g_{2y} g_{3x} g_{4x} g_{5y}]. \end{aligned} \quad (\text{B18})$$

Here  $P'$  is the term that sums to zero by symmetry. We can introduce the notations  $\Delta = P - P'$  and  $\Pi_x, \Pi_y, \Pi_{xy}$  (definition of the last three quantities will be clear from the next equation) and write

$$\Delta = \Pi_x q_x^4 + \Pi_y q_y^4 + \Pi_{xy} q_x^2 q_y^2 \quad (\text{B19})$$

Now, if it happens so that

$$\sum_{\{\mathbf{g}\}} F[\{\mathbf{g}\}] \Pi_x = \sum_{\{\mathbf{g}\}} F[\{\mathbf{g}\}] \Pi_y = \frac{1}{2} \sum_{\{\mathbf{g}\}} F[\{\mathbf{g}\}] \Pi_{xy} , \quad (\text{B20})$$

where  $F[\{\mathbf{g}\}]$  is the coefficient appearing on the first line of (B16), then we would obtain the result  $\sigma_4^{(c)} = \beta q^4$ . This is what we can expect to happen in triangular lattices with  $C_6$  symmetry. However, there is no general or obvious reason why the second equality in (B20) should hold in the case of  $C_4$  symmetry, and there are sufficient grounds to believe that it does not. As a result, the function  $\Delta(\mathbf{q})$  and the summation result  $\sum_{\{\mathbf{g}\}} F[\{\mathbf{g}\}] \Delta(\mathbf{q})$  are not circularly-symmetric.

We can write the result for  $\sigma_4^{(c)}$  in a form similar to

that used in lower orders if we account for the identity  $q_x^4 + q_y^4 = q^4 - 2q_x^2 q_y^2$ . Then

$$\begin{aligned} \sigma_4^{(c)} = & 8 [q^4 - 2(q_x^2 q_y^2)] \\ & \times \left[ \langle b | T(T \circ G_x) T(T \circ G_x) T | b \rangle + (G_x \rightarrow G_y) \right] \\ & + q_x^2 q_y^2 \\ & \times \left[ \langle b | T(T \circ G_x) T(T \circ G_y) T | b \rangle + (G_x \leftrightarrow G_y) \right. \\ & + \langle b | T(T \circ G_{xy}) T(T \circ G_{xy}) T | b \rangle + (G_{xy} \rightarrow G_{yx}) \\ & \left. + \langle b | T(T \circ G_{xy}) T(T \circ G_{yx}) T | b \rangle + (G_{xy} \leftrightarrow G_{yx}) \right] . \end{aligned} \quad (\text{B21})$$