PHYSICAL REVIEW E, VOLUME 64, 056604

Optical pulse propagation in nonlinear photonic crystals

N. A. R. Bhat and J. E. Sipe
Department of Physics, University of Toronto, 60 St. George Street, Toronto, Ontario, Canada M5S 1A7

(Received 23 April 2001; published 15 October 2001)

We present a formalism for optical pulse propagation in nonlinear photonic crystals of arbitrary dimensionality. Using a multiple-scale analysis, we derive the dynamical nonlinear Schrödinger equation obeyed by the envelope function modulating an underlying Bloch function. Effective coefficients appear in that equation characterizing the effects of Kerr nonlinearity, linear gain or loss, and material dispersion. They depend on how the underlying Bloch function ''samples'' these effects in the photonic crystal, and require for their calculation a specification of these effects throughout the photonic crystal, and the calculated bandstructure of the photonic crystals in the linear, nondispersive limit. We show that wave packets from different bands can experience significantly modified effective material properties.

DOI: 10.1103/PhysRevE.64.056604 PACS number(s): 42.70.Qs, 42.65.Sf

I. INTRODUCTION

As codes become available for the calculation of photonic band structures, and the linear optical properties of photonic crystals become better understood, attention can shift to the nonlinear optical properties of these structures. Already there have been studies of nonlinear effects by numerical simulations of the full nonlinear Maxwell equations [1,2]. While for special problems this may remain the theoretical approach of choice, it is limited in that the underlying physics is not immediately apparent from the results of the simulation. As well, the loss of superposition in nonlinear optics means that any such simulation yields only limited insight into the general character of nonlinear propagation.

A different approach has been common for many years in the study of the nonlinear optical properties of one-dimensional photonic crystals, or ''gratings.'' In the presence of a weak nonlinearity, effective field equations that describe the dynamics of envelope functions are derived from the nonlinear Maxwell equations [3]. At frequencies near and within the band gap these can take the form of coupled mode equations, while at mid-band frequencies it is a nonlinear Schrödinger equation that describes the evolution of the appropriate envelope function. In fact, a nonlinear Schrödinger equation treatment can be applied even at a band edge, or within a band gap, if a pulse is not too short [4]. Properties such as the group velocity and group velocity dispersion, and the strength and spatial distribution of the nonlinearity, appear as parameters of these effective field equations. Once they are derived, a body of mathematical work devoted to the characterization of their solutions can be called on to obtain insight into the general nature of nonlinear light propagation. This approach has led to the prediction and description of gap and Bragg solitons [5,6], the understanding of various scenarios for all-optical switching [7], the description of the appearance of modulational instability [8], as well as the prediction of bistability and chaotic behavior in certain structures [9].

Work along these lines for higher-dimensional photonic crystals has already appeared [10], where coupled mode equations were applied to study nonlinear field structures in two-dimensional structures. But those equations were derived in a heuristic way analogous to that usually used for weak one-dimensional gratings [11]. In higher-dimensional photonic crystals, where the refractive index difference between components can be large, this approach is both theoretically suspect and in practice insufficient, in that it does not yield parameters that adequately capture the consequences of the spatial distribution of any underlying absorption or nonlinearity. For the realistic description of nonlinear propagation in actual photonic crystals, a better approach is required.

Once again earlier work on one-dimensional structures indicates a way to proceed. Some years ago [12] it was shown how effective field equations could be derived even in the presence of strong refractive index variations. Employing the exact Bloch functions of the corresponding linear problem in the absence of any material dispersion, nonlinearity, or gain or loss, a multiple-scales approach that treats these latter effects as ''small''—in a sense that can be made precise—leads to the derivation of effective field equations. The derivation yields parameters involving sums over photonic bands of various coefficients; their physical significance is identified using results from a photonic $k\cdot p$ theory that parallels such a theory for electrons. Recently this approach has been generalized to one-dimensional birefringent structures [13].

It is such an attack that is the subject of this paper, where we address higher-dimensional photonic crystals. We begin with the exact Bloch functions of the photonic crystal, which we assume have been found with the neglect of any material dispersion, gain or loss, or nonlinearity. This defines what is essentially our ''unperturbed'' system. Our goal here is to treat the in-band propagation problem, at frequencies removed from a possible photonic band gap, and where any band degeneracies can be neglected. We seek a solution of the full nonlinear problem by modulating a Bloch function with an envelope function and deriving the dynamical equation for that envelope function. We only consider a Kerr nonlinearity, linear loss or gain, and small material dispersion, but various generalizations would follow immediately. A multiple-scales approach identifies the relative size of the different perturbing effects, and using it we derive a generalized nonlinear Schrödinger equation. From this we can immediately identify much of the physics of light propagation through the structure, such as an effective n_2 nonlinear coefficient and an effective loss or gain coefficient, as well as a group velocity and group velocity dispersion. All of these depend crucially on how the underlying Bloch function ‘‘samples’’ the appropriate perturbing effect in the photonic crystal. The nonlinear Schrödinger equation itself can be used for a more detailed treatment of pulse propagation.

The analysis here is qualitatively more difficult than that in the one-dimensional case because of the need to respect the divergence Maxwell equations,

$$\begin{aligned} \nabla\cdot\mathbf{D}&=0,\\ \nabla\cdot\mathbf{B}&=0, \end{aligned} \tag{1}$$

that the full displacement field **D** and magnetic field **B** must satisfy. It has already been pointed out [14] that these constraints make the $k\cdot p$ theory more complicated for photonic crystals than it is for electronic crystals. As well, to ensure that these equations are exactly satisfied by the inevitably approximate multiple-scales results, it is convenient not to work with the electromagnetic fields themselves as the basic fields of the theory, but rather with two vector potentials. As is usual in multiple-scales treatments, much of the effort is devoted to setting up the problem in a formalism that leads to the easy application of a multiple-scales approach. In Sec. II we identify the Bloch functions of the linear, nondispersive Maxwell equations, and cast those equations in a matrix form that considerably simplifies the $k\cdot p$ treatment, the results of which play an essential role in the multiple-scales analysis. In Sec. III we introduce the vector potentials and set up the full Maxwell equations for our problem. The multiple scales derivation is presented in Sec. IV, and in Sec. V we explicitly consider nonlinearity, material dispersion and gain or loss at one particular level of scaling. In Sec. VI our envelope function equation is derived, and the physical significance of the different parameters is identified. In Sec. VII we calculate these for a well-studied photonic crystal and illustrate how they can vary through a band. The road to various generalizations is indicated in Sec. VIII, as well as our conclusions.

II. BLOCH FUNCTIONS

Bloch functions of photonic crystals are most often found by identifying the stationary solutions of the so-called ‘‘master equation,’’ which is second order in time. However, for our purposes it is more convenient to work with equations that are first order in time. In this section we describe the Bloch functions of the first order Maxwell equations, connecting them with the Bloch functions of the master equation. Next, we introduce a matrix formulation of Maxwell’s equations that will be convenient for later analytic manipulations. Finally, we extend the $k\cdot p$ expansion of solid state physics to this formalism.

A. Maxwell equations

The macroscopic electromagnetic fields satisfy two dynamical equations,

$$\begin{aligned} \dot{\mathbf{D}}&=\nabla\times\mathbf{H},\\ \dot{\mathbf{B}}&=-\nabla\times\mathbf{E}, \end{aligned} \tag{2}$$

which must be solved subject to initial constraints (1) and the appropriate constitutive relations. In this section we take those to be

$$\mathbf{B}(\mathbf{r},t)=\mu_0\mathbf{H}(\mathbf{r},t), \tag{3}$$

$$\mathbf{D}(\mathbf{r},t)=\varepsilon_0 n^2(\mathbf{r})\mathbf{E}(\mathbf{r},t), \tag{4}$$

where μ_0 and ε_0 are, respectively, the permeability and permittivity of free space, and the local index of refraction $n(\mathbf{r})$ is assumed real. For stationary solutions of the form $\mathbf{E}(\mathbf{r},t)=\mathbf{E}(\mathbf{r})\exp(-i\omega t)$, etc., (1),(2) then reduce to

$$\begin{aligned} -i\omega\varepsilon_0 n^2(\mathbf{r})\mathbf{E}(\mathbf{r})&=\nabla\times\mathbf{H}(\mathbf{r}),\\ i\omega\mu_0\mathbf{H}(\mathbf{r})&=\nabla\times\mathbf{E}(\mathbf{r}), \end{aligned} \tag{5}$$

and

$$\begin{aligned} \nabla\cdot[n^2(\mathbf{r})\mathbf{E}(\mathbf{r})]&=0,\\ \nabla\cdot\mathbf{H}(\mathbf{r})&=0. \end{aligned} \tag{6}$$

We divide the solutions of Eqs. (5) and (6) into two types, type P_1 with $\omega\neq 0$ and type P_2 with $\omega=0$.

For P_1 solutions we need only require that the equations (5) be satisfied, since the divergence equations (6) then follow immediately. From Eqs. (5) it is clear that if $(\mathbf{H}_I(\mathbf{r}),\mathbf{E}_I(\mathbf{r}))$ is a solution with frequency ω_I, then $(\mathbf{H}_I^*(\mathbf{r}),\mathbf{E}_I^*(\mathbf{r}))$ is a solution with frequency $-\omega_I$, a result of time reversal symmetry. This has important consequences for the nature of stationary solutions for photonic crystals, where $n(\mathbf{r})=n(\mathbf{r}+\mathbf{R})$ for any lattice vector $\mathbf{R}$, that we identify below.

For a photonic crystal, Bloch’s theorem guarantees that the stationary solutions can be chosen to be of the form

$$\begin{aligned} \mathbf{H}_{m\mathbf{k}}(\mathbf{r})&=\mathbf{h}_{m\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}},\\ \mathbf{E}_{m\mathbf{k}}(\mathbf{r})&=\mathbf{e}_{m\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}, \end{aligned} \tag{7}$$

where the crystal wave vector **k** lies in the first Brillouin zone, m is a band index, and $\mathbf{h}_{m\mathbf{k}}(\mathbf{r})=\mathbf{h}_{m\mathbf{k}}(\mathbf{r}+\mathbf{R})$, $\mathbf{e}_{m\mathbf{k}}(\mathbf{r})=\mathbf{e}_{m\mathbf{k}}(\mathbf{r}+\mathbf{R})$ for any lattice vector **R**. Using Eq. (7) in the complex conjugate of Eq. (5) it follows that, associated with each solution $(\mathbf{H}_{m\mathbf{k}}(\mathbf{r}),\mathbf{E}_{m\mathbf{k}}(\mathbf{r}))$ with frequency $\omega_{m\mathbf{k}}>0$, there is another solution $(-\mathbf{h}_{m\mathbf{k}}^*(\mathbf{r})e^{-i\mathbf{k}\cdot\mathbf{r}},\mathbf{e}_{m\mathbf{k}}^*(\mathbf{r})e^{-i\mathbf{k}\cdot\mathbf{r}})$ with the same frequency; it is therefore proportional to one of the stationary solutions with crystal wave vector $-\mathbf{k}$. We can link these two by adopting a standard phase convention and putting $\mathbf{h}_{m(-\mathbf{k})}(\mathbf{r})=-\mathbf{h}_{m\mathbf{k}}^*(\mathbf{r})$ and $\mathbf{e}_{m(-\mathbf{k})}(\mathbf{r})=\mathbf{e}_{m\mathbf{k}}^*(\mathbf{r})$. Then $(\mathbf{H}_{m(-\mathbf{k})}(\mathbf{r}),\mathbf{E}_{m(-\mathbf{k})}(\mathbf{r}))=(-\mathbf{H}_{m\mathbf{k}}^*(\mathbf{r}),\mathbf{E}_{m\mathbf{k}}^*(\mathbf{r}))$ is the solution associated with crystal wave vector $-\mathbf{k}$ and frequency $\omega_{m\mathbf{k}}$. Finally, adopting another phase convention by defining $\mathbf{h}_{\bar{m}\mathbf{k}}(\mathbf{r})=-\mathbf{h}_{m\mathbf{k}}(\mathbf{r})$ and $\mathbf{e}_{\bar{m}\mathbf{k}}(\mathbf{r})=\mathbf{e}_{m\mathbf{k}}(\mathbf{r})$, and where we put $(\mathbf{H}_{\bar{m}\mathbf{k}}(\mathbf{r}),\mathbf{E}_{\bar{m}\mathbf{k}}(\mathbf{r}))=(\mathbf{h}_{\bar{m}\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}},\mathbf{e}_{\bar{m}\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}})$, we see that

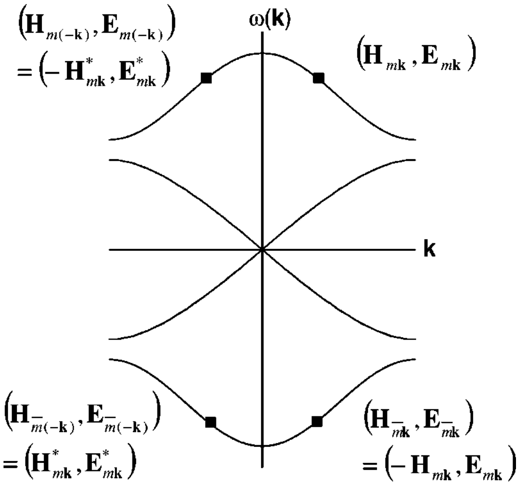

FIG. 1. Schematic diagram of the relationship between solutions at $\pm\omega$ and $\pm\mathbf{k}$.

$(\mathbf{H}_{\bar{m}\mathbf{k}}(\mathbf{r}),\mathbf{E}_{\bar{m}\mathbf{k}}(\mathbf{r}))=(\mathbf{H}^*_{m(-\mathbf{k})}(\mathbf{r}),\mathbf{E}^*_{m(-\mathbf{k})}(\mathbf{r}))$ is the solution of Eq. (5) corresponding to crystal wave vector $\mathbf{k}$ and frequency $-\omega_{m\mathbf{k}}$, which we define as $\omega_{\bar{m}\mathbf{k}}$.

In summary, if $(\mathbf{H}_{m\mathbf{k}},\mathbf{E}_{m\mathbf{k}})$ is a solution of Eqs. (5) with crystal wave vector $\mathbf{k}$ and frequency $\omega_{m\mathbf{k}}$, then: $(\mathbf{H}_{\bar{m}\mathbf{k}},\mathbf{E}_{\bar{m}\mathbf{k}})=(\mathbf{H}^*_{m(-\mathbf{k})},\mathbf{E}^*_{m(-\mathbf{k})})=(-\mathbf{H}_{m\mathbf{k}},\mathbf{E}_{m\mathbf{k}})$ is the corresponding solution with crystal wave vector $\mathbf{k}$ and frequency $\omega_{\bar{m}\mathbf{k}}=-\omega_{m\mathbf{k}}$; $(\mathbf{H}_{m(-\mathbf{k})},\mathbf{E}_{m(-\mathbf{k})})=(\mathbf{H}^*_{\bar{m}\mathbf{k}},\mathbf{E}^*_{\bar{m}\mathbf{k}})=(-\mathbf{H}^*_{m\mathbf{k}},\mathbf{E}^*_{m\mathbf{k}})$ is the corresponding solution with crystal wave vector $-\mathbf{k}$ and frequency $\omega_{m(-\mathbf{k})}=\omega_{m\mathbf{k}}$; $(\mathbf{H}_{\bar{m}(-\mathbf{k})},\mathbf{E}_{\bar{m}(-\mathbf{k})})=(-\mathbf{H}_{m(-\mathbf{k})},\mathbf{E}_{m(-\mathbf{k})})=(\mathbf{H}^*_{m\mathbf{k}},\mathbf{E}^*_{m\mathbf{k}})$ is the corresponding solution with crystal wave vector $-\mathbf{k}$ and frequency $\omega_{\bar{m}(-\mathbf{k})}=\omega_{\bar{m}\mathbf{k}}$. These relations are indicated schematically in Fig. 1.

Since we can find $\mathbf{E}_{m\mathbf{k}}(\mathbf{r})$ from $\mathbf{H}_{m\mathbf{k}}(\mathbf{r})$ and $\omega_{m\mathbf{k}}$ using the first of Eqs. (5), the specification of $\mathbf{H}_{m\mathbf{k}}(\mathbf{r})$ as $\mathbf{k}$ ranges over a half of the Brillouin zone that does not include $-\mathbf{k}$ if it includes $\mathbf{k}$ (such as the region $k_z \geqslant 0$) is sufficient to identify both the solutions associated with the $\omega_{m\mathbf{k}}$ and the $\omega_{\bar{m}\mathbf{k}}$ over the entire Brillouin zone. Those $\mathbf{H}_{m\mathbf{k}}(\mathbf{r})$ are often found by solving for the eigenfunctions of the master equation,

$$\nabla\times\left[\frac{1}{n^2(\mathbf{r})}\nabla\times\mathbf{H}_{m\mathbf{k}}(\mathbf{r})\right]=\frac{\omega^2_{m\mathbf{k}}}{c^2}\mathbf{H}_{m\mathbf{k}}(\mathbf{r}), \tag{8}$$

which follows immediately from Eq. (5).

It follows from Eq. (8) that two eigenfunctions with different eigenvalues are orthogonal, as are of course two eigenfunctions with different crystal wave vectors. We choose degenerate eigenfunctions at the same wave vector so that they are mutually orthogonal, and normalize the eigenfunctions according to

$$\frac{\mu_0}{\varepsilon_0}\int\frac{d^3r}{\Omega}\mathbf{H}^*_{m\mathbf{k}}(\mathbf{r})\cdot\mathbf{H}_{m'\mathbf{k}'}(\mathbf{r})=\delta_{mm'}\delta_{\mathbf{k}\mathbf{k}'}, \tag{9}$$

where Ω is the normalization volume. It then follows from Eqs. (5) that

$$\int\frac{d^3r}{\Omega}n^2(\mathbf{r})\mathbf{E}^*_{m\mathbf{k}}(\mathbf{r})\cdot\mathbf{E}_{m'\mathbf{k}'}(\mathbf{r})=\delta_{mm'}\delta_{\mathbf{k}\mathbf{k}'}, \tag{10}$$

and that the $\mathbf{E}_{m\mathbf{k}}(\mathbf{r})$ are dimensionless.

Turning to type P_2 solutions, we have both $\nabla\cdot\mathbf{H}=0$ and, from the first of Eqs. (5), $\nabla\times\mathbf{H}=\mathbf{0}$. The only solutions of these equations subject to periodic boundary conditions over our normalization volume are those of uniform $\mathbf{H}$. From the remaining equations (5) and (6) we find

$$\begin{gathered}\nabla\times\mathbf{E}(\mathbf{r})=\mathbf{0},\\ \nabla\cdot[n^2(\mathbf{r})\mathbf{E}(r)]=\mathbf{0}.\end{gathered} \tag{11}$$

Typically there are only solutions of these equations at $\mathbf{k}=\mathbf{0}$. But in any case the conditions on $\mathbf{H}(\mathbf{r})$ and $\mathbf{E}(\mathbf{r})$ decouple for type P_2 solutions. Arbitrary solutions of type P_2 can be written as linear combinations of type P_2 solutions with $\mathbf{E}=\mathbf{0}$ and type P_2 solutions with $\mathbf{H}=\mathbf{0}$. We return to these in the next section.

B. Matrix eigenvalue equation

While the master equation (8) is the eigenvalue equation that is probably most often used to find the photonic Bloch functions, for analytic manipulations it is easier to work with a matrix eigenvalue equation. The two equations (5) can be written as

$$\mathsf{M}\times\mathbf{\Phi}=\omega\mathsf{n}\mathbf{\Phi}, \tag{12}$$

where we have introduced the column of complex vector fields

$$\mathbf{\Phi}=\begin{pmatrix}\frac{1}{2}\sqrt{n(\mathbf{r})}\mathbf{E}(\mathbf{r})+\frac{i}{2}\sqrt{\frac{\mu_0}{\varepsilon_0 n(\mathbf{r})}}\mathbf{H}(\mathbf{r})\\ \frac{1}{2}\sqrt{n(\mathbf{r})}\mathbf{E}(\mathbf{r})-\frac{i}{2}\sqrt{\frac{\mu_0}{\varepsilon_0 n(\mathbf{r})}}\mathbf{H}(\mathbf{r})\end{pmatrix}, \tag{13}$$

the matrix n is just the unit matrix times the index of refraction,

$$\mathsf{n}=\begin{pmatrix}n(\mathbf{r}) & 0\\ 0 & n(\mathbf{r})\end{pmatrix}, \tag{14}$$

and the Hermitian matrix operator M is given by

$$\mathsf{M}=\begin{pmatrix}c\nabla & -\frac{c\nabla n(\mathbf{r})}{2n(\mathbf{r})}\\ \frac{c\nabla n(\mathbf{r})}{2n(\mathbf{r})} & -c\nabla\end{pmatrix}. \tag{15}$$

Here and in equations below the combination of vector cross product and matrix multiplication is handled in the obvious way; for example, writing

$$\mathbf{\Phi}=\begin{pmatrix}\boldsymbol{\phi}^+(\mathbf{r})\\ \boldsymbol{\phi}^-(\mathbf{r})\end{pmatrix}, \tag{16}$$

we have

$$\mathsf{M}\times\mathbf{\Phi}\equiv\begin{pmatrix} c\nabla\times\boldsymbol{\phi}^{+}(\mathbf{r})-\dfrac{c\nabla n(\mathbf{r})}{2n(\mathbf{r})}\times\boldsymbol{\phi}^{-}(\mathbf{r}) \\ \dfrac{c\nabla n(\mathbf{r})}{2n(\mathbf{r})}\times\boldsymbol{\phi}^{+}(\mathbf{r})-c\nabla\times\boldsymbol{\phi}^{-}(\mathbf{r}) \end{pmatrix}. \tag{17}$$

We divide the solutions of Eq. (12) into those of type T, for which $\omega\neq 0$, and type L for which $\omega=0$.

For a photonic crystal, the case of interest, M and n inherit the periodicity of the index of refraction, and hence we can seek solutions of Eq. (12) in Bloch form

$$\mathbf{\Phi}_{\alpha\mathbf{k}}=\mathbf{U}_{\alpha\mathbf{k}}e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{18}$$

where α labels the different solutions at a given $\mathbf{k}$,

$$\mathbf{U}_{\alpha\mathbf{k}}=\begin{pmatrix}\mathbf{u}^{+}_{\alpha\mathbf{k}}(\mathbf{r})\\ \mathbf{u}^{-}_{\alpha\mathbf{k}}(\mathbf{r})\end{pmatrix}, \tag{19}$$

and $\mathbf{u}_{\alpha\mathbf{k}}(\mathbf{r})=\mathbf{u}_{\alpha\mathbf{k}}(\mathbf{r}+\mathbf{R})$ for any lattice constant $\mathbf{R}$. Since the equation (12) is equivalent to equations (5), there is a one-to-one relation between P_1 solutions and T solutions. We thus have T solutions $\mathbf{\Phi}_{mT\mathbf{k}}$ of Eq. (12) with $\omega=\omega_{m\mathbf{k}}$, and T solutions $\mathbf{\Phi}_{\bar{m}T\mathbf{k}}$ of Eq. (12) with $\omega=\omega_{\bar{m}\mathbf{k}}$, where

$$\mathbf{U}_{\tilde{m}T\mathbf{k}}=\begin{pmatrix}\dfrac{1}{2}\sqrt{n(\mathbf{r})}\mathbf{e}_{\tilde{m}\mathbf{k}}(\mathbf{r})+\dfrac{i}{2}\sqrt{\dfrac{\mu_0}{\varepsilon_0 n(\mathbf{r})}}\mathbf{h}_{\tilde{m}\mathbf{k}}(\mathbf{r})\\ \dfrac{1}{2}\sqrt{n(\mathbf{r})}\mathbf{e}_{\tilde{m}\mathbf{k}}(\mathbf{r})-\dfrac{i}{2}\sqrt{\dfrac{\mu_0}{\varepsilon_0 n(\mathbf{r})}}\mathbf{h}_{\tilde{m}\mathbf{k}}(\mathbf{r})\end{pmatrix}, \tag{20}$$

here and below we use $\tilde{m}$ to indicate either m or $\bar{m}$, and we use the subscript T to distinguish between the type T solutions and the type L solutions, denoted $\mathbf{\Phi}_{\tilde{m}L\mathbf{k}}$, that we discuss below. Note that $\mathbf{\Phi}_{\bar{m}T\mathbf{k}}=\overline{\mathbf{\Phi}}_{mT(-\mathbf{k})}$, where for every $\mathbf{\Phi}$ of the form (16) we define

$$\overline{\mathbf{\Phi}}\equiv\begin{pmatrix}[\boldsymbol{\phi}^{-}(\mathbf{r})]^{*}\\ [\boldsymbol{\phi}^{+}(\mathbf{r})]^{*}\end{pmatrix}. \tag{21}$$

The type L solutions satisfy $\mathsf{M}\times\mathbf{\Phi}=\mathbf{0}$, or

$$\nabla\times\mathbf{H}=\mathbf{0},$$

$$\nabla\times\mathbf{E}=\mathbf{0}. \tag{22}$$

Hence, as for P_2 solutions, we can separately consider L solutions for which $\mathbf{E}=\mathbf{0}$ and those for which $\mathbf{H}=\mathbf{0}$. But while there will be $\mathbf{\Phi}_{\tilde{m}L\mathbf{k}}$ solutions that correspond to P_2 solutions, there will be others as well, since the electric and magnetic fields associated with L solutions need not satisfy Eq. (6) while those of P_2 solutions must. The presence of such extra, unphysical solutions of eigenvalue equations associated with photonic crystal problems is not unusual; for example, the master equation (8) has unphysical solutions at $\omega=0$ that do not satisfy $\nabla\cdot\mathbf{H}=0$ [14]. As in the analysis of that equation, we will typically not need to explicitly construct these unphysical solutions. But it will be important to identify their form and, at least in principle, enumerate them.

We now turn to the L solutions. There are a host of these solutions characterized by any given crystal momentum $\mathbf{k}$, which all satisfy equation (12) with eigenvalue zero. Nonetheless, to maintain a consistent notation below, we label these by subscripts $mL\mathbf{k}$, where m simply labels the different L solutions with crystal momentum $\mathbf{k}$. Consider first the L solutions for which the associated $\mathbf{E}=\mathbf{0}$, which we denote by $\mathbf{\Phi}'_{mL\mathbf{k}}$. We can form a complete set of these solutions by taking

$$\mathbf{\Phi}'_{mL\mathbf{k}}=\frac{i}{\sqrt{2n(\mathbf{r})}}\mathbf{L}_{m\mathbf{k}}(\mathbf{r})\begin{pmatrix}1\\ -1\end{pmatrix}, \tag{23}$$

where for each crystal wave vector $\mathbf{k}$ and each reciprocal lattice vector $\mathbf{G}_m$ we have a longitudinal field

$$\mathbf{L}_{m\mathbf{k}}(\mathbf{r})=\upsilon(\mathbf{k}+\mathbf{G}_m)e^{i(\mathbf{k}+\mathbf{G}_m)\cdot\mathbf{r}}, \tag{24}$$

where for a general vector $\mathbf{V}$ we have set

$$\upsilon(\mathbf{V})=\mathbf{V}/|\mathbf{V}|. \tag{25}$$

As $\mathbf{k}$ ranges over the first Brillouin zone and $\mathbf{G}_m$ ranges over the reciprocal lattice, $\mathbf{k}+\mathbf{G}_m$ ranges over all of reciprocal space; thus the $\mathbf{L}_{m\mathbf{k}}(\mathbf{r})$ form a basis set for expanding all longitudinal functions. Uniform solutions ($\mathbf{k}=\mathbf{G}_m=\mathbf{0}$) are a special case; these are both longitudinal and transverse, and instead of one basis function we in fact have three. We can take these to be, for example, $\mathbf{L}_{(x)\mathbf{00}}(\mathbf{r})=\hat{\mathbf{x}}$, $\mathbf{L}_{(y)\mathbf{00}}(\mathbf{r})=\hat{\mathbf{y}}$, and $\mathbf{L}_{(z)\mathbf{00}}(\mathbf{r})=\hat{\mathbf{z}}$. The $\mathbf{\Phi}'_{(i)\bar{0}L\mathbf{0}}$ columns associated with the fields $\mathbf{L}_{(i)\mathbf{00}}(\mathbf{r})$ correspond to P_2 solutions; typically the remaining L solutions with vanishing $\mathbf{E}$ are not solutions of the full set of Maxwell equations.

To construct the L solutions for which the associated $\mathbf{H}=\mathbf{0}$, we take

$$\mathbf{\Phi}''_{mL\mathbf{k}}=\sqrt{\frac{n(\mathbf{r})}{2}}\mathbf{J}_{m\mathbf{k}}(\mathbf{r})\begin{pmatrix}1\\ 1\end{pmatrix}, \tag{26}$$

where, at a given $\mathbf{k}$, we assemble the $\mathbf{J}_{m\mathbf{k}}(\mathbf{r})$ by including first the electric fields of any type P_2 solutions, orthogonalized and normalized according to Eq. (10). We add to these a set of solutions at each crystal wave vector $\mathbf{k}$ that can be obtained by starting with an $\mathbf{L}_{m\mathbf{k}}(\mathbf{r})$, Gram-Schmidt orthogonalizing according to Eq. (10) to any of the type P_2 solutions at that $\mathbf{k}$, and to any other irrotational solutions at that $\mathbf{k}$ already so identified. Carrying this procedure to the exhaustion of the $\mathbf{L}_{m\mathbf{k}}(\mathbf{r})$, a complete set of irrotational solutions orthogonalized according to Eq. (10) can, in principle, be identified. Of these, only the $\mathbf{J}_{m\mathbf{k}}(\mathbf{r})$ of the P_2 solutions are solutions to the full Maxwell equations (5) and (6).

Instead of the solutions $\mathbf{\Phi}'_{mL\mathbf{k}}$ and $\mathbf{\Phi}''_{mL\mathbf{k}}$, for which $\mathbf{E}=\mathbf{0}$ and $\mathbf{H}=\mathbf{0}$, respectively, it is more convenient to work with columns that, like $\mathbf{\Phi}_{\tilde{m}T\mathbf{k}}$, have both an associated $\mathbf{E}$ and $\mathbf{H}$ field. We therefore introduce

$$\boldsymbol{\Phi}_{mL\mathbf{k}} \equiv \frac{1}{\sqrt{2}}(\boldsymbol{\Phi}'_{mL\mathbf{k}}+\boldsymbol{\Phi}''_{mL\mathbf{k}}),$$
$$\boldsymbol{\Phi}_{\bar{m}L\mathbf{k}} \equiv \overline{\boldsymbol{\Phi}}_{mL(-\mathbf{k})}. \tag{27}$$

All the columns $\boldsymbol{\Phi}_{\tilde{m}S\mathbf{k}}$ thus identified, where $\tilde{m}$ is an m or an $\bar{m}$, and S is T or L, are orthonormalized using n as a metric

$$\int \frac{d^3r}{\Omega} \boldsymbol{\Phi}^{\dagger}_{\tilde{m}S\mathbf{k}} \cdot \mathsf{n} \boldsymbol{\Phi}_{\tilde{m}'S'\mathbf{k}'} = \delta_{\tilde{m}\tilde{m}'}\delta_{SS'}\delta_{\mathbf{kk}'}, \tag{28}$$

and we have defined the row

$$\boldsymbol{\Phi}^{\dagger}_{\tilde{m}S\mathbf{k}} = ((\boldsymbol{\phi}^{+}_{\tilde{m}S\mathbf{k}})^{*} \;\; (\boldsymbol{\phi}^{-}_{\tilde{m}S\mathbf{k}})^{*}), \tag{29}$$

etc. The orthonormalization condition (28) is easily confirmed using the orthogonality conditions (9) and (10), the fact that the integral of the dot product of a (nonuniform) longitudinal field and a (nonuniform) transverse field vanishes, and the conditions given after Eq. (6).

For convenience we will denote $(\tilde{m}S)$ by α. Likewise, if $\tilde{m}=m$ and $S=T$ we let $\omega_{\alpha\mathbf{k}}$ indicate $\omega_{m\mathbf{k}}$, while if $\tilde{m}=\bar{m}$ and $S=T$ we let $\omega_{\alpha\mathbf{k}}$ indicate $\omega_{\bar{m}\mathbf{k}}$; if $S=L$ we put $\omega_{\alpha\mathbf{k}}=0$. Finally, for $\mathbf{k}'=\mathbf{k}$ the condition (28) leads to an orthonormality condition for the $\mathbf{U}_{\alpha\mathbf{k}}$,

$$\int_{cell} \frac{d^3r}{\Omega_{cell}} \mathbf{U}^{\dagger}_{\alpha\mathbf{k}} \cdot \mathsf{n} \mathbf{U}_{\alpha'\mathbf{k}} = \delta_{\alpha\alpha'}, \tag{30}$$

where $\delta_{\alpha\alpha'}=\delta_{\tilde{m}\tilde{m}'}\delta_{SS'}$, Ω_{cell} is the volume of the unit cell, and the integral ranges over the unit cell.

Finally, we note from (13) and (18) and the following equations we have, for $S=T$,

$$\mathbf{E}_{\alpha\mathbf{k}}(\mathbf{r}) = \frac{\boldsymbol{\phi}^{+}_{\alpha\mathbf{k}}(\mathbf{r})+\boldsymbol{\phi}^{-}_{\alpha\mathbf{k}}(\mathbf{r})}{\sqrt{n(\mathbf{r})}},$$
$$\mathbf{H}_{\alpha\mathbf{k}}(\mathbf{r}) = -i\sqrt{\frac{\varepsilon_0 n(\mathbf{r})}{\mu_0}}[\boldsymbol{\phi}^{+}_{\alpha\mathbf{k}}(\mathbf{r})-\boldsymbol{\phi}^{-}_{\alpha\mathbf{k}}(\mathbf{r})], \tag{31}$$

where the functions $\mathbf{E}_{\tilde{m}T\mathbf{k}}(\mathbf{r})$ and $\mathbf{H}_{\tilde{m}T\mathbf{k}}(\mathbf{r})$ are simply the functions $\mathbf{E}_{\tilde{m}\mathbf{k}}(\mathbf{r})$ and $\mathbf{H}_{\tilde{m}\mathbf{k}}(\mathbf{r})$ identified by Eq. (7) and the following discussion. For $S=L$ we take Eq. (31) to be the definitions of $\mathbf{E}_{\alpha\mathbf{k}}(\mathbf{r})$ and $\mathbf{H}_{\alpha\mathbf{k}}(\mathbf{r})$. Then from Eqs. (5) for $S=T$, and Eq. (22) for $S=L$, we have

$$-i\omega_{\alpha\mathbf{k}}\varepsilon_0 n^2(\mathbf{r})\mathbf{E}_{\alpha\mathbf{k}}(\mathbf{r}) = \boldsymbol{\nabla}\times\mathbf{H}_{\alpha\mathbf{k}}(\mathbf{r}),$$
$$i\omega_{\alpha\mathbf{k}}\mu_0\mathbf{H}_{\alpha\mathbf{k}}(\mathbf{r}) = \boldsymbol{\nabla}\times\mathbf{E}_{\alpha\mathbf{k}}(\mathbf{r}), \tag{32}$$

for both types of solutions. Further, the conditions (9) and (10) that hold for $\mathbf{E}_{m\mathbf{k}}(\mathbf{r})$ and $\mathbf{H}_{m\mathbf{k}}(\mathbf{r})$ also hold, by construction, for all the $\mathbf{E}_{\tilde{m}S\mathbf{k}}(\mathbf{r})$ and $\mathbf{H}_{\tilde{m}S\mathbf{k}}(\mathbf{r})$. That is, we have

$$\frac{\mu_0}{\varepsilon_0}\int \frac{d^3r}{\Omega}\mathbf{H}^{*}_{\alpha\mathbf{k}}(\mathbf{r})\cdot\mathbf{H}_{\alpha'\mathbf{k}'}(\mathbf{r}) = \delta_{\alpha\alpha'}\delta_{\mathbf{kk}'} \tag{33}$$

$$\int \frac{d^3r}{\Omega} n^2(\mathbf{r})\mathbf{E}^{*}_{\alpha\mathbf{k}}(\mathbf{r})\cdot\mathbf{E}_{\alpha'\mathbf{k}'}(\mathbf{r}) = \delta_{\alpha\alpha'}\delta_{\mathbf{kk}'}.$$

C. $k\cdot p$ equations

An advantage of working with the matrix eigenvalue equation (12) is that the $k\cdot p$ expansion takes a much simpler form than if, say, one works with the master equation (8) [14]. Using the Bloch form (18) in the matrix eigenvalue equation, we find the equation that the periodic part $\mathbf{U}_{\alpha\mathbf{k}}$ of that Bloch function must satisfy,

$$\mathsf{H}(\mathbf{k})\times\mathbf{U}_{\alpha\mathbf{k}} = \omega_{\alpha\mathbf{k}}\mathsf{n}\mathbf{U}_{\alpha\mathbf{k}}, \tag{34}$$

where $\mathsf{H}(\mathbf{k})\equiv\mathsf{M}+\mathsf{v}(\mathbf{k})$, and

$$\mathsf{v}(\mathbf{k}) \equiv \begin{pmatrix} ic\mathbf{k} & 0 \\ 0 & -ic\mathbf{k} \end{pmatrix}. \tag{35}$$

We now expand in the usual way about a given wave vector $\mathbf{k}_0$,

$$[\mathsf{H}(\mathbf{k}_0)+\mathsf{v}(\boldsymbol{\kappa})]\times\mathbf{U}_{\alpha\mathbf{k}} = \omega_{\alpha\mathbf{k}}\mathsf{n}\mathbf{U}_{\alpha\mathbf{k}}, \tag{36}$$

where $\boldsymbol{\kappa}\equiv\mathbf{k}-\mathbf{k}_0$. Although generalization is easily done, for simplicity we here choose an $(\alpha\mathbf{k})$ such that the indicated band is nondegenerate; this of course requires that $S=T$. Then we can expand $\omega_{\alpha\mathbf{k}}$ about $\mathbf{k}_0$,

$$\omega_{\alpha\mathbf{k}} = \omega_{\alpha\mathbf{k}_0}+\kappa^i\omega^{i(1)}_{\alpha\mathbf{k}_0}+\kappa^i\kappa^j\omega^{ij(2)}_{\alpha\mathbf{k}_0}+O(\boldsymbol{\kappa}^3), \tag{37}$$

where the $\omega^{i(1)}_{\alpha\mathbf{k}_0}$ and $\omega^{ij(2)}_{\alpha\mathbf{k}_0}$ are expansion coefficients, and the superscripts indicate Cartesian components that are summed over when repeated. We use the $\mathbf{U}_{\beta\mathbf{k}_0}$ as a basis for expanding $\mathbf{U}_{\alpha\mathbf{k}}$,

$$\mathbf{U}_{\alpha\mathbf{k}} = \mathbf{U}_{\alpha\mathbf{k}_0}+\kappa^i{\sum_{\beta}}' a^{i(1)}_{\beta}\mathbf{U}_{\beta\mathbf{k}_0}+\kappa^i\kappa^j{\sum_{\beta}}' a^{ij(2)}_{\beta}\mathbf{U}_{\beta\mathbf{k}_0} + O(\boldsymbol{\kappa}^3), \tag{38}$$

where the $a^{i(1)}_{\beta}$ and $a^{ij(2)}_{\beta}$ are expansion coefficients, and we only sum over the bands as indicated; the prime denotes that the α band is omitted from the sum. This will simplify the following derivation, at the cost of foregoing normalization (30) of the $\mathbf{U}_{\alpha\mathbf{k}}$; that can always be done later and, since our main interest is in extracting the $\omega^{i(1)}_{\alpha\mathbf{k}_0}$ and $\omega^{ij(2)}_{\alpha\mathbf{k}_0}$, it will not be a concern. We do assume that the $\mathbf{U}_{\beta\mathbf{k}_0}$ satisfy the normalization condition (30).

Inserting Eqs. (37) and (38) into Eq. (36), we collect terms according to their order in $\boldsymbol{\kappa}$. The zeroth order terms give Eq. (34) with $\mathbf{k}$ replaced by $\mathbf{k}_0$. The first order terms, when simplified using the eigenvalue equation (34) for $\mathbf{U}_{\beta\mathbf{k}_0}$, yield

$$\omega^{i(1)}_{\alpha\mathbf{k}_0} = v^{i}_{\alpha\alpha}(\mathbf{k}_0) \tag{39}$$

when dotted into $\mathbf{U}^{\dagger}_{\alpha\mathbf{k}_0}$ and integrated over a unit cell using Eq. (30), and

$$a^{i(1)}_{\beta} = -\frac{v^{i}_{\beta\alpha}(\mathbf{k}_0)}{\omega_{\beta\mathbf{k}_0} - \omega_{\alpha\mathbf{k}_0}} \tag{40}$$

when dotted into one of the $\mathbf{U}^{\dagger}_{\beta\mathbf{k}_0}$, with $\beta \neq \alpha$, and integrated over a unit cell. Here we have defined

$$\mathbf{v}_{\alpha'\alpha}(\mathbf{k}) \equiv -i\int_{cell}\frac{d^3r}{\Omega_{cell}}\mathbf{U}^{\dagger}_{\alpha'\mathbf{k}} \times \mathsf{V}\mathbf{U}_{\alpha\mathbf{k}} \tag{41}$$

for arbitrary α', α, and $\mathbf{k}$, and

$$\mathsf{V} \equiv \begin{pmatrix} c & 0 \\ 0 & -c \end{pmatrix}. \tag{42}$$

Similarly, the second order terms yield

$$\omega^{ij(2)}_{\alpha\mathbf{k}_0} = -\sum_{\beta}{}' \frac{v^{j}_{\alpha\beta}(\mathbf{k}_0)v^{i}_{\beta\alpha}(\mathbf{k}_0)}{\omega_{\beta\mathbf{k}_0} - \omega_{\alpha\mathbf{k}_0}} \tag{43}$$

when dotted into $\mathbf{U}^{\dagger}_{\alpha\mathbf{k}_0}$ and integrated over a unit cell using Eq. (30), after substituting Eq. (40) in the result. From Eq. (37) and its derivatives about $\boldsymbol{\kappa} = \mathbf{0}$, we can use Eqs. (39) and (43) to identify the general expressions

$$\frac{\partial \omega_{\alpha\mathbf{k}}}{\partial k^i} = v^{i}_{\alpha\alpha}(\mathbf{k}), \tag{44}$$

and

$$\frac{\partial^2 \omega_{\alpha\mathbf{k}}}{\partial k^i \partial k^j} = -2\sum_{\beta}{}' \frac{v^{i}_{\alpha\beta}(\mathbf{k})v^{j}_{\beta\alpha}(\mathbf{k})}{\omega_{\beta\mathbf{k}} - \omega_{\alpha\mathbf{k}}}, \tag{45}$$

for a band α that is nondegenerate at $\mathbf{k}$; these are the main results of this section. Working out the expression (41) in terms of the photonic band functions (31), we find

$$\mathbf{v}_{\alpha'\alpha}(\mathbf{k}) = \frac{1}{2\varepsilon_0}\int_{cell}\frac{d^3r}{\Omega_{cell}}[\mathbf{E}^{*}_{\alpha'\mathbf{k}}(\mathbf{r}) \times \mathbf{H}_{\alpha\mathbf{k}}(\mathbf{r}) + \mathbf{E}_{\alpha\mathbf{k}}(\mathbf{r}) \times \mathbf{H}^{*}_{\alpha'\mathbf{k}}(\mathbf{r})]. \tag{46}$$

The velocity matrix element between two L states vanishes, since both the electric and magnetic fields in those states are longitudinal. Quite generally the matrix elements between two T states, and between a T state and an L state, will be nonzero. The diagonal matrix element

$$\mathbf{v}_{\alpha\alpha}(\mathbf{k}) = \frac{1}{\varepsilon_0}\int_{cell}\frac{d^3r}{\Omega_{cell}}\mathrm{Re}[\mathbf{E}^{*}_{\alpha\mathbf{k}}(\mathbf{r}) \times \mathbf{H}_{\alpha\mathbf{k}}(\mathbf{r})] \tag{47}$$

is proportional to the time average of the Poynting vector averaged over the Bloch state, as might be expected since it is the group velocity (44) associated with that Bloch state. However, this simple relation between the group velocity and the average Poynting vector in a photonic crystal does not seem to have been realized before. Finally, we note that the sum in the expression (45) for the group velocity dispersion involves both physical T states and unphysical L states. As in the corresponding expression for the group velocity dispersion that can be derived from the master equation (8) [14], the part of that sum involving unphysical states can be written in terms of physical states, yielding an expression involving a sum over only physical states. We do not present that here because it will not be of use to us. Of course, in most practical applications of the $k\cdot p$ method it is only bands β that are close in frequency to the band α of interest that will make a significant contribution to the sum (45).

III. DYNAMICAL EQUATIONS

We now set up the full dynamical field equations. Instead of the simple constitutive relation (4) for the displacement field, we now consider the more general form

$$\mathbf{D}(\mathbf{r},t) = \varepsilon_0 n^2(\mathbf{r})\mathbf{E}(\mathbf{r},t) + \mathbf{P}(\mathbf{r},t), \tag{48}$$

where the additional polarization $\mathbf{P}(\mathbf{r},t)$ describes the nonlinear response as well as material dispersion and gain or loss in the media. We will turn to the expression for $\mathbf{P}(\mathbf{r},t)$ in the following section. Using the constitutive relations (3) and (48) in the Maxwell curl equations (2), we find

$$\begin{gathered}\varepsilon_0 n^2(\mathbf{r})\dot{\mathbf{E}}(\mathbf{r},t) = \boldsymbol{\nabla} \times \mathbf{H}(\mathbf{r},t) - \dot{\mathbf{P}}(\mathbf{r},t),\\ \mu_0\dot{\mathbf{H}}(\mathbf{r},t) = -\boldsymbol{\nabla} \times \mathbf{E}(\mathbf{r},t)\end{gathered} \tag{49}$$

as the dynamical equations that the fields $\mathbf{E}(\mathbf{r},t)$ and $\mathbf{H}(\mathbf{r},t)$ satisfy. But these equations must be solved subject to the initial conditions (1). Particularly in the context of the multiple-scales analysis that we will introduce, these initial conditions are not easy to implement; it is generally easier to work with dynamical equations that are not so restricted. We can do this by introducing potentials that automatically guarantee Eq. (1) in the following way. For any reasonable, nonuniform initial fields $\mathbf{B}$ and $\mathbf{D}$ satisfying Eq. (1), we can introduce potentials $\mathbf{A}$ and $\mathbf{N}$,

$$\begin{gathered}\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A},\\ \mathbf{D} = -\boldsymbol{\nabla} \times \mathbf{N}\end{gathered} \tag{50}$$

that so describe those fields. Then it is easy to confirm that, for fields $\mathbf{B}$ and $\mathbf{D}$ given by Eq. (50) at later times, the Maxwell equations (2) are satisfied, along with the constitutive relations (3) and (48), if the potentials $\mathbf{A}$ and $\mathbf{N}$ satisfy the dynamical equations

$$\begin{gathered}\varepsilon_0 n^2(\mathbf{r})\dot{\mathbf{A}}(\mathbf{r},t) = \boldsymbol{\nabla} \times \mathbf{N}(\mathbf{r},t) + \mathbf{P}(\mathbf{r},t),\\ \mu_0\dot{\mathbf{N}}(\mathbf{r},t) = -\boldsymbol{\nabla} \times \mathbf{A}(\mathbf{r},t).\end{gathered} \tag{51}$$

These we will take as our basic dynamical equations; clearly there are no subsidiary conditions that initial values of $\mathbf{A}$ and $\mathbf{N}$ must satisfy. Once the potentials are found at later times from Eqs. (51), the electromagnetic fields themselves can be obtained from Eqs. (50) and the constitutive relations (3) and (48).

It is convenient to write the dynamical equations (51) in matrix form as

$$i\mathsf{n}\frac{\partial\boldsymbol{\Psi}}{\partial t}=\mathsf{M}\times\boldsymbol{\Psi}+i\boldsymbol{\Xi}, \qquad (52)$$

where the column $\boldsymbol{\Psi}$ of complex vector fields is given by

$$\boldsymbol{\Psi}=\begin{pmatrix}\frac{1}{2}\sqrt{n(\mathbf{r})}\mathbf{A}(\mathbf{r},t)+\frac{i}{2}\sqrt{\frac{\mu_0}{\varepsilon_0 n(\mathbf{r})}}\mathbf{N}(\mathbf{r},t)\\ \frac{1}{2}\sqrt{n(\mathbf{r})}\mathbf{A}(\mathbf{r},t)-\frac{i}{2}\sqrt{\frac{\mu_0}{\varepsilon_0 n(\mathbf{r})}}\mathbf{N}(\mathbf{r},t)\end{pmatrix}, \qquad (53)$$

and the column

$$\boldsymbol{\Xi}=\boldsymbol{\xi}(\mathbf{r},t)\begin{pmatrix}1\\1\end{pmatrix}, \qquad (54)$$

where

$$\xi(\mathbf{r},t)\equiv\frac{\mathbf{P}(\mathbf{r},t)}{2\varepsilon_0\sqrt{n(\mathbf{r})}} \qquad (55)$$

is purely real. Note that not any arbitrary column of complex vector fields,

$$\boldsymbol{\Psi}=\begin{pmatrix}\boldsymbol{\psi}^{+}(\mathbf{r},t)\\ \boldsymbol{\psi}^{-}(\mathbf{r},t)\end{pmatrix} \qquad (56)$$

will represent real potentials $\mathbf{A}(\mathbf{r},t)$ and $\mathbf{N}(\mathbf{r},t)$ according to Eq. (53); however, given an arbitrary column $\boldsymbol{\Psi}$ we can construct a column $\boldsymbol{\Psi}+\overline{\boldsymbol{\Psi}}$ representing real potentials, where the physical conjugate $\overline{\boldsymbol{\Psi}}$ of a column vector (56) is defined according to

$$\overline{\boldsymbol{\Psi}}=\begin{pmatrix}[\boldsymbol{\psi}^{-}(\mathbf{r},t)]^{*}\\ [\boldsymbol{\psi}^{+}(\mathbf{r},t)]^{*}\end{pmatrix}. \qquad (57)$$

(cf. the corresponding expression (21) for eigenfunctions of M). Typically we will write a column (53) as $\boldsymbol{\Psi}_{+}+\boldsymbol{\Psi}_{-}$, where $\boldsymbol{\Psi}_{-}\equiv\overline{\boldsymbol{\Psi}}_{+}$. Defining the operator

$$\pounds\equiv i\mathsf{n}\frac{\partial}{\partial t}-\mathsf{M}\times, \qquad (58)$$

we see that the physical conjugate of $\pounds\boldsymbol{\Psi}_{+}$ is $-\pounds\boldsymbol{\Psi}_{-}$. Hence, if we likewise have $\boldsymbol{\Xi}=\boldsymbol{\Xi}_{+}+\boldsymbol{\Xi}_{-}$ and satisfy

$$\pounds\boldsymbol{\Psi}_{+}=i\boldsymbol{\Xi}_{+}, \qquad (59)$$

then the dynamical equations (52) will be satisfied. It is this last equation that we will reduce by a multiple-scales analysis.

IV. MULTIPLE SCALES

In order to reduce Eq. (59) to a simpler form, we use the asymptotic method of multiple scales. We seek to scale the physical parameters in Eq. (59) to a small parameter η, then keep terms in Eq. (59) up to a given order in η. By satisfying Eq. (59) to successively higher orders in η, we will better capture the dynamics of the exact solution, in an asymptotic sense.

A. Fields

In the absence of nonlinearity, material dispersion, and gain or loss, the general the solution of Eq. (59) is of the form

$$\boldsymbol{\Psi}_{+}=a\sum_{\beta,\mathbf{k}}f_{\beta\mathbf{k}}\boldsymbol{\Phi}_{\beta\mathbf{k}}e^{-i\omega_{\beta\mathbf{k}}t}, \qquad (60)$$

where $f_{\beta\mathbf{k}}$ are dimensionless coefficients used to expand arbitrary initial conditions in terms of the complete set of eigencolumns $\boldsymbol{\Phi}_{\beta\mathbf{k}}$, and where *a* carries the units of $\boldsymbol{\Psi}_{+}$. By construction, the subsequent dynamics are linear, so that in the presence of nonlinearity, it is not possible to form an exact solution using Eq. (60). One way of proceeding is to solve the nonlinear eigenproblem [2]; however, this approach does not give propagating pulse solutions, and the loss of linear superposition precludes a general solution of the form (60). Even in the absence of nonlinearity an expansion (60) may not be the best way to identify the nature of the solution. So we construct approximate solutions of Eq. (59) by replacing the constant coefficients $f_{\beta\mathbf{k}}$ in Eq. (60) with functions $f_{\beta\mathbf{k}}(\mathbf{r},t)$ that vary slowly in space and time. Furthermore, to keep track of the ''slowness,'' we explicitly separate different length and time scales in the problem, by introducing a small parameter η and writing functions as

$$f(\mathbf{r},t)=F(x,\eta x,\eta^2 x\ \ldots,y,\eta y,\eta^2 y\ \ldots,z,\eta z,\eta^2 z\ \ldots;t,\eta t,\eta^2 t,\ \ldots)\equiv F(\mathbf{r}_0,\mathbf{r}_1,\mathbf{r}_2,\ \ldots;t_0,t_1,t_2,\ \ldots), \qquad (61)$$

where *F* is assumed to vary equally significantly as each of its arguments varies over a range *d*, or a period τ. These are chosen to be the shortest length and time scales in the problem; *d*, for example, is taken as the size of a lattice constant and τ will be identified below. Then the ranges and periods $d^{(p)}=d/\eta^p$ and $\tau^{(p)}=\tau/\eta^p$ for $p=0,1,2\cdots$ define the multiple scales of the problem. Derivatives are given by, for example,

$$\frac{\partial f(\mathbf{r},t)}{\partial t}=\frac{\partial F}{\partial t_0}+\eta\frac{\partial F}{\partial t_1}+\eta^2\frac{\partial F}{\partial t_2}+\cdots. \qquad (62)$$

We will be looking at situations in which one of the $f_{\alpha\mathbf{k}}(\mathbf{r},t)$ will be larger than the rest. This will be called the principal component, and we take α to identify one of the physical $(S=T)$ solutions $\mathbf{\Phi}$ of the matrix equation (12) discussed in Sec. II; we further identify the shortest time scale of the problem to be the period associated with $\omega_{\alpha\mathbf{k}}$, $\tau=2\pi/\omega_{\alpha\mathbf{k}}$. Other components, labeled by β and called companion components, will have amplitudes smaller by a factor of η. The case of two principal components has been considered in one dimension [12], but for our purposes here one principal component will suffice. Then we seek an approximate solution of the form, $\mathbf{\Psi}=\mathbf{\Psi}_{+}+\mathbf{\Psi}_{-}$, where $\mathbf{\Psi}_{-}=\overline{\mathbf{\Psi}}_{+}$, and

$$\mathbf{\Psi}_{+}=a\left(f_{\alpha\mathbf{k}}(\mathbf{r},t)\mathbf{\Phi}_{\alpha\mathbf{k}}+\sum_{\beta}{}' f_{\beta\mathbf{k}}(\mathbf{r},t)\mathbf{\Phi}_{\beta\mathbf{k}}\right)e^{-i\omega_{\alpha\mathbf{k}}t_0}, \tag{63}$$

where the prime denotes exclusion of the principal, or α band. To capture the amplitude scaling of the principal (α) and companion $(\beta\neq\alpha)$ components, we take

$$\begin{aligned} f_{\alpha\mathbf{k}}(\mathbf{r},t)&=F^{(0)}_{\alpha\mathbf{k}}(\mathbf{r}_1,\mathbf{r}_2,\ldots;t_1,t_2,\ldots)\\ f_{\beta\mathbf{k}}(\mathbf{r},t)&=\eta F^{(1)}_{\beta\mathbf{k}}(\mathbf{r}_1,\mathbf{r}_2,\ldots;t_1,t_2,\ldots)\\ &\quad+\eta^2F^{(2)}_{\beta\mathbf{k}}(\mathbf{r}_1,\mathbf{r}_2,\ldots;t_1,t_2,\ldots). \end{aligned} \tag{64}$$

Note that these slowly varying quantities have no dependence on $\mathbf{r}_0$ or t_0, and that the amplitudes of companion components are smaller than that of the principal component by a factor of η. Hereafter the arguments of the F will be kept implicit. To identify η, we look at the range Λ over which $f_{\alpha\mathbf{k}}(\mathbf{r},t=0)$ varies, and set

$$\eta=\frac{gd}{\Lambda}, \tag{65}$$

where g is a factor of order unity that we will set later to guarantee consistency. This identification of η means that the first variation of $\mathbf{\Psi}_{+}$ through space, above and beyond its dependence over a distance d through the Bloch function, is through its dependence on $\mathbf{r}_1$.

From the dependence of $\mathbf{\Psi}$ on $\mathbf{N}$ and $\mathbf{A}$ (53), we see that this $\mathbf{\Psi}_{+}$ leads to an

$$\mathbf{N}(\mathbf{r},t)=a\left(F^{(0)}_{\alpha\mathbf{k}}\mathbf{H}_{\alpha\mathbf{k}}(\mathbf{r})+\sum_{\beta}{}'\,[\eta F^{(1)}_{\beta\mathbf{k}}+\eta^2F^{(2)}_{\beta\mathbf{k}}+\cdots]\mathbf{H}_{\beta\mathbf{k}}(\mathbf{r})\right)e^{-i\omega_{\alpha\mathbf{k}}t_0}+\text{c.c.}, \tag{66}$$

where we have used Eq. (31) for the photonic band functions associated with the $\mathbf{\Phi}_{\gamma\mathbf{k}}$. We can now determine $\mathbf{D}(\mathbf{r},t)$ from its expression (50) in terms of the potential $\mathbf{N}(\mathbf{r},t)$, and then use the constitutive relation (48) to identify $\mathbf{E}(\mathbf{r},t)$. In doing so we anticipate a result of the following section that $\mathbf{P}(\mathbf{r},t)$ will be of order η^2; we then find

$$\mathbf{E}(\mathbf{r},t)=\mathbf{E}_{+}(\mathbf{r},t)+\text{c.c.}=\mathcal{E}(\mathbf{r},t)e^{-i\omega_{\alpha\mathbf{k}}t_0}+\text{c.c.}, \tag{67}$$

where

$$\mathcal{E}(\mathbf{r},t)=i\omega_{\alpha\mathbf{k}}aF^{(0)}_{\alpha\mathbf{k}}\mathbf{E}_{\alpha\mathbf{k}}(\mathbf{r})+O(\eta) \tag{68}$$

contains no dependence on t_0. Here we have used the relations (32) that the fields $\mathbf{E}_{\alpha\mathbf{k}}(\mathbf{r})$ and $\mathbf{H}_{\alpha\mathbf{k}}(\mathbf{r})$ satisfy. From this result we see that $\omega_{\alpha\mathbf{k}}a$ is a characteristic size of the electric field; we henceforth denote it by

$$e_0\equiv\omega_{\alpha\mathbf{k}}a. \tag{69}$$

We can also identify

$$\begin{aligned} \frac{\partial\mathcal{E}(\mathbf{r},t)}{\partial t}&=ie_0\eta\frac{\partial F^{(0)}_{\alpha\mathbf{k}}}{\partial t_1}\mathbf{E}_{\alpha\mathbf{k}}(\mathbf{r})+O(\eta^2),\\ \frac{\partial^2\mathcal{E}(\mathbf{r},t)}{\partial t^2}&=O(\eta^2), \end{aligned} \tag{70}$$

expressions that will be used in a following section.

B. Multiple-scales expansion

We now construct an expansion of $\pounds\mathbf{\Psi}_{+}$ in powers of η,

$$\pounds\mathbf{\Psi}_{+}=(\pounds\mathbf{\Psi}_{+})^{(0)}+\eta(\pounds\mathbf{\Psi}_{+})^{(1)}+\eta^2(\pounds\mathbf{\Psi}_{+})^{(2)}+\cdots \tag{71}$$

for the $\mathbf{\Psi}_{+}$ given above Eq. (63). To construct this we need the corresponding expansions of the terms that appear in $\pounds$ (58). These are simply

$$\frac{\partial}{\partial t}=\frac{\partial}{\partial t_0}+\eta\frac{\partial}{\partial t_1}+\eta^2\frac{\partial}{\partial t_2}+\cdots \tag{72}$$

and

$$\mathsf{M}=\mathsf{M}^{(0)}+\eta\mathsf{M}^{(1)}+\eta^2\mathsf{M}^{(2)}+\cdots, \tag{73}$$

where

$$\mathsf{M}^{(0)}=\begin{pmatrix} c\boldsymbol{\nabla}^{(0)} & -\dfrac{c\boldsymbol{\nabla}^{(0)}n(\mathbf{r}_0)}{2n(\mathbf{r}_0)}\\ \dfrac{c\boldsymbol{\nabla}^{(0)}n(\mathbf{r}_0)}{2n(\mathbf{r}_0)} & -c\boldsymbol{\nabla}^{(0)} \end{pmatrix} \tag{74}$$

and, for $j>0$,

$$\mathsf{M}^{(j)}=\begin{pmatrix} c\boldsymbol{\nabla}^{(j)} & 0\\ 0 & -c\boldsymbol{\nabla}^{(j)} \end{pmatrix}=\mathsf{V}\boldsymbol{\nabla}^{(j)}, \tag{75}$$

where $\boldsymbol{\nabla}^{(j)}$ denotes the gradient with respect to $\mathbf{r}_j$. Collecting all the terms and using the fact that the $\mathbf{\Phi}_{\gamma\mathbf{k}}$ are eigenfunctions of the operator $\mathsf{M}^{(0)}$, we find $(\pounds\mathbf{\Psi}_{+})^{(0)}=0$ while

$$a^{-1}e^{i\omega_{\alpha\mathbf{k}}t_0}(\pounds\boldsymbol{\Psi}_+)^{(1)}=i\frac{\partial F_{\alpha\mathbf{k}}^{(0)}}{\partial t_1}\mathsf{n}\boldsymbol{\Phi}_{\alpha\mathbf{k}}+\mathsf{V}\boldsymbol{\Phi}_{\alpha\mathbf{k}}\times\boldsymbol{\nabla}^{(1)}F_{\alpha\mathbf{k}}^{(0)}$$
$$+\sum_{\beta}{}'\,(\omega_{\alpha\mathbf{k}}-\omega_{\beta\mathbf{k}})F_{\beta\mathbf{k}}^{(1)}\mathsf{n}\boldsymbol{\Phi}_{\beta\mathbf{k}}, \tag{76}$$

and

$$a^{-1}e^{i\omega_{\alpha\mathbf{k}}t_0}(\pounds\boldsymbol{\Psi}_+)^{(2)}=i\frac{\partial F_{\alpha\mathbf{k}}^{(0)}}{\partial t_2}\mathsf{n}\boldsymbol{\Phi}_{\alpha\mathbf{k}}+\mathsf{V}\boldsymbol{\Phi}_{\alpha\mathbf{k}}\times\boldsymbol{\nabla}^{(2)}F_{\alpha\mathbf{k}}^{(0)}$$
$$+\sum_{\beta}{}'\left(i\frac{\partial F_{\beta\mathbf{k}}^{(1)}}{\partial t_1}\mathsf{n}\boldsymbol{\Phi}_{\beta\mathbf{k}}+\mathsf{V}\boldsymbol{\Phi}_{\beta\mathbf{k}}\right.$$
$$\left.\times\boldsymbol{\nabla}^{(1)}F_{\beta\mathbf{k}}^{(1)}\right)+\sum_{\beta}{}'\,(\omega_{\alpha\mathbf{k}}$$
$$-\omega_{\beta\mathbf{k}})F_{\beta\mathbf{k}}^{(2)}\mathsf{n}\boldsymbol{\Phi}_{\beta\mathbf{k}}. \tag{77}$$

In the absence of any nonlinearity, dispersion, or gain or loss, our equation (59) would simply require that $(\pounds\boldsymbol{\Psi}_+)^{(1)}$ and $(\pounds\boldsymbol{\Psi}_+)^{(2)}$ vanish. The results for this special case can be extracted from the more general analysis we present below to take into account a nonvanishing $\boldsymbol{\Xi}$.

V. NONLINEARITY, MATERIAL DISPERSION, AND GAIN OR LOSS

We now turn to the general form of the constitutive relations that we adopt, and identify the $\mathbf{P}(\mathbf{r},t)$ that appears in the relation (48) between $\mathbf{D}(\mathbf{r},t)$ and $\mathbf{E}(\mathbf{r},t)$ that we have assumed above. Instead of the constitutive relations (3) and (4) we adopt the relations

$$\mathbf{B}(\mathbf{r},t)=\mu_0\mathbf{H}(\mathbf{r},t),$$
$$\mathbf{D}(\mathbf{r},t)=\mathbf{D}_L(\mathbf{r},t)+\mathbf{P}_{NL}(\mathbf{r},t), \tag{78}$$

where $\mathbf{P}_{NL}(\mathbf{r},t)$ is a nonlinear contribution to the polarization that we will discuss shortly, and the linear response is taken to be of the form

$$\mathbf{D}_L(\mathbf{r},t)=\int_{-\infty}^{t}\varepsilon(\mathbf{r},t-t')\mathbf{E}(\mathbf{r},t')dt'. \tag{79}$$

The form of the linear response is general enough to include material dispersion and gain or loss, and could be easily extended to include material birefringence; in photonic crystals, $\varepsilon(\mathbf{r},t)=\varepsilon(\mathbf{r}+\mathbf{R},t)$ for any lattice vector $\mathbf{R}$.

A. Linear response

We introduce Fourier transforms of our fields in the usual way,

$$\mathbf{E}(\mathbf{r},t)=\int_{-\infty}^{+\infty}\frac{d\omega}{2\pi}\mathbf{E}(\mathbf{r},\omega)e^{-i\omega t}, \tag{80}$$

etc., but use $\Delta\omega$ to denote the Fourier transform variable of $\mathcal{E}(\mathbf{r},t)$. Using $\mathbf{D}_{L+}(\mathbf{r},t)$ to denote the portion of $\mathbf{D}_L(\mathbf{r},t)$ obtained by using only $\mathbf{E}_+(\mathbf{r},t)$ in Eq. (79), taking the Fourier transform yields

$$\mathbf{D}_{L+}(\mathbf{r},\omega_{\alpha\mathbf{k}}+\Delta\omega)=\varepsilon(\mathbf{r},\omega_{\alpha\mathbf{k}}+\Delta\omega)\mathcal{E}(\mathbf{r},\Delta\omega). \tag{81}$$

Now since $\mathcal{E}(\mathbf{r},t)$ only contains time dependence through variables t_j for $j>0$, the Fourier components $\Delta\omega$ that are important in $\mathcal{E}(\mathbf{r},\Delta\omega)$ satisfy $|\Delta\omega|\ll\omega_{\alpha\mathbf{k}}$. Hence we are led to expand

$$\varepsilon(\mathbf{r},\omega_{\alpha\mathbf{k}}+\Delta\omega)=\varepsilon(\mathbf{r},\omega_{\alpha\mathbf{k}})+(\Delta\omega)\varepsilon'(\mathbf{r})+\frac{1}{2}(\Delta\omega)^2\varepsilon''(\mathbf{r})$$
$$+\cdots, \tag{82}$$

where $\varepsilon'(\mathbf{r})\equiv\partial\varepsilon(\mathbf{r},\omega_{\alpha\mathbf{k}})/\partial\omega_{\alpha\mathbf{k}}$, $\varepsilon''(\mathbf{r})\equiv\partial^2\varepsilon(\mathbf{r},\omega_{\alpha\mathbf{k}})/\partial\omega_{\alpha\mathbf{k}}^2$, etc. We now introduce a nominal real refractive index $n(\mathbf{r})$ by putting

$$\varepsilon(\mathbf{r},\omega_{\alpha\mathbf{k}})=\varepsilon_0 n^2(\mathbf{r})+\varepsilon_C(\mathbf{r}), \tag{83}$$

where the ‘‘correction’’ permittivity $\varepsilon_C(\mathbf{r})$ is given by

$$\varepsilon_C(\mathbf{r})=\mathrm{Re}[\varepsilon(\mathbf{r},\omega_{\alpha\mathbf{k}})]-\varepsilon_0 n^2(\mathbf{r})+i\,\mathrm{Im}[\varepsilon(\mathbf{r},\omega_{\alpha\mathbf{k}})]. \tag{84}$$

The obvious choice for $n(\mathbf{r})$ is to set $\mathrm{Re}[\varepsilon(\mathbf{r},\omega_{\alpha\mathbf{k}})]-\varepsilon_0 n^2(\mathbf{r})=0$, but to describe electro-optic modifications of the dielectric constant it may be convenient to choose a different nominal index $n(\mathbf{r})$. Using Eqs. (82)–(84) in the expression (81) for $\mathbf{D}_{L+}(\mathbf{r},\omega)$, we inverse Fourier transform to identify $\mathbf{D}_{L+}(\mathbf{r},t)$, partially integrating to eliminate the powers of $\Delta\omega$; adding in the complex conjugate leads to the result

$$\mathbf{D}_L(\mathbf{r},t)=\varepsilon_0 n^2(\mathbf{r})\mathbf{E}(\mathbf{r},t)+\mathbf{P}_L(\mathbf{r},t), \tag{85}$$

where

$$\mathbf{P}_L(\mathbf{r},t)=\varepsilon_C(\mathbf{r})\mathcal{E}(\mathbf{r},t)e^{-i\omega_{\alpha\mathbf{k}}t}+i\varepsilon'(\mathbf{r})e^{-i\omega_{\alpha\mathbf{k}}t}\frac{\partial\mathcal{E}(\mathbf{r},t)}{\partial t}$$
$$-\frac{1}{2}\varepsilon''(\mathbf{r})e^{-i\omega_{\alpha\mathbf{k}}t}\frac{\partial^2\mathcal{E}(\mathbf{r},t)}{\partial t^2}+\cdots+\mathrm{c.c.} \tag{86}$$

B. Nonlinear response

Turning to the nonlinear response, we assume the nonlinearity is weak and due to a third-order response that, at frequencies of interest, is far off resonance. Then an appropriate model for the nonlinear polarization $\mathbf{P}_{NL}(\mathbf{r},t)$ is

$$P_{NL}^{i}(\mathbf{r},t)=\varepsilon_0\chi_3^{ijkl}(\mathbf{r})E^j(\mathbf{r},t)E^k(\mathbf{r},t)E^l(\mathbf{r},t), \tag{87}$$

where the third-order susceptibility is purely real, is unchanged under any permutation of the Cartesian components, and exhibits the periodicity of the photonic crystal, $\chi_3^{ijkl}(\mathbf{r})$

$=\chi_3^{ijkl}(\mathbf{r}+\mathbf{R})$, where **R** is any lattice vector. Using the expression (67) for the electric field $\mathbf{E}(\mathbf{r},t)$, this expression reduces to

$$P^i_{NL}(\mathbf{r},t)=3\varepsilon_0\chi_3^{ijkl}(\mathbf{r})\mathcal{E}^j(\mathbf{r},t)\mathcal{E}^k(\mathbf{r},t)[\mathcal{E}^l(\mathbf{r},t)]^*e^{-i\omega_{\alpha\mathbf{k}}t}+\text{c.c.}, \tag{88}$$

when we neglect third harmonic generation terms. These Fourier components of $\mathbf{P}_{NL}(\mathbf{r},t)$, centered at $3\omega_{\alpha\mathbf{k}}$ and $-3\omega_{\alpha\mathbf{k}}$ are typically non-phase-matched and, in any case, are usually at frequencies high enough where absorption is sufficient that significant fields at these frequencies are not generated. Exceptions to this have been considered in one-dimensional structures [15], but will not be considered here.

C. Scaling

We can now identify the $\mathbf{P}(\mathbf{r},t)$ in our posited expression (48) as

$$\mathbf{P}(\mathbf{r},t)=\mathbf{P}_L(\mathbf{r},t)+\mathbf{P}_{NL}(\mathbf{r},t), \tag{89}$$

with the linear and nonlinear contributions given by Eqs. (86) and (88), respectively. In order to write $\mathbf{P}(\mathbf{r},t)$ in powers of η, we need to scale the magnitudes of $\varepsilon_C(\mathbf{r})$, $\varepsilon'(\mathbf{r})$, $\varepsilon''(\mathbf{r})$, and $\chi_3^{ijkl}(\mathbf{r})$ to powers of η. For the linear response we assume

$$\begin{gathered}\varepsilon_C(\mathbf{r})=\eta^2\hat{\varepsilon}_C(\mathbf{r}),\\ \varepsilon'(\mathbf{r})=\eta\frac{\hat{\varepsilon}'(\mathbf{r})}{\omega_{\alpha\mathbf{k}}},\\ \varepsilon''(\mathbf{r})=\eta\frac{\hat{\varepsilon}''(\mathbf{r})}{\omega^2_{\alpha\mathbf{k}}},\end{gathered} \tag{90}$$

where $\hat{\varepsilon}_C(\mathbf{r})$, $\hat{\varepsilon}'(\mathbf{r})$, and $\hat{\varepsilon}''(\mathbf{r})$ have dimensions of permittivity and are assumed to be on the order of the largest linear media permittivity, or smaller. Using these expressions in our result (86) for $\mathbf{P}_L(\mathbf{r},t)$, along with the expressions (68) and (70) for $\mathcal{E}(\mathbf{r},t)$ and its derivatives, we find that our expansion for $\mathbf{P}_L(\mathbf{r},t)$ begins with order η^2,

$$\mathbf{P}_L(\mathbf{r},t)=\eta^2\mathbf{P}_L^{(2)}(\mathbf{r},t)+\cdots, \tag{91}$$

where

$$\mathbf{P}_L^{(2)}(\mathbf{r},t)=\left[i\omega_{\alpha\mathbf{k}}aF^{(0)}_{\alpha\mathbf{k}}\hat{\varepsilon}_C(\mathbf{r})-a\frac{\partial F^{(0)}_{\alpha\mathbf{k}}}{\partial t_1}\hat{\varepsilon}'(\mathbf{r})\right]\mathbf{E}_{\alpha\mathbf{k}}(\mathbf{r})e^{-i\omega_{\alpha\mathbf{k}}t}+\text{c.c.} \tag{92}$$

Turning to the nonlinear polarization, the strength of the nonlinear polarization is characterized by $\chi_3^{ijkl}(\mathbf{r})e_0^2$, and here we will assume the scaling

$$\chi_3^{ijkl}(\mathbf{r})e_0^2=\eta^2\gamma^{ijkl}(\mathbf{r}), \tag{93}$$

where $\varepsilon_0\gamma^{ijkl}(\mathbf{r})$ has units of permittivity and is assumed to be on the order of the largest linear media permittivity, or smaller. Using Eq. (93) in our result (88) for $\mathbf{P}_{NL}(\mathbf{r},t)$, we find that the expansion for that term also begins with order η^2,

$$\mathbf{P}_{NL}(\mathbf{r},t)=\eta^2\mathbf{P}^{(2)}_{NL}(\mathbf{r},t)+\cdots, \tag{94}$$

where

$$P^{i(2)}_{NL}(\mathbf{r},t)=3i\omega_{\alpha\mathbf{k}}a\varepsilon_0F^{(0)}_{\alpha\mathbf{k}}|F^{(0)}_{\alpha\mathbf{k}}|^2\gamma^{ijkl}(\mathbf{r})E^j_{\alpha\mathbf{k}}(\mathbf{r})E^k_{\alpha\mathbf{k}}(\mathbf{r})\times[E^l_{\alpha\mathbf{k}}(\mathbf{r})]^*e^{-i\omega_{\alpha\mathbf{k}}t}+\text{c.c.} \tag{95}$$

Combining the linear (92) and nonlinear (95) contributions to $\mathbf{P}(\mathbf{r},t)$ Eq. (89), we see that $\boldsymbol{\Xi}$ (54) is indeed of the form $\boldsymbol{\Xi}_+ + \boldsymbol{\Xi}_-$, where $\boldsymbol{\Xi}_-=\overline{\boldsymbol{\Xi}}_+$, with

$$\boldsymbol{\Xi}_+=\eta^2\boldsymbol{\Xi}_+^{(2)}+\cdots=\eta^2\boldsymbol{\xi}_+^{(2)}\begin{pmatrix}1\\1\end{pmatrix}+\cdots \tag{96}$$

and

$$\begin{aligned}a^{-1}e^{i\omega_{\alpha\mathbf{k}}t}\xi_+^{i(2)}={}&\frac{i\omega_{\alpha\mathbf{k}}}{2\varepsilon_0}F^{(0)}_{\alpha\mathbf{k}}\frac{\hat{\varepsilon}_C(\mathbf{r})E^i_{\alpha\mathbf{k}}(\mathbf{r})}{\sqrt{n(\mathbf{r})}}\\ &-\frac{1}{2\varepsilon_0}\frac{\partial F^{(0)}_{\alpha\mathbf{k}}}{\partial t_1}\frac{\hat{\varepsilon}'(\mathbf{r})E^i_{\alpha\mathbf{k}}(\mathbf{r})}{\sqrt{n(\mathbf{r})}}\\ &+\frac{3}{2}i\omega_{\alpha\mathbf{k}}F^{(0)}_{\alpha\mathbf{k}}|F^{(0)}_{\alpha\mathbf{k}}|^2\\ &\times\frac{\gamma^{ijkl}(\mathbf{r})E^j_{\alpha\mathbf{k}}(\mathbf{r})E^k_{\alpha\mathbf{k}}(\mathbf{r})[E^l_{\alpha\mathbf{k}}(\mathbf{r})]^*}{\sqrt{n(\mathbf{r})}}.\end{aligned} \tag{97}$$

VI. ENVELOPE FUNCTION EQUATIONS

We are now in a position to implement the dynamical equations (59) to order η^2. Since $\boldsymbol{\Xi}_+^{(1)}=0$, the order η equations are simply $(\pounds\boldsymbol{\Psi}_+)^{(1)}=0$. Returning to the expression (76) for $(\pounds\boldsymbol{\Psi}_+)^{(1)}$, we first dot the order η equation into $\boldsymbol{\Phi}^\dagger_{\alpha\mathbf{k}}$ and integrate over $\mathbf{r}_0$ to find

$$\frac{\partial F^{(0)}_{\alpha\mathbf{k}}}{\partial t_1}+v^i_{\alpha\alpha}(\mathbf{k})\frac{\partial F^{(0)}_{\alpha\mathbf{k}}}{\partial r_1^i}=0, \tag{98}$$

while dotting that equation into $\boldsymbol{\Phi}^\dagger_{\beta\mathbf{k}}$, for $\beta\neq\alpha$, yields

$$F^{(1)}_{\beta\mathbf{k}}=-\frac{iv^i_{\beta\alpha}(\mathbf{k})}{(\omega_{\alpha\mathbf{k}}-\omega_{\beta\mathbf{k}})}\frac{\partial F^{(0)}_{\alpha\mathbf{k}}}{\partial r_1^i}. \tag{99}$$

Here we have used the orthogonality conditions (28) as well as the expression (41) for the velocity matrix element $\mathbf{v}_{\alpha'\alpha}(\mathbf{k})$.

The order η^2 equations, $(\pounds\boldsymbol{\Psi}_+)^{(2)}=i\Xi_+^{(2)}$, can be dealt with in the same way. Dotting into $\boldsymbol{\Phi}^\dagger_{\beta\mathbf{k}}$, for $\beta\neq\alpha$, and integrating yields the equations for $F^{(2)}_{\beta\mathbf{k}}$ for $\beta\neq\alpha$, which we will not write down here. Dotting into $\mathbf{\Phi}^{\dagger}_{\alpha\mathbf{k}}$ and integrating, we obtain an equation that can be readily simplified. Using the expression (99) for $F^{(1)}_{\beta\mathbf{k}}$, we find the resulting expression involves a sum that can be identified as the group velocity dispersion (45). We can also use Eq. (98) to write $\partial F^{(1)}_{\alpha\mathbf{k}}/\partial t_1$ in terms of the $\partial F^{(1)}_{\alpha\mathbf{k}}/\partial r^i_1$. Our efforts yield

$$\frac{\partial F^{(0)}_{\alpha\mathbf{k}}}{\partial t_2}+v^{i}_{\alpha\alpha}(\mathbf{k})\frac{\partial F^{(0)}_{\alpha\mathbf{k}}}{\partial r^i_2}-\frac{i}{2}\frac{\partial^2\omega_{\alpha\mathbf{k}}}{\partial k^i\partial k^j}\frac{\partial^2 F^{(0)}_{\alpha\mathbf{k}}}{\partial r^i_1\partial r^j_1}=\frac{i\omega_{\alpha\mathbf{k}}\sigma F^{(0)}_{\alpha\mathbf{k}}}{\eta^2}-\frac{\mu}{\eta}\frac{\partial F^{(0)}_{\alpha\mathbf{k}}}{\partial t_1}+\frac{i\alpha e_0^2}{\eta^2}F^{(0)}_{\alpha\mathbf{k}}|F^{(0)}_{\alpha\mathbf{k}}|^2, \quad (100)$$

where we have defined coefficients σ, μ, and α in terms of our original physical quantities,

$$\sigma=\frac{1}{2}\int_{cell}\frac{d^3r}{\Omega_{cell}}\frac{\varepsilon_C(\mathbf{r})}{\varepsilon_0}|\mathbf{E}_{\alpha\mathbf{k}}(\mathbf{r})|^2,$$

$$\mu=\frac{1}{2}\omega_{\alpha\mathbf{k}}\int_{cell}\frac{d^3r}{\Omega_{cell}}\frac{\varepsilon'(\mathbf{r})}{\varepsilon_0}|\mathbf{E}_{\alpha\mathbf{k}}(\mathbf{r})|^2, \quad (101)$$

$$\alpha=\frac{3}{2}\omega_{\alpha\mathbf{k}}\int_{cell}\frac{d^3r}{\Omega_{cell}}\chi_3^{ijkl}(\mathbf{r})[E^i_{\alpha\mathbf{k}}(\mathbf{r})]^*E^j_{\alpha\mathbf{k}}(\mathbf{r})E^k_{\alpha\mathbf{k}}(\mathbf{r})\times[E^l_{\alpha\mathbf{k}}(\mathbf{r})]^*.$$

Each of Eqs. (98), (99), and (100) must be consistent with our scaling assumptions. That is, the $F^{(i)}_{\alpha\mathbf{k}}$ and $F^{(i)}_{\beta\mathbf{k}}$ must all be of order unity (or less), and vary equally significantly as each of its spatial and temporal arguments vary over d and τ, respectively. A simple analysis here follows such an examination in the one-dimensional case [12]. In particular, we must satisfy the two derivative conditions $|\partial F^{(0)}_{\alpha\mathbf{k}}/\partial t_1|$, $|\partial F^{(0)}_{\alpha\mathbf{k}}/\partial t_2|\leqslant|F^{(0)}_{\alpha\mathbf{k}}/\tau|=\omega_{\alpha\mathbf{k}}|F^{(0)}_{\alpha\mathbf{k}}|/(2\pi)$ and the condition that $|F^{(1)}_{\beta\mathbf{k}}|$ be of order unity or less. Since $|\partial F^{(0)}_{\alpha\mathbf{k}}/\partial r^i_1|\simeq|F^{(0)}_{\alpha\mathbf{k}}|/gd$ and $|\partial^2F^{(0)}_{\alpha\mathbf{k}}/\partial r^i_1\partial r^i_1|\simeq|F^{(0)}_{\alpha\mathbf{k}}|/g^2d^2$ Eq. (65), the first of the derivative conditions, together with Eq. (98), requires that we set

$$g\geqslant\left|\frac{2\pi}{\omega_{\alpha\mathbf{k}}d}v_{\alpha\alpha}(\mathbf{k})\right|. \quad (102)$$

The third condition, together with Eq. (99), requires that

$$g\geqslant\left|\frac{v_{\beta\alpha}(\mathbf{k})}{(\omega_{\alpha\mathbf{k}}-\omega_{\beta\mathbf{k}})}\frac{1}{d}\right| \quad (103)$$

for $\beta\neq\alpha$; this condition is clearly best satisfied if all bands β are remote in frequency from the band associated with the principal component $F^{(0)}_{\alpha\mathbf{k}}$. Finally, the second derivative condition, together with Eq. (100), requires that

$$g^2\geqslant\pi\left|\frac{1}{\omega_{\alpha\mathbf{k}}d^2}\frac{\partial^2\omega_{\alpha\mathbf{k}}}{\partial k^i\partial k^j}\right| \quad (104)$$

for arbitrary i and j. Like the requirement (103), this is most easily satisfied if all bands other than α are ''remote.'' Of course, these requirements must be satisfied while still maintaining

$$\eta\ll 1. \quad (105)$$

In fact, the requirements (102)–(104) are overly conservative, because in equation (65) we implicitly took $f_{\alpha\mathbf{k}}(\mathbf{r},t=0)$ to vary over a single length Λ in all Cartesian directions; by setting out separate lengths Λ_x, Λ_y, and Λ_z over which $f_{\alpha\mathbf{k}}(\mathbf{r},t=0)$ varies in the indicated directions, the above conditions can be made less restrictive for some pulses. We do not write out those more general conditions here.

Recalling the definition of Λ and η [see Eq. (65)], we refer to the requirements (102)–(105) as *kinematical consistency conditions.* Given the pulse specification at $t=0$, and the photonic band structure, it is easy to check whether or not they are satisfied. Typically, the more remote in frequency a band is from the other bands, the shorter the pulse can be while still satisfying these conditions. If they hold at $t=0$, then the dynamical equations (98) and (100) can be considered a good approximate description of the initial dynamics for $t>0$. But as t increases we also require that the solution of Eqs. (98) and (100) leads to an evolving $F^{(0)}_{\alpha\mathbf{k}}$ that remains of order unity or less. This much more stringent *dynamical consistency condition* is discussed in Sec. VIII. For the rest of this section we assume that all these consistency conditions are satisfied.

We can amalgamate our dynamical equations at orders η^0, η^1, and η^2 by using

$$\frac{\partial F^{(0)}_{\alpha\mathbf{k}}}{\partial t}=\eta\frac{\partial F^{(0)}_{\alpha\mathbf{k}}}{\partial t_1}+\eta^2\frac{\partial F^{(0)}_{\alpha\mathbf{k}}}{\partial t_2}+\cdots, \quad (106)$$

and likewise for the gradients of $F^{(0)}_{\alpha\mathbf{k}}$; we combine Eqs. (98) and (100) to find, to order η^2,

$$\frac{\partial F^{(0)}_{\alpha\mathbf{k}}}{\partial t}+(1-\mu)v^i_{\alpha\alpha}(\mathbf{k})\frac{\partial F^{(0)}_{\alpha\mathbf{k}}}{\partial r^i}-\frac{1}{2}i\frac{\partial^2\omega_{\alpha\mathbf{k}}}{\partial k^i\partial k^j}\frac{\partial^2F^{(0)}_{\alpha\mathbf{k}}}{\partial r^i\partial r^j}-i\omega_{\alpha\mathbf{k}}\sigma F^{(0)}_{\alpha\mathbf{k}}-i\alpha e_0^2F^{(0)}_{\alpha\mathbf{k}}|F^{(0)}_{\alpha\mathbf{k}}|^2=0. \quad (107)$$

Making a change of variables to a moving frame with $\overline{t}\equiv t$ and $\overline{\mathbf{r}}\equiv\mathbf{r}-(1-\mu)\mathbf{v}_{\alpha\alpha}(\mathbf{k})t$ and dropping the overbars, we obtain

$$\frac{\partial f_{\alpha\mathbf{k}}}{\partial t}-\frac{1}{2}i\frac{\partial^2\omega_{\alpha\mathbf{k}}}{\partial k^i\partial k^j}\frac{\partial^2 f_{\alpha\mathbf{k}}}{\partial r^i\partial r^j}-i\omega_{\alpha\mathbf{k}}\sigma f_{\alpha\mathbf{k}}-i\alpha e_0^2 f_{\alpha\mathbf{k}}|f_{\alpha\mathbf{k}}|^2=0 \quad (108)$$

a nonlinear Schrödinger equation for the envelope function of the principal component.

We now consider the physics associated with this equation. It describes a wave packet moving at a velocity $(1-\mu)\mathbf{v}_{\alpha\alpha}(\mathbf{k})$ relative to the lab frame, so from Eq. (101) it is clear that μ characterizes the change in group velocity, due to material dispersion, from that of the band structure characterized by a nondispersive refractive index $n(\mathbf{r})$. We can confirm the nature of the coefficient μ, as well as the other coefficients we consider below, by looking at the limit of a uniform medium. In that limit $n(\mathbf{r})$ is uniform, $n(\mathbf{r})=n$, and the Bloch functions reduce to plane waves that we take to be linearly polarized. Recalling the normalization conditions (9) and (10), the fields are

$$\mathbf{E}_{\alpha\mathbf{k}}(\mathbf{r})=\frac{\hat{\mathbf{e}}_{\alpha}e^{i\mathbf{k}\cdot\mathbf{r}}}{n},$$

$$\mathbf{H}_{\alpha\mathbf{k}}(\mathbf{r})=\frac{\hat{\mathbf{h}}_{\alpha}e^{i\mathbf{k}\cdot\mathbf{r}}}{c\mu_0}, \tag{109}$$

where $k=|\mathbf{k}|=n\omega/c$, and $\hat{\mathbf{e}}_{\alpha}\times\hat{\mathbf{h}}_{\alpha}=\hat{\mathbf{k}}\equiv\mathbf{k}/k$. In this limit we find $\mathbf{v}_{\alpha\alpha}(\mathbf{k})=c\hat{\mathbf{k}}/n$ from the general expression (47), as expected, and from Eq. (101) we find

$$(1-\mu)\mathbf{v}_{\alpha\alpha}(\mathbf{k})\rightarrow\frac{c}{n}\left(1-\frac{1}{2}\omega\frac{\varepsilon'}{\varepsilon_0 n^2}\right)\hat{\mathbf{k}}. \tag{110}$$

The *exact* group velocity in a uniform medium is determined by the dispersion relation $\omega=ck\sqrt{\varepsilon_0/\varepsilon(\omega)}$. For a uniform medium with an $\varepsilon(\omega)$ characterized by Eq. (82), we find a group velocity magnitude $d\omega/dk=(c/n)(1+\varepsilon'\omega/(2n^2\varepsilon_0))^{-1}$. Our limit (110) agrees with this to lowest order in the dispersion coefficient $\omega\varepsilon'/\varepsilon_0$, as would be expected from our scaling (90). Of course, in a photonic crystal $\mathbf{v}_{\alpha\alpha}(\mathbf{k})$ will not be simply uniform in magnitude as $\mathbf{k}$ varies, even neglecting material dispersion, and μ will also be a complicated function of $\mathbf{k}$ as Bloch functions sample the underlying material dispersion differently at different $\mathbf{k}$, according to Eq. (101). Nonetheless, since in a uniform medium μ is typically much less than unity and represents a small correction to the group velocity, and because it is $(1-\mu)$ that appears as a multiplicative correction to $\mathbf{v}_{\alpha\alpha}(\mathbf{k})$, for typical photonic crystals material dispersion should represent a small correction as well. Of course, material dispersion will enter to higher orders than η^2 if the derivation leading to Eq. (108) is extended to a higher order; or, if material dispersion is physically more important for a given problem than the scaling (90) identifies, a more appropriate scaling would have to be adopted. In either case a correction to the group velocity dispersion term—the term involving $\partial^2\omega_{\alpha\mathbf{k}}/\partial k^i\partial k^j$ in Eq. (108)—due to material dispersion will arise.

But we pass over these matters for now and neglect material dispersion for the rest of our discussion in this section, setting $\varepsilon'(\mathbf{r})=0$. In the limit of a one-dimensional photonic crystal, where we take $\hat{\mathbf{z}}$ to be both the direction of propagation and the direction of crystal periodicity, equation (108) reduces to

$$\frac{\partial f_{\alpha\mathbf{k}}}{\partial t}-\frac{1}{2}i\omega''_{\alpha\mathbf{k}}\frac{\partial^2 f_{\alpha\mathbf{k}}}{\partial z^2}-i\omega_{\alpha\mathbf{k}}\sigma f_{\alpha\mathbf{k}}-i\alpha e_0^2 f_{\alpha\mathbf{k}}|f_{\alpha\mathbf{k}}|^2=0, \tag{111}$$

where $\omega''_{\alpha\mathbf{k}}=\partial^2\omega_{\alpha\mathbf{k}}/\partial(k^z)^2$ is the group velocity dispersion. This result is in agreement with earlier work [12], but extended to include the effective index modification term involving σ, which is discussed below. In the higher-dimensional equation (108), the group velocity dispersion term actually includes effects due to both group velocity dispersion and diffraction. In a uniform medium for which $\omega=c|\mathbf{k}|/n$ only diffraction in fact survives; if we choose a wave packet centered at $\mathbf{k}=k_0\hat{\mathbf{z}}$, we find

$$-\frac{1}{2}i\frac{\partial^2\omega_{\alpha\mathbf{k}}}{\partial k^i\partial k^j}\frac{\partial^2 f_{\alpha\mathbf{k}}}{\partial r^i\partial r^j}\rightarrow-\frac{ic}{2nk_0}\left(\frac{\partial^2 f_{\alpha\mathbf{k}}}{\partial x^2}+\frac{\partial^2 f_{\alpha\mathbf{k}}}{\partial y^2}\right), \tag{112}$$

which will lead to the spreading in the xy plane of a wave packet propagating in the $\hat{\mathbf{z}}$ direction. For a true photonic crystal, of course, in general $\partial^2\omega_{\alpha\mathbf{k}}/\partial(k^z)^2$ will not vanish at $\mathbf{k}=k_0\hat{\mathbf{z}}$, and that term will describe the group velocity dispersion that is absent in a uniform medium with no material dispersion. As well, of course, there is no guarantee that the tensor $\partial^2\omega_{\alpha\mathbf{k}}/\partial k^i\partial k^j$ has as one its principal axes the direction in which the pulse is propagating; hence effects involving group velocity dispersion and diffraction in a complicated way can easily arise.

To consider the physical significance of the other terms in our nonlinear Schrödinger equation, we note that for the form (66)–(68) assumed for our fields we have

$$\langle\mathbf{E}(\mathbf{r},t)\times\mathbf{H}(\mathbf{r},t)\rangle=2\varepsilon_0 e_0^2|F^{(0)}_{\alpha\mathbf{k}}|^2\mathbf{v}_{\alpha\alpha}(\mathbf{k})=2|\Psi(\mathbf{r},t)|^2\hat{\mathbf{s}}$$

$$\equiv I(\mathbf{r},t)\hat{\mathbf{s}} \tag{113}$$

to lowest order, where the brackets $\langle\rangle$ indicate both a spatial average over a unit cell and a time average over the period associated with the frequency $\omega_{\alpha\mathbf{k}}$; here we have defined

$$\Psi(\mathbf{r},t)\equiv e_0\sqrt{\varepsilon_0 v_{\alpha\alpha}(\mathbf{k})}F^{(0)}_{\alpha\mathbf{k}}, \tag{114}$$

while $v_{\alpha\alpha}(\mathbf{k})$ denotes the magnitude of $\mathbf{v}_{\alpha\alpha}(\mathbf{k})$ and $\hat{\mathbf{s}}$ its direction. In terms of this new field $\Psi(\mathbf{r},t)$ our nonlinear Schrödinger equation is

$$\frac{\partial\Psi(\mathbf{r},t)}{\partial t}+v^i_{\alpha\alpha}(\mathbf{k})\frac{\partial\Psi(\mathbf{r},t)}{\partial r^i}=\frac{1}{2}i\frac{\partial^2\omega_{\alpha\mathbf{k}}}{\partial k^i\partial k^j}\frac{\partial^2\Psi(\mathbf{r},t)}{\partial r^i\partial r^j}+i\omega_{\alpha\mathbf{k}}\sigma\Psi(\mathbf{r},t)+i\frac{\alpha}{\varepsilon_0 v_{\alpha\alpha}(\mathbf{k})}\Psi(\mathbf{r},t)|\Psi(\mathbf{r},t)|^2, \tag{115}$$

where as mentioned above we have put $\mu=0$.

To concentrate on the effect of the nonlinear term in this equation, for the moment we neglect both the σ term, which results from a perturbation of the linear optical properties, and the group velocity dispersion term. Then, considering propagation in the $\hat{\mathbf{s}}$ direction, for a field $\Psi(\mathbf{r},t)$ that is independent of time we find that equation (115) reduces to

$$v_{\alpha\alpha}(\mathbf{k})\frac{\partial\Psi(s)}{\partial s}=i\frac{\alpha}{\varepsilon_0 v_{\alpha\alpha}(\mathbf{k})}\Psi(s)|\Psi(s)|^2. \tag{116}$$

The assumption that $\Psi(\mathbf{r},t)$ is independent of time implies that the full electric field is oscillating at the Bloch frequency $\omega_{\alpha\mathbf{k}}$. Now we write the solution of Eq. (116) in terms of an effective nonlinear refractive index coefficient n_2,

$$\Psi(s)=\Psi(0)\exp[i(n_2 I)(\omega_{\alpha\mathbf{k}}s/c)], \tag{117}$$

where

$$n_2=\frac{1}{2\varepsilon_0 c}\frac{\alpha}{\omega_{\alpha\mathbf{k}}}\left(\frac{c}{v_{\alpha\alpha}(\mathbf{k})}\right)^2. \tag{118}$$

The effects of the photonic crystal structure are apparent here. First, the coefficient α characterizes how the Bloch function modulated by the envelope function $\Psi(\mathbf{r},t)$ is ''samples'' the underlying nonlinearity. Second, there are two factors of $[v_{\alpha\alpha}(\mathbf{k})/c]^{-1}$. The first can be traced back to the form (115) of the nonlinear Schrödinger equation that arises if the amplitude of the energy flux $I(\mathbf{r},t)$ [essentially $\Psi(\mathbf{r},t)$] is used as the fundamental field. It arises because the nonlinear coefficient n_2 is referenced to the energy flux while the nonlinearity itself depends on the electric field strength; for a given energy flux, the electric field strength is greater the smaller the group velocity. The second factor of $[v_{\alpha\alpha}(\mathbf{k})/c]^{-1}$ appears through the solution of Eq. (116), and arises physically because the smaller the group velocity the more time fields spend experiencing a nonlinearity while propagating over a fixed length s.

Only the second of these factors appears if we consider the effect of a modified linear dielectric constant on the propagation of light. Returning to Eq. (115) and now neglecting the group velocity dispersion and the nonlinearity, in place of Eq. (116) we find

$$v_{\alpha\alpha}(\mathbf{k})\frac{\partial\Psi(s)}{\partial s}=i\omega_{\alpha\mathbf{k}}\sigma\Psi(s), \tag{119}$$

and introducing the modification $\tilde{n}$ to the effective index $[c/v_{\alpha\alpha}(\mathbf{k})]$ by writing the solution of this equation as

$$\Psi(s)=\Psi(0)\exp[i\tilde{n}(\omega_{\alpha\mathbf{k}}s/c)], \tag{120}$$

we identify

$$\tilde{n}=\sigma\left(\frac{c}{v_{\alpha\alpha}(\mathbf{k})}\right). \tag{121}$$

Note that σ and $\tilde{n}$ can both be complex if the modification $\varepsilon_C(\mathbf{r})$ to the linear dielectric constant is complex; the real part of $\tilde{n}$ is associated with a change in effective phase velocity, while the imaginary part of $\tilde{n}$ describes the effect of gain or loss on propagation.

We close this section by confirming that our expressions (118) and (121) for n_2 and $\tilde{n}$, respectively, reduce to the correct results in the limit of a uniform, nondispersive medium. Recalling that $v_{\alpha\alpha}(\mathbf{k})=c/n$ in this limit, using the Bloch functions (109) appropriate here, and evaluating α and σ from their definitions (101), we find

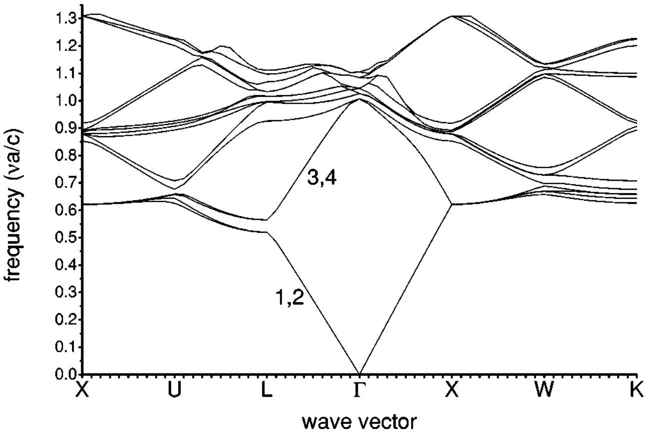

FIG. 2. Band structure of a close-packed fcc photonic crystal consisting of air spheres in PMMA, plotted along high symmetry lines. The frequency is ν, and a is the distance between the centers of neighboring spheres. For clarity, bands 1 through 4 between Γ and L are labeled.

$$\begin{aligned} n_2 &\rightarrow \frac{3\chi_3^{eeee}}{4\varepsilon_0 c n^2}, \\ \tilde{n} &\rightarrow \frac{\varepsilon_C}{2\varepsilon_0 n}, \end{aligned} \tag{122}$$

where we have taken both χ_3^{eeee}, the indicated component of χ_3^{ijkl} associated with only the direction $\hat{\mathbf{e}}_\alpha$, and ε_C, the linear modification of the dielectric constant, to be uniform. The first of Eq. (122) is the well-known relation between n_2 and χ_3 in a uniform medium, and the second is, for $\varepsilon_C/\varepsilon_0 \ll 1$ [see equation (90)], the expected correction to the index of refraction $\sqrt{n^2+\varepsilon_C/\varepsilon_0}\approx n+\varepsilon_C/2\varepsilon_0 n$ due to a linear modification in the dielectric constant.

VII. NUMERICAL EXAMPLES

In this section, we calculate and interpret the effective nonlinearity, and gain or loss coefficients, of a typical photonic crystal. While the above expressions can be applied to crystals of any dimensionality, here we consider the three dimensional example of an fcc close-packed poly(methyl methacrylate) (hereafter PMMA) inverse opal. PMMA has a refractive index of 1.49 and is taken to have a χ_3 nonlinear susceptibility tensor characteristic of an isotropic material. The band structure and photonic Bloch modes for the nondispersive, linear, lossless limit were calculated using a frequency domain eigensolver [16], with results shown in Fig. 2. We consider a pulse propagating in the (111) direction, which in practice often corresponds to coupling at normal incidence. Care must be taken since the high symmetry of the (111) direction leads to degenerate bands. Although we did not treat degenerate bands in the formalism, a generalization of the argument leading to Eq. (99) shows that if $\mathbf{v}_{\alpha\beta}(\mathbf{k})\cdot\nabla f_{\alpha\mathbf{k}}$ is sufficiently small everywhere for all bands β degenerate with band α, then the nonlinear Schrödinger dynamics of equation (115) will be valid for band α. One way of satisfying this condition is to consider systems for

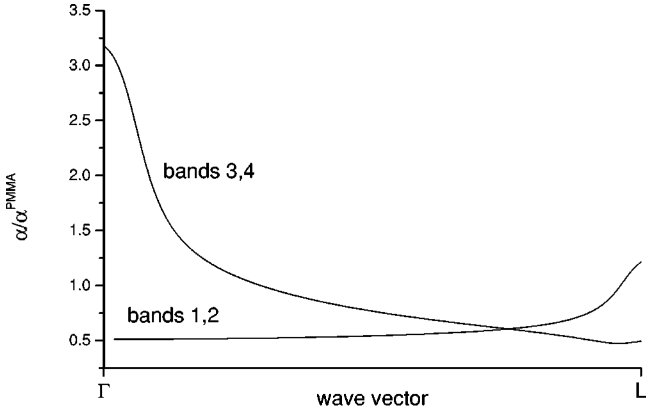

FIG. 3. Effective Kerr coefficient α vs wave vector along (111) direction for close-packed air spheres in PMMA fcc crystal, plotted in units of α^{PMMA}, the Kerr coefficient of uniform bulk PMMA.

which both $\nabla f_{\alpha\mathbf{k}}\times\mathbf{v}_{\alpha\alpha}(\mathbf{k})$ is sufficiently small everywhere and $\mathbf{v}_{\alpha\beta}(\mathbf{k})\cdot\mathbf{v}_{\alpha\alpha}(\mathbf{k})$ is zero for all β. The first condition can be satisfied at least initially if the transverse width of the pulse is sufficiently large; we assume we are dealing with such pulses. The second can be satisfied if the modes within the degenerate eigenspace are suitably chosen; the modes can be so chosen in the examples that follow. The general derivation of pulse dynamics with a degenerate principal component requires a more thorough treatment, which we plan to include in a future publication.

In the calculations below, we consider a pulse characterized by a principal component associated with one of the lowest four bands. These consist of a doubly degenerate dielectric band and a doubly degenerate air band. We note that the crystal has a reflection symmetry plane that contains the L point and the z axis [17]; for the sake of specificity we choose the eigenstates of this reflection as bases within each of the degenerate eigenspaces, and refer to the states with reflection eigenvalues ± 1 as the evenlike and oddlike states, respectively. This choice of states ensures that $\mathbf{v}_{\alpha\beta}(\mathbf{k})\cdot\mathbf{v}_{\alpha\alpha}(\mathbf{k})$ is zero for all β. In fact, the lowest four bands exemplify a special case for which $\mathbf{v}_{\alpha\beta}(\mathbf{k})\cdot\mathbf{v}_{\alpha\alpha}(\mathbf{k})$ vanishes for any choice of orthogonal modes within the degenerate eigenspaces.

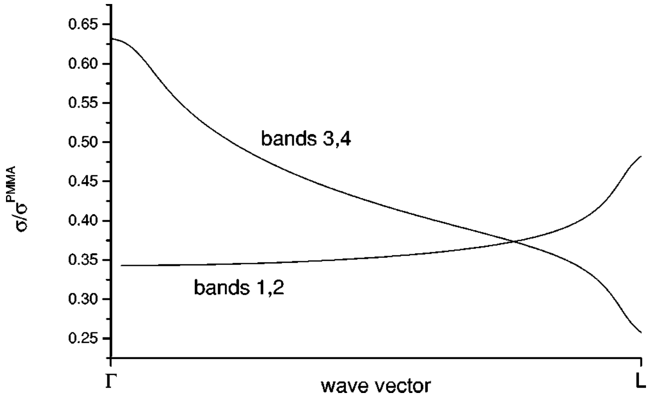

FIG. 4. Effective gain or loss coefficient σ vs wave vector along (111) direction for close-packed air spheres in PMMA fcc crystal, plotted in units of σ^{PMMA}, the value for bulk uniform PMMA.

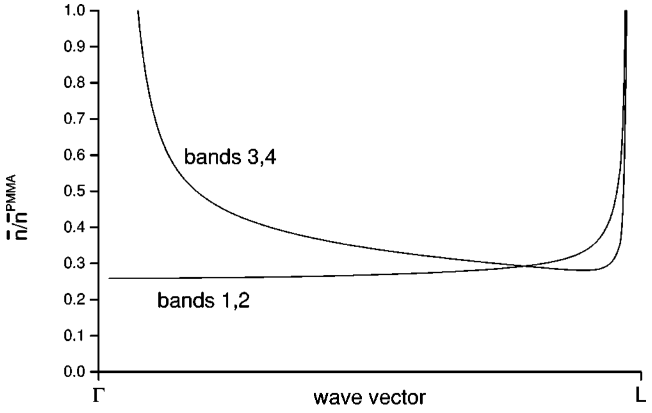

FIG. 5. Effective gain or loss index $\tilde{n}$ vs wave vector along (111) direction for close-packed air spheres in PMMA fcc crystal, plotted in units of $\tilde{n}^{PMMA}$, the value for bulk PMMA.

In order to separate the effects due to reduced group velocities from those due to the concentration of the fields within high nonlinearity, or high gain or loss media, we plot both the mode sampling coefficients α and σ (101), and the effective index coefficients n_2 (118) and $\tilde{n}$ (121). Due to symmetry, bands 1 and 2 have identical such coefficients, as do bands 3 and 4. In Fig. 3, we see the difference between a dielectric band and an air band. As $\mathbf{k}$ nears the band edge at L, in the lower bands the energy is concentrated more and more in the dielectric, while in the upper bands the energy is concentrated more and more in the air. As a result, the effective Kerr coefficient α of the dielectric band increases and that of the air band decreases. In addition, at wave vectors where the bands flatten out, the Bloch functions acquire the nature of standing waves rather than traveling waves, leading to an increase in the integral of the fourth power of the Bloch function for a fixed normalization; this results in a slight recovery in the α coefficient of the upper band before it reaches the L point. Note that as the upper band approaches Γ it becomes both degenerate with higher bands and near to even higher bands; here the kinematic consistency conditions would require a longer and longer pulse to justify the scaling introduced in our derivation to be appropriate. In Fig. 4 we observe somewhat similar qualitative features for σ, as we might expect. In Fig. 5, we plot $\tilde{n}$. The factor $c/v_{\alpha\alpha}(\mathbf{k})$ leads to divergences in $\tilde{n}$ as the bands flatten out, reflecting the enhancement of the accumulation of phase and of gain or loss as light propagates more slowly through the crystal. In Fig. 6 we plot the nonlinear index n_2. Here the $c^2/v_{\alpha\alpha}^2(\mathbf{k})$ factor characterizing both the propagation time and field enhancement effects due to ''slow light'' completely dominates the behavior of n_2 as the bands flatten out.

VIII. CONCLUSIONS

We have described pulse envelope function dynamics in photonic crystals, considering χ_3 nonlinearities, as well as linear material dispersion, gain and loss. The main results are equations (108) and (115). Equation (108) is a dynamical nonlinear Schrödinger equation for the pulse envelope func-

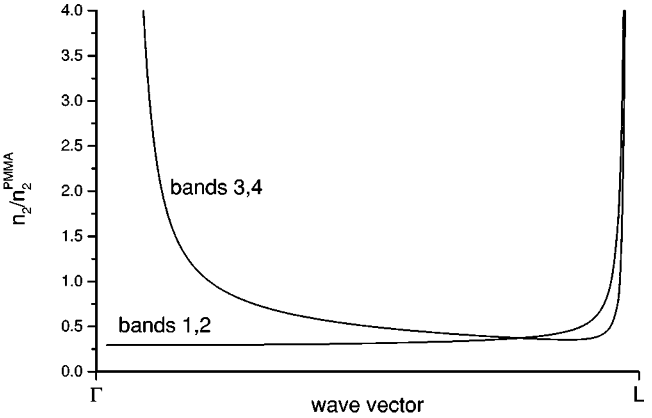

FIG. 6. Effective nonlinear index n_2 vs wave vector along (111) direction for close-packed air spheres in PMMA fcc crystal, plotted in units of n_2^{PMMA}, the value for bulk PMMA. We see that n_2 is greatly enhanced near band edges, where group velocity effects dominate.

tion. The equation is derived using a multiple-scale analysis, in which the pulse is approximated by an envelope function modulating a Bloch function as the carrier wave. Different terms in the equation clearly represent different physical phenomena such as group velocity, group velocity dispersion, diffraction, self-phase modulation, carrier frequency shift, and gain or loss. The coefficients of the different terms depend on the underlying photonic crystal through integrals of powers of the modulated Bloch function over the unit cell of the crystal in the linear, lossless, nondispersive limit. Thus we can identify the degree to which a physical effect modifies the propagation of a pulse by calculating the way the carrier photonic Bloch function spatially samples the appropriate property of the underlying crystal constituents. This approach leverages the ease of calculation of the band structure of the simplified crystal, for the Bloch functions in the absence of nonlinearity, loss and dispersion can then be used to determine the consequences of precisely those effects on pulse propagation. As an example, we have cast the dynamical equation (108) into equation (115) for the energy flux amplitude, allowing us to identify and interpret an effective nonlinear index n_2, as well as an effective complex shift of the linear index $\tilde{n}$, which could account for gain or loss as well as electro-optic shifts in the linear index. The effective indices can vary significantly as one considers different carrier Bloch functions even within the same band. We have identified and described the roles played by the varying group velocity, field concentration, and standing wave nature of the Bloch function as one scans through the Brillouin zone. Additionally, we have identified kinematical and dynamical consistency conditions that serve as quantitative measures of the applicability of the nonlinear Schrödinger dynamics, suggest when new physics warrants inclusion in the analysis and indicate when different scalings are required.

A number of important generalizations will proceed from this work. Immediate generalization to the case of a principal component from a degenerate band will yield dynamics that are characteristic of the crossover to higher-dimensional crystals, providing a qualitative departure from the more familiar dynamics of one-dimensional Bragg gratings. Furthermore, while we have restricted our analysis in this work to the case of a single principal component, a study of the case of two principal components—that of two excited pulses with different carrier modes—will also yield characteristically higher-dimensional dynamics. The study of two pulses at the same frequency with either identical or differing wave vectors, and the study of two pulses at the same wave vector but with slightly different frequencies, promise new physics. In one dimension, the latter case is used to describe coupling between modes at the upper and lower band edges *via* the coupled mode equations, which yield gap soliton solutions. We plan to treat the generalization to three dimensions in a future publication.

Further generalizations will involve the inclusion of other physical effects and scalings, such as underlying material birefringence, defect modes of the photonic crystal, nearly degenerate modes, χ_2 phenomena, and higher-order dispersion and nonlinearity. The latter two effects will inevitably play a crucial role in the nonlinear optics of higher-dimensional photonic crystals, even if the single principal component assumption made in this paper is valid. For while we can expect that over some length and time scales—as identified by the consistency conditions—the nonlinear Schrödinger equation will provide a good description of typical nonlinear propagation and even gap soliton phenomena, in general there is no analog of the stability of soliton solutions of the one-dimensional Schrödinger equation in higher dimensions. In such photonic crystals unchecked self-focusing can lead to the inevitable breakdown of the nonlinear Schrödinger equation, and the consequent necessity of including higher dispersion and nonlinearity to properly describe the physics. The approach developed here provides a straightforward way to identify the higher-order terms of physical relevance.

In summary, we have presented a general approach for the derivation of nonlinear dynamical equations that describe the propagation of optical pulses in nonlinear photonic crystals. As an example we have considered the nonlinear Schrödinger equation limit appropriate for the envelope function modulating a single, nondegenerate photonic Bloch function. But the approach is much more general than that, and typically will allow the determination of coefficients that characterize the nonlinearity, dispersion, and gain or loss of the effective field or fields identified in terms of the sampling of the underlying physical properties of the photonic crystal components by the modulated Bloch function or functions. Thus, even before a detailed solution of the appropriate nonlinear dynamical equations is addressed, much of the physics of the nonlinear propagation, and its dependence on photonic crystal properties, can be immediately identified.

ACKNOWLEDGMENTS

This work was supported by the National Science and Engineering Research Council of Canada, the Government of Ontario, and the Walter C. Sumner Foundation.

[1] P. Tran, Phys. Rev. B **52**, 10 673 (1995).
[2] V. Lousse and J. P. Vigneron, Phys. Rev. E **63**, 027602 (2001).
[3] C. M. de Sterke and J. E. Sipe, in *Progress in Optics*, edited by E. Wolf (North-Holland, Amsterdam, 1994), Vol. XXXIII, pp. 203–260.
[4] C. M. de Sterke and J. E. Sipe, Phys. Rev. A **38**, 5149 (1988).
[5] B. J. Eggleton, R. E. Slusher, C. M. de Sterke, P. A. Krug, and J. E. Sipe, Phys. Rev. Lett. **76**, 1627 (1996).
[6] D. Taverner, N. G. R. Broderick, D. J. Richardson, R. I. Laming, and M. Ibsen, Opt. Lett. **23**, 328 (1998).
[7] M. Asobe, Opt. Fiber Technol.: Mater., Devices Syst. **3**, 142 (1997).
[8] B. J. Eggleton, C. M. de Sterke, and R. E. Slusher, Opt. Commun. **149**, 267 (1998).
[9] C. M. de Sterke and J. E. Sipe, Phys. Rev. A **42**, 2858 (1991).
[10] N. Aközbek and S. John, Phys. Rev. E **57**, 2287 (1998).
[11] J. E. Sipe, L. Poladian, and C. M. de Sterke, J. Opt. Soc. Am. A **11**, 1307 (1994).
[12] C. M. de Sterke, D. G. Salinas, and J. E. Sipe, Phys. Rev. E **54**, 1969 (1996).
[13] S. Pereira and J. E. Sipe, Phys. Rev. E **62**, 5745 (2000).
[14] J. E. Sipe, Phys. Rev. E **62**, 5672 (2000).
[15] C. M. de Sterke and J. E. Sipe, Phys. Rev. A **39**, 5163 (1989).
[16] S. G. Johnson and J. D. Joannopoulos, Opt. Express **8**, 173 (2001).
[17] A. P. Cracknell, in *Numerical Data and Functional Relationships in Science and Technology*, edited by K.-H. Hellwege and J. L. Olsen, Landolt-Börnstein, New Series, Group III, Vol. 13, Pt. c (Springer-Verlag, Berlin, 1984).