Hamiltonian formulation of the nonlinear coupled mode equations

Suresh Pereira and J. E. Sipe
Department of Physics, University of Toronto, Toronto, M5S 1A7, Canada
(Received 11 April 2002; published 28 August 2002)

We derive a canonical Hamiltonian formulation of the nonlinear coupled mode equations (CME) that govern the dynamics of pulse propagation in a one-dimensional, periodic Kerr medium when the frequency content of the pulse is in the vicinity of a photonic band gap, and sufficiently narrow relative to a carrier frequency. The Hamiltonian is equal to the energy in the electromagnetic field. We show that even for large photonic band gaps (25% of the Bragg frequency), the CME give an excellent approximation to the dispersion relation of the linear, periodic medium. This suggests that two- and three-dimensional photonic band-gap materials, which necessarily have large index contrasts, might be effectively described by a set of generalized CME.

DOI: 10.1103/PhysRevE.66.026606 PACS number(s): 03.50.De, 42.65.−k

I. INTRODUCTION

Investigators of optical pulse propagation in one-dimensional, isotropic, Kerr nonlinear media, with a periodic variation in the linear index of refraction, often use a set of heuristic nonlinear coupled mode equations (CME) [1–4]. In the derivation of these equations it is typically assumed that the index contrast of the periodic variation is very small relative to the average effective index of refraction. With the introduction of photonic band-gap (PBG) materials which, at least in two or three dimensions, require the use of strong index contrasts [5,6], the nonlinear CME have recently been derived using the underlying Bloch functions of the periodic medium as an expansion basis, which allows for the treatment of gratings with higher index contrasts [7,8].

In this paper, we present a canonical Hamiltonian formulation of the nonlinear CME in one dimension. By *canonical* we mean that our Hamiltonian can be used to derive the exact equations of motion using canonical commutation relations, *and* that it is numerically equal to the energy of the (nonlinear) electromagnetic field. The CME that we present are equivalent to those derived earlier for strong gratings [7,8], but our approach has two distinct advantages. First, the Hamiltonian formulation allows for ease in quantizing the theory; this we do not do here, but defer to a later publication. Second, the Hamiltonian formulation aids in the identification of symmetries and their relation to conserved quantities at the effective field level. In particular, we use the reduced Hamiltonian to identify two more conserved quantities: the momentum, associated with space-translation symmetry, and a conserved charge associated with phase-translation symmetry [9].

It is hoped that the convenience and physical insight of a coupled-mode approach will carry over to higher-dimensional PBG materials, but there are two reasons why it might not: first, the dispersion relation of a two- or three-dimensional PBG material is often quite complicated, so that there may be no regions where the nonlinear CME would be applicable; second, even were the dispersion relation sufficiently well-behaved that use of the CME could be considered, the large index contrasts involved might cast the validity of the *linear* predictions of the CME into doubt. A full analysis of the three-dimensional problem has yet to be done. In this paper, we investigate the predictions of the linear CME in the presence of a strong index contrast for a one-dimensional system. The linear CME make a very clear prediction for the dispersion relation of the periodic medium [1,8]. We compare this prediction with the exact dispersion relation for a periodic media with very large photonic band gaps. Even when the width of the photonic band gap is 25% of the Bragg frequency, we find that the coupled mode equations predict the group velocity and group velocity dispersion of the system to within about 10% in the region where they would be expected to hold. This agreement between the CME and the exact solutions suggests that the CME will remain useful as a heuristic guide to nonlinear pulse-propagation experiments even in the presence of very strong index contrasts.

An outline of the rest of the paper is as follows. In Sec. II, we rewrite Maxwell's equations for a one-dimensional, periodic, Kerr nonlinear medium in a canonical Hamiltonian formulation, using a dual field introduced earlier by others [9–12]. In Sec. III, we define effective fields, in terms of which we write a reduced Hamiltonian that generates the nonlinear CME. In Sec. IV, we rewrite the reduced Hamiltonian in terms of fields that have a familiar interpretation as traveling waves. In Sec. V, we compare the approximate dispersion relation predicted by the CME to the exact dispersion relation of the periodic system. In Sec. VI we conclude.

II. CANONICAL FORMULATION OF MAXWELL'S EQUATIONS

We begin by introducing a dual field, $\Lambda(z,t)$ [9–11], that satisfies

$$\partial_z \Lambda = D, \quad \partial_t \Lambda = -H, \tag{1}$$

where D and H are the electric flux density and magnetic field strength familiar from Maxwell's equations. We assume that the medium of interest is nonmagnetic ($\mu=\mu_0$, the permeability of free space) and periodic with period d. In the absence of nonlinearity, we can determine the Bloch functions of the system using the usual ansatz [13],

$$\Lambda(z,t) \propto \theta_\mu(z) e^{-i\omega_\mu t} + \text{c.c.}, \tag{2}$$

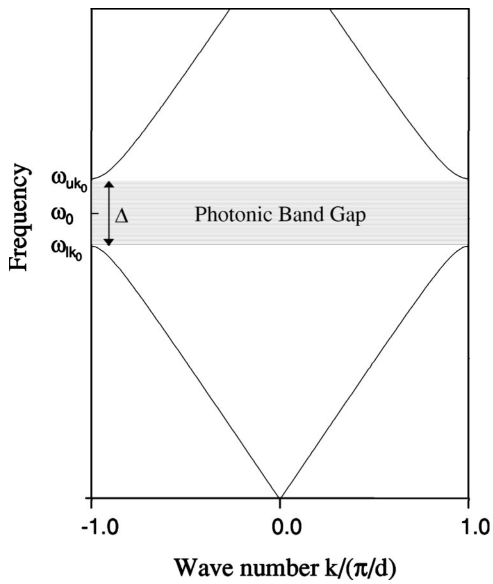

FIG. 1. Sketch of a dispersion relation for a one-dimensional, periodic medium. Identified on the graph are the Bragg frequency and Bragg wave number of the system. We label the quantities ω_0, ω_{uk_0}, ω_{lk_0}, and Δ used in the text. We show the lowest-order photonic band gap, so that $l=1$, $u=2$, and $k_0=\pi/d$.

where c.c. stands for "complex conjugate." From Bloch's theorem [13], we can write our Bloch functions in terms of a discrete band index *m*, and a reduced wave number, *k*, with $-\pi/d<k\leqslant\pi/d$, so for the θ_μ we have

$$\theta_{mk}(z)=u_{mk}(z)e^{ikz}, \tag{3}$$

where the u_{mk} have the periodicity of the lattice, $u_{mk}(z)=u_{mk}(z+d)$. We note that $\omega_{mk}=\omega_{m(-k)}$, so we can choose our Bloch functions such that $\theta_{mk}(z)=\theta^*_{m(-k)}$. We normalize the Bloch functions via

$$\int_{-L/2}^{L/2}\theta^*_{mk}(z)\,\theta_{m'k'}(z)dz=N\delta_{mm'}\delta_{kk'}, \tag{4}$$

where L is a normalization length, and where we have chosen the normalization constant $N=L/d$, which is then identified as the number of unit cells in the normalization length. This choice of normalization means that our wave numbers take on only discrete values, and that the difference between two adjacent wave numbers is $2\pi/L$. A typical dispersion relation in this reduced-wave-number scheme is sketched in Fig. 1. Shown in the figure are the quantities ω_0, ω_{u0}, ω_{l0}, and Δ, as well as a photonic band gap of the system, all of which will be discussed later in the text.

In the presence of nonlinearity, we use the Bloch functions of the underlying periodic medium as an expansion basis for the dual field,

$$\Lambda(z,t)=\sum_m\sum_{k=-\pi/d}^{\pi/d}\sqrt{\frac{\hbar}{2N\mu_0\omega_{mk}}}[a_{mk}(t)\,\theta_{mk}(z)+\text{c.c.}], \tag{5}$$

where the amplitudes, $a_{mk}(t)$, are the classical analogue of raising and lowering operators. We now specialize to a Kerr nonlinear medium, and assume that the nonlinearity is weak: $\varepsilon_0\chi^{(3)}E^3\ll\varepsilon E$, where $\chi^{(3)}$ and ε, the nonlinear susceptibility and linear permittivity of the medium, are assumed to be periodic: $\chi^{(3)}(z)=\chi^{(3)}(z+d)$ and $\varepsilon(z)=\varepsilon(z+d)$. Under these approximations, Maxwell's equations can be written as [9]

$$\mathsf{H}=\mathsf{H}_L+\mathsf{H}_{NL} \tag{6}$$

with

$$\mathsf{H}_L=\sum_{mk}\hbar\omega_{mk}|a_{mk}|^2, \tag{7}$$

$$\mathsf{H}_{NL}=-\frac{\hbar^2\varepsilon_0}{16N^2\mu_0^2}\int_{-L/2}^{L/2}dz\frac{\chi^{(3)}(z)}{\varepsilon^4(z)}\times\left[\prod_{i=1}^{4}\sum_{m_ik_i}\frac{(a_{m_ik_i}\theta'_{m_ik_i}+\text{c.c.})}{\sqrt{\omega_{m_ik_i}}}\right],$$

where H_L is the portion of the full Hamiltonian that generates the linear dynamics of the electromagnetic field and H_{NL} is the portion that generates the nonlinear dynamics. Here H is numerically equal to the energy of the electromagnetic field in the presence of the material medium. The mode amplitudes satisfy commutation relations

$$[a_{mk}(t),a^\dagger_{m'k'}(t)]=\delta_{mm'}\delta_{kk'}, \tag{8}$$

$$[a_{mk}(t),a_{m'k'}(t)]=0,$$

and canonical equations of motion [9],

$$i\frac{da_{mk}}{dt}=\frac{1}{\hbar}[a_{mk},\mathsf{H}], \tag{9}$$

which give

$$\frac{da_{mk}}{dt}=-i\omega_{mk}a_{mk}+i\frac{\hbar\varepsilon_0}{4N^2\mu_0^2}\int_{-L/2}^{L/2}dz\frac{\theta'^*_{mk}\chi^{(3)}}{\sqrt{\omega_{mk}}\varepsilon^4}\times\left[\prod_{i=1}^{3}\sum_{m_ik_i}\frac{(a_{m_ik_i}\theta'_{m_ik_i}+\text{c.c.})}{\sqrt{\omega_{m_ik_i}}}\right], \tag{10}$$

where we have suppressed the z dependence of $\theta'_{mk}(z)$, $\chi^{(3)}(z)$, and $\varepsilon(z)$, and the time dependence of $a_{mk}(t)$.

III. THE COUPLED MODE EQUATIONS

In this section, we subject our Hamiltonian (6) to a series of approximations that are relevant to pulses that we wish to describe by the nonlinear coupled mode equations [7]. We assume that the frequency content of the pulse is entirely contained in the two bands $m=u,l$, where $m=u$ refers to the band just above the given photonic band gap, and $m=l$ refers to the band just below the gap. Because of this assumption, we can consider a Hamiltonian that includes contributions only from the $m=l,u$ bands without introducing any appreciable error into our expression for the total energy in the electromagnetic field. We also assume that the frequencies and wave numbers in the pulse of interest are close to a photonic band gap of the system, where the meaning of "close" will be made more precise in the following. To characterize the band gap, we define the gap width, Δ, and the Bragg frequency, ω_0, which is the frequency at the center of the gap,

$$\Delta\equiv(\omega_{uk_0}-\omega_{lk_0}), \tag{11}$$

$$\omega_0\equiv\frac{1}{2}(\omega_{uk_0}+\omega_{lk_0}),$$

where ω_{uk_0} is the frequency at the upper edge of the band gap, ω_{lk_0} is the frequency at the lower edge of the band gap, and k_0 can take on the value 0 or π/d depending on the photonic band gap of interest. The quantities ω_0, ω_{uk_0}, ω_{lk_0}, and Δ are indicated in Fig. 1. The photonic band gap in the figure is the lowest-order gap, so that $l=1$, $u=2$, and $k_0=\pi/d$.

We now make the transition from the mode amplitudes a_{lk} and a_{uk} to new mode amplitudes, g_{lk} and g_{uk}, that are more amenable to a coupled mode equation formulation. These new mode amplitudes are defined via the Bogoliubov transformation

$$a_{uk}=\gamma_k g_{uk}+i\beta_k g_{lk}, \tag{12}$$

$$a_{lk}=\gamma_k g_{lk}+i\beta_k g_{uk},$$

where γ_k and β_k are assumed to be real. We show below that our theory is valid when $|\beta_k|\ll|\gamma_k|$ so that from Eqs. (12) it is clear that $g_{l(u)k}$ is composed mostly of $a_{l(u)k}$ with only a small component of $a_{u(l)k}$. This field definition is consistent with previous analyses of the CME [7,8] where it has been shown that the fields in the CME are mixtures of fields that are contained in the upper and lower bands of the dispersion relation. We impose canonical commutation relations on the g_{uk} and g_{lk},

$$[g_{lk},g_{lk'}^{\dagger}]=[g_{uk},g_{uk'}^{\dagger}]=\delta_{kk'}, \tag{13}$$

with all other commutators vanishing. Since the a_{uk} and a_{lk} also satisfy such commutation relations (8), this leads to a restriction on the values of the γ_k and β_k,

$$\gamma_k^2+\beta_k^2=1. \tag{14}$$

When we write H_L (7) in terms of the g_{uk} and g_{lk} using Eqs. (12) we find a reduced Hamiltonian, H_L^R, that generates the linear dynamics of the electromagnetic field,

$$\begin{aligned}\mathsf{H}_L^R=&\sum_k\{\hbar\omega_{uk}\gamma_k^2+\hbar\omega_{lk}\beta_k^2\}g_{uk}g_{uk}^{\dagger}+\sum_k\{\hbar\omega_{lk}\gamma_k^2\\&+\hbar\omega_{uk}\beta_k^2\}g_{lk}g_{lk}^{\dagger}-i\hbar\sum_k\beta_k\gamma_k\{\omega_{uk}-\omega_{lk}\}\\&\times(g_{uk}g_{lk}^{\dagger}-g_{uk}^{\dagger}g_{lk}).\end{aligned} \tag{15}$$

In order to construct a reduced Hamiltonian that will derive the appropriate coupled mode equations, we assume that the frequencies in the upper and lower bands in the dispersion relation are symmetric about the Bragg frequency, ω_0, so that

$$\omega_{uk}=\omega_0+\frac{\Delta}{2}+f(K), \tag{16}$$

$$\omega_{lk}=\omega_0-\frac{\Delta}{2}-f(K),$$

where we have introduced the wave-number detuning

$$K\equiv k-k_0. \tag{17}$$

The form of $f(K)$ can be determined in the following manner. We write

$$\begin{aligned}\omega_{uk}=&\left(\omega_0+\frac{\Delta}{2}\right)+\left(\frac{\partial\omega_{uk}}{\partial k}\bigg]_{K=0}\right)K\\&+\frac{1}{2}\left(\frac{\partial^2\omega_{uk}}{\partial k^2}\bigg]_{K=0}\right)K^2+\cdots.\end{aligned} \tag{18}$$

Because we are evaluating the derivatives at the band edge, we know that the first derivative of ω_{uk} is zero. In the Appendix, we show that

$$\frac{\partial^2\omega_{uk}}{\partial k^2}\bigg]_{K=0}\simeq2\frac{|v_g|^2}{\Delta}[1+O(\eta)], \tag{19}$$

where we have introduced the smallness parameter $\eta=\Delta/\omega_0$, and where

$$v_g=\frac{c^2}{2\omega_0}\int_0^d\frac{1}{n^2(z)}\left[\left(\frac{\partial\theta_{uk_0}^*}{\partial z}\right)\theta_{lk_0}-\theta_{uk_0}^*\left(\frac{\partial\theta_{lk_0}}{\partial z}\right)\right]dz \tag{20}$$

plays the role that a "velocity matrix element" would in a theory of electrons in a periodic potential [14]. To the same order

$$\frac{\partial^2\omega_{lk}}{\partial k^2}\bigg]_{K=0}\simeq-2\frac{|v_g|^2}{\Delta}[1+O(\eta)]. \tag{21}$$

Our goal is to derive a simplified form for the Hamiltonian (15) accurate to order η^2. To this end we now place a restriction on the maximum allowable value of K by asserting that $K_{\max}/(\pi/d)=O(\eta^2)$, a choice we discuss in more detail after Eq. (25) below. We truncate Eq. (18) after the second derivative term; this is justified because the dispersion relation is even about k_0, so that the odd derivatives vanish, and because the largest value of the $2N$th term in the expansion (18) will be proportional to $[K_{\max}/(\pi/d)]^{2N}=O(\eta^{4N})$ and can thus be ignored at the level of perturbation that we are considering here. Comparing the form of $f(K)$ [Eq. (16)] to Eq. (18) and (19), we find that to lowest order

$$f(K)=\frac{|v_g|^2K^2}{\Delta}. \tag{22}$$

We note that the quantity $v_g\simeq\omega_0/(\pi/d)$, so that $f(K_{\max})/\Delta=O(\eta^2)$.

We next impose the condition

$$\omega_{uk}\gamma_k^2+\omega_{lk}\beta_k^2=\omega_{uk_0}, \tag{23}$$

$$\omega_{lk}\gamma_k^2+\omega_{uk}\beta_k^2=\omega_{lk_0}$$

on γ_k and β_k; because of Eq. (14), this is consistent with Eqs. (16). Then, using Eqs. (16) in Eqs. (23), and recalling the normalization condition on γ_k and β_k (14), we find

$$\gamma_k=\sqrt{\frac{\Delta+f(K)}{\Delta+2f(K)}},\quad \beta_k=\sqrt{\frac{f(K)}{\Delta+2f(K)}}. \tag{24}$$

Given the definition (22) of $f(K)$, it is clear that $\beta_k/\gamma_k=O(\eta)$. Using these definitions of γ_k and β_k (24) and the expression for $f(K)$ (22) in the expression for the reduced Hamiltonian (15), we find

$$\begin{aligned}\mathsf{H}_L^R=&\hbar\omega_0\sum_K\{g_{uK}g_{uK}^\dagger+g_{lK}g_{lK}^\dagger\}\\&+\hbar\frac{\Delta}{2}\sum_K\{g_{uK}g_{uK}^\dagger-g_{lK}g_{lK}^\dagger\}\\&-i\hbar v_g\sum_K K(g_{uK}g_{lK}^\dagger-g_{uK}^\dagger g_{lK})+O(\eta^3),\end{aligned} \tag{25}$$

where we have used a shorthand notation wherein $g_{uK}=g_{u(k_0+K)}$ and $g_{lK}=g_{l(k_0+K)}$. The second term in H_L^R is $O(\eta)$ with respect to the first term. Given our choice that $K_{\max}/(\pi/d)=O(\eta^2)$, the third term is $O(\eta^2)$ with respect to the first term, because $v_gK/\omega_0\simeq K_{\max}/(\pi/d)=O(\eta^2)$. Of course, were we to choose $K_{\max}/(\pi/d)=O(\eta)$, then the third term in H_L^R would be the same order as the second term, and the form of the Hamiltonian would seem to be unchanged. However, strictly speaking, the K in the third term should be replaced by $K[1+f(K)/\Delta]^{1/2}$. To ignore the dynamics described by $f(K)/\Delta$, we require $(f(K)/\Delta)=O(\eta)$, which, in turn, requires $K_{\max}/(\pi/d)=O(\eta^{3/2})$. Were we to choose this restriction on $K_{\max}$, then we could still recover the same Hamiltonian (25), but the terms in H_L^R that we ignore would be $O(\eta^2)$ rather than $O(\eta^3)$. So if we are considering only H_L^R, as we do in Sec. V, then the condition $K_{\max}/(\pi/d)=O(\eta^{3/2})$ would still allow us to reduce Eq. (15) to the desired Eq. (25) form. However, in the Appendix we show that we require $K_{\max}/(\pi/d)\leqslant O(\eta^2)$ for the nonlinear portion of the Hamiltonian to be easily tractable.

We now turn to the consideration of pulse dynamics using our Hamiltonian formulation. To do so, we build an effective field as the Fourier superposition of the $g_{mK}(t)$ [8],

$$g_m(z,t)=\sqrt{\frac{1}{L}}\sum_K g_{mK}(t)e^{iKz}, \tag{26}$$

in terms of which we find a reduced Hamiltonian density that generates the linear field dynamics,

$$\begin{aligned}\mathcal{H}_L^R(z,t)=&\hbar\omega_0(g_u^\dagger g_u+g_l^\dagger g_l)+\hbar\left(\frac{\Delta}{2}\right)(g_u^\dagger g_u-g_l^\dagger g_l)\\&-\frac{\hbar v_g}{2}\left(g_l^\dagger\frac{\partial g_u}{\partial z}-g_u\frac{\partial g_l}{\partial z}+\text{c.c.}\right).\end{aligned} \tag{27}$$

Using Eq. (26), the equal-time canonical commutation relations, defined in Eq. (13) for the mode amplitudes, extend to the effective fields as

$$[g_m(z,t),g_n^\dagger(z',t)]=\delta_{mn}\delta(z-z'), \tag{28}$$

where to use the Dirac δ function we have assumed that we are in the $L\to\infty$ limit, and that both z and z' are in the same normalization length. The equations of motion (9) become

$$i\frac{\partial g_m(z,t)}{\partial t}=\frac{1}{\hbar}[g_m(z,t),\mathsf{H}^R]. \tag{29}$$

We now consider the portion of the reduced Hamiltonian density associated with the nonlinear dynamics of the electromagnetic field. We use the smallness parameter $\eta\simeq\Delta/\omega_0$ to quantify the strength of the terms in H_{NL} (7). We first assume that the largest nonlinear terms in the Hamiltonian are $O(\eta^2)$ with respect to the largest linear terms. In the Appendix, we show that when $K_{\max}/(\pi/d)=O(\eta^2)$, we can write $\theta_{mk}=\theta_{mk_0}e^{iKz}+O(\eta)$. With this restriction on $K_{\max}$ we find that $\omega_{m(k_0+K)}=\omega_{mk_0}[1+O(\eta^3)]$, and that $a_{mk}=g_{mk}[1+O(\eta^2)]$. Since the strength of the nonlinearity is already $O(\eta^2)$, we find that

$$\mathsf{H}_{\text{NL}}^R=-\frac{\hbar^2\varepsilon_0}{16\mu_0^2}\int_{-L/2}^{L/2}dz\,\frac{\chi^{(3)}(z)}{\varepsilon^4(z)}\left[\prod_{i=1}^4\sum_{K_i}\left(\frac{g_{uK_i}e^{iK_iz}\theta'_{uk_0}}{\sqrt{\omega_{uk_0}}}+\frac{g_{lK_i}e^{iK_iz}\theta'_{lk_0}}{\sqrt{\omega_{lk_0}}}+\text{c.c.}\right)\right]+O(\eta^3), \tag{30}$$

where the $K_i = k_i - k_0$ are wave-number detunings. We ignore terms in Eq. (30) that have either zero or four complex conjugates, since they represent third-harmonic generation. We also ignore terms in Eq. (30) with one or three complex conjugates, because the dynamics that they describe will have only a negligible effect on the pulse dynamics. That is, terms with one or three complex conjugates will, roughly speaking, couple fields oscillating at $e^{-i\omega_0 t}$ to fields oscillating at $e^{+i\omega_0 t}$. Because of the rapid oscillation, the effect of this coupling is small, and is usually ignored by application of the rotating wave approximation (RWA) [15]. If those terms are kept here, it is straightforward to show that they can be eliminated by a multiple scales analysis for $\eta \ll 1$ that rigorously effects the RWA; we do not explicitly do this.

We still have to account for terms in Eq. (30) with two complex conjugates, which we do in the following manner. From Bloch's theorem we can write

$$\theta'_{pk_0}(z) = v_{pk_0}(z) e^{ik_0 z}, \tag{31}$$

where $p = u$ or l, and where $v_{pk_0}(z)$ is periodic with the lattice. The z-dependent portion of the integrand of a general term in (30) with two complex conjugates will be

$$I_{pqrs} = \int_{-L/2}^{L/2} \left(\frac{\chi^{(3)}}{\varepsilon^4} v_{pk_0} v^{\dagger}_{qk_0} v_{rk_0} v^{\dagger}_{sk_0} \right) \times e^{i(K_1 - K_2 + K_3 - K_4)} dz. \tag{32}$$

We can expand the portion in the parentheses on the left-hand side of Eq. (32) as a Fourier series because all quantities are periodic with the lattice. We then find

$$I_{pqrs} = \sum_n q^n_{pqrs} \int_{-L/2}^{L/2} e^{i(K_1 - K_2 + K_3 - K_4 + n(2\pi/d))} dz, \tag{33}$$

where the Fourier expansion coefficient is

$$q^n_{pqrs} = \frac{1}{d} \int_0^d \frac{\chi^{(3)}}{\varepsilon^4} \theta'_{pk_0} \theta'^{\dagger}_{qk_0} \theta'_{rk_0} \theta'^{\dagger}_{sk_0} e^{-in(2\pi/d)z} dz. \tag{34}$$

The integral on the right-hand side of Eq. (33) will vanish unless

$$K_1 - K_2 + K_3 - K_4 - n(2\pi/d) = 0. \tag{35}$$

Since we have stipulated that $K_i/(\pi/d) = O(\eta^2)$, this condition (35) can only be satisfied when $n = 0$, so that

$$I_{pqrs} = q^0_{pqrs} \int_{-L/2}^{L/2} e^{-i(K_1 - K_2 + K_3 - K_4)} dz. \tag{36}$$

In fact, only certain values of q^0_{pqrs} will be nonzero. This is because we are working at the band edge, so that the Bloch functions and their derivatives can be written as real functions with definite parity, with θ'_{uk_0} of opposite parity to θ'_{lk_0}. If we then assume that the quantity $\chi^{(3)}(z)/\varepsilon^4(z)$ is even, then the integrand of Eq. (34) will be odd about the point $z = d/2$, and will hence vanish, for q^0_{uuul}, q^0_{lllu}, and any permutations of these. This assumption can easily be relaxed, but adopting it we find that the reduced Hamiltonian density associated with H^R_{NL} is

$$\mathcal{H}^R_{NL}(z,t) = -\frac{\hbar}{2}\{\alpha_{uuuu}|g_u|^4 + 4\alpha_{uull}|g_u|^2|g_l|^2\} - \frac{\hbar}{2}\{\alpha_{ulul} g_u^2 g_l^{\dagger 2} + \alpha_{lulu} g_l^2 g_u^{\dagger 2} + \alpha_{llll}|g_l|^4\}, \tag{37}$$

where we have defined

$$\alpha_{pqrs} = \frac{3\hbar\varepsilon_0}{4\mu_0^2 d} \int_0^d dz \frac{\chi^{(3)}}{\varepsilon^4} \frac{(\theta'_{pk_0}\theta'^{\dagger}_{qk_0}\theta'_{rk_0}\theta'^{\dagger}_{sk_0})}{\sqrt{\omega_{pk_0}\omega_{qk_0}\omega_{rk_0}\omega_{sk_0}}}. \tag{38}$$

IV. TRAVELING-WAVE BASIS

In the preceding section, we constructed a reduced Hamiltonian density in terms of effective fields, $g_{u/l}(z,t)$, that were built as Fourier superpositions of Bloch functions at the band edges of a photonic band gap. It is well known that at the band edges the underlying Bloch functions are standing waves, and the $g_{u/l}(z,t)$ are thus effective fields associated with them. In this section, we convert the Hamiltonian formulation from the fields $g_{u/l}(z,t)$, to the fields $G_{\pm}(z,t)$, which are associated with traveling waves, and which satisfy the familiar coupled mode equations.

We start by defining [7]

$$G_{\pm}(z,t) = \frac{(g_l(z,t) \mp i g_u(z,t))}{\sqrt{2}}. \tag{39}$$

Using these new effective fields in Eq. (27), we find the reduced Hamiltonian densities

$$\mathcal{H}^R_L(z,t) = \hbar\omega_0(|G_+|^2 + |G_-|^2) - \frac{\hbar\Delta}{2}(G_+G_-^{\dagger} + G_-G_+^{\dagger}) - i\frac{\hbar v_g}{2} \times \left(G_+^{\dagger}\frac{\partial G_+}{\partial z} - G_-^{\dagger}\frac{\partial G_-}{\partial z} - \text{c.c.} \right), \tag{40}$$

and from Eq. (37) we find

$$\mathcal{H}^R_{NL}(z,t) = -\frac{\hbar}{2}\alpha_0\{|G_+|^4 + |G_-|^4 + 4|G_+|^2|G_-|^2\} - \hbar\alpha_1(G_+G_-^{\dagger} + G_-G_+^{\dagger})|G_+|^2 - \hbar\alpha_1(G_+G_-^{\dagger} + G_-G_+^{\dagger})|G_-|^2 - \frac{\hbar}{2}\alpha_2\{G_+^2G_-^{\dagger 2} + G_-^2G_+^{\dagger 2}\}, \tag{41}$$

where

$$\alpha_0 \equiv \tfrac{1}{4}\{\alpha_{uuuu} + 2\alpha_{uull} + \alpha_{llll}\}, \tag{42}$$

$$\alpha_1 \equiv \tfrac{1}{4}\{-\alpha_{uuuu}+\alpha_{llll}\},$$

$$\alpha_2 \equiv \tfrac{1}{4}\{\alpha_{uuuu}-6\alpha_{uull}+\alpha_{llll}\},$$

where α_{pqrs} is given by Eq. (38). This gives us a reduced Hamiltonian,

$$\mathsf{H}^R=\int_{-L/2}^{L/2}\{\mathcal{H}_L^R(z,t)+\mathcal{H}_{NL}^R(z,t)\}dz, \tag{43}$$

with commutation relations

$$[G_\pm(z,t),G_\pm^\dagger(z',t)]=\delta(z-z'), \tag{44}$$

where all other commutation relations are zero, and the Heisenberg equations of motion,

$$i\hbar\frac{\partial G_\pm}{\partial t}=[G_\pm,H]. \tag{45}$$

The coupled mode equations given by Eqs. (43) and (45) are

$$\begin{aligned}0=&\,i\frac{\partial G_\pm}{\partial t}\pm iv_g\frac{\partial G_\pm}{\partial z}-\omega_0G_\pm+\frac{\Delta}{2}G_\mp\\&+\alpha_0(|G_\pm|^2+2|G_\mp|^2)G_\pm\\&+\alpha_1(|G_\pm|^2+|G_\mp|^2)G_\mp+\alpha_1(G_\pm G_\mp^\dagger+G_\mp G_\pm^\dagger)G_\pm\\&+\alpha_2G_\mp^2G_\pm^\dagger.\end{aligned} \tag{46}$$

In these equations, the parameter v_g has the familiar interpretation of the group velocity in the absence of a grating. The form of these equations is the same as those presented earlier by de Sterke *et al.* [7,16]. However, there the field amplitudes were envelope functions that directly modulated the Bloch functions of the underlying medium, and a canonical formalism based on those field amplitudes cannot easily be constructed. In a previous paper [9], we have discussed the connection between the effective fields used in our Hamiltonian approach and the envelope functions used by de Sterke *et al*.

With the reduced Hamiltonian (41) in hand, we can now investigate the conserved quantities of the system. The required procedure is similar to that outlined in an earlier paper [9], so we simply present the results. In addition to the Hamiltonian, we find the following two conserved quantities:

$$\mathsf{P}=\frac{i\hbar v_g}{2c}\int_{-L/2}^{L/2}\left(G_+^\dagger\frac{\partial G_+}{\partial z}+G_-^\dagger\frac{\partial G_-}{\partial z}-\text{c.c.}\right)dz, \tag{47}$$

$$\mathsf{Q}=\hbar\omega_0\int_{-L/2}^{L/2}\{|G_+(z,t)|^2+|G_-(z,t)|^2\}dz.$$

The quantity P is the conserved momentum associated with translational invariance and the quantity Q is the conserved charge associated with phase translation invariance. We note that the underlying periodic system does not possess space translation invariance, but at the level of the effective fields such an invariance is, indeed, obtained.

V. DISCUSSION

We now consider the coupled mode equations in the absence of nonlinearity,

$$0=\frac{i}{v_g}\frac{\partial G_+}{\partial t}+i\frac{\partial G_+}{\partial z}-\frac{\omega_0}{v_g}G_++\kappa G_-, \tag{48}$$

$$0=\frac{i}{v_g}\frac{\partial G_-}{\partial t}-i\frac{\partial G_-}{\partial z}-\frac{\omega_0}{v_g}G_-+\kappa G_+,$$

where

$$\kappa\equiv\frac{\Delta}{2v_g}. \tag{49}$$

It is well known [1,8] that these coupled mode equations (48) give a definite prediction for the dispersion relation of the periodic system in the vicinity of the photonic band gap:

$$\omega(K)=\omega_0\pm v_g\sqrt{K^2+\kappa^2}, \tag{50}$$

$$\omega'(K)\equiv\frac{\partial\omega(K)}{\partial K}=\pm v_g\frac{K}{\sqrt{K^2+\kappa^2}},$$

$$\omega''(K)\equiv\frac{\partial^2\omega(K)}{\partial K^2}=\pm v_g\frac{\kappa^2}{(K^2+\kappa^2)^{3/2}},$$

where the $(+)$ sign refers to frequencies above the gap, and the $(-)$ sign to those below the gap.

Although the form of the CME presented here is equivalent to the heuristic CME derived elsewhere [1], the parameters ω_0, κ, and v_g are taken from the true band structure of the system. For strong index contrasts, the usual heuristic expressions [1] for those quantities are inadequate. In the remainder of this paper, we investigate how effectively the strong-grating CME reproduce the properties of the linear dispersion relation: phase velocity, group velocity, and group velocity dispersion. We consider systems in which the band-gap width is such a large fraction of the Bragg frequency that the validity of Δ/ω_0 as a smallness parameter is called into question. Nevertheless, we show that even when $\Delta/\omega_0 \simeq 0.25$, the CME give an excellent approximation to the linear properties of the structure. Furthermore, the CME remain remarkably valid over a larger range of K values.

In our simulations, we consider the index profile shown in Fig. 2, where in a unit cell of width d, a given portion, s, has index n_h, and the remainder has index n_l,

$$n(z)=\begin{cases}n_l, & -d/2<z<-s/2\\ n_h, & -s/2<z<s/2\\ n_l, & +s/2<z<d/2.\end{cases}$$

For a given frequency ω, the corresponding wave number, k, can be determined by the transcendental equation [17]

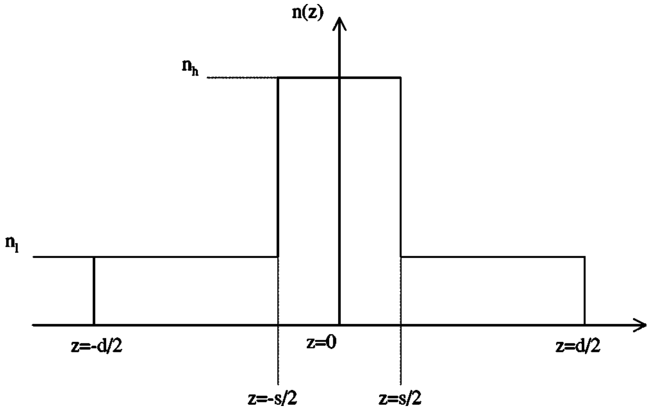

FIG. 2. Index profile throughout a unit cell of the periodic medium used in the simulations. An analytical solution exists for both the Bloch functions and the dispersion relation for a periodic medium with this unit cell.

$$\cos(kd)=\cos[p(d-s)]\cos(qs)-\frac{1}{2}\left(\frac{p^2+q^2}{qp}\right)\times\sin[p(d-s)]\sin(qs), \tag{51}$$

where we have defined $p\equiv n_l\omega/c$ and $q\equiv n_h\omega/c$. The Bloch functions, which are needed to evaluate v_g, can be determined by a simple transfer-matrix technique [17]. For a given fill fraction $F\equiv s/d$, we consider values of n_l and n_h in the following manner. We fix the lower and upper band indices to be $l=1,u=2$; that is, we investigate the lowest-order band gap associated with the system. We then vary n_l and n_h until we achieve a target Bragg frequency, ω_0^T, and a target band-gap width, Δ^T. For different fill fractions, the Bloch functions at the lower and upper band edge will be different, and hence the value of v_g will be different. We consider the two lowest-order bands because the main reason why the dispersion relation predicted by the CME deviates from the exact dispersion relation is that the CME do not include the effects of higher-order bands. As will be seen, the CME give an excellent approximation to the lower band, $l=1$, because the expression (19) for the group velocity dispersion is less affected by other photonic band gaps in the system. For the upper band, the curvature in the band that is induced by the higher-order photonic band gaps is more marked, and so the dispersion relation predicted by the CME deviates more noticeably from the exact dispersion relation.

For the remainder of the paper, we fix the target Bragg frequency to be $\omega_0^T=2c/d$, where c is the speed of light. For a small target band-gap width, $\Delta^T\leqslant 10^{-4}\omega_0^T$, this value of ω_0^T would correspond to a medium with an average index equal to $\pi/2$. In our simulations, we let the target band-gap width be 10%, 25%, and 50% of ω_0^T. For a fill fraction $F=0.005$, and $\Delta^T=0.5\omega_0^T$, we find that $n_l=1.26$ and $n_h=15.7$. This value of n_h would likely not be physically realizable, but we stress that these simulations are intended to investigate the validity of the CME, not the feasibility of designing photonic band-gap structures. In fact, such a fill fraction represents something of a worst-case scenario, because an index profile with $F=0.005$ contains an immense number of higher-order Fourier components, so higher-order bands interfere with the efficacy of the CME in the upper band. If we simulate structures with the same value of ω_0^T and Δ^T, but with a fill fraction closer to $F=0.5$, then the CME give a much better approximation to the exact dispersion relation. In much of what follows we use $F=0.005$, in order to demonstrate that the CME give a very robust description of the linear dynamics of the electromagentic field; but at the end of this section we verify that using $F=0.5$ makes the predicted dispersion relation of the CME much more accurate.

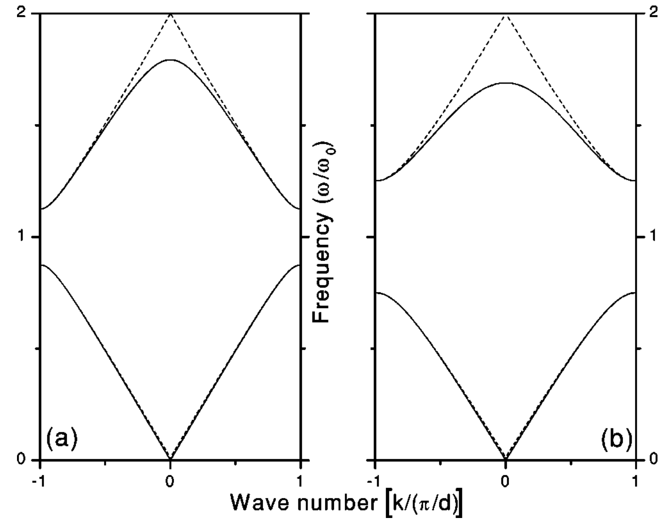

FIG. 3. Exact dispersion relation (solid line) and CME prediction (dashed line) for a periodic system with $F=0.005$. In (a) the gap width is 25% of the Bragg frequency; in (b) the gap width is 50% of the Bragg frequency. In both cases the CME give an excellent approximation to the exact dispersion relation for frequencies below the Bragg frequency. For frequencies above the Bragg frequency the agreement is excellent for wave numbers very close to the Bragg wave number, k_0. Away from the Bragg wave number the exact dispersion relation curves in order to account for the next higher photonic band gap. The CME do not account for this extra curvature.

In Fig. 3, we plot the exact dispersion relation (solid line) and the CME prediction (dashed line) for structures with $F=0.005$, and with $\Delta^T=0.25\omega_0$ [Fig. 3(a)] and $\Delta^T=0.5\omega_0$ [Fig. 3(b)]. The value of ω is normalized to ω_0, and the value of k is normalized to π/d. For frequencies below ω_0, the dispersion relation predicted by the CME is virtually indistinguishable from the exact dispersion relation. For frequencies above ω_0, the exact value diverges from the CME prediction, because the existence of the next higher-order photonic band-gap (centered about $\omega=2\omega_0$) is not built into the coupled mode equations. The true dispersion relation has to curve downwards, because it must account for this next higher photonic band gap. Nevertheless, the CME give a very good fit to the dispersion relation for $\omega\leqslant 1.5\omega_0$ when $\Delta^T=0.25\omega_0$, and for $\omega\leqslant 1.35\omega_0$ when $\Delta^T=0.5\omega_0$.

In addition to giving an excellent approximation to the dispersion relation, the CME accurately predict the value of the complex wave number inside the photonic band gap. In-

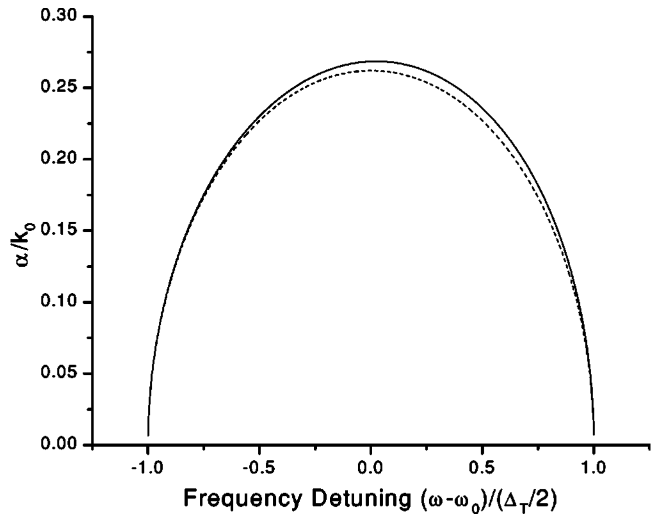

FIG. 4. Imaginary part of the wave number for frequencies inside the photonic band gap for a structure with $F=0.005$, and a gap width that is 50% of the Bragg frequency. The CME approximation (dashed line) agrees very closely with the exact value (solid line).

side the gap we expect the wave number to be $k=k_0+i\zeta$. Inverting the expression for $\omega(K)$ (50) we find that for frequencies that are inside the photonic band gap the CME predict

$$i\zeta(\omega)=\pm\frac{1}{v_g}\sqrt{(\omega-\omega_0)^2-v_g^2\kappa^2}. \tag{52}$$

In Fig. 4, we compare the CME prediction of $\zeta(\omega)$ (dashed line) to that given by the analytical expression for $F=0.005$ and $\Delta^T=0.5\omega_0$. Again the agreement is excellent; the peak value of $\zeta(\omega)$ predicted by the CME differs by only 2.5% relative to the exact value. The asymmetry between frequencies above and below the Bragg frequency that was seen in Fig. 3 is again evident in Fig. 4, but the effect is much more slight.

In Fig. 5, we compare the values of $\omega'(K)$ and $\omega''(K)$ predicted by the coupled mode equations (crosses) to those given by the exact dispersion relation (solid line) of the rectangular index profile, as a function of frequency. We use $\Delta^T=0.25\omega_0^T$ and $F=0.005$. Because we expect our coupled mode equations to be valid only when $K/(\pi/d)$ is small, we plot the values of $\omega'(K)$ and $\omega''(K)$ for $-0.2<K/(\pi/d)<0.2$. The CME give an excellent approximation to the exact values.

We have mentioned that a fill fraction $F=0.005$ is something of a worst-case scenario. In order to verify this, we define a group velocity dispersion deviation coefficient

$$R(K)\equiv\left|\frac{\omega''(K)\rfloor_{\text{exact}}-\omega''(K)\rfloor_{\text{CME}}}{\omega''(K)\rfloor_{\text{exact}}}\right|\times 100\%. \tag{53}$$

We then fix $\Delta^T=0.25\omega_0^T$, and determine the value of $R_{\max}$ for $-0.2<K/(\pi/d)<0.2$ for various fill fractions F. We find that although R_{MAX} is about 15% for $F=0.005$, it drops to as little as 1% for $F=0.5$. This is because for $F=0.5$ the

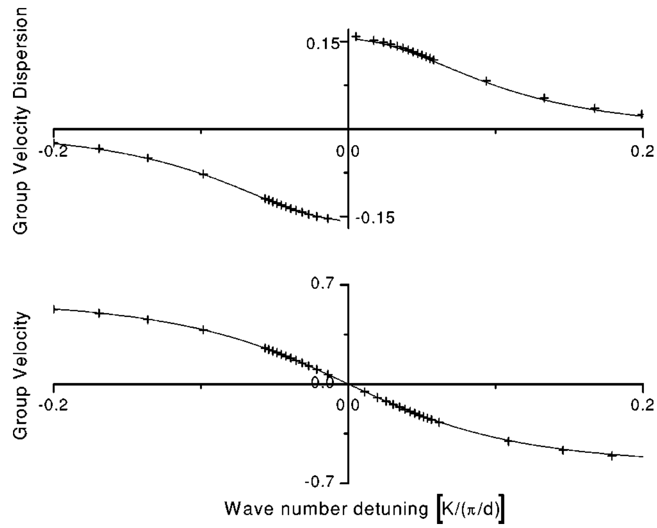

FIG. 5. Group velocity and group velocity dispersion as a function of wave number for the exact dispersion relation (solid line) and the CME (crosses). The structure has $F=0.005$ and a gap width that is 25% of the Bragg frequency. For wave number detunings that are relatively small $(K/(\pi/d)<0.2)$, the CME give an excellent approximation.

Fourier component of the periodic index variation that is responsible for the next higher photonic band gap is not too strong, so that the extra curvature needed for the band to open does not interefere as much with frequencies in the vicinity of the first photonic band gap.

VI. CONCLUSION

We have presented a canonical Hamiltonian formulation of Maxwell's equations in the presence of a nonlinear, periodic Kerr medium. The Hamiltonian is written in terms of mode amplitudes that modulate the Bloch functions of the linear medium. We have shown that if the electromagnetic field is composed of frequencies within or near a photonic band gap of the system, then a reduced Hamiltonian, written in terms of effective fields, can be used to describe light propagation in the system. The Hamiltonian is equal to the energy in the field, and can easily be quantized. The equations of motion that are generated by the reduced Hamiltonian are the nonlinear CME.

We have investigated the effectiveness with which the CME approximate the dispersion relation of the underlying periodic medium in the absence of nonlinearity. It was shown that even for large index contrasts, which lead to the opening of photonic band gaps with widths up to 25% of the Bragg frequency, the CME give an excellent approximation to the dispersion relation both within the photonic band gap and for a large range of frequencies outside the band gap. Since the CME are applicable to one-dimensional systems with a large index contrast, they might remain a useful heuristic tool for the investigation of pulse propagation in two- and three-dimensional photonic crystals. The method used in this paper to derive the nonlinear CME can easily be extended to higher-dimensional systems, and will thus be useful in such an investigation.

ACKNOWLEDGMENTS

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada, and Photonics Research Ontario.

APPENDIX: $\boldsymbol{k}\cdot\boldsymbol{p}$ FOR THE DUAL FIELD

The purpose of this appendix is twofold: first, we derive the expressions (19) and (21) for the group velocity dispersion given in the text; second, we justify the statements made before Eq. (30) concerning the behavior of quantities such as ω_{mk} and θ_{mk} near a band gap. We start by using Eqs. (2) and (3) in Maxwell's equations in order to generate a set of equations typical of those used in $\boldsymbol{k}\cdot\boldsymbol{p}$ analyses [14],

$$0=(\hat{V}_k-\mu_0\omega_{mk}^2)u_{mk}, \tag{A1}$$

$$0=\left[\frac{\partial\hat{V}_k}{\partial k}-\mu_0\frac{\partial(\omega_{mk}^2)}{\partial k}\right]u_{mk}+[\hat{V}_k-\mu_0\omega_{mk}^2]\frac{\partial u_{mk}}{\partial k},$$

$$0=\left[\frac{\partial^2\hat{V}_k}{\partial k^2}-\mu_0\frac{\partial^2(\omega_{mk}^2)}{\partial k^2}\right]u_{mk}+2\left[\frac{\partial\hat{V}_k}{\partial k}-\mu_0\frac{\partial(\omega_{mk}^2)}{\partial k}\right]\frac{\partial u_{mk}}{\partial k}+[\hat{V}_k-\mu_0\omega_{mk}^2]\frac{\partial^2 u_{mk}}{\partial k^2},$$

where the second and third equations in Eqs. (A1) are the k derivatives of the first equation, and where we have defined the differential operator

$$\hat{V}_k=-\frac{1}{\varepsilon}\left(\frac{\partial}{\partial z}+ik\right)^2+\frac{\varepsilon_z}{\varepsilon^2}\left(\frac{\partial}{\partial z}+ik\right) \tag{A2}$$

with $\varepsilon_z=d\varepsilon/dz$. Multiplying the second equation in Eqs. (A1) by u_{pk}^*, and integrating over d, the length of one unit cell gives

$$\left\langle u_{pk}\left|\frac{\partial u_{mk}}{\partial k}\right.\right\rangle dz=-\frac{\left\langle u_{pk}\left|\frac{\partial\hat{V}_k}{\partial k}\right|u_{mk}\right\rangle}{\mu_0(\omega_{pk}^2-\omega_{mk}^2)}, \tag{A3}$$

where we have defined a notation for the overlap integrals,

$$\langle A|\hat{O}|B\rangle=\int_0^d A^*(\hat{O}B)dz,$$

where $\hat{O}$ is a differential operator. In writing down Eq. (A3) we have assumed that u_{mk} is orthogonal to $\partial u_{mk}/\partial k$, as is usual in $\boldsymbol{k}\cdot\boldsymbol{p}$ analyses. The only consequence of this is that the u_{mk} will in general not be normalized in the same way as the u_{mk_0}, and thus the θ_{mk} will not be normalized in the same way as the θ_{mk_0}. We adopt this convention *only* in this appendix, and not in the text. Our first purpose in this appendix is to derive expressions for the group velocity dispersion at the wave number k_0, which will be expressed in terms of Bloch functions at k_0, so the normalization of the other Bloch functions is ultimately unimportant. Our second purpose involves rewriting $\theta_{m(k_0+K)}\simeq\theta_{mk_0}e^{iKz}$, so that the normalization *is* important. However, we will show that our approximation to the Bloch functions at wave number k_0+K is correctly normalized to the desired order in perturbation theory.

Multiplying the third equation in Eqs. (A1) by u_{mk}^* and integrating over d we find an expression for the group velocity dispersion (GVD),

$$\frac{\partial^2\omega_{mk}}{\partial k^2}=-\frac{1}{\omega_{mk}}\sum_{q\neq m}\left(\frac{(\omega_{qk}+\omega_{mk})^2}{(\omega_{qk}^2-\omega_{mk}^2)}v_{qm}v_{mq}\right)+\left\langle u_{mk}\left|\frac{1}{\mu_0\varepsilon\omega_{mk}}\right|u_{mk}\right\rangle-\frac{1}{\omega_{mk}}\left(\frac{\partial\omega_{mk}}{\partial k}\right)^2, \tag{A4}$$

where we have defined a group velocity matrix element

$$v_{mn}(k)=\frac{1}{\mu_0}\left[\frac{\left\langle u_{mk}\left|\frac{\partial\hat{V}_k}{\partial k}\right|u_{nk}\right\rangle}{(\omega_{mk}+\omega_{nk})}\right]. \tag{A5}$$

At the band edge, the quantity $\partial\omega_{mk}/\partial k$ vanishes so that, for the upper band,

$$\frac{\partial^2\omega_{uk_0}}{\partial k^2}=\frac{1}{\omega_{uk_0}}\left\{2\left(\frac{\omega_0}{\Delta}\right)v_{lu}v_{ul}-\sum_{q\neq u,l}\left(\frac{(\omega_{qk_0}+\omega_{uk_0})^2}{(\omega_{qk_0}^2-\omega_{uk_0}^2)}v_{qu}v_{uq}\right)+\left\langle u_{uk_0}\left|\frac{1}{\mu_0\varepsilon}\right|u_{uk_0}\right\rangle\right\}. \tag{A6}$$

Because we are assuming that $\Delta\ll\omega_0$, and that other bands in the system are distant from the band of interest, the first term in the brackets will be much larger than the second. It will also be much larger than the third, as can be seen by noting the following. The quantity $\varepsilon(z)\geqslant 1$, so

$$\left\langle u_{uk_0}\left|\frac{1}{\mu_0\varepsilon}\right|u_{uk_0}\right\rangle<c^2\langle u_{uk_0}|u_{uk_0}\rangle=c^2.$$

The quantity $v_{ul}v_{lu}$ is on the order of c^2, so the ratio of the third term to the first term is approximately Δ/ω_0. We define a smallness parameter $\eta=\Delta/\omega_0$ so that

$$\frac{\partial^2\omega_{uk_0}}{\partial k^2}=2\frac{v_{lu}v_{ul}}{\Delta}[1+O(\eta)], \tag{A7}$$

where we have used the fact that $\omega_0/\omega_{uk_0}\simeq 1$. In the text we have defined $v_g \equiv -iv_{ul}(k_0)$.

We now turn to the second issue of this appendix. We start by expanding an arbitrary $u_{mk}(z)$ as a linear combination of the u's at wave number k_0,

$$u_{mk}(z)=\sum_n \gamma_{mk}^{nk_0} u_{nk_0}(z), \tag{A8}$$

where the $\gamma_{mk}^{nk_0}$ are not functions of z. We evaluate $\gamma_{mk}^{nk_0}$ by taking a Taylor series,

$$u_{lk}=u_{lk_0}+K\left(\frac{\partial u_{lk}}{\partial k}\bigg\rfloor_{K=0}\right)+\cdots, \tag{A9}$$

$$u_{uk}=u_{uk_0}+K\left(\frac{\partial u_{uk}}{\partial k}\bigg\rfloor_{K=0}\right)+\cdots,$$

where we have defined $K\equiv k-k_0$. We now place a restriction on the maximum allowable value of K in our theory, $K_{\max}/(\pi/d)=O(\eta^2)$, which allows us to ignore the K^2 and higher-order terms in the expansion (A9). An expression for $\partial u_{mk}/\partial k\rfloor_{K=0}$ can be found by writing

$$\frac{\partial u_{mk}}{\partial k}\bigg\rfloor_{K=0}=\sum_q a_q u_{qk_0}$$
$$=-\frac{1}{\mu_0}\sum_q\left[\frac{\left\langle u_{qk_0}\left|\frac{\partial \hat{V}_k}{\partial k}\right|u_{mk_0}\right\rangle}{(\omega_{qk_0}^2-\omega_{mk_0}^2)}\right]u_{qk_0}, \tag{A10}$$

where we have used Eq. (A3). We assume that the value of the overlap integrals, $\langle u_{qk_0}|\partial\hat{V}_k/\partial k|u_{mk_0}\rangle$, will be of roughly the same order, in which case terms with $q\neq u,l$ will be $O(\eta)$ with respect to terms with $q=u,l$. Because of this we can write, to lowest order in η,

$$u_{lk}\simeq\gamma_{lk}^{lk_0}u_{lk_0}+\gamma_{lk}^{uk_0}u_{uk_0}, \tag{A11}$$

$$u_{uk}\simeq\gamma_{uk}^{lk_0}u_{lk_0}+\gamma_{uk}^{uk_0}u_{uk_0}$$

with

$$\gamma_{lk}^{lk_0}=1,\quad \gamma_{uk}^{lk_0}=-iv_g\frac{K}{\Delta}, \tag{A12}$$

$$\gamma_{uk}^{uk_0}=1,\quad \gamma_{lk}^{uk_0}=-iv_g\frac{K}{\Delta}.$$

This means that a Bloch function in the upper band, say, can be written as

$$\theta_{uk}=\sum_b \gamma_{uk}^{bk_0}\theta_{bk_0}e^{iKz}\simeq\left(\theta_{uk_0}-iv_g\frac{K}{\Delta}\theta_{lk_0}\right)e^{iKz}. \tag{A13}$$

Now, since v_g is of order $\omega_0/(\pi/d)$, and because we have assumed that $K_{\max}/(\pi/d)=O(\eta^2)$, we find that $v_g K_{\max}/\Delta \simeq O(\eta)$, so we can write

$$\theta_{uk}=\theta_{uk_0}e^{iKz}+O(\eta), \tag{A14}$$

$$\theta_{lk}=\theta_{lk_0}e^{iKz}+O(\eta).$$

These expressions for θ_{uk} and θ_{lk} are used to determine the portion of the Hamiltonian that generates the nonlinear dynamics. Although θ_{uk} and θ_{lk} are not, strictly speaking, normalized according to Eq. (4), they are normalized to $O(\eta)$.

[1] C.M. de Sterke and J.E. Sipe, in *Progress in Optics*, edited by E. Wolf (North-Holland, Amsterdam, 1994), Vol. 33, pp. 203–260.
[2] W. Samir, S.J. Garth, and C. Past, J. Opt. Soc. Am. B **11**, 64 (1994).
[3] B.J. Eggleton *et al.*, Opt. Commun. **149**, 267 (1998).
[4] B.J. Eggleton, C.M. de Sterke, and R.E. Slusher, J. Opt. Soc. Am. B **14**, 2980 (1997).
[5] S. John, Phys. Rev. Lett. **65**, 2486 (1987).
[6] E. Yablonovitch, Phys. Rev. Lett. **65**, 2059 (1987).
[7] C.M. de Sterke, D.G. Salinas, and J.E. Sipe, Phys. Rev. E **54**, 1969 (1996).
[8] Suresh Pereira and J.E. Sipe, Phys. Rev. E **62**, 5745 (2000).
[9] Suresh Pereira and J.E. Sipe, Phys. Rev. E **65**, 046601 (2002).
[10] M. Hillery and L.D. Mlodinow, Phys. Rev. A **30**, 1860 (1984).
[11] P.D. Drummond, Phys. Rev. A **42**, 6845 (1990).
[12] P.D. Drummond, in *Recent Developments in Quantum Optics*, edited by R. Inguva (Plenum, New York, 1993), pp. 65–76.
[13] N.W. Ashcroft and N.D. Mermin, *Solid State Physics* (Saunders College, Philadelphia, 1976).
[14] M. Lax, *Symmetry Principles in Solid State and Molecular Physics* (Wiley, New York, 1974).
[15] Dietrich Marcuse, *Theory of Dielectric Optical Waveguides* (Academic, New York, 1974).
[16] Our Eq. (46) should be compared with Eq. (102) in Ref. [7]. However, the equation in [7] contains a typographical error in the nonlinear term modulated by α_2. The equation in this paper is the correct equation.
[17] D.R. Smith *et al.*, J. Opt. Soc. Am. B **10**, 314 (1993).