Light propagation through birefringent, nonlinear media with deep gratings

Suresh Pereira and J. E. Sipe
Department of Physics, University of Toronto, 60 St. George Street, Toronto, Ontario, Canada M5S 1A7
(Received 2 February 2000; revised manuscript received 25 April 2000)

We present a theory that includes birefringence in the description of one-dimensional photonic band-gap materials with a Kerr nonlinearity. The Bloch functions in the absence of nonlinearity completely characterize the linear problem, for deep as well as shallow gratings, and the method of multiple scales is used to include the effects of nonlinearity and finite optical pulse length. We derive two sets of equations appropriate in different frequency regimes, a set of coupled mode equations and a set of coupled nonlinear Schrödinger equations; we investigate the connections between these equations and where their regimes of validity overlap. Finally, we use our results to describe energy exchange between polarization modes in a birefringent medium.

PACS number(s): 42.79.Dj, 42.81.Gs

I. INTRODUCTION

In recent years, much effort has been devoted to the study of one-dimensional photonic band-gap materials in the presence of a Kerr nonlinearity [1–6]. A great deal of the experimental work in this field has concentrated on fiber Bragg gratings, which typically have refractive index variations on the order of 10^{-4} [7–9]. With such small index variations, it is reasonable to apply the heuristic coupled mode equations, or the appropriate nonlinear Schrödinger equation, to analyze experimental results [6,7]. However, index changes as high as 0.04 have been reported in fibers [10], and experiments employing etched semiconductor wave guides have been proposed [11]. These systems have sufficiently large index contrasts so as to cast doubt on the validity of the heuristic coupled mode equations. In a recent paper, a coupled mode theory was developed that accounts for strong gratings, in which the index contrast varies over a significant fraction of the average background index, with a Kerr nonlinearity [5].

In this paper we seek to extend the strong grating, nonlinear coupled mode formalism to include birefringence. Although optical fibers are nominally isotropic, the process of writing a grating introduces a birefringence on the order of 10^{-6} [12]. Birefringence has the effect of separating the photonic band gaps of the two polarizations so that, in certain frequency ranges, light of one polarization can propagate freely while the other is blocked. This has immediate deleterious consequences for proposed devices based on circular polarization, where the linearly polarized signals are mixed. The robustness of nonlinear effects, such as soliton formation and propagation through grating structures, has yet to be studied in the presence of birefringence. The dynamics here can be expected to be more complicated than in a bare optical fiber. In a bare optical fiber the two polarizations have different group velocities, but can be considered to suffer equal dispersion [13]; this is not generally valid in the presence of a grating. In addition to fiber experiments, semiconductor wave guides with a $\chi^{(3)}$ nonlinearity, which possess TE and TM modes with different group velocities and dispersions, and which can have large index contrasts, have been studied experimentally [11]. Although our formalism is strictly one dimensional, it provides a qualitative insight into the properties of such structures. We note, too, that experiments in the literature, such as the all optical AND gate demonstrated by Taverner *et al*. [9], require a coupled mode formalism for their convenient analysis, as will other experiments aimed at exploiting polarization and nonlinearity.

In a previous paper we reported weak-grating coupled mode equations for pulses in a nonlinear, birefringent, periodic medium [14]. This paper includes derivations for three sets of equations: weak- and strong-grating coupled mode equations (CME), and coupled nonlinear Schrödinger equations (CNLSE) in the presence of birefringence. We use Bloch theory to characterize the linear, birefringent problem, and the method of multiple scales to include the nonlinearity and finite pulse width. Both the birefringence and nonlinearity are assumed to be weak, in a sense to be made precise below.

The outline of this paper is as follows. In Sec. II we discuss the linear properties of a one-dimensional, birefringent, periodic medium. In Sec. III we introduce the method of multiple scales, which we then use in Sec. IV to derive a set of coupled nonlinear Schrödinger equation, and in Sec. V to derive a set of of coupled mode equations. In Sec. VI we discuss the connection between the nonlinear Schrödinger equations and the coupled mode equations, and their respective regions of validity. In Sec. VII we present numerical simulations that verify the connections discussed in Sec. VI. We also investigate the phenomenon of nonlinear energy exchange in the context of the CNLSE, and present an analytic model to describe that effect.

II. LINEAR EQUATIONS AND BASIS FUNCTIONS

We begin with the linear Maxwell equations in the presence of a dielectric tensor that is a function of only one Cartesian component, $\varepsilon=\varepsilon(z)$. We assume that the (x,y) coordinates can be chosen such that for all z the tensor is diagonal, $\varepsilon=\mathrm{diag}(\varepsilon_{xx}(z),\varepsilon_{yy}(z))$. Neglecting magnetic effects by setting the permeability μ equal to that of free space, $\mu=\mu_0$, we can then define indices of refraction associated with polarization along the x and y axes, $n_i^2(z)=\varepsilon_{ii}(z)/\varepsilon_0$, where ε_0 is the permittivity of free space and where, for the remainder of the text, the index i runs over x

and y. We seek fields $\mathbf{E}(\mathbf{r},t),\mathbf{H}(\mathbf{r},t)$ that depend only on the coordinate z. To proceed, we introduce local mode amplitudes $A^{\pm}_{x,y}$ [5] according to

$$A^{\pm}_{x}=\frac{1}{2}\sqrt{\frac{n_x(z)}{n_0}}\left[E_x(z,t)\pm Z_0\frac{H_y(z,t)}{n_x(z)}\right], \tag{1}$$

$$A^{\pm}_{y}=\frac{1}{2}\sqrt{\frac{n_y(z)}{n_0}}\left[E_y(z,t)\mp Z_0\frac{H_x(z,t)}{n_y(z)}\right],$$

where n_0 is a reference refractive index and $Z_0=(\mu_0/\varepsilon_0)^{1/2}$ is the impedance of free space. Using Eq. (1) in Maxwell's equations we can derive the differential equations that the $A^{\pm}_i$ fields satisfy,

$$i\mathbf{n}_i\frac{\partial\mathbf{A}_i}{\partial t}=\mathbf{M}_i\cdot\mathbf{A}_i, \tag{2}$$

with the column vectors

$$\mathbf{A}_i=\begin{bmatrix}A^{+}_i(z,t)\\ A^{-}_i(z,t)\end{bmatrix}, \tag{3}$$

the matrix differential operators

$$\mathbf{M}_i=\begin{bmatrix}-ic\dfrac{\partial}{\partial z} & \dfrac{1}{2}\left(\dfrac{\partial[\ln n_i(z)]}{\partial z}\right)\\ -\dfrac{1}{2}\left(\dfrac{\partial[\ln n_i(z)]}{\partial z}\right) & ic\dfrac{\partial}{\partial z}\end{bmatrix}, \tag{4}$$

where c is the speed of light in vacuum, and the index matrices

$$\mathbf{n}_i=\begin{bmatrix}n_i(z) & 0\\ 0 & n_i(z)\end{bmatrix}. \tag{5}$$

The similarity between our Eqs. (2) and those of de Sterke *et al.* [5] allows us to proceed in a manner analogous to theirs, except for the complication of having both x and y polarized fields. The idea is to assume an harmonic time dependence $e^{-i\omega_{\mu i}t}$ for the $\mathbf{A}_i$ fields, and then formally solve for the z dependence in terms of the eigenvectors $\mathbf{\Psi}_{\mu i}$ of the matrix $\mathbf{n}_i^{-1}\mathbf{M}_i$.

Periodic structures. To find the $\mathbf{\Psi}_{\mu i}$ the particular $\varepsilon(z)$ must be specified. Since we assume $\varepsilon(z)$ is periodic with period d, $\varepsilon(z+d)=\varepsilon(z)$, we can use Bloch's theorem to construct the $\mathbf{\Psi}_{\mu i}$ [15]. To connect with other literature it is convenient to write the $\mathbf{\Psi}_{\mu i}$ in terms of the corresponding solutions $\phi_{\mu i}(z)$ for the electric field itself which satisfy [5]

$$-c^2\frac{\partial^2\phi_{\mu i}(z)}{\partial z^2}=\omega^2_{\mu i}n_i^2(z)\phi_{\mu i}(z), \tag{6}$$

where $\omega_{\mu i}$ is considered positive, and are of the form

$$\phi_{mi}(k;z)=e^{ikz}u_{mi}(k;z), \tag{7}$$

where $u_{mi}(k;z+d)=u_{mi}(k;z)$; that is, the $u_{mi}(k;z)$ have the periodicity of the lattice. Note that the index μ has been replaced by a discrete band index m and a reduced wave number k, $(-\pi/d<k\leqslant\pi/d)$. If we seek $\phi_{mi}(k;z)$ that satisfy periodic boundary conditions over a normalization length L, then k must be of the form $2\pi p/L$ where p is an integer. We denote the associated eigenfrequencies $\omega_{mi}(k)$. For each polarization the Bloch functions are orthogonal through the metric $n_i^2(z)$,

$$\int_0^L\phi^{*}_{m'i}(k';z)n_i^2(z)\phi_{mi}(k;z)dz=N\delta_{m'm}\delta_{k'k}, \tag{8}$$

where the normalization constant $N=L/d$ has been chosen to facilitate passage to the $L\to\infty$ limit. In terms of the $\phi_{mi}(k;z)$ we find [5]

$$\mathbf{\Psi}_{mi}(k)=\begin{bmatrix}\psi^{+}_{mi}(k;z)\\ \psi^{-}_{mi}(k;z)\end{bmatrix}, \tag{9}$$

with

$$\psi^{\pm}_{mi}(k;z)=\frac{1}{2}\left[\sqrt{n_i(z)}\phi_{mi}(k;z)\right.$$
$$\left.\mp\frac{ic}{\omega_{mi}(k)}\frac{1}{\sqrt{n_i(z)}}\frac{\partial\phi_{mi}(k;z)}{\partial z}\right]. \tag{10}$$

Properties of the dispersion relation such as group velocity and group velocity dispersion at a given m,k point, for a given polarization, can be determined via the ''$k\cdot p$ expansion'' [5]. The use of the $\mathbf{\Psi}_{mi}(k)$ is preferred over the use of the usual $\phi_{mi}(k;z)$ because the former leads to a much simpler $k\cdot p$ expansion and a much simpler implementation of a multiple scales analysis. We here simply give the key results. The velocity matrix element $v_{pq(i)}(k)$, at wave number k associated with bands p and q and associated with polarization i is defined as

$$\frac{v_{pq(i)}(k)}{c}=-\frac{ic}{2}\left(\frac{1}{\omega_{pi}(k)}+\frac{1}{\omega_{qi}(k)}\right)$$
$$\times\int_0^d\phi^{*}_{pi}(k;z)\frac{\partial\phi_{qi}(k;z)}{\partial z}dz. \tag{11}$$

The group velocity and group velocity dispersion are given by

$$\omega'_{mi}(k)\equiv\frac{d\omega_{mi}(k)}{dk}=v_{mm(i)}(k) \tag{12}$$

and

$$\omega''_{mi}(k)\equiv\frac{d^2\omega_{mi}(k)}{dk^2}=-2\sum_{\tilde{p}\neq m}\frac{v_{m\tilde{p}(i)}(k)v_{\tilde{p}m(i)}(k)}{\omega_{\tilde{p}i}(k)-\omega_{mi}(k)}. \tag{13}$$

We note that the sum in Eq. (13) goes over positive and negative frequencies [16].

III. NONLINEARITY AND MULTIPLE-SCALE ANALYSIS

Having characterized the linear problem in the presence of birefringence, we now turn to the inclusion of nonlinear-

ity. In the presence of a nonlinear polarization $\mathbf{P}^{\mathrm{NL}}(\mathbf{r},t)=\hat{x}P_x^{\mathrm{NL}}(z,t)+\hat{y}P_y^{\mathrm{NL}}(z,t)$, the Maxwell equations become

$$i\mathbf{n}_i(z)\frac{\partial \mathbf{A}_i}{\partial t}=\mathbf{M}_i\cdot\mathbf{A}_i+\mathbf{B}_i, \tag{14}$$

where

$$\mathbf{B}_i=-\frac{i}{2\varepsilon_o\sqrt{n_0 n_i(z)}}\frac{\partial P_i^{NL}}{\partial t}\begin{bmatrix}1\\1\end{bmatrix}. \tag{15}$$

To describe the nonlinear polarization we adopt a nondispersive Kerr model

$$P_i(\mathbf{r},t)=\varepsilon_o\chi_{ijkl}^{(3)}E_j(\mathbf{r},t)E_k(\mathbf{r},t)E_l(\mathbf{r},t), \tag{16}$$

with $i,j,k,l=x,y$. It is clear that this form of the nonlinear polarization will couple the $\mathbf{A}_i$ vectors. Of course, the Kerr model (16) is only a reasonable assumption if the intensities involved are not large enough that higher order susceptibilities need also be included. We refer to this as the weakly nonlinear regime.

Multiple-scale analysis. We would like to use the nonlinear equation (14) to treat pulses described by envelope functions that are slowly varying in time and space relative to a carrier frequency and lattice period respectively. One method of carefully accounting for the effects of a ''weak nonlinearity'' and ''slowly varying'' pulses is the method of multiple scales [5,6]. This method requires the introduction of several time and space scales via a smallness parameter $\eta\ll 1$. One can then write a typical function as

$$f(z,t)=F(z,\eta z,\eta^2 z,\ldots;t,\eta t,\eta^2 t,\ldots), \tag{17}$$

where F is assumed to vary equally significantly as each of its spatial arguments varies over a given range l, and each of its temporal arguments varies over a given period τ. The multiple scales of the problem are defined by

$$z_p=\eta^p z, \tag{18}$$

$$t_p=\eta^p t.$$

Using Eq. (18) in Eq. (17) we find

$$f(z,t)=F(z_0,z_1,z_2,\ldots;t_0,t_1,t_2,\ldots) \tag{19}$$

and

$$\frac{\partial f}{\partial z}=\frac{\partial F}{\partial z_0}+\eta\frac{\partial F}{\partial z_1}+\eta^2\frac{\partial F}{\partial z_2}+\cdots, \tag{20}$$

$$\frac{\partial f}{\partial t}=\frac{\partial F}{\partial t_0}+\eta\frac{\partial F}{\partial t_1}+\eta^2\frac{\partial F}{\partial t_2}+\cdots.$$

For our purposes, the characteristic length scale l is the lattice period and the characteristic time scale is $\tau=2\pi/\omega_0$, where ω_0 is on the order of a typical carrier frequency. These quantities represent the shortest length and fastest time scales in the problem. One can see from Eq. (18) that the z_p,t_p account for field variations over successively longer length and time scales.

For illustrative purposes, consider a trial solution of our nonlinear equations of the form:

$$\mathbf{A}_i=\left\{f_{pi}(k_i;z,t)\boldsymbol{\Psi}_{pi}(k_i)+\sum_{c\neq p}f_{ci}(k_i;z,t)\boldsymbol{\Psi}_{ci}(k_i)\right\}e^{-i\omega_{pi}(k_i)t}+\text{c.c.}, \tag{21}$$

where the p subscript indexes a large principal component with band index p and wave vector k and the c subscript indexes smaller companion components with band index c and wave vector k. If f_{pi} is not varying over too short a distance and the nonlinear effects are not too strong, in a sense to be made more precise below, then we would expect that even the nonlinear Maxwell equations could be approximately satisfied by having $f_{pi}(k_i;z,t)$ acquire a time dependence that involves variations on the order of time scales long compared to $1/\omega_{pi}(k_i)$. Of course, small corrections must be expected to this description, which we see below can be described by adding small amplitudes of other Bloch functions $\boldsymbol{\Psi}_{ci}(k_i)$.

To implement this strategy, we write

$$f_{pi}(k;z,t)=aF_{pi}^{(0)}(k;z_1,z_2,\ldots;t_1,t_2,\ldots), \tag{22}$$

$$f_{ci}(k;z,t)=a\sum_{j=1}\eta^j F_{ci}^{(j)}(k;z_1,z_2,\ldots;t_1,t_2,\ldots.).$$

The quantity ''a'' has been introduced to characterize a typical amplitude of the fields; it is set such that the $F_{qi}^{(i)}(k)$ are dimensionless and of order unity.

To set up Eq. (14) for a multiple scales analysis, we cast it in terms of these newly defined variables z_n,t_n. This can be done quite generally, without specifying whether there is one or more principal component in $\mathbf{A}_i$. We find

$$i\mathbf{n}_i\cdot\left\{\sum_{j=1}^{\infty}\eta^j\frac{\partial}{\partial t_j}\right\}\mathbf{A}_i=\left\{\mathbf{M}_i^{(0)}-\mathbf{V}\sum_{j=1}^{\infty}\eta^j\frac{\partial}{\partial z_j}\right\}\cdot\mathbf{A}_i+\mathbf{B}_i, \tag{23}$$

where

$$\mathbf{V}=\begin{bmatrix}c&0\\0&-c\end{bmatrix}, \tag{24}$$

and

$$\mathbf{M}_i^{(0)}=\begin{bmatrix}-ic\dfrac{\partial}{\partial z_0}&\dfrac{1}{2}\left(\dfrac{\partial[\ln n_i(z_0)]}{\partial z_0}\right)\\-\dfrac{1}{2}\left(\dfrac{\partial[\ln n_i(z_0)]}{\partial z_0}\right)&ic\dfrac{\partial}{\partial z_0}\end{bmatrix}. \tag{25}$$

We would like to solve Eq. (23) in successive powers of η, so we must characterize the nonlinearity in terms of η. To do so, we set a typical component $\chi_{ijkl}^{(3)}(z)$ equal to $\chi_{\mathrm{NL}}\gamma(z)$, where $\gamma(z)$ is of order unity and dimensionless. Then the quantity $\chi_{\mathrm{NL}}a^2$ can be considered to characterize the ''strength'' of the nonlinearity. If the value $\chi_{\mathrm{NL}}a^2$ is of order η^s with $s=1,2,\ldots$, then the intensity index of the nonlin-

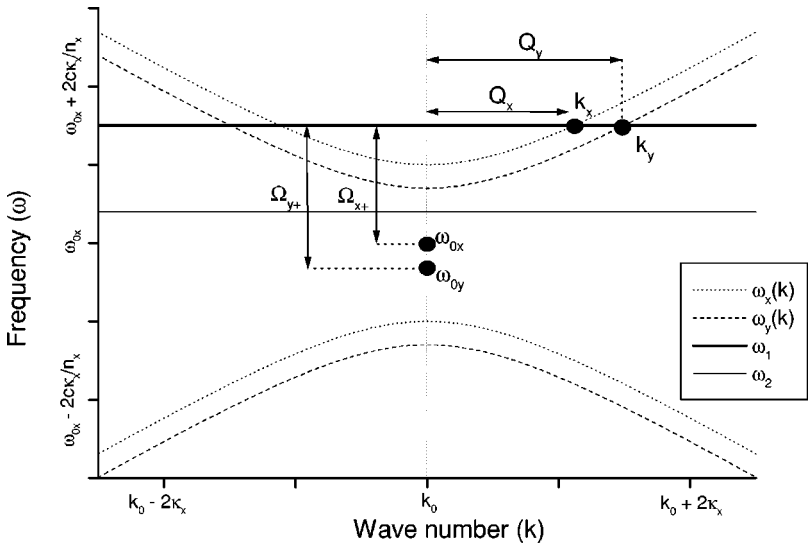

FIG. 1. Dispersion relations in the vicinity of the Bragg wave vector. There are two situations: (a) A carrier frequency ω_1, for which there is one principal component for each polarization. The frequency gives a different wave-vector for each polarization, which accounts for birefringence. (b) A carrier frequency ω_2 in the bandgap; here one would use two principal components. In this case the pulse is carried by the average of the Bragg frequencies shown, which accounts for the birefringence. The mismatch between the incident frequency and the Bragg frequencies can be included in the slowly varying amplitudes. The quantities Q_x, Q_y and Ω_{x+}, Ω_{y+} are detuning parameters used in Sec. VI. The quantity κ_x is the grating strength parameter, defined in Sec. V.

earity is s, and the leading term in $\mathbf{B}_i$ will be of order η^s. Although the solution to Eq. (23) can, in general, be pushed to higher powers of η, it is not reasonable to take the analysis past the intensity index. This is because the nonlinearity has itself been approximated; to include higher orders of η we would have to include higher susceptibilities in Eq. (16).

IV. ONE PRINCIPAL COMPONENT; S=2

For a pulse that is not too short, with carrier wave vector away from the center or edges of the band structure, we seek a description in terms of one principal component for each polarization; a sufficiently long pulse, with a correspondingly narrow frequency content, can be detuned at the band edge, or even slightly within the gap and still be reasonably described by one principal component [17]. The birefringence introduces a wave number (k) mismatch between the two polarizations, each of which is carried at the same frequency (ω), as shown in (a) of Fig. 1. We write our fields as

$$\mathbf{A}_i = \left\{ f_{pi}(k_i;z,t)\boldsymbol{\Psi}_{pi}(k_i) + \sum_{c\neq p} f_{ci}(k_i;z,t)\boldsymbol{\Psi}_{ci}(k_i) \right\} e^{-i\omega_{pi}(k_i)t_0} + \text{c.c.}, \quad (26)$$

where $\omega_{px}(k_x)=\omega_{py}(k_y)$. We stress that although the carrier frequency $\omega_{pi}(k_i)$ is the same for both polarizations, the derivatives will, in general, be unequal.

Equations describing light in periodic, Kerr-nonlinear media are often presented in terms of the electric field or a similar quantity [6]. We here opt to rewrite our $\mathbf{A}$ field in terms of quantities directly comparable to power, because this is the most readily accessible experimental quantity. Using the form of the $\mathbf{A}_i$ fields in the definition of the Poynting vector

$$\mathbf{S}=\mathbf{E}(z,t)\times\mathbf{H}(z,t), \quad (27)$$

we find, using the velocity matrix elements, and group velocity expressions given by Eqs. (11) and (12), that we can express the time and space average of the Poynting vector to $O(\eta^1)$ as

$$(S_{AV})_i = \frac{2n_0}{Z_0 cd}|f_{pi}(k_i;z,t)|^2\omega'_{pi}(k_i). \quad (28)$$

This equation (28) suggests a field definition

$$X(z,t)\equiv\sqrt{\frac{2n_0\omega'_{px}(k_x)\mathcal{A}_{\text{eff}}}{Z_0 cd}}f_{px}(k_x;z,t), \quad (29)$$

$$Y(z,t)\equiv\sqrt{\frac{2n_0\omega'_{py}(k_y)\mathcal{A}_{\text{eff}}}{Z_0 cd}}f_{py}(k_y;z,t),$$

where $\mathcal{A}_{\text{eff}}$ is an effective cross-sectional area in the (x,y) plane associated with the problem. The X and Y fields are defined such that $|X|^2$ is the power in the x-polarized field and $|Y|^2$ is the power in the y-polarized field.

To deal with the nonlinearity, we assume here an intensity index $s=2$, which means that our nonlinearity enters the equations at the same scale as the grating group velocity dispersion. Under this assumption our eigenvalue equation (23) becomes

$$i\mathbf{n}_i\cdot\left\{\frac{\partial}{\partial t_0}+\eta^1\frac{\partial}{\partial t_1}+\eta^2\frac{\partial}{\partial t_2}\right\}\mathbf{A}_i = \left\{\mathbf{M}_i^{(0)}-\mathbf{V}\eta^1\frac{\partial}{\partial z_1}-\mathbf{V}\eta^2\frac{\partial}{\partial z_2}\right\}\cdot\mathbf{A}_i+\mathbf{B}_i, \quad (30)$$

where the nonlinear term $\mathbf{B}_i$ enters at order η^2. Note that to order η^0 Eq. (30) is satisfied because at that order one simply recovers the linear eigenvalue equation (2). To complete the analysis we collect terms first in η^1 and then η^2, which gives us two sets of equations for each polarization. By combining these equations we can extract a set of CNLSE in a manner analogous to that presented by de Sterke *et al.* [5]. We find, to order η^1,

$$i\frac{\partial X}{\partial t_1}=-i\omega'_{px}(k_x)\frac{\partial X}{\partial z_1}, \quad (31)$$

and similar for Y. From the η^2 order equations we find

TABLE I. Coefficients for the CNLSE. The y values are determined by interchanging $x \leftrightarrow y$.

α^x_{SPM}	$\frac{3}{4}\frac{Z_0 c d \omega_{px}}{\omega'_{px} A_{\mathrm{eff}}} \int_0^d dz_0 \chi_{xxxx}(z) \lvert u_{px}(k_x;z_0)\rvert^4$
α^x_{CPM}	$\frac{3}{4}\frac{Z_0 c d \omega_{py}}{\omega'_{py} A_{\mathrm{eff}}} \int_0^d dz_0 \chi_{xxyy}(z) \lvert u_{px}(k_x;z_0)\rvert^2 \lvert u_{py}(k_y;z_0)\rvert^2$
α^x_{PC}	$\frac{3}{4}\frac{Z_0 c d \omega_{py}}{\omega'_{py} A_{\mathrm{eff}}} \int_0^d dz_0 \chi_{xxyy}(z) u^2_{py}(k_y;z_0) u^{*2}_{px}(k_x;z_0)$

$$i\frac{\partial X}{\partial t_2} = -i\omega'_{px}(k_x)\frac{\partial X}{\partial z_2} - \frac{1}{2}\omega''_{px}(k_x)\frac{\partial^2 X}{\partial z_1^2} + \left(\sqrt{\frac{2n_0\omega'_{px}\mathcal{A}_{\mathrm{eff}}}{Z_0 c d}}\right)\frac{1}{\eta^2 N}\int_0^L dz_0 \boldsymbol{\Psi}^\dagger_{pk_x}\cdot\mathbf{B}_x e^{i\omega_{px}(k_x)t_0}. \tag{32}$$

The quantity $\mathbf{B}_i$ is defined in Eq. (15), but we only need to write the electric-field contributions to $\mathbf{B}_i$ to order η^0 to keep Eq. (32) self-consistent; recall that the nonlinear susceptibility is of order η^2, so the last term in Eq. (32) will be of order η^0. The form of the nonlinear susceptibility has been chosen as that of an isotropic medium, but in principle any $\chi^{(3)}$ tensor could be used. We note, though, that the birefringence is considered small because of a limitation imposed by our method, discussed after Eq. (33). Thus, since the effect of the nonlinearity itself is already considered small, the deviations in $\chi^{(3)}$ due to lack of isotropy will typically be of the next lowest order in η, and hence can be ignored. The overlap integral in Eq. (32) is evaluated as

$$\left(\sqrt{\frac{2n_0\omega'_{px}\mathcal{A}_{\mathrm{eff}}}{Z_0 c d}}\right)\int_0^d \boldsymbol{\Psi}^\dagger_{mx}\cdot\mathbf{B}_x e^{i\omega_{px}t_0}dz_0 = -\alpha^x_{\mathrm{SPM}}\lvert X\rvert^2 X - \alpha^x_{\mathrm{CPM}}\lvert Y\rvert^2 X - \alpha^x_{\mathrm{PC}} X^* Y^2 e^{2i(k_y-k_x)z_0}, \tag{33}$$

where we note that the quantity $e^{2i(k_y-k_x)z_0}$ has not been integrated because we assume that $(k_y-k_x)=\eta K$, where K is of the order of the average wave number $(k_x+k_y)/2$. In this case $2i(k_y-k_x)z_0 = 2i\eta K z_0 = 2iKz_1$. Since z_1 and z_0 are considered to be independent variables the quantity $e^{2i(k_y-k_x)z_0}$ remains. The value of the coefficients α are given in Table I.

We now relate our scaled derivatives to full time and space derivatives. Assembling Eqs. (20), (31), (32), (33), and noting that the equations for the Y fields can be derived by interchanging $x \leftrightarrow y$ in the preceding, we obtain the following coupled nonlinear Schrödinger equations:

$$0 = i\frac{\partial X}{\partial t} + i\omega'_{px}(k_x)\frac{\partial X}{\partial z} + \frac{1}{2}\omega''_{px}(k_x)\frac{\partial^2 X}{\partial z^2} + \{\alpha^x_{\mathrm{SPM}}\lvert X\rvert^2 + \alpha^x_{\mathrm{CPM}}\lvert Y\rvert^2\}X + \alpha^x_{\mathrm{PC}} Y^2 X^* e^{i\lambda z}, \tag{34}$$

$$0 = i\frac{\partial Y}{\partial t} + i\omega'_{py}(k_y)\frac{\partial Y}{\partial z} + \frac{1}{2}\omega''_{py}(k_y)\frac{\partial^2 Y}{\partial z^2} + \{\alpha^y_{\mathrm{SPM}}\lvert Y\rvert^2 + \alpha^y_{\mathrm{CPM}}\lvert X\rvert^2\}Y + \alpha^y_{\mathrm{PC}} X^2 Y^* e^{-i\lambda z}.$$

The quantity

$$\lambda = 2(k_y - k_x) \tag{35}$$

characterizes the birefringence in the system. The coefficients α are so subscripted because α_{SPM} accounts for self-phase-modulation; α_{CPM} accounts for cross-phase modulation; and α_{PC} accounts for phase conjugation. We note that equations similar to (34) have been studied extensively in the literature [13,18–21].

V. TWO PRINCIPAL COMPONENTS; S=1

We now turn to describing pulses whose carrier frequencies are in or very close to a photonic bandgap, either at the band center or the band edge. In (b) of Fig. 1, we show the case where the frequency of the pulse is within the photonic bandgap. The pulse can, however, be detuned outside the bandgap and still be well described by the theory presented here (see Sec. VI). As discussed above, this situation often requires the use of two principal components in the description of our fields. We set the reference wave number to be the same for the x and y polarizations; the frequency mismatch between the Bragg frequencies of the two polarizations accounts for the birefringence. We find that a derivation of the coupled mode equations only requires us to carry our results through to order η^1, so we simply write our $\mathbf{A}$ fields as

$$\mathbf{A}_i(z,t) = \{f_{ui}(z_1,z_2,\ldots;t_1,t_2,\ldots)\boldsymbol{\Psi}_{ui}(k_0) + f_{li}(z_1,z_2,\ldots;t_1,t_2,\ldots)\boldsymbol{\Psi}_{li}(k_0)\}e^{-i\bar{\omega}t_0} + \mathrm{c.c.} + O(\eta^2), \tag{36}$$

where k_0 is the wave vector at the band edge (assumed in Fig. 1) or band center. The quantities f_{ui} and f_{li} modulate Bloch functions associated with the upper and lower band of the given polarization i, respectively; both are principal components in the sense defined above. The carrier frequency, common to both polarizations, $\bar{\omega} = \frac{1}{2}(\omega_{0x}+\omega_{0y})$, is the average of the Bragg frequencies of the two polarizations, $\omega_{0i} = \frac{1}{2}(\omega_{ui}+\omega_{li})$, where ω_{ui} is the frequency associated with the lowest point of the upper band and ω_{li} is the frequency associated with the highest point of the lower band on the dispersion relation (Fig. 1).

By using our expression for $\mathbf{A}_i$ (36) in the matrix equation (23) we find, to order η^1,

$$i\frac{\partial f_{ui}}{\partial t_1} = \sigma_{ui} f_{ui} - i v_{ul(i)}\frac{\partial f_{li}}{\partial z_1} + \frac{1}{\eta}\int_0^d \boldsymbol{\Psi}^\dagger_{ui}\cdot\mathbf{B}_i e^{iw_i t_0}, \tag{37}$$

$$i\frac{\partial f_{li}}{\partial t_1} = \sigma_{li} f_{li} - i v_{lu(i)}\frac{\partial f_{ui}}{\partial z_1} + \frac{1}{\eta}\int_0^d \boldsymbol{\Psi}^\dagger_{ui}\cdot\mathbf{B}_i e^{iw_i t_0},$$

where we have used the definitions $\eta\sigma_{l/ui} \equiv \omega_{l/ui} - \bar{\omega}$, where σ_{li} and σ_{ui} are of order ω_{0i}; this is equivalent to assuming

that the bandgap is small relative to the carrier frequency, i.e., $(\omega_{ui}-\omega_{li})/\omega_{0i}\ll 1$. Note that we can satisfy this condition and still have a strong grating in the sense we have discussed here. Since we are only carrying the calculation to order η, Eq. (37) become

$$i\frac{\partial f_{ui}}{\partial t}-v_{gi}\frac{\partial f_{li}}{\partial z}-(\omega_{ui}-\overline{\omega})f_{ui}+\theta_{ui}(z,t;f_{ux},f_{uy},f_{lx},f_{ly})=0, \tag{38}$$

$$i\frac{\partial f_{li}}{\partial t}+v_{gi}\frac{\partial f_{ui}}{\partial z}-(\omega_{li}-\overline{\omega})f_{li}+\theta_{li}(z,t;f_{ux},f_{uy},f_{lx},f_{ly})=0,$$

where

$$v_{gi}\equiv iv_{lu(i)}=-iv_{ul(i)}, \tag{39}$$

and the function $\theta(z,t;f_{ux},f_{uy},f_{lx},f_{ly})$ represents the complicated overlap integrals in Eq. (37).

We now introduce [22]

$$G_{x\pm}=(f_{lx}\mp if_{ux})/\sqrt{2}, \tag{40}$$

$$G_{y\pm}=(f_{ly}\mp if_{uy})/\sqrt{2}.$$

From the definition of the Poynting vector (27) and using Maxwell's equations we find, to $O(\eta^0)$

$$S^i_{AV}=\frac{2v_{gi}n_0}{Z_0cd}\{|G_{i\pm}|^2-|G_{i\mp}|^2\}. \tag{41}$$

This expression suggests a definition

$$X_\pm\equiv\sqrt{\frac{2v_{gx}n_0\mathcal{A}_{\rm eff}}{Z_0cd}}G_{x\pm}e^{+i\delta t/4}, \tag{42}$$

$$Y_\pm\equiv\sqrt{\frac{2v_{gy}n_0\mathcal{A}_{\rm eff}}{Z_0cd}}G_{y\pm}e^{-i\delta t/4},$$

where $\mathcal{A}_{\rm eff}$ has been defined following Eq. (29), and where the exponential factor $e^{\pm i\delta t/4}$ has been included in anticipation of the form of the final equations, with

$$\delta=2(\omega_{0x}-\omega_{0y}). \tag{43}$$

This is equivalent to using the Bragg frequencies ω_{0i} to carry the $X_\pm, Y_\pm$ fields. These new fields are travelling waves normalized such that the quantities $(|X_+|^2-|X_-|^2),(|Y_+|^2-|Y_-|^2)$ represent the power in each polarization. Using the definitions (40) and (42) in Eq. (38), and evaluating the overlap integrals, we can write our full coupled mode equations

$$\begin{aligned}0=&\frac{i}{v_{gx}}\frac{\partial X_\pm}{\partial t}\pm i\frac{\partial X_\pm}{\partial z}+\kappa_xX_\mp+\alpha^x_0\{|X_\pm|^2+2|X_\mp|^2\}X_\pm+\alpha^x_1\{|X_\pm|^2+|X_\mp|^2\}X_\mp+\alpha^x_1\{X_\pm X^*_\mp+X_\mp X^*_\pm\}X_\pm+\alpha^x_2X^2_\pm X^*_\mp\\&+\beta^x_0\{|Y_\pm|^2+|Y_\mp|^2\}X_\pm+\beta^x_1\{|Y_\pm|^2+|Y_\mp|^2\}X_\mp+\beta^x_2\{Y^*_\pm Y_\mp+Y^*_\mp Y_\pm\}X_\pm+\beta^x_4X_\mp Y^*_\pm Y_\mp+\beta^x_5X_\mp Y^*_\mp Y_\pm\\&+[(\gamma^x_0Y^2_\pm+2\gamma^x_2Y_\pm Y_\mp+\gamma^x_4Y^2_\mp)X^*_\pm+(\gamma^x_1\{Y^2_\pm+Y^2_\mp\}+2\gamma^x_3Y_\pm Y_\mp)X^*_\mp]e^{i\delta t}.\end{aligned} \tag{44}$$

The appropriate equations for the $Y_\pm$ can be found by interchanging $X\leftrightarrow Y$ in Eq. (44) and changing $\delta\rightarrow-\delta$. In these equations the value

$$\kappa_i=\frac{\omega_{ui}-\omega_{li}}{2v_{gi}} \tag{45}$$

accounts for the grating strength, and δ accounts for the strength of the intrinsic birefringence.

The coupling coefficients $\{\alpha,\beta,\gamma\}$ have a rather involved definition. We start by defining

$$\alpha^x_{pqrs}=\frac{3}{16}\frac{\overline{\omega}Z_0cd}{v^2_{gx}A_{\rm eff}}\int_0^d\chi_{xxxx}(z)\phi^*_{px}(z)\phi_{qx}(z)\phi^*_{rx}(z)\phi_{sx}(z)dz, \tag{46}$$

$$\beta^x_{pqrs}=\frac{1}{8}\frac{\overline{\omega}Z_0cd}{v_{gy}v_{gx}A_{\rm eff}}\int_0^d(3\chi_{xxyy}(z))\times\phi^*_{px}(z)\phi_{qx}(z)\phi^*_{ry}(z)\phi_{sy}(z)dz,$$

$$\gamma^x_{pqrs}=\frac{1}{16}\frac{\overline{\omega}Z_0cd}{v_{gy}v_{gx}A_{\rm eff}}\int_0^d(3\chi_{xxyy}(z))\times\phi^*_{px}(z)\phi_{qy}(z)\phi^*_{rx}(z)\phi_{sy}(z)dz.$$

The indices p,q,r,s can take on the values l,u, that is, they index the upper and lower bands. Notice that in the definitions of β^x_{pqrs} and γ^x_{pqrs} the values of the Bloch functions in the integral alternate between x and y. The coefficients of the $X_\pm$ coupled mode equations (44) are shown in Table II. The y values of the coefficients can be found by switching $x\leftrightarrow y$ in Eq. (46), and in Table II.

Weak grating limit of the CME. Many fiber gratings have small index contrasts, which allows us to simplify the coupled mode equations (44) by considering a weak grating of the form

$$n_i(z)=\overline{n_i}+\delta n_i\cos(2k_0z), \tag{47}$$

where $\overline{n_i}$ is the background index, δn_i is the index modulation with $\delta n_i\ll\overline{n_i}$, and k_0 is the wave number that defines

TABLE II. x coefficients for the CME. The y values are determined by interchanging $x \leftrightarrow y$.

Coeff.	Value	Weak grating
α_0^x	$(\alpha^x_{uuuu}+2\alpha^x_{uull}+\alpha^x_{llll})$	$\frac{3}{4}\frac{\bar{\omega}\bar{\chi}^{(3)}Z_0}{\bar{n}_x^2 c\mathcal{A}_{\rm eff}}$
α_1^x	$(-\alpha^x_{uuuu}+\alpha^x_{llll})$	0
α_2^x	$(\alpha^x_{uuuu}-6\alpha^x_{uull}+\alpha^x_{llll})$	0
β_0^x	$(\beta^x_{uuuu}+\beta^x_{uull}+\beta^x_{lluu}+\beta^x_{llll})$	$\frac{1}{2}\frac{\bar{\omega}\bar{\chi}^{(3)}Z_0}{\bar{n}_x\bar{n}_y c\mathcal{A}_{\rm eff}}$
β_1^x	$(-\beta^x_{uuuu}-\beta^x_{uull}+\beta^x_{lluu}+\beta^x_{llll})$	0
β_2^x	$(-\beta^x_{uuuu}+\beta^x_{uull}-\beta^x_{lluu}+\beta^x_{llll})$	$\frac{1}{2}\frac{\bar{\omega}\bar{\chi}^{(3)}Z_0}{\bar{n}_x\bar{n}_y c\mathcal{A}_{\rm eff}}$
β_3^x	$(\beta^x_{uuuu}-\beta^x_{uull}-\beta^x_{lluu}+\beta^x_{llll})$	0
β_4^x	$(\beta^x_3-4\beta^x_{lulu})$	0
β_5^x	$(\beta^x_3+4\beta^x_{lulu})$	0
γ_0^x	$(\gamma^x_{uuuu}-\gamma^x_{ulul}-\gamma^x_{lulu}+\gamma^x_{llll}+4\gamma^x_{lulu})$	$\frac{1}{4}\frac{\bar{\omega}\bar{\chi}^{(3)}Z_0}{\bar{n}_x\bar{n}_y c\mathcal{A}_{\rm eff}}$
γ_1^x	$(-\gamma^x_{uuuu}+\gamma^x_{ulul}-\gamma^x_{lulu}+\gamma^x_{llll})$	0
γ_2^x	$(-\gamma^x_{uuuu}-\gamma^x_{ulul}+\gamma^x_{lulu}+\gamma^x_{llll})$	0
γ_3^x	$(\gamma^x_{uuuu}+\gamma^x_{ulul}+\gamma^x_{lulu}+\gamma^x_{llll})$	$\frac{1}{4}\frac{\bar{\omega}\bar{\chi}^{(3)}Z_0}{\bar{n}_x\bar{n}_y c\mathcal{A}_{\rm eff}}$
γ_4^x	$(\gamma^x_{uuuu}-\gamma^x_{ulul}-\gamma^x_{lulu}+\gamma^x_{llll}-4\gamma^x_{lulu})$	0

the band edge. In the presence of a weak grating, the Bloch functions at the band edge can be evaluated, and normalized via Eq. (8),

$$\phi_{ui}(k_0;z)=i\sqrt{\frac{2}{\bar{n}_i^2 d}}\sin(k_0 z), \tag{48}$$

$$\phi_{li}(k_0;z)=\sqrt{\frac{2}{\bar{n}_i^2 d}}\cos(k_0 z).$$

If we use these forms for the Bloch functions and assume a uniform nonlinearity, then many of the coefficients in the coupled mode equations (44) are identically zero. We confirm, using Eq. (11), that in this limit the quantity v_{gi} is simply equal to the group velocity in the absence of the grating, $v_{gi}=c/\bar{n}_i$. With this in mind we rewrite Eq. (44) as

$$\begin{aligned}0=&\,i\frac{\bar{n}_x}{c}\frac{\partial X_\pm}{\partial t}\pm i\frac{\partial X_\pm}{\partial z}+\kappa_x X_\mp+\alpha_x\{|X_\pm|^2+2|X_\mp|^2\}X_\pm\\&+\beta_x\{|Y_\pm|^2+|Y_\mp|^2\}X_\pm+\beta_x X_\mp Y_\mp^* Y_\pm\\&+\gamma_x\{X_\pm^* Y_\pm^2+2X_\mp^* Y_\pm Y_\mp\}e^{i\delta t},\end{aligned} \tag{49}$$

with

$$\alpha_x=\frac{3}{4}\frac{\bar{\omega}\bar{\chi}^{(3)}Z_0}{\bar{n}_x^2 c A_{\rm eff}}, \tag{50}$$

$$\beta_x=\frac{1}{2}\frac{\bar{\omega}\bar{\chi}^{(3)}Z_0}{\bar{n}_x\bar{n}_y c A_{\rm eff}},$$

$$\gamma_x=\frac{1}{4}\frac{\bar{\omega}\bar{\chi}^{(3)}Z_0}{\bar{n}_x\bar{n}_y c A_{\rm eff}}.$$

The grating coefficient is

$$\kappa_i=\frac{1}{2}\frac{\delta n_i}{\bar{n}_i}\frac{\pi}{d}. \tag{51}$$

Again, the $Y_\pm$ equations can be found by switching $x\leftrightarrow y$ and $\delta\rightarrow-\delta$ in Eqs. (49) and (50), from which we note that $\beta_x=\beta_y$ and $\gamma_x=\gamma_y$.

For a very weak birefringence, where $\bar{n}_x\approx\bar{n}_y$, the coefficients in Eq. (50) are in the ratio $\{\alpha:\beta:\gamma\}=\{3:2:1\}$. In the stationary limit these equations agree with those given by Samir *et al.* [23].

VI. CONNECTING THE CNLSE AND THE CME

In the previous sections we derived two types of equations: a set of coupled nonlinear Schrödinger equations, typically valid outside the bandgap, and a set of nonlinear coupled mode equations, typically valid within or near the bandgap. As we will see in this section, the coupled mode equations make very definite predictions about the dispersion relation and the Bloch functions of the periodic system. When these predicted Bloch functions and dispersion relation deviate from the true values of the system, then the approximations that have been used to derive the coupled mode equations have broken down; this allows us to determine the limits of validity of the equations. On the other hand, the nonlinear Schrödinger equation relies on the local properties of the dispersion relation, so if the nonlinearity is sufficiently small it should always be valid as long as one is sufficiently far away from a bandgap, or other portion of the dispersion relation with significant higher-order curvature. If the frequency content of a pulse is very narrow, then higher order dispersion will have little effect, so the Schrödinger equation should be valid at the band edge and even slightly inside the band gap.

A further point to be discussed is how the solutions to the nonlinear Schrödinger equation relate to those of the coupled mode equations. Understanding this allows us to identify the range where either approach could be used, an important goal because although the coupled mode equations are easily solvable via numerical techniques, they are difficult to solve analytically. As mentioned, there is a great deal of work in the literature on equations similar to our CNLSE [13,18–21], so if we understand how solutions of the CNLSE are related to solutions of the CME, then the CNLSE literature becomes available to aid in the investigation of birefringence phenomena near the gap. Specifically we want to know how to relate the two CNLSE fields X and Y, to the four CME fields $X_\pm$ and $Y_\pm$, and we want to get a sense of how close to the gap we must be before the CNLSE cease to effectively describe the problem. Our method is to start with the weak grating nonlinear coupled mode equations and perform a further multiple scales analy-

sis to derive the nonlinear Schrödinger equations. The use of the weak grating equations simplifies the mathematics, and does not significantly affect the final results, for reasons discussed below. The method involved follows closely the analysis of de Sterke and Sipe [6], except that in the present case the nonlinear terms are much more involved, so we only sketch the results.

We define

$$\mathbf{F}_x=\begin{bmatrix}X_+(z,t)\\ X_-(z,t)\end{bmatrix},\quad \mathbf{F}_y=\begin{bmatrix}Y_+(z,t)\\ Y_-(z,t)\end{bmatrix}, \tag{52}$$

with which the linear portion of the coupled mode equations can be written as

$$\left[i\sigma^3\frac{\partial}{\partial z}+i\frac{\overline{n_i}}{c}\sigma^0\frac{\partial}{\partial t}+\sigma^1\kappa_i\right]\mathbf{F}_i=0, \tag{53}$$

where we have used the Pauli spin matrices

$$\sigma^1=\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad \sigma^3=\begin{pmatrix}1&0\\0&-1\end{pmatrix}, \tag{54}$$

and the unit matrix σ^0. We seek solutions of Eq. (53) of the form $\mathbf{F}_i=\mathbf{f}_i e^{-i(\Omega_{i\pm}t-Q_iz)}$, where the wave vector detuning is $Q_i=k_i-k_0$. If the full frequency $\omega_i>\omega_{0i}$, then we call the detuning parameter Ω_{i+} and otherwise we call it Ω_{i-}, with $\Omega_{i\pm}=\omega_i-\omega_{0i}$. The $\Omega_\pm$ are associated with the upper and lower band via the dispersion relation

$$\frac{\overline{n_i}}{c}\Omega_{i\pm}(Q)=\pm\sqrt{\kappa_i^2+Q^2}, \tag{55}$$

which follows from substituting the $\mathbf{F}_i$ in Eq. (53). From the dispersion relation, the group velocity, and group velocity dispersion are

$$\Omega'_{i\pm}(Q)\equiv\frac{d\Omega_{i\pm}}{dQ}=\left(\frac{c}{\overline{n_i}}\right)^2\frac{Q}{\Omega_{i\pm}},$$

$$\Omega''_{i\pm}(Q)\equiv\frac{d^2\Omega_{i\pm}}{dQ^2}=\left(\frac{c}{\overline{n_i}}\right)^2\frac{1}{\Omega_{i\pm}}[1-\rho_i^2(Q)], \tag{56}$$

where $\rho_i(Q)=\overline{n_i}\Omega'_{i+}(Q)/c$ is the ratio of the group velocity at a given wave vector for a point in the *upper* band, relative to the group velocity in the absence of the grating. The eigenvectors have the form

$$\mathbf{f}_i^{(+)}(Q)=\begin{bmatrix}f_{i+}^{(+)}\\ f_{i-}^{(+)}\end{bmatrix}=\frac{1}{\sqrt{2}}\begin{bmatrix}\sqrt{1+\rho_i(Q)}\\ -\sqrt{1-\rho_i(Q)}\end{bmatrix}, \tag{57}$$

$$\mathbf{f}_i^{(-)}(Q)=\begin{bmatrix}f_{i+}^{(-)}\\ f_{i-}^{(-)}\end{bmatrix}=\frac{1}{\sqrt{2}}\begin{bmatrix}\sqrt{1-\rho_i(Q)}\\ \sqrt{1+\rho_i(Q)}\end{bmatrix},$$

where the $f_i^{(\pm)}(Q)$ are associated with the $\Omega_{i\pm}$, respectively.

From these eigenvectors one can extract the Bloch functions of the periodic structure, in the coupled mode equations limit. Comparing to the form of the Bloch functions (7), we find

$$\phi_{(\pm)i}(k;z)=\frac{1}{\sqrt{2d\overline{n_i^2}}}[\sqrt{1\pm\rho_i}\mp\sqrt{1\mp\rho_i}e^{-2ik_0z}]e^{ikz}, \tag{58}$$

where the factor $1/\sqrt{2d\overline{n_i^2}}$ has been included for proper normalization via Eq. (8). The function multiplying e^{ikz} can be identified as $u_{(\pm)i}(k;z)$.

If we include the nonlinearity, then we can write the coupled mode equations as

$$\left[i\sigma^3\frac{\partial}{\partial z}+i\frac{\overline{n_i}}{c}\sigma^0\frac{\partial}{\partial t}+\sigma^1\kappa_i\right]\mathbf{F}_i+\mathbf{N}_i=\mathbf{0}, \tag{59}$$

where $\mathbf{N}_i$ is the nonlinear term that follows immediately from Eq. (49). For simplicity we concentrate on detuning into the upper band $\Omega_{i+}(Q_i)$. We represent our field vector $\mathbf{F}_i$ as being *mostly* in the upper band, but with a small component in the lower band. We start by writing the field vectors as

$$\mathbf{F}_i=\frac{1}{\sqrt{\rho_i}}[\eta a_i(z_n;t_n)\mathbf{f}_i^{(+)}(Q_i)+\eta^2 b_{i2}(z_n;t_n)\mathbf{f}_i^{(-)}(Q_i)+\eta^3 b_{i3}(z_n;t_n)\mathbf{f}_i^{(-)}(Q_i)+\cdots]e^{-i\Omega_{i+}(Q_i)t_0}e^{iQ_iz_0}, \tag{60}$$

where we have introduced the multiple scales variables z_n, t_n as in Eq. (18). The upper-band component a_i dominates the expansion of $\mathbf{F}_i$, and hence plays the role of a principal component; the b_i terms are companion components. The numerical value of the detunings, $\Omega_{+i}(Q_i)$ and Q_i, will be different for each polarization, but in each case we are detuning to the same frequency ω, as shown in (a) of Fig. 1. The normalization factor $1/\sqrt{\rho_i}$ has been introduced so that the envelope functions a_i will be directly related to power. Since the nonlinearity involves only cubic-type terms or higher, we can write

$$\mathbf{N}_x=\eta^3\mathbf{N}_{x3}+\cdots. \tag{61}$$

To evaluate $\mathbf{N}_{x3}$ we combine Eqs. (52), (57), (60), from which it is apparent that to lowest order in η

$$X_\pm=\pm\frac{\eta}{\sqrt{2}}\left(\sqrt{\frac{1\pm\rho_x}{\rho_x}}\right)a_x(z_n;t_n)e^{-i\Omega_{x+}(Q_x)t}e^{+iQ_xz}, \tag{62}$$

$$Y_\pm=\pm\frac{\eta}{\sqrt{2}}\left(\sqrt{\frac{1\pm\rho_y}{\rho_y}}\right)a_y(z_n;t_n)e^{-i\Omega_{y+}(Q_y)t}e^{+iQ_yz}.$$

Then, using Eq. (62) in Eq. (61), we find

$$\mathbf{N}_x e^{+i\Omega_{x+}(Q_x)t_0} e^{-iQ_x z} = \frac{\alpha_x |a_x|^2 a_x}{2\rho_x\sqrt{\rho_x}}(3\sigma^0 - \rho_x\sigma^3)\mathbf{f}_x^{(+)} + \frac{\beta_x |a_y|^2 a_x}{\rho_y\sqrt{\rho_x}}\left(\sigma^0 - \frac{\sqrt{1-\rho_y^2}}{2}\sigma^1\right)\mathbf{f}_x^{(+)} + \frac{\gamma_x}{\rho_y\sqrt{\rho_x}} a_y^2 a_x^* e^{+i\lambda z}\left(\frac{1}{2}(\sigma^0 + \rho_y\sigma^3) - \sqrt{1-\rho_y^2}\sigma^1\right)\mathbf{f}_x^{(+)}, \tag{63}$$

where λ is the birefringence parameter quoted earlier (35). Note that to order η^1 the forward and backward going fields $X_\pm$ are associated with the multiple scales envelope function a_x. This means that, were we to use the strong grating equations, the form of $\mathbf{N}_{3i}$ would be the same, but the values of the coefficients would change. However, since the values of the weak grating Bloch functions are known, it is straightforward to compare the nonlinear Schrödinger equation derived from the weak grating CME, to the weak grating CNLSE.

Using this nonlinear operator in Eq. (6.17) of de Sterke and Sipe [6] allows us to write down the CNLSE:

$$0 = i\frac{\partial a_x}{\partial t} + i\Omega'_{x+}\frac{\partial a_x}{\partial z} + \frac{1}{2}\Omega''_{x+}\frac{\partial^2 a_x}{\partial z^2} + \alpha^x_{\mathrm{SPM}}|a_x|^2 a_x + \alpha^x_{\mathrm{CPM}}|a_y|^2 a_x + \alpha^x_{pc} a_y^2 a_x^* e^{i\lambda z}, \tag{64}$$

$$0 = i\frac{\partial a_y}{\partial t} + i\Omega'_{y+}\frac{\partial a_y}{\partial z} + \frac{1}{2}\Omega''_{y+}\frac{\partial^2 a_y}{\partial z^2} + \alpha^y_{\mathrm{SPM}}|a_y|^2 a_y + \alpha^y_{\mathrm{CPM}}|a_x|^2 a_y + \alpha^y_{pc} a_x^2 a_y^* e^{-i\lambda z},$$

where Ω'_{i+}, and Ω''_{i+} are the group velocity and group velocity dispersion at the given detuning (56), and the nonlinear coefficients are

$$\alpha^x_{\mathrm{SPM}} = \frac{c}{\bar{n}_x}\alpha_x\frac{[3-\rho_x^2]}{2\rho_x}, \tag{65}$$

$$\alpha^x_{\mathrm{CPM}} = \frac{c}{\bar{n}_x}\beta_x\left\{\frac{2+\sqrt{(1-\rho_x^2)(1-\rho_y^2)}}{2\rho_y}\right\},$$

$$\alpha^x_{\mathrm{PC}} = \frac{c}{\bar{n}_x}\gamma_x\left\{\frac{(1+\rho_x\rho_y)+2\sqrt{(1-\rho_x^2)(1-\rho_y^2)}}{2\rho_y}\right\}.$$

The coefficients (65) lead to the concept of an *effective nonlinearity* because their values are dependent on Q, the detuning from the Bragg wave vector.

To connect Eq. (64) to the CNLSE given by Eq. (34), we recall that both the $(X_\pm, Y_\pm)$ fields used by the CME, and the (X,Y) fields used by the CNLSE are normalized such that their square moduli represent power. If we wish to connect the CNLSE and CME fields we require that

$$|X|^2 = |X_+|^2 - |X_-|^2 = |a_x|^2 \tag{66}$$

and similar for Y. We have used Eq. (62) for a_x. Hence, our fields X,Y and a_x, a_y are equivalent. Using the Bloch functions (58) we can show that the coefficients given above Eq. (65) agree with those in Table I.

TABLE III. Parameters used in numerical simulations.

Index of refraction ($\bar{n}_x$)	1.50
Index modulation (δn)	1.67×10^{-4}
Birefringence ($\bar{n}_y - \bar{n}_x$)	2×10^{-6}
Nonlinear index (n_2; W/cm^2)	2.3×10^{-16}
Bragg wavelength (nm)	1052.00

VII. NUMERICAL SIMULATIONS

The simulations are intended to illustrate two points. First, we demonstrate the validity of the CNLSE approximation with respect to the CME approximation, as discussed in Sec. VI. Second, we investigate the effect of energy exchange between the two polarizations, which may be of importance in the development of new devices. For the sample calculations, we used paramaters of a typical optical fiber, given in Table III.

A. Comparing the CNLSE and CME

To compare the CNLSE and the CME equations, we consider a pulse propagating through a grating with parameters given in Table III, using each set of equations. We solve the CNLSE by a split-step Fourier technique: At each time step the linear portion of the equations are solved in the Fourier domain, while the nonlinear portions are solved using a fourth order Runge-Kutta integration scheme [19]; we solve Eqs. (64) in a frame travelling with the average velocity of the two pulses. The CME are solved using a collocation algorithm [24].

To define a frequency control parameter, we first define a total band-gap width

$$\delta\bar{\omega} = (\omega_{0x} - \omega_{0y}) + \frac{\delta n}{\bar{n}_x}\bar{\omega},$$

where it has been assumed that $\bar{n}_x < \bar{n}_y$, so that $(\omega_{0x} - \omega_{0y}) > 0$. In terms of these we define the frequency detuning

$$\Delta = \frac{\omega - \bar{\omega}}{\delta\bar{\omega}},$$

where ω is the carrier frequency of the pulse.

We start with simulations using the values of Δ given in Table III. The initial intensity was 1.10 G W/cm^2 in each polarization, the initial pulse was a Gaussian with a full width at half maximum (FWHM) pulse width 200 ps, and was chosen such that the initial frequency content to the pulse did not extend into the gap. Table IV compares the velocities observed by the CME and the CNLSE for the x polarization after 3000 ps of simulation time, from which it can be seen that both algorithms predict the same velocity even very close to the band gap.

TABLE IV. Comparison of velocities between the CME and CNLSE algorithms.

Δ	CME	CNLSE
1.05	0.306	0.313
1.10	0.420	0.425
1.20	0.555	0.555
1.30	0.625	0.625
1.40	0.700	0.700
1.50	0.740	0.740

Figure 2 compares the pulse shapes of the *x* polarization, after 2000 ps of simulation time for $\Delta=1.20$ and $\Delta=1.10$. It can be seen that although the two algorithms agree very closely for $\Delta=1.20$, at $\Delta=1.10$ the differences are more marked. Figure 3 compares the total energy in the *y* polarized pulse as a function of time for $\Delta=1.20,1.10$ using both numerical techniques. The data at $\Delta=1.20$ is much more consistent than at $\Delta=1.10$. The initial ringing for the CME data is a consequence of our initial conditions [24], but it damps out quite quickly.

There are two reasons why the derivation in Sec. VI would fail. First, the CME includes all orders of dispersion, while the CNLSE includes only second order dispersion. Second, the CNLSE derivation assumes that there is but slight build up of reflected waves which, as one nears the band gap, is decreasingly accurate. To quantify the effect of the first objection, we calculate the quantity

$$\Omega'''_{+i}=\frac{d\Omega''_{+i}}{dQ}=-3\frac{\Omega'_{i+}\Omega''_{i+}}{\Omega_{i+}}, \tag{67}$$

in terms of which the expressions for second- and third-order dispersion lengths, assuming a Gaussian pulse, are [25]

$$L_{D2}=\frac{T^2_{\mathrm{FWHM}}\Omega'^3_{i+}}{2.772|\Omega''_{i+}|}, \quad L_{D3}=\frac{T^3_{\mathrm{FWHM}}\Omega'^4_{i+}}{4.615|\Omega'''_{i+}|}, \tag{68}$$

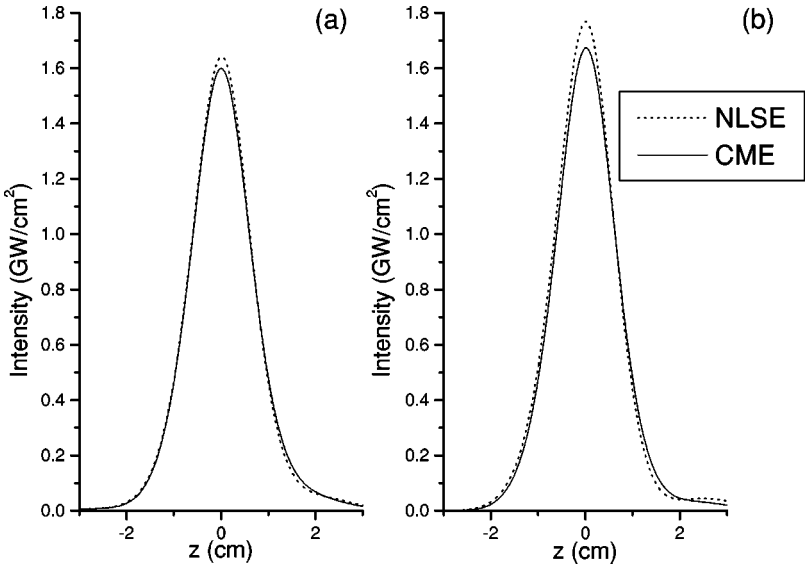

FIG. 2. Comparison of the CME and NLSE y polarization pulse profiles after 2000 ps of simulation time for (a) $\Delta=1.20$ and (b) $\Delta=1.10$. It is evident in (b) that the two algorithms are giving different results.

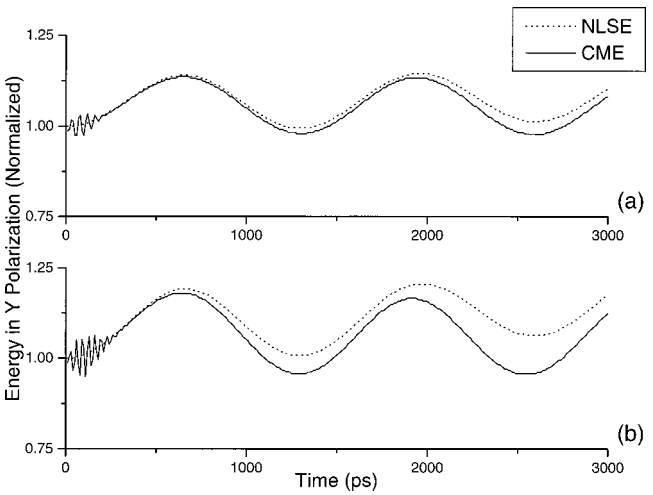

FIG. 3. Comparison of the energy in the y polarization as predicted by the CME and NLSE for (a) $\Delta=1.20$ and (b) $\Delta=1.10$. The plots are normalized such that at t=0 the energy is 1.0. Although neither detuning gives an exact agreement, the divergence for $\Delta=1.10$ is significantly greater.

where T_{FWHM} is the pulse width; if $L_{D3}\approx L_{D2}$, then third-order dispersion effects become important. We thus have a criterion on T_{FWHM} that

$$T_{\mathrm{FWHM}}\gg\frac{5}{\Omega_{i+}}, \tag{69}$$

for third-order effects to be ignored [26].

To quantify the second limitation we note [25] that a Gaussian pulse with a given T_{FWHM} has a frequency width

$$\omega_{1/e}=1.665/T_{\mathrm{FWHM}}. \tag{70}$$

Thus, for a given carrier frequency ω, the frequency spectrum of the pulse will extend into the band gap if

$$(\omega-\omega_{1/e})<(\overline{\omega}+\delta\overline{\omega}). \tag{71}$$

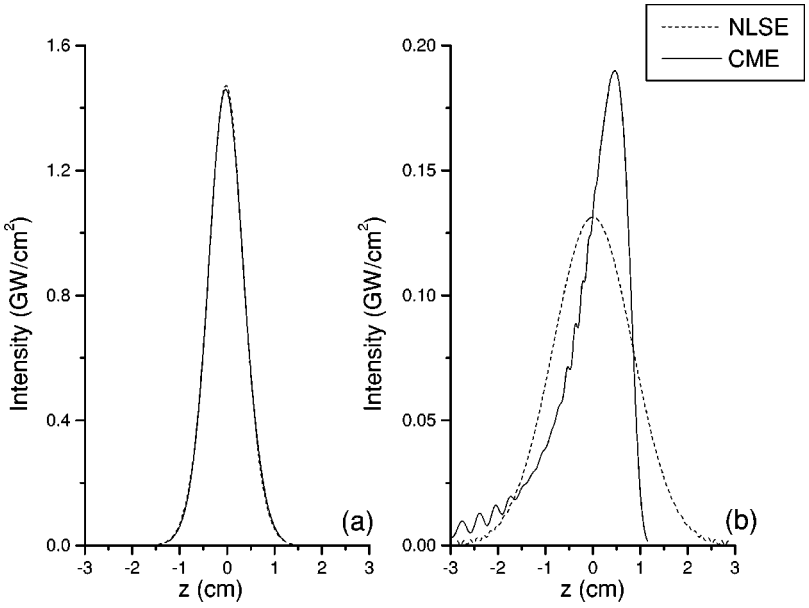

FIG. 4. Pulse profiles after 500 ps simulation time for initial pulse widths of (a) 50 ps and (b) 10 ps. It is clear that the 10 ps pulse is experiencing a great deal of higher order dispersion under the CME.

However, as the pulse frequency nears the gap, it will, of course, experience higher-order dispersion as well as, eventually, reflection, so that this criterion is not completely distinct from the one presented in the preceding paragraph.

We present simulations to underscore the first objection. We use a grating with the physical parameters in Table III, and a pulse with initial intensity 1.50 G W/cm^2 and detuning $\Delta = 2.00$. We concentrate on a single polarization, since birefringence is incidental to the higher order dispersion. Using the criterion (69) we find that $T_{\mathrm{FWHM}} \gg 12.5$ ps. In Fig. 4 we plot the simulated pulse profile after 500 ps of simulation time using both the CME and CNLSE for a T_{FWHM} of 50 and 10 ps. It can be seen that the 10 ps pulse experiences a great deal of higher-order dispersion. We note that only a small amount of reflected waves build up in this simulation, so that the second objection is irrelevant. We note, too, that since we have not attempted to simulate a soliton, the self-phase modulation will tend to increase the frequency spectrum of the pulse, so that eventually the results of the two integrations must diverge.

B. Energy exchange between the polarizations

We now consider the nonlinear energy exchange between polarizations in the context of the CNLSE. Referring to Fig. 3, we see that for the parameters being used the energy in each polarization follows, to some approximation, a cosinusoidal pattern. We present here a simple model to explain this for the case of light initially polarized at 45° to the principal axes, and having a power and pulse width similar to that of a soliton in an isotropic medium. The full dynamics of energy exchange are quite complicated, so we here concentrate on determining the period of oscillation, and the amount of energy transferred during the first oscillation period.

To discuss energy exchange we define the quantities

$$q_x(t) = \int_{-\infty}^{+\infty} \frac{|X(z,t)|^2}{v_x} dz, \tag{72}$$

$$q_y(t) = \int_{-\infty}^{+\infty} \frac{|Y(z,t)|^2}{v_y} dz,$$

and we note that

$$q_T = q_x + q_y, \tag{73}$$

is a constant of the motion. We recall that our fields are normalized such that $|X|^2, |Y|^2$ represent power, but Eq. (64) is best suited for integration in time, so that the natural quantity to calculate is that in Eq. (72). The factors of $v_x = \Omega'_{x+}$, $v_y = \Omega'_{y+}$, Eq. (64), are necessary to make q_x, q_y energies.

Using the form for the energy in the x polarized pulse (72) we can calculate

$$\frac{\partial q_x}{\partial t} = -i \frac{\alpha^x_{pc}}{v_x} \int_{-\infty}^{+\infty} [Y^2 X^{*2} e^{i\lambda z} - \text{c.c.}] dz, \tag{74}$$

$$\frac{\partial^2 q_x}{\partial t^2} = -\frac{\alpha^x_{\mathrm{PC}}}{v_x} \int_{-\infty}^{+\infty} [\xi - 2\alpha^x_{\mathrm{PC}}(|X|^2 - |Y|^2)] \times (Y^2 X^{*2} e^{i\lambda z} + \text{c.c.}) dz + 4 \frac{\alpha^{x2}_{\mathrm{PC}}}{v_x} \times \int |X|^2 |Y|^2 (|X|^2 - |Y|^2) dz, \tag{75}$$

where

$$\xi = \frac{(v_x + v_y)}{2} \lambda.$$

For small birefringences, $\lambda \ll 1$, so $e^{i\lambda z} \approx \cos(\lambda z) \approx 1$. We write our fields X and Y as

$$X(z,t) = X_r(z,t) e^{i\gamma_x t}, \tag{76}$$

$$Y(z,t) = Y_r(z,t) e^{i\gamma_y t}, \tag{77}$$

where the functions $X_r(z,t)$, $Y_r(z,t)$ are real, and the $\gamma_{x,y}$ describe phase accumulation, which here is assumed independent of space, and linear in time. Using these approximations in Eqs. (74) and (75) we find

$$\frac{\partial q_x}{\partial t} = -2\frac{\alpha^x_{\rm PC}}{v_x}\int_{-\infty}^{+\infty} Y_r^2 X_r^{*2} \sin(\gamma t)\,dz = a\,\sin(\gamma t), \quad (78)$$

$$\frac{\partial^2 q_x}{\partial t^2} \approx -2\frac{\alpha^x_{\rm PC}}{v_x}\int_{-\infty}^{+\infty} [\xi - 2\alpha^x_{\rm PC}(X_r^2 - Y_r^2)](Y_r^2 X_r^{*2})\cos(\gamma t)\,dz$$

$$= b\,\cos(\gamma t), \quad (79)$$

where a and b have been defined implicitly, and where

$$\gamma = 2(\gamma_y - \gamma_x). \quad (80)$$

Note that in going from Eqs. (75)–(79) we have ignored the third term in Eq. (75). This is because we are only interested in the dynamics during the first oscillation period, during which this term is small for the parameters of interest. We now assume a cosinusoidal variation for $q_x(t)$ and $q_y(t)$:

$$q_x(t) = \overline{q_x}(1 - \sigma\cos\gamma t), \quad (81)$$

$$q_y(t) = \overline{q_y}\left(1 + \frac{\overline{q_x}}{\overline{q_y}}\sigma\cos\gamma t\right),$$

where, by conservation of energy, $\overline{q_y} = q_T - \overline{q_x}$, and where γ is the period, and σ is the strength, of the oscillation.

By comparing the derivatives of the ansatz (81) with the approximation (78),(79), we can identify the two quantities

$$\gamma = \frac{b}{a}, \quad \sigma = \frac{a^2}{\overline{q_x} b}. \quad (82)$$

To evaluate the quantities a and b, we make the further approximation that only the *amplitude* of the real part of the pulse changes, not the spatial profile. In the following we assume a sech profile for our pulses

$$X_r(z,t) = \overline{X}(z)\,\mathrm{sech}\left(\frac{t}{T_0}\right)e^{i\gamma z}, \quad (83)$$

$$Y_r(z,t) = \overline{Y}(z)\,\mathrm{sech}\left(\frac{t}{T_0}\right)e^{i\gamma_y z}. \quad (84)$$

Using these fields in the approximate expressions (78) and (79) for the time derivatives, we can evaluate the expected value of both σ and γ, the former as the root of a quadratic equation, and the latter as a function of σ:

$$0 = 7.87\alpha^x_{\rm PC}\overline{Y}_0^2\sigma^2 + \sigma\left(\frac{4}{3}\alpha^x_{\rm PC}\overline{Y}_0^2 - \xi\right) - \frac{4}{3}\alpha^x_{pc}\overline{Y}_0^2,$$

$$\gamma = \xi - \frac{16}{5}\alpha^x_{\rm PC}\overline{Y}_0^2\{\sigma - \sigma^2\}, \quad (85)$$

where Y_0 is the initial amplitude of the pulse. The validity of these approximations is demonstrated in Fig. 5 for a detuning

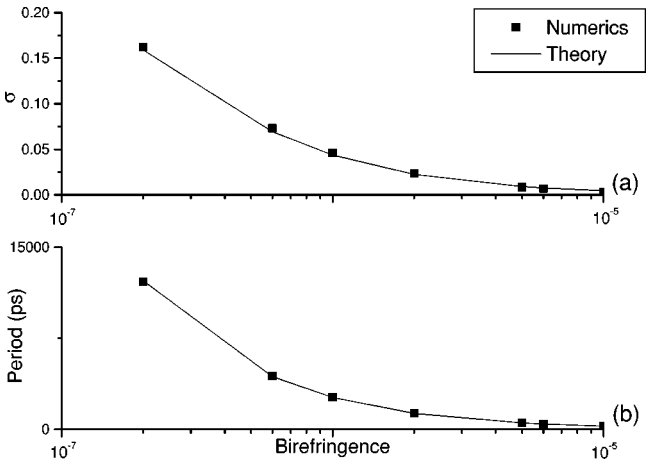

FIG. 5. Comparison of the period and strength of the nonlinear energy exchange in the CNLSE *versus* the theory presented in the paper, for a detuning of $\Delta = 1.50$.

$\Delta = 1.5$, initial peak intensity of 0.55 G W/cm^2 in each polarization, and various values of birefringence. At lower detunings there is an increased tendency to pulse separation; and, since the nonlinearity is higher, the approximation that the fields pick up spatially independent phase is no longer good. Higher energy exchange values σ, correspond to higher periods, which means that devices wishing to exploit the energy exchange would have to trade off between device length and amount of energy exchanged.

VIII. CONCLUSION

We have discussed the propagation of light through a strong grating structure in the presence of birefringence and a Kerr nonlinearity. The effect of the birefringence is to separate the photonic band gaps associated with the two polarizations. Far from the photonic band gaps, and even near the gaps if the pulses are not too short, the electromagnetic field can be well described by two coupled nonlinear Schrödinger equations, one associated with each polarization. Here the situation is somewhat similar to propagation in a one-dimensional (1D) structure without a grating, with the dispersion due to the underlying material medium. But the grating structure is richer in two respects. First, the two polarization modes can have both different group velocities and different group velocity dispersions, whereas in uniform 1D structures differences in the group velocity dispersions can typically be neglected. Second, the effective nonlinearity is a function of the carrier frequency of the pulse, since it depends on how the appropriate Bloch function samples the distribution of nonlinearity in the underlying medium.

At carrier frequencies close to the gap or within the gap, the electromagnetic field is described by two sets of coupled mode equations. The analog of a coupled mode description in the absence of birefringence, here there is one pair of equations for each polarization. For a range of parameters either set of equations can be used, and we identified the conditions required for this and confirmed them with numerical examples. We also showed how the nonlinear Schrödinger equations could be used to provide a simple understanding of one scenario of energy exchange between the two polarizations. It is clear from this, as well as from the

general form of the equations we derive, that whole new regimes of nonlinear phenomena can appear when birefringence exists in 1D photonic band-gap structures, including all-optical switching geometries that have no analog in isotropic structures. Thus the derivation of the sets of equations we presented here is of interest not only in its own right, but as a starting point for addressing what to date is the largely unexplored territory of birefringent, nonlinear photonic band-gap structures.

ACKNOWLEDGMENTS

This work was supported by the National Science and Engineering Research Council of Canada, and by Photonics Research Ontario.

[1] H. G. Winful, J. H. Marburger, and E. Garmire, Appl. Phys. Lett. **35**, 379 (1979).
[2] W. Chen and D. L. Mills, Phys. Rev. Lett. **58**, 160 (1987).
[3] C. M. de Sterke and J. E. Sipe, Phys. Rev. A **38**, 5149 (1988).
[4] S. Radic, N. George, and G. P. Agrawal, J. Opt. Soc. Am. B **12**, 671 (1995).
[5] C. M. de Sterke, D. G. Salinas, and J. E. Sipe, Phys. Rev. E **54**, 1969 (1996).
[6] 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.
[7] B. J. Eggleton, C. M. de Sterke, and R. E. Slusher, J. Opt. Soc. Am. B **16**, 587 (1999).
[8] B. J. Eggleton *et al.*, Phys. Rev. Lett. **76**, 1627 (1996).
[9] D. Taverner *et al.*, Opt. Lett. **20**, 246 (1998).
[10] V. Mizrahi *et al.*, Appl. Phys. Lett. **63**, 1727 (1993).
[11] N. G. R. Broderick *et al.*, *Quantum Electronics and Laser Science Conference, OSA Technical Digest* (Optical Society of America, Washington DC, 1999), p. 9.
[12] T. Erdogan and V. Mizrahi, J. Opt. Soc. Am. B **11**, 2100 (1994).
[13] C. R. Menyuk, Opt. Lett. **12**, 614 (1987).
[14] S. Pereira and J. E. Sipe, Opt. Express **3**, 418 (1998).
[15] P. S. J. Russell, J. Mod. Opt. **38**, 1599 (1991).
[16] From Bloch's theorem we write $\mathbf{\Psi}_{mi}(k)=\mathbf{h}_{mi}(k)e^{ikz}$, where the periodic dependence of $\mathbf{\Psi}_{mi}(k)$ and $\mathbf{h}_{mi}(k)$ on z is suppressed. In the $k\cdot p$ method $\mathbf{h}_{mi}(K+\delta k)$ is expanded in terms of the $\mathbf{h}_{mi}(K)$. Clearly these include all the $\mathbf{h}_{mi}(K)$. But note that if $\mathbf{\Psi}_{mi}(K)$ is an eigenvector of $\mathbf{n}_i^{-1}\cdot\mathbf{M}_i$ with eigenvalue $\omega_{mi}(K)$, then $\mathbf{\Psi}^*_{mi}(K)$ is an eigenvector of $\mathbf{n}_i^{-1}\cdot\mathbf{M}_i$ with eigenvalue $-\omega_{mi}(K)$. Hence, since $\omega_{mi}(-K)=\omega_{mi}(K)$, $\mathbf{\Psi}^*_{mi}(-K)$ is an eigenvector of $\mathbf{n}_i^{-1}\cdot\mathbf{M}_i$ with eigenvalue $-\omega_{mi}(K)\equiv\omega_{\bar{m}i}(K)$, and is characterized by a crystal momentum K. We define $\mathbf{h}_{\bar{m}i}(K)$ according to $\mathbf{\Psi}^*_{mi}(-K)=\mathbf{h}_{\bar{m}i}(K)e^{iKz}$. When we wish to expand $\mathbf{h}_{ni}(k)$ we must use the full set $\mathbf{h}_{\tilde{m}i}(K)$, where the range of possible $\tilde{m}$ includes the m and $\bar{m}$. This was not done explicitly by de Sterke *et al.* [5].
[17] C. M. de Sterke and J. E. Sipe, Phys. Rev. A **42**, 550 (1990).
[18] Y. Silberberg and Y. Barad, Opt. Lett. **20**, 246 (1995).
[19] N. Akhmediev and J. M. Soto-Crespo, Phys. Rev. E **49**, 5742 (1994).
[20] J. M. Soto-Crespo, N. Akhmediev, and A. Ankiewicz, J. Opt. Soc. Am. B **12**, 1100 (1991).
[21] Y. Chen, Phys. Rev. E **57**, 3542 (1998).
[22] C. M. de Sterke, Phys. Rev. E **48**, 4136 (1993).
[23] W. Samir, S. J. Garth, and C. Pask, J. Opt. Soc. Am. B **11**, 64 (1994).
[24] C. M. de Sterke, K. R. Jackson, and B. D. Robert, J. Opt. Soc. Am. B **8**, 403 (1991).
[25] G. P. Agrawal, *Non-Linear Fiber Optics* (Academic, San Diego, 1989).
[26] G. Lenz, B. J. Eggleton, and N. Litchinitser, J. Opt. Soc. Am. B **15**, 715 (1998).