Coupled modes and the nonlinear Schrödinger equation

C. Martijn de Sterke and J. E. Sipe
Department of Physics and Ontario Laser and Lightwave Research Centre, University of Toronto, Toronto, Ontario, Canada M5S 1A7
(Received 25 September 1989; revised manuscript received 2 February 1990)

We study the interaction of co- and contrapropagating modes in the presence of an intensity-dependent refractive index, and find that, under suitable conditions, these processes can be described by the nonlinear Schrödinger equation.

I. INTRODUCTION

The interaction between different waveguide modes is often described by a set of coupled-mode equations.[1] Essential in this description is that two branches of the dispersion relation of the modes, which would cross in the absence of an interaction, experience an anticrossing under the influence of a *finite* interaction. The nature of the interaction and the guiding medium is irrelevant in this description—the coupled-mode equations therefore apply to a variety of guided wave phenomena, which include elasto-optic and electro-optic interactions, as well as interactions mediated by surface corrugations, and through the proximity effect as in directional couplers, in both waveguide and fiber geometries. As first pointed out by Yariv,[1] two cases with quite different properties are distinguished in coupled-mode theory. In the first of these, the *contrapropagating* case, the two modes travel in opposite directions, for example, when a waveguide mode is reflected off a (nearly) Bragg-matched grating structure.[1] In the second case, the modes travel in the same direction as, for example, in directional couplers. In this, the *copropagating* case, the two modes can exchange energy periodically.[1]

The difference in properties between co- and contrapropagation is related to the way in which the interaction lifts the degeneracy between the modes: For contrapropagating modes the anticrossing leads to a forbidden *energy gap*,[2] while it leads to a forbidden *wave-vector gap*[2] when the modes copropagate. In the present paper we study how these mode interactions are modified in the presence of an intensity-dependent refractive index. The coupled-mode equations for this nonlinear case have recently been shown to have solitary-wave solutions.[3-5] In the present paper, however, we study the applicability of the nonlinear Schrödinger equation to describe these nonlinear phenomena. Using an envelope function approach we show that this equation holds in the limiting case that the Bloch function at one side of the stop gap (of either kind) dominates the rapid spatial (for contrapropagating modes), or temporal (for copropagating modes) dependence of the electric field. This requirement can be satisfied for contrapropagating modes by just suitably limiting the bandwidth of the radiation source. One can thus use the nonlinear Schrödinger equation to describe Bragg reflection of a nearly monochromatic wave off a periodic structure. For copropagating modes, one would have to limit the *spatial* bandwidth suitably, which is far less practical.

The comparisons of the soliton solution to the nonlinear Schrödinger equation and the rather complicated exact solitary wave solutions to the general nonlinear coupling problem demonstrate that, in the proper limit, these two solutions are identical. However, since the nonlinear Schrödinger equation is integrable, it is the preferred description in the domain where it is valid.

We have previously used an envelope function approach to describe Bragg reflection in the presence of a nonlinearity.[6-9] Starting with a wave equation for the electromagnetic field, that treatment did not presuppose a weak linear coupling between the modes due to the periodicity in the dielectric constant.[6,7] However, in the case of linear coupling being sufficiently weak that coupled mode theory can be used, the present treatment is in fact more general since the nature of the linear interaction between the modes does not enter into the discussion. Indeed, one can begin with the coupled-mode equations with the coupling parameters set phenomenologically, or found from experiment, and does not have to go back to an underlying wave equation. Whereas our previous work[6-9] made explicit use of the Bloch functions of the periodic grating, the present treatment is based only on the coupled-mode equations and their associated eigenvectors. A one-to-one correspondence exists between these eigenvectors of the (linearized) equations and the Bloch functions bordering the stop gap under consideration, so that in a suitable limit the two treatments are equivalent.

The organization of this paper is as follows. In Sec. II we first briefly give the solitary-wave solutions to the coupled contrapropagating mode equations. We then demonstrate the applicability of the nonlinear Schrödinger equation to the same problem, and compare the solutions to this equation to the solitary-wave solutions. In Sec. III we demonstrate that this equation can easily be adjusted to apply to copropagating modes as well. In Sec. IV we argue that in practice the validity of the nonlinear Schrödinger equation to copropagating modes is quite limited. We then briefly summarize our conclusions.

II. CONTRAPROPAGATING MODES

For two contrapropagating modes one obtains the set of equations[1,3,4,10,11]

$$+i\frac{\partial \mathcal{E}_+}{\partial z}+i\frac{\eta}{c}\frac{\partial \mathcal{E}_+}{\partial t}+\kappa\mathcal{E}_- e^{-2i\nu z}+\Gamma_s|\mathcal{E}_+|^2\mathcal{E}_+ + 2\Gamma_\times|\mathcal{E}_-|^2\mathcal{E}_+ = 0\ ,$$

$$-i\frac{\partial \mathcal{E}_-}{\partial z}+i\frac{\eta}{c}\frac{\partial \mathcal{E}_-}{\partial t}+\kappa\mathcal{E}_+ e^{+2i\nu z}+\Gamma_s|\mathcal{E}_-|^2\mathcal{E}_- + 2\Gamma_\times|\mathcal{E}_+|^2\mathcal{E}_- = 0\ , \tag{2.1}$$

where c is the speed of light, c/η is the group velocity of the two modes, κ is a coupling coefficient which gives the strength of the linear interaction, ν is a detuning parameter, and $\Gamma_{s,\times}$ are nonlinear coupling coefficients which are proportional to various overlap integrals of the modes. The coefficients Γ_s describe the self-phase modulation of the modes, whereas $\Gamma_\times$ describes cross-phase modulation. Further, $\mathcal{E}_\pm$ denote slowly varying electric field amplitudes (in space and time) of the forward (+) and backward (−) traveling modes. Note that Eqs. (2.1) imply that, except for traveling in opposite directions, these modes are identical. The analysis is far more complicated if this condition is not satisfied, and we do not discuss this case here. For a specific example, in the case of a planar optical waveguide, the $\mathcal{E}_\pm$ bear the following relation to the actual electric field $\mathbf{E}(x,\mathbf{R};t)$:[11]

$$\mathbf{E}(x,\mathbf{R};t)=[\mathcal{E}_+(\mathbf{R};t)\mathbf{f}_+(x;\omega_0)e^{i\mathbf{k}_0\cdot\mathbf{R}}+\mathcal{E}_-(\mathbf{R};t)\mathbf{f}_-(x;\omega_0)e^{-i\mathbf{k}_0\cdot\mathbf{R}}]e^{-i\omega_0 t}+\text{c.c.}, \tag{2.2}$$

where $\mathbf{f}_\pm$ are properly normalized waveguide modes traveling in opposite directions, which for a planar waveguide depend on single coordinate x only; ω_0 and $\mathbf{k}_0$ are the temporal and spatial center frequencies of the field; and $\mathbf{R}=y\hat{y}+z\hat{z}$ denotes the coordinates in the waveguide plane [note that if the $\mathcal{E}_\pm$ satisfy Eqs. (2.1), the field does not depend on y]. It should be mentioned that relations similar to Eq. (2.2) can be written down for fibers, or for any other system which an be described by coupled-mode equations as Eqs. (2.1).

We now remove the exponential factors in Eq. (2.1) by introducing the functions $\mathcal{F}_\pm$ through the relations[1]

$$\mathcal{E}_\pm=\mathcal{F}_\pm e^{i\nu[\mp z+(c/\eta)t]}\ , \tag{2.3}$$

and find

$$+i\frac{\partial \mathcal{F}_+}{\partial z}+i\frac{\eta}{c}\frac{\partial \mathcal{F}_+}{\partial t}+\kappa\mathcal{F}_- +\Gamma_s|\mathcal{F}_+|^2\mathcal{F}_+ + 2\Gamma_\times|\mathcal{F}_-|^2\mathcal{F}_+ = 0\ ,$$

$$-i\frac{\partial \mathcal{F}_-}{\partial z}+i\frac{\eta}{c}\frac{\partial \mathcal{F}_-}{\partial t}+\kappa\mathcal{F}_+ +\Gamma_s|\mathcal{F}_-|^2\mathcal{F}_- + 2\Gamma_\times|\mathcal{F}_+|^2\mathcal{F}_- = 0\ . \tag{2.4}$$

Before studying the properties of the full set of nonlinear equations we first consider the limiting case in which the amplitude of the forward and backward traveling waves is small enough to neglect the nonlinear terms. The properties of the thus obtained linearized set are quite important as they can aid in the understanding of the full set of equations. Specifically, an eigenvector analysis of the linear system proves to be a very convenient starting point for describing the properties of the nonlinear system. Thus neglecting now the nonlinear terms we rewrite Eqs. (2.4) as

$$\begin{bmatrix} -i\frac{\partial}{\partial z} & -\kappa \\ -\kappa & +i\frac{\partial}{\partial z} \end{bmatrix}\mathcal{F}=i\frac{\eta}{c}\frac{\partial}{\partial t}\mathcal{F}\ , \tag{2.5}$$

where $\mathcal{F}$ is the column vector with elements $\mathcal{F}_+$ and $\mathcal{F}_-$. We can find the dispersion relation associated with these equations by making the substitution $\mathcal{F}\propto e^{i(kz-\Omega t)}$, so that we obtain

$$\mathcal{M}\mathbf{v}\equiv\begin{bmatrix} +k & -\kappa \\ -\kappa & -k \end{bmatrix}\mathbf{v}=\frac{\eta\Omega}{c}\mathbf{v}\ , \tag{2.6}$$

which can readily be solved to give

$$\frac{\eta\Omega_\pm}{c}=\pm(k^2+\kappa^2)^{1/2}\ ,\qquad \mathbf{v}_\pm=\begin{bmatrix} \kappa \\ k\mp(k^2+\kappa^2)^{1/2} \end{bmatrix}\ . \tag{2.7}$$

We see from these expressions that no solutions for the angular frequency can be found in the range $-\kappa<\eta\Omega/c<+\kappa$, so that no traveling-wave solutions in this range are allowed. In the case when the coupled-mode equations [Eqs. (2.1)] result from a periodic variation in the (effective) dielectric constant, this is the well-known (energy) stop gap associated with Bragg reflection off a periodic structure (cf. Fig. 1). It can be shown that, within the approximations that lead to the derivation of Eqs. (2.1) from the full Maxwell equations, the two eigenvectors $\mathbf{v}_\pm$ in Eq. (2.7) correspond to the two Bloch functions bordering the stop gap under consideration. Note, however, that in this treatment all other Bloch functions of the periodic structure are disregarded, in contrast to our earlier treatment of this problem,[6,7] thus significantly reducing the level of complexity.

We now return to the fully nonlinear set of equations [Eqs. (2.4)]. It has been pointed out recently that in the limit in which the self-phase modulation terms vanish ($\Gamma_s=0$), Eqs. (2.4) are identical to those of the *massive Thirring model* used in quantum field theory.[3,4,12] This set of equations is integrable, and its soliton solutions are well known.[12] Equations (2.4), on the other hand, appear to be nonintegrable[4] and therefore do not have soliton solutions. However, solitary-wave solutions to Eqs. (2.4) (with $\Gamma_s\neq 0$) have been found by suitably transforming the soliton solutions to the massive Thirring model.[4] Below, we demonstrate that, within certain limits to be discussed, solutions to Eqs. (2.4) can also be found using

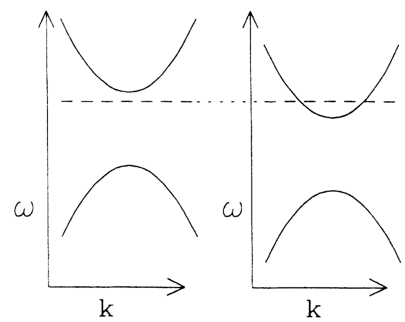

FIG. 1. Illustration of the origin of gap solitons for a positive nonlinearity. At low-field strengths the center frequency falls inside the stop gap (left). At high-field strengths, such as around the maximum of the soliton, the refractive indices increase due to the nonlinearity, so that the dispersion relation shifts down in energy, the center frequency locally tunes out of the gap, and traveling-wave solutions are allowed (right). If the parameter $\delta \ll 1$, as in the figure, the required field strength for this to happen is minimal. A similar argument holds when the nonlinearity is negative.

the nonlinear Schrödinger equation. Within these limits the use of the nonlinear Schrödinger equation is obviously the method of choice since it is an integrable equation. We finish this section by comparing the exact solitary-wave solutions to the soliton solutions of the nonlinear Schrödinger equation in the regime where it is valid.

Following Aceves and Wabnitz[4] we write the solitary-wave solutions to Eqs. (2.4) as

$$\mathcal{F}_\pm = \alpha \tilde{\mathcal{F}}_\pm e^{i\eta(\theta)} , \tag{2.8}$$

where the $\tilde{\mathcal{F}}_\pm$ are the soliton solutions to the massive Thirring model, which, in our notation, read

$$\begin{aligned} \tilde{\mathcal{F}}_+ &= \pm \left[\pm \frac{\kappa}{2\Gamma_\times} \right]^{1/2} \frac{1}{\Delta} \sin\delta \, e^{\pm i\sigma} \operatorname{sech}(\theta \mp i\delta/2) , \\ \tilde{\mathcal{F}}_- &= - \left[\pm \frac{\kappa}{2\Gamma_\times} \right]^{1/2} \Delta \sin\delta \, e^{\pm i\sigma} \operatorname{sech}(\theta \pm i\delta/2) . \end{aligned} \tag{2.9}$$

The choice of the signs in these equations is determined by that of the relative signs of the linear and nonlinear coupling coefficients. Further,[12]

$$\begin{aligned} \theta &= \kappa \sin\delta \left[z - \frac{c}{\eta} v t \right] \Big/ (1-v^2)^{1/2} , \\ \sigma &= \kappa \cos\delta \left[v z - \frac{c}{\eta} t \right] \Big/ (1-v^2)^{1/2} , \end{aligned} \tag{2.10}$$

where the dimensionless quantity v is given by[12]

$$v = (1-\Delta^4)/(1+\Delta^4) . \tag{2.11}$$

Finally, we have defined

$$\begin{aligned} \alpha &= (1+R_+ + R_-)^{-1/2} , \\ e^{i\eta(\theta)} &= \left[-\frac{e^{2\theta}+e^{\pm i\delta}}{e^{2\theta}+e^{\mp i\delta}} \right]^{(R_+ - R_-)/(1+R_+ + R_-)} , \end{aligned} \tag{2.12}$$

where

$$R_\pm = \frac{\Gamma_s}{4\Gamma_\times} \frac{(1 \pm v)^2}{1-v^2} . \tag{2.13}$$

In these equations Δ and δ $(0<\delta<\pi)$ are free parameters. According to Eqs. (2.10) and (2.11) the parameter v (and thus Δ) determines the soliton velocity, while δ determines the position in the stop gap. The value $\delta=\pi/2$ denotes a solution with center frequency in the middle of the gap, whereas $\delta \rightarrow 0(\pi)$ denotes solutions near the top (bottom), depending on whether the coupling coefficients are positive or negative. This can be verified by using Eqs. (2.9) in Eq. (2.3) and identifying the full frequency dependence. To understand why the maximum field strength increases with δ, it is important to realize that the solitons exists because, through the nonlinearity, the electric field shifts the photonic band structure such that the center frequency locally no longer falls within the stop gap, but corresponds to an allowed band.[6,7] Since a positive (negative) nonlinearity shifts the photonic band structure down (up) in energy, the lowest field strength is required if the field is tuned close to the upper (lower) edge of the stop gap.[6,7] This is illustrated in Fig. 1. Numerical work by Aceves and Wabnitz[4] has shown that the solutions Eqs. (2.8)–(2.13) do not have all the properties commonly associated with solitons: collisions between two such solutions seem to give rise to more than just phase shifts. These solutions should therefore be denoted as "solitary waves."[4]

We next solve Eq. (2.4) using an approximate method which makes use of an envelope function approach. Before doing so, it is important to realize that the $\mathcal{F}_\pm$ are slowly varying by virtue of a synchronous approximation which has been applied before Eqs. (2.1) can be written down. In this approximation only terms which are (almost) phase matched are retained.[1] Whereas the actual fields vary at a rate ω_0 [Eq. (2.2)] of about 10^{15} s^{-1}, the time dependence of the $\mathcal{F}_\pm$ extends over the range of influence of the mode interaction, which roughly corresponds to the size of the stop gap and is thus several orders of magnitude below optical frequencies. In our subsequent analysis we introduce an even slower varying set of variables which typically vary on the scale of a fraction of a stop gap. Since we have three different levels of variables, it is important to distinguish the various parameters—the group velocity c/η in Eqs. (2.1) and (2.4) equals the slope of the dispersion curves of the two modes in the *absence* of the linear interaction, and differ from the slope and curvature of the dispersion curves in the *presence* of that interaction, to be introduced below.

As we saw before, the solutions to the linearized set of equations can be written as a suitable linear combination of the eigenvectors (normal modes) $\mathbf{v}_\pm$ each multiplied by the plane-wave factor $e^{i(kz-\Omega_\pm t)}$, where Ω and k are related by Eqs. (2.7). Returning to the fully nonlinear system of Eqs. (2.4), it is clear that at any given point in space and time we can write the solution to the general system as a linear combination of the $\mathbf{v}_\pm$ with their associated plane-wave factors as well, but the amplitudes will *not* be constant now. If we assume, however, that these amplitudes vary on much slower scales than the factors $e^{i(kz-\Omega t)}$, we can use a variety of mathematical techniques to separate the slowly varying amplitudes from the rapidly varying plane-wave factors, giving rise to equations for the slowly varying amplitudes only. However, since the ensuing equations do not appear to be simpler than the original coupled-mode equations,[7] and since we have solitary-wave solutions to the coupled-mode equations [Eqs. (2.8)–(2.13)], this general approach is not very useful. The scheme of separating the slowly and rapidly varying field components can, however, give useful results if we restrict ourselves as follows. We assume that one of the eigenvectors and its associated plane-wave factor dominate the rapidly varying field components. Physically, this means that the frequency content of the electric field overlaps only one of the edges of the stop gap significantly.[6,7] The results of such an approach thus only applies to pulses whose temporal width is much larger than that associated with stop gap. For a planar optical waveguide with a typical inverse coupling coefficient of about a millimeter, this excludes pulses shorter than a few tens of picoseconds, whereas for an optical fiber with a typical inverse coupling coefficient of a meter it excludes pulses shorter than a few tens of nanoseconds. Note, however, that the resulting solitons travel much slower than the speed of light,[9] and the spatial extent of pulses of duration τ in the structure is thus much less than $c\tau/\eta$.

In separating the rapidly varying from the more slowly varying field components we use the *method of multiple scales.* This technique has been described before (see, e.g., Refs. 6 and 7), so we will not show a detailed calculation here. To use the method, a set of coordinates describing variations on different time and length scales is introduced through

$$z=z_0+\mu z_1+\mu^2 z_2+\cdots \qquad t=t_0+\mu t_1+\mu^2 t_2+\cdots \qquad (\mu \ll 1)\ , \tag{2.14}$$

and the variation of a function on the different scales z_n, t_n, is treated as if those quantities were independent variables. Further, the solution is approximated by the expression

$$\mathcal{F}=(\mu a(z_1,z_2,\ldots;t_1,t_2,\ldots)|\pm\rangle + \mu^2 b(z_1,z_2,\ldots;t_1,t_2,\ldots)|\mp\rangle)e^{i(kz_0-\Omega_\pm t_0)}, \tag{2.15}$$

where a ket notation is used, $|+\rangle$ referring to $\mathbf{v}_+$ of Eq. (2.7), and the $|-\rangle$ refer to $\mathbf{v}_-$. The slowly varying envelope functions a and b are, at least for now, completely arbitrary. Note that the upper (lower) sign in Eq. (2.15) corresponds to a wave package centered near the top (bottom) of the stop gap. Since $\mu \ll 1$, we assume in writing Eq. (2.15) that, consistent with the discussion above, $\mathbf{v}_+$ ($\mathbf{v}_-$) dominates the rapid variations of the $\mathcal{F}$. Substituting Eqs. (2.14) and (2.15) into Eqs. (2.4) we find that these equations are satisfied up to order μ. Collecting next all terms proportional to μ^2 it is found that

$$\left[-\frac{\eta}{c}\Omega_\pm|\mp\rangle+\mathcal{M}|\mp\rangle\right]b=i\frac{\partial a}{\partial z_1}\hat{\sigma}_3|\pm\rangle+i\frac{\eta}{c}\frac{\partial a}{\partial t_1}|\pm\rangle\ , \tag{2.16}$$

where $\hat{\sigma}_3$ is the third of the Pauli matrices [diag(1, −1)] and $\mathcal{M}$ was defined in Eq. (2.6). We analyze Eq. (2.16) by mapping it onto the two eigenvectors. Taking first the inner product with $\mathbf{v}_\pm$ leads to

$$\left[\frac{c}{\eta}\frac{\langle\pm|\hat{\sigma}_3|\pm\rangle}{\langle\pm|\pm\rangle}\right]\frac{\partial a}{\partial z_1}+\frac{\partial a}{\partial t_1}=0\ . \tag{2.17}$$

It is straightforward to show from Eq. (2.7) that the expression within brackets in Eq. (2.17) equals the group velocity $\Omega'=(\partial\Omega_\pm/\partial k)$. The slowly varying function $a(z_1,z_2,\ldots;t_1,t_2,\ldots)$ thus cannot depend on z_1 and t_1 independently, but only on the linear combination $z_1-\Omega' t_1$. This means that, to this order, the envelope travels with the group velocity. We stress that this is the group velocity in the *presence* of the interaction (thus not c/η). We now introduce the new variables

$$\zeta_1=z_1-\Omega' t_1\ , \qquad \tau_1=t_1\ , \tag{2.18}$$

so that $a=a(\zeta_1;z_2,\ldots;t_2,\ldots)$.

Mapping Eq. (2.16) onto the eigenvector $\mathbf{v}_\mp$, it is found that the envelope function b does not depend on τ_1 either, and that

$$b(\zeta_1;z_2,\ldots;t_2,\ldots)=\mp i\frac{c}{\eta}\frac{1}{\Omega_+-\Omega_-}\frac{\langle\mp|\hat{\sigma}_3|\pm\rangle}{\langle\mp|\mp\rangle}\frac{\partial a}{\partial\zeta_1}\ . \tag{2.19}$$

We have found expressions similar to Eqs. (2.17) and (2.19) previously,[6] but in this earlier work, our starting point was the wave equation for the electric field with a nonlinear periodic dielectric function. The present work is based on the coupled-mode equations [Eqs. (2.1), which, as discussed in the Introduction, are valid under less restrictive conditions.

We now finally consider the terms proportional to μ^3 and map the resulting expression onto $\mathbf{v}_\pm$. It is then found that

$$i\left[\frac{\partial a}{\partial t_2}+\Omega'\frac{\partial a}{\partial z_2}\right]+\left[\pm\frac{c^2}{\eta^2}\frac{|\langle+|\hat{\sigma}_3|-\rangle|^2}{\langle-|-\rangle\langle+|+\rangle(\Omega_+-\Omega_-)}\right]\times\frac{\partial^2 a}{\partial\zeta_1^2}+\Xi\langle\pm|\pm\rangle|a|^2a=0\ , \tag{2.20}$$

where the nonlinear coefficient is given by

$$\frac{2\eta}{c}\Xi=(2\Gamma_\times+\Gamma_s)-(2\Gamma_\times-\Gamma_s)\left[\frac{\eta}{c}\Omega'\right]^2\ . \tag{2.21}$$

Equation (2.20), which is the nonlinear Schrödinger equation, has been widely studied and is known to be integrable.[13] The factor within brackets in the second term in this equation equals half the group velocity dispersion

$\Omega'' \equiv d^2\Omega_\pm/dk^2$. Introducing the variables ζ_2, τ_2 in the same way as ζ_1, τ_1 in Eq. (2.18), and assuming that the eigenvectors are normalized to unity [cf. Eq. (2.7)], allows us to rewrite Eq. (2.20) as

$$i\frac{\partial a}{\partial \tau_2} + \frac{1}{2}\Omega'' \frac{\partial^2 a}{\partial \zeta_1^2} + \Xi |a|^2 a = 0 \ . \tag{2.22}$$

Having solved Eq. (2.22) for a, we can then find b from Eq. (2.19) and $\mathcal{F}$ from Eq. (2.15). Equation (2.22) is similar to results of previous investigations in which we studied the Bragg reflection of a waveguide mode off a grating structure.[6-9] As discussed, however, the present work is more general. In the remainder we drop the subscripts of ζ_1 and τ_2 and we suppress all coordinates ζ_n, τ_n for $n > 2$, so that $a = a(\zeta, \tau)$.[6]

The one-soliton solutions to the nonlinear Schrödinger equation reads[13]

$$a(\zeta,\tau) = \left[\frac{C_1^2}{\Omega''\Xi}\right]^{1/2} e^{i(C_2\zeta - C_2^2\tau/2 + C_1^2\tau/2)/\Omega''} \times \operatorname{sech}\left[\frac{C_1}{\Omega''}(\zeta - C_2\tau)\right] , \tag{2.23}$$

where the velocities C_1 and C_2 can in principle be chosen freely. But since a is an envelope function, we must have $C_1, C_2 \ll c/\eta$. Since ζ, τ themselves are coordinates in a moving frame [Eq. (2.18)], we see from Eq. (2.23) that the soliton velocity is the sum of Ω' and C_2. Physically this means that the total velocity is made up of that of the eigenvectors (Ω') and that of the envelope (C_2). It is straightforward to show that, to second order in the velocities, the total electric field does not depend on the way in which the total velocity is distributed.

We next compare in this low-velocity limit the approximate solutions following from our envelope approach to the more complicated solitary-wave solutions of the full problem [Eqs. (2.8)–(2.13)]. This matter has also been discussed briefly by Aceves and Wabnitz.[4] Since the way the total velocity is distributed is irrelevant, we may take $k=0$, so that $\Omega'=0$ [thus $\zeta = z$ and $\tau = t$, Eq. (2.18)], $\Omega_\pm = \pm(c/\eta)\kappa$, $\Omega'' = \pm c/(\eta\kappa)$, and $\mathbf{v}_\pm^T = (1/\sqrt{2})(1, \mp 1)$. If we then further make the substitutions $C_1 = cp/\eta$ and $C_2 = cq/\eta$, so that the envelope function approach requires that $p, q \ll 1$, and use Eqs. (2.15) and (2.19), we find that

$$\mathcal{F}(z,t) = \left[\frac{\pm\kappa}{2\Gamma_\times + \Gamma_s}\right]^{1/2} p e^{\pm i\phi_1} \operatorname{sech}(\phi_2) \left[\begin{bmatrix} 1 \\ \mp 1 \end{bmatrix} + \left[\frac{q}{2} \pm i\frac{p}{2}\tanh(\phi_2)\right] \begin{bmatrix} 1 \\ \pm 1 \end{bmatrix} \right] , \tag{2.24}$$

where

$$\phi_1 = q\kappa z - \frac{c}{\eta}\kappa t + \frac{c\kappa}{2\eta}(p^2 - q^2)t \ ,$$
$$\phi_2 = p\kappa \left[z - \frac{c}{\eta} q t \right] . \tag{2.25}$$

Within its range of validity ($p, q \ll 1$), Eqs. (2.24) and (2.25) are identical to the exact solitary-wave solutions found above. To see this we associate δ, v, with p, q, respectively, and expand Eqs. (2.8)–(2.13) uniformly to second order in these parameters. However, since z, t can be arbitrarily large, the solutions in Eqs. (2.24) and (2.25) fail at large spatial and temporal distances from the soliton maximum. The requirement $v \ll 1$ means, of course, that the soliton velocity be much less than the group velocities of the interacting modes. Further, Eqs. (2.9) show that $\delta \ll 1$ is the low-field limit of the exact solutions. According to the discussion in the paragraph below Eq. (2.13), this corresponds to the situation in which the system can tune itself most easily out of the gap (see Fig. 1).

III. COPROPAGATING CASE

The interaction between two copropagating modes in the presence of nonlinearity can be described by a set of equations which differ from Eqs. (2.1) only by the relative signs of the spatial derivatives and the different group velocities of the modes involved,[11]

$$i\frac{\partial}{\partial Z}\mathcal{E}_1 + i\frac{\eta_1}{c}\frac{\partial}{\partial T}\mathcal{E}_1 + \kappa e^{-2i\nu Z}\mathcal{E}_2 + \Gamma_1|\mathcal{E}_1|^2\mathcal{E}_1 + \Gamma_c|\mathcal{E}_2|^2\mathcal{E}_1 = 0 \ ,$$
$$i\frac{\partial}{\partial Z}\mathcal{E}_2 + i\frac{\eta_2}{c}\frac{\partial}{\partial T}\mathcal{E}_2 + \kappa e^{+2i\nu Z}\mathcal{E}_1 + \Gamma_1|\mathcal{E}_2|^2\mathcal{E}_2 + \Gamma_c|\mathcal{E}_1|^2\mathcal{E}_2 = 0 \ , \tag{3.1}$$

in notation similar to that in Eq. (2.1). For the specific case of a planar optical waveguide the relation between the actual electric field and $\mathcal{E}_{1,2}$ is again given by Eq. (2.2). By introducing the new coordinates $z = Z$, $t = T - \bar{\eta}Z/c$, and a new set of envelope functions by $\mathcal{E}_{1,2} = \mathcal{F}_{1,2}e^{\mp i\nu(z \mp ct/\eta)}$, where $\eta_1 = \bar{\eta} + \eta$, $\eta_2 = \bar{\eta} - \eta$, we find from Eq. (3.1) the set of coupled equations

$$i\frac{\partial}{\partial z}\mathcal{F}_1 + i\frac{\eta}{c}\frac{\partial}{\partial t}\mathcal{F}_1 + \kappa\mathcal{F}_2 + \Gamma_1|\mathcal{F}_1|^2\mathcal{F}_1 + \Gamma_c|\mathcal{F}_2|^2\mathcal{F}_1 = 0 \ ,$$
$$i\frac{\partial}{\partial z}\mathcal{F}_2 - i\frac{\eta}{c}\frac{\partial}{\partial t}\mathcal{F}_2 + \kappa\mathcal{F}_1 + \Gamma_1|\mathcal{F}_2|^2\mathcal{F}_2 + \Gamma_c|\mathcal{F}_1|^2\mathcal{F}_2 = 0 \ . \tag{3.2}$$

As pointed out by others before,[5] this set is identical to Eqs. (2.4) for contrapropagating modes, but with the roles of z and ct/η reversed. Equations (2.8)–(2.13) thus apply to the present case as well. They yield solitary-wave solutions to Eqs. (3.2) which have been called *resonance solitons*.[5] Our present interest, however, is again the possible applicability of the nonlinear Schrödinger equation. Starting with the linearized version of Eqs. (3.2), one can find a set of eigenvectors and eigenvalues, just as in Sec. II. Since the roles of z and t (and thus those of Ω and k) are reversed, however, one now finds that certain values for k are forbidden, indicating a k gap.[2] In the same way as in Sec. II one can find that the coefficient of the eigenvectors satisfies the nonlinear Schrödinger equation. The condition for its validity is again that one of the Bloch functions bordering the (k) gap dominates the rapid variations of the electric field. We thus conclude that, in the limit described above, resonance solitons can be described by the nonlinear Schrödinger equation. The reader is referred to Ref. 5 for a description of some of the properties of these electric field solutions.

We now finally investigate if the nonlinear Schrödinger equation applies to the well-known situation in which two copropagating modes periodically exchange energy.[1] To do this we consider Eqs. (3.2) with $\mathcal{F}_{1,2} \propto e^{ikz}$ [cf. Eq. (2.6)], which leads to the (unnormalized) eigenvectors $\mathbf{v}_{1,2}^T = (1, \pm 1)$. These correspond to states in which each of the copropagating modes carry equal amounts of energy—the situation in which one of the modes carries most of the energy must thus correspond to the superpositions $\mathbf{v}_1 \pm \mathbf{v}_2$. In the periodic exchange of energy between two copropagating modes, the total state of the linear system must thus beat periodically between these two superpositions. We thus conclude that in the process of periodic energy exchange between copropagating modes, in which both eigenvectors are equally important, the nonlinear Schrödinger equation is *not* valid.

IV. DISCUSSION AND CONCLUSIONS

The different properties of linearly interacting co- and contrapropagating modes was traced back by Yariv to a different form of the conservation law of energy flux for these two cases.[1] In turn, this distinction causes quite different behavior in the presence of a nonlinearity. The crucial assumption, which leads in Secs. II and III to the nonlinear Schrödinger equation for the envelope function, is that a single eigenvector and its associated plane-wave factor dominate the more rapidly varying behavior of the linearized system. This condition can be satisfied for two interacting contrapropagating modes by confining the frequency content of the radiation to the suitable part of the stop gap. In principle, this can be accomplished straightforwardly since the (temporal) frequency content of the field is determined directly by the source. For two copropagating modes, on the other hand, one should suitably limit the wave-vector content of the field. This is far more difficult to accomplish as the spatial frequency content of the field follows from the frequency content and the material response.

We do not discuss the applications of our findings here. We have done so before, for the specific case of Bragg reflection of a periodic grating,[8,9] where the nonlinear Schrödinger equation was used to describe "gap solitons" (Refs. 6, 8, and 14) and "self-localized light" (Refs. 3 and 9). We plan to return to additional applications of this theory in a future publication.

In conclusion, we have shown that if the temporal frequency content of the signal significantly overlaps only one of the edges of the stop gap, the interaction of two similar counterpropagating modes in the presence of a intensity-dependent refractive index can be described by the nonlinear Schrödinger. A similar conclusion can be drawn for copropagating modes, but the required narrow spatial frequency content is hard to achieve. The field profile in this limit is identical to that of the solitary-wave solutions to the coupled-mode equations, which can be found from the solutions to the massive Thirring model. Because of the wide applicability of the coupled-mode equations, these conclusions hold for a wide variety of geometries and interaction mechanisms.

ACKNOWLEDGMENTS

This work was supported in part by the Ontario Laser and Lightwave Research Centre.

[1]A. Yariv, IEEE J. Quantum Electron. **QE-9**, 919 (1973).
[2]H. Kogelnik and C. V. Shank, J. Appl. Phys. **43**, 2327 (1972).
[3]D. N. Christodoulides and R. I. Joseph, Phys. Rev. Lett. **62**, 1746 (1989).
[4]A. B. Aceves and S. Wabnitz, Phys. Lett. A **141**, 37 (1989).
[5]S. Wabnitz, Opt. Lett. **14**, 871 (1989).
[6]C. M. de Sterke and J. E. Sipe, Phys. Rev. A **38**, 5149 (1988).
[7]C. M. de Sterke and J. E. Sipe, Phys. Rev. A **39**, 5163 (1989).
[8]C. M. de Sterke and J. E. Sipe, J. Opt. Soc. Am. B **6**, 1722 (1989).
[9]C. M. de Sterke and J. E. Sipe, Opt. Lett. **14**, 871 (1989).
[10]H. G. Winful, Appl. Phys. Lett. **46**, 527 (1985).
[11]C. M. de Sterke and J. E. Sipe, J. Opt. Soc. Am. A **7**, 636 (1990).
[12]D. J. Kaup and A. C. Newell, Lett. Nuovo Cimento **20**, 325 (1977).
[13]R. K. Dodd, J. C. Eilbeck, J. D. Gibbon, and H. C. Morris, *Solitons and Nonlinear Wave Equations* (Academic, London, 1982).
[14]W. Chen and D. L. Mills, Phys. Rev. Lett. **58**, 160 (1987).