

## Extensions and generalizations of an envelope-function approach for the electrodynamics of nonlinear periodic structures

C. Martijn de Sterke and J. E. Sipe

*Department of Physics, University of Toronto, Toronto, Ontario, Canada M5S 1A7  
and Ontario Laser and Lightwave Research Centre, Toronto, Ontario, Canada M5S 1A7*

(Received 27 December 1988)

Based on an envelope-function approach developed in a previous publication [de Sterke and Sipe, Phys. Rev. A **38**, 5149 (1988)], we give criteria for the experimental observation of gap solitons. Moreover, the theory is extended in several respects—it is shown that in the limit of harmonic time dependence, our treatment is equivalent to the effective-mass approximation in condensed-matter physics. Further, we derive equations of motion for the fields under more general conditions than before, including third-harmonic generation by the optical nonlinearity. By use of Hamiltonian methods it is shown that third-harmonic generation in periodic structures has quite different properties than in homogeneous materials.

### I. INTRODUCTION

In a previous paper,<sup>1</sup> to be referred herein as I, we presented an envelope-function formalism to describe the electrodynamics of nonlinear periodic structures. This method allowed us to give a comprehensive analytic description of phenomena which were previously observed in computer experiments by Chen and Mills.<sup>2,3</sup> Among these phenomena is the remarkable property that, for certain values of the intensity, these nonlinear periodic structures can be transparent to electromagnetic radiation with a frequency falling within one of the stop gaps. This is in sharp contrast to the properties of linear periodic stacks, which reflect such radiation very strongly. As a part of their computer experiments, Chen and Mills<sup>2,3</sup> also calculated the electric field profiles within the stack, and showed that the envelope function of the field can attain a hyperbolic secant shape. For this reason, they introduced the term *gap soliton* to refer to such field profiles.<sup>2</sup> Our analysis in I showed that there is indeed a soliton mechanism underlying the transmissivity properties of the nonlinear periodic stacks, so that this nomenclature is very appropriate. In our previous paper our analysis was focused directly on a description of the effects observed by Chen and Mills. In the present paper, we consider the matter in a broader perspective and also study the limitations of the present approach more systematically.

The physical mechanism behind the remarkable transmissivity properties of nonlinear periodic stacks can be briefly explained. The incoming radiation is chosen to have a frequency falling within one of the stop gaps of the structure, but very close to one of the edges. In the linear limit, only exponentially growing and decaying envelope functions are possible and, for a finite stack, an exponentially decaying envelope function results, leading to a very low transmissivity. But if the intensity of the radiation is high enough in a region inside the stack, the structure can be tuned locally out of the stop gap. In order to accomplish this, the nonlinearity must have the appropri-

ate sign. Exponentially growing and decaying solutions to the left and right of this region can then be connected in the high-intensity region to form a self-consistent solution which is finite everywhere. A series of such high-intensity regions, connected by qualitatively linearly behaving regions where the envelope is first decaying exponentially from one high intensity and then rising to meet another, can, of course, also occur. Our analysis in I showed that the envelope function in this case is given by one of the Jacobi elliptic functions. The periodicity of these functions assures that the field itself is then periodic, which, in turn, implies that the stack is transparent if this period fits an integer number of times in the stack length. This behavior is illustrated in Fig. 1.

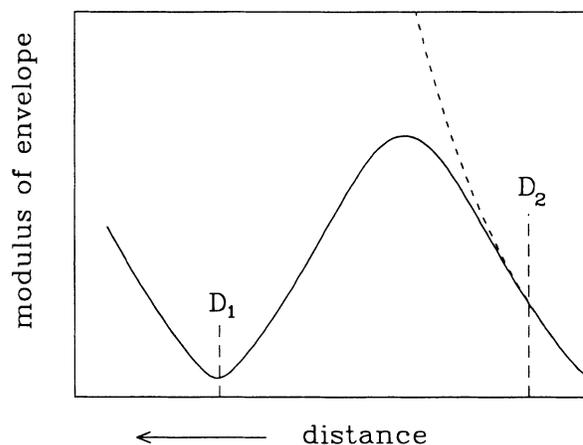


FIG. 1. Solid line: modulus of the envelope function within the nonlinear stack at a given wavelength and energy flow. Dashed line: same, but in a simple linear stack. Notice that the field at the back surface of the structure is fixed by the boundary conditions. For stack length  $D_2$  the nonlinearity has negligible influence. For length  $D_1$  the nonlinear stack is perfectly transparent in contrast with its linear counterpart.

In I we considered envelope functions which modulate a Bloch function of the periodic structure. This approach, which is often used in condensed-matter physics,<sup>4</sup> differs from that used for most optical problems (e.g., for distributed feedback lasers<sup>5</sup> and corrugated waveguides<sup>6</sup>), in which it is customary to describe the electromagnetic field by envelope functions modulating forward-traveling and backward-scattered waves. The treatment in terms of Bloch functions, however, has appeared to be more fruitful in the present problem. We found previously that considering only a single Bloch function and its associated envelope function gives a good qualitative and quantitative description of the computer experiments of Chen and Mills.<sup>2,3</sup> The reason for this is that, by choosing the frequency of the incoming radiation to be very close to one of the edges of a stop gap, as described in the previous paragraph, only the Bloch function associated with that edge has to be included explicitly. In actual experiments, for example, in optical fibers, the stop gaps may be too small for such a treatment to be sufficient, so that the Bloch functions associated with both edges of the stop gap may have to be included. In the present paper we derive equations appropriate to that situation. A similar generalization also allows us to consider the possible use of these nonlinear periodic structures for third-harmonic generation.

Apart from generalizations of our previous treatment, we also present some consequences of the theory. In the present paper we show that, in the limit of harmonic time dependence and tuning to a stop gap, our treatment gives identical results as those following from a method which is very similar to the well-understood effective-mass treatment from condensed-matter physics.<sup>4</sup> This result allows us to better understand the approximations leading to our envelope-function approach.

We further show that the equations for the envelope functions can be derived from an effective Hamiltonian. The physical interpretation of the conserved quantities in this description gives additional insight into the properties of the nonlinear periodic structures. The equations which we derived in I can be obtained from a two-dimensional rotationally symmetric effective Hamiltonian. The conservation of angular momentum in such Hamiltonians corresponds then to the conservation of the Poynting vector in the original physical problem. The Hamiltonian itself has a less obvious interpretation, but we show that the minimization of the action integral, which directly leads to the Lagrangian formalism in classical mechanics, is equivalent to the minimization of the time-averaged free energy associated with the electromagnetic radiation in the nonlinear periodic structure. We demonstrate that the description of third-harmonic generation can also be cast into a Hamiltonian form. In this case, one has to consider an effective Hamiltonian describing two rotationally symmetric systems which, through a nonlinear coupling, can exchange angular momentum, corresponding to the exchange of energy between the fundamental and its third harmonic in the nonlinear stack. We next contrast this Hamiltonian with a Hamiltonian from which the equations for third-harmonic generation in a homogeneous slab can be de-

rived. The differences between these two suggest that third-harmonic generation in these two environments exhibit some very important differences. As an example of this, we show that the phase-matching condition in the periodic structures is less strict than in nonlinear homogeneous slabs.<sup>7,8</sup>

The organization of this paper is as follows. We first give a brief overview of the most important equations we derived in I. We also present various criteria for the validity of this treatment and for the experimental observation of the gap solitons. In Sec. III we then show that in the limit of harmonic time dependence, our equations can also be derived through a standard effective-mass treatment. We then extend our analysis in Sec. IV to include multiple Bloch functions explicitly. In Sec. V we present the Hamiltonian description of some of the properties of the nonlinear periodic stacks, and some of the results are discussed in Sec. VI. Appendixes present details we require of the band structure in the Kogelnik approximation<sup>5</sup> and some comments on the form of the electromagnetic free energy we adopt.

## II. OVERVIEW AND CONSEQUENCES OF PREVIOUS RESULTS

In this section we provide an overview of the key results in I, and use them to establish the conditions necessary to observe soliton effects in periodic dielectric stacks with a Kerr-type nonlinearity. The analysis starts with the wave equation for the electric field  $E(x,t)$ , with a driving term due to the Kerr nonlinearity<sup>9</sup>

$$\begin{aligned} -c^2 \frac{\partial^2}{\partial x^2} E(x,t) + \epsilon(x) \frac{\partial^2}{\partial t^2} E(x,t) \\ = -4\pi \chi^{(3)}(x) \frac{\partial^2}{\partial t^2} [E(x,t)]^3, \quad (2.1) \end{aligned}$$

where  $c$  is the speed of light in vacuum,  $\chi^{(3)}(x)$  is the nonlinear susceptibility, and  $\epsilon(x)$  is the periodic dielectric function. The only restriction on  $\chi^{(3)}(x)$  is that it has the same period as the dielectric function. The method of multiple scales<sup>10</sup> is now used to derive an equation for the envelope function only. In order to do so, different time scales,  $t_\alpha = \mu^\alpha t$  ( $\mu \ll 1$ ,  $\alpha = 0, 1, 2, \dots$ ), and length scales,  $x_\alpha = \mu^\alpha x$ , are introduced, where  $x_0$  is the length scale over which the dielectric function is periodic and  $t_0$  is the fastest time scale in the problem, e.g., the inverse of the carrier frequency. All new variables are subsequently considered to be independent. Further, the electric field is written in an asymptotic series  $E = \mu e_1 + \mu^2 e_2 + \dots$ . This expansion for the field, as well as the new time and space variables, are introduced in the wave equation [Eq. (2.1)], and terms with equal powers of  $\mu$  are collected. The thus obtained equations are solved, starting with the terms proportional to  $\mu$ , until the desired level of approximation is obtained. Finally, then, the expansion parameter  $\mu$  is set equal to unity.

In carrying out this procedure we have to make an ansatz regarding the first term in the asymptotic series for the electric field. The simplest such ansatz, and the one used in I, is

$$e_l(x,t) = a(x,t)\varphi_m(x)e^{-i\omega_m t} + \text{c.c.}, \quad (2.2)$$

where  $\varphi_m(x)$  is a Bloch function and  $a(x,t)$  is an as yet undefined envelope function, which varies on a much slower length scale; c.c. denotes complex conjugation. The Bloch functions of the stack form a complete, orthogonal set, but because these functions satisfy the wave equation, rather than the Schrödinger equation, the orthogonality relations contain the dielectric function as a kernel, thus

$$\langle l|\epsilon|m\rangle \equiv \int_0^L \varphi_l^*(x)\epsilon(x)\varphi_m(x)dx = \delta_{l,m}, \quad (2.3)$$

where  $L$  is a normalization length. In I we found that the first two terms in the asymptotic series provide an adequate description of the field, leading to an expression for the electric field

$$E(x,t) = \left[ a(z,t)\varphi_m(x) + \sum_{l \neq m} \Lambda_{l,m} d \left[ \frac{\partial a(z,t)}{\partial z} \right] \varphi_l(x) \right] \times e^{-i\omega_m t} + \text{c.c.}, \quad (2.4)$$

in agreement with Eq. (2.2). In Eq. (2.4),  $d$  is the period of the dielectric function and  $\omega_m$  and  $\omega_l$  are the angular frequencies associated with the Bloch functions  $\varphi_m$  and  $\varphi_l$ , respectively. Further, the dimensionless coupling coefficient  $\Lambda_{l,m}$  between two Bloch functions is

$$\Lambda_{l,m} = \frac{2ic}{d} \frac{\langle l|\Omega|m\rangle}{\omega_l^2 - \omega_m^2}, \quad (2.5)$$

with the operator  $\Omega$  given by

$$\Omega = -ic \frac{d}{dx}. \quad (2.6)$$

The coordinate  $z$  is a spatial coordinate in a frame moving with the group velocity  $\omega'_m$  associated with the state  $m$ , thus  $z \equiv x - \omega'_m t$ . A major result of I is that  $\bar{a}(z,t)$ , which is defined through  $a(z,t) = \sqrt{L}\bar{a}(z,t)$ , satisfies the nonlinear Schrödinger equation

$$i \frac{\partial \bar{a}}{\partial t} + \frac{1}{2} \omega_m'' \frac{\partial^2 \bar{a}}{\partial z^2} + \alpha_m |\bar{a}|^2 \bar{a} = 0, \quad (2.7)$$

subject to the appropriate boundary conditions. In this equation,  $\omega_m''$  is the group-velocity dispersion associated with  $\varphi_m$  and  $\alpha_m$  is the effective nonlinearity of the modulated Bloch state,

$$\alpha_m = 6\pi\omega_m L \int_0^L \chi^{(3)}(x) |\varphi_m(x)|^4 dx. \quad (2.8)$$

In agreement with the expectations, the first term in the expression for the electric field in Eq. (2.4) is the dominant one. The much smaller, second term, is however necessary, for example, to calculate the energy flow through the system and thus has to be retained. To denote these contributions to the electric field, therefore, we use the terminology *principal terms* and *companion terms* in a slight generalization of the notation in I. Although we shall see in Sec. IV that the ansatz in Eq. (2.2) is inappropriate in certain situations, it is sufficient to

give a correct description of the results of Chen and Mills. Under more general circumstances we have to write the principal component of the electric field as the sum of two or more terms.

In I we limited ourselves to electromagnetic fields which have harmonic time dependence, with envelopes which are at rest in space. The latter restriction implies that the state  $m$  has a vanishing group velocity, and, therefore, that this state borders a stop gap. It also implies that  $\varphi_m$  can be chosen real (see also Appendix A). The form of  $\Lambda_{l,m}$  in Eq. (2.6) prescribes that  $m$  only couple to states  $l$  with identical crystal momentum. This means in the present situation that  $\varphi_l$  is also real and that  $l$  borders a stop gap as well. Since  $z$  is the spatial coordinate in the frame traveling with the group velocity  $\omega'_m$ , our restrictions allow us to replace  $z$  by  $x$  in Eqs. (2.4) and (2.7).

The restriction to harmonic time dependence implies that we look for envelope functions of the form

$$\bar{a}(x,t) = \psi(x)e^{-i\delta t}. \quad (2.9)$$

Using this expression, Eq. (2.7) can be rewritten as

$$\delta\psi + \frac{1}{2} \omega_m'' \frac{d^2\psi}{dx^2} + \alpha_m |\psi|^2 \psi = 0. \quad (2.10)$$

Under the two restrictions mentioned above, the electric field can now finally be expressed in  $\psi$  as

$$E(x,t) = \left[ \psi(x) [\sqrt{L}\varphi_m(x)] + \sum_{l \neq m} \Lambda_{l,m} d \left[ \frac{d\psi(x)}{dx} \right] [\sqrt{L}\varphi_l(x)] \right] \times e^{-i\omega t} + \text{c.c.}, \quad (2.11)$$

where  $\omega = \omega_m + \delta$  is the angular frequency of the incoming radiation. The angular frequency  $\delta$ , which is much smaller than either  $\omega$  or  $\omega_m$ , can thus be interpreted as a detuning between the frequencies of the incoming radiation and that associated with the state  $m$ .

The general solutions to Eq. (2.10) can be written in terms of Jacobi elliptic functions. In the limiting case in which the envelope function  $\psi$  is real, however, the solution is simply

$$\psi(x) = A \operatorname{sech}(Bx), \quad (2.12)$$

where the parameters  $A$  and  $B$  are defined as

$$A = (-2\delta/\alpha_m)^{1/2}, \quad (2.13a)$$

$$B = (-2\delta/\omega_m'')^{1/2}. \quad (2.13b)$$

Equation (2.12) represents a soliton at rest in space, with a height  $A$  and with a width of about  $1/B$ . In I we showed that the field given in Eqs. (2.11) and (2.12) has the same properties as the gap solitons found in numerical studies by Chen and Mills.<sup>2,3</sup> Because the reality of the envelope given by Eq. (2.12), and the fact that we have restricted ourselves to Bloch functions with this same property (see Appendix A), we find that the energy flow associated with the field in Eqs. (2.11) with (2.12) vanishes. For small, but finite, values of the energy flow,

however, the envelope function still resembles the functional form in Eq. (2.12) quite closely. In particular, under those circumstances, the height and the width of the envelope function of the electric field are still approximately given by  $A$  and  $1/B$ . This allows us to formulate a criterion which must be satisfied in order to observe the gap soliton experimentally.

In order to observe the soliton and its related effects (described in I), it is necessary that the length of the stack  $D$  is larger than the width of the soliton, or

$$D \gtrsim 1/B . \quad (2.14)$$

Through Eqs. (2.13) we can express this inequality in terms of the group-velocity dispersion, the effective nonlinearity and the maximum amplitude  $A$  of the envelope function, as

$$\alpha_m A^2 D^2 \gtrsim \omega_m'' . \quad (2.15)$$

We now make use of the fact that Bloch functions are of order unity, so that we can define a maximum time-averaged Poynting vector amplitude  $S_d$  in the medium by

$$\frac{c\bar{n}}{2\pi} A^2 = S_d , \quad (2.16)$$

where  $\bar{n}$  is the average linear refractive index of the stack. The fact that the Bloch functions are of order unity also allows us to approximate by  $\alpha_m \approx 6\pi\omega_m\chi^{(3)}$ , where  $\chi^{(3)}$  is an average nonlinearity [cf. Eq. (2.8)]. Defining the nonlinear correction to the index of refraction to be  $n^{(2)}S_d$ , we easily find that  $\chi^{(3)} = c\bar{n}^2 n^{(2)} / (12\pi^2)$ . These substitutions give then

$$\omega_m S_d \bar{n} n^{(2)} D^2 \gtrsim \omega_m'' . \quad (2.17)$$

To simplify this expression we now use the Kogelnik approximation.<sup>5</sup> This approximation, which is valid in the limit of small (linear) refractive-index variations, allows us to find approximate Bloch functions and dispersion relations (see Appendix A). Within this approximation we find the group-velocity dispersion to be  $\omega_m'' \approx 2c^2 / \bar{n}(\Delta n)\omega_m$  [Eq. (A12)], where  $(\Delta n)$  is the first Fourier component of the (linear) refractive-index profile [Eq. (A2)]. The final expression is now obtained by writing  $\omega_m = \pi c / d\bar{n}$ , since  $m$  is at the edge of the Brillouin zone, giving rise to the condition

$$(\Delta n)(n^{(2)}S_d)\mathcal{N}^2 \gtrsim 1 , \quad (2.18)$$

where  $\mathcal{N}$  is the number of periods in the stack. Since  $S_d$  is the maximum energy flux in the material and we want to avoid optical damage, we can take  $S_d$  to be the nonlinear change in the refractive index just below the damage threshold.

It was mentioned above that the estimate in Eq. (2.18) only holds if the energy flow is small enough so that the actual envelope function is given by hyperbolic secant in Eq. (2.12) to a reasonable approximation. This notion can be quantified by the condition that the maximum amplitude of the actual envelope function, which was denoted by  $\sqrt{I_+}$  [see Eq. (3.21) in I], does not differ very much from the value  $A$  for the soliton in Eq. (2.12), or

$$I_+ \approx A^2 . \quad (2.19)$$

By using the definition of  $I_+$  in I and the application of the Kogelnik approximation, it is found that Eq. (2.19) can be rewritten as a condition on the energy flow

$$\frac{\bar{n}}{(\Delta n)} (n^{(2)}S) \lesssim \left[ \frac{\delta}{\omega_{\text{gap}}} \right]^2 , \quad (2.20)$$

where  $\omega_{\text{gap}}$  is the size of the gap which borders the state  $\varphi_m$ . Notice that the right-hand side of this equation is much smaller than unity.

Equation (2.18) gives a necessary condition for the experimental observation of the gap soliton and its related effects. It is not sufficient, however, since, given the length of the structure and the energy flow, the detuning  $\delta$  should have a certain minimum value as well. If  $\delta$  were very small,  $B$  would be very small as well, and the exponential growth of the envelope function, starting from the back surface, would not be rapid enough to reach high-field intensities within the length of the stack. The above discussion avoids this matter altogether, since the detuning is eliminated in the derivation leading to Eq. (2.18). Consequently, the detuning is implicitly assumed to be chosen in an optimal way. The following argument gives us a relation between  $\delta$  and  $\mathcal{N}$ , and between  $\delta$  and  $S_d$ . An optimum detuning can be estimated by approximating the envelope function in the wings of the field distribution by a simple exponential, which is correct in the limit in which the nonlinearity vanishes. In I it was demonstrated that the amplitude of the envelope function is about  $\sqrt{I_m}$  at the back of the structure. We further know that the nonlinearity starts to play a role when the amplitude of the envelope reaches the value  $A$ . The optimum value of the detuning of the structure can thus be deduced from

$$\sqrt{I_m} e^{BD} \approx A . \quad (2.21)$$

If  $\delta$  were larger than the value following from this condition, the maximum field strength would be larger than necessary, thus causing unnecessary optical damage. If  $\delta$  were smaller, on the other hand, the field strength would not grow rapidly enough within the stack for the nonlinearity to be important. By Eqs. (3.12), (3.22), and (4.8) in I, we know that  $I_m$  is proportional to the energy flow. Again using the Kogelnik approximation, Eq. (2.21) can be rewritten as

$$\mathcal{N} \approx \left[ \frac{1}{4\pi} \frac{c}{\delta d} \frac{1}{(\Delta n)} \right]^{1/2} \ln \left[ \frac{16}{3\pi^2} \frac{\delta d}{c} \bar{n}^3 \frac{1}{n^{(2)}S} \right] . \quad (2.22)$$

This complicated  $\delta$  dependence is due to the fact that  $\delta$  enters both  $A$  and  $B$ .

Finally, we address the validity of the envelope function in the form of Eq. (2.11), with a single principal term. In Sec. VI of I, we presented a criterion [Eqs. (6.4) and (6.5)] for this validity. In the form as given in I it is not very practical, however, and we have rewritten it in a form which is similar to Eq. (2.20) using the various definitions in I and using the Kogelnik approximation. It is found that

$$\frac{\bar{n}}{(\Delta n)}(n^{(2)}S) \ll 1. \quad (2.23)$$

This means that the refractive-index change due to the nonlinearity should be much smaller than the linear index differences in the stack. The inequality in Eq. (2.23) is very stringent, and our numerical experience indicates that it should be interpreted in terms of large factors (at least about 100 or 1000 or so). Further, we have found that for practical situations, the conditions in Eq. (2.20) and (2.23) on the energy flow are essentially similar. We thus see that Eqs. (2.18) and Eq. (2.22) assure that our gap soliton and its associated phenomena can be observed experimentally. These conditions, however, were derived under the condition that the energy flow is small enough to satisfy Eq. (2.20) [or, equivalently, Eq. (2.23)].

### III. EQUIVALENCE TO THE EFFECTIVE-MASS FORMALISM

In this section we show the equivalence of our envelope-function approach to the effective-mass approximation,<sup>4,11</sup> in the limit of harmonic time dependence and vanishing group velocity. There are several reasons for stressing this relation. First, the effective-mass approximation is a well-understood tool in solid-state physics, and the experience with this formalism helps in the application of the envelope-function approach to problems in nonlinear optics. More important, the derivation of the effective-mass formalism is quite different than the method of multiple scales we employed in I. For this reason, it sheds a different light on the limitations of the envelope-function approach, and on the validity of the results. Actually, the derivation of the envelope-function equation using the effective-mass formalism requires that the envelope function should vary slowly on the scale of three lattice constants, rather than on the scale of a single one. This result is not obtained when using the method of multiple scales. Finally, the effective-mass formalism points the way to more general electromagnetic fields for which a description with more than a single Bloch function is required.

The derivation of the equations of the effective-mass approximation has been widely published.<sup>4,11</sup> These derivations, however, are based on the Schrödinger equation, whereas our present interest lies in the wave equation for the electric field [Eq. (2.1)]. The two derivations are nevertheless quite similar, so we only give the major steps. Instances where the quantum-mechanics treatment differs from the present wave treatment are highlighted.

The effective-mass approximation makes use of a set of functions  $\chi_{n,k}(x)$ , which are defined in terms of Bloch functions at a convenient location  $k_0$  in the Brillouin zone,<sup>4,11</sup>

$$\chi_{n,k}(x) = e^{i(\Delta k)x} \varphi_{n,k_0}(x), \quad (3.1)$$

where

$$(\Delta k) = k - k_0. \quad (3.2)$$

The parameter  $k_0$  is selected such that the dispersion curve has an extremal point at this position in the Brillouin zone (BZ). In our simple one-dimensional geometry, this implies that  $k_0$  is either in the middle of the Brillouin zone or at an edge. In more general, three-dimensional geometries, however,  $k_0$  can be anywhere in the Brillouin zone. This choice of  $k_0$  assures that the Bloch functions associated with this position have a definite parity.<sup>4</sup> Note that this choice of  $k_0$  is also consistent with our assumption of vanishing group velocity.

Just like the Bloch functions, the set of functions  $\chi_{n,k}(x)$  forms a complete, orthogonal set, with the dielectric function as a kernel [cf. Eq. (2.3)]; thus

$$\begin{aligned} \int_0^L \chi_{n,k}^*(x) \epsilon(x) \chi_{m,q}(x) dx &\equiv [nk | \epsilon | mq] \\ &= \delta_{n,m} \delta(k - q) \end{aligned} \quad (3.3)$$

in simplified notation.<sup>12</sup> From now on, the square brackets refer to matrix elements with respect to  $\chi_{n,k}(x)$ , whereas the usual bra-ket notation is reserved for matrix elements between Bloch functions.

We now consider the wave equation [Eq. (2.1)] again and neglect third-harmonic generation by the optical nonlinearity. In the limit of harmonic time dependence, it can be written as

$$-c^2 \frac{d^2 E(x)}{dx^2} - 12\pi\omega^2 \chi^{(3)}(x) |E(x)|^2 E(x) = \epsilon(x) \omega^2 E(x). \quad (3.4)$$

The total electric field is now expanded in terms of the functions  $\chi_{n,k}(x)$ , thus<sup>4,11,12</sup>

$$E(x) = \sum_n \int_{\text{BZ}} dq A_n(q) \chi_{n,q}(x). \quad (3.5)$$

Substituting this expansion into the wave equation, left multiplying by  $\chi_{m,k}^*$ , and integrating over the normalization length  $L$  it is then found that the functions  $A_n(q)$  satisfy

$$\begin{aligned} &[\omega_{m,k_0}^2 + c^2(\Delta k)^2 \langle m | m \rangle] A_m(k) \delta_{n,m} \delta(k - q) + \sum_{\substack{n \\ n \neq m}} [2c(\Delta k) \langle m | \Omega | n \rangle + c^2(\Delta k)^2 \langle m | n \rangle] A_n(k) \delta(k - q) \\ &+ \frac{2\omega^2}{\omega_m} \sum_{n_1} \sum_{n_2} \sum_{n_3} \int dq_1 \int dq_2 \int dq_3 A_{n_1}(q_1) A_{n_2}^*(q_2) A_{n_3}(q_3) \alpha_{m,n_1,n_2,n_3} \delta(k - q_1 + q_2 - q_3) \\ &= \omega^2 A_m(k) \delta_{n,m} \delta(k - q), \end{aligned} \quad (3.6)$$

where  $\alpha_{m,n_1,n_2,n_3}$  is an effective nonlinearity, very much like  $\alpha_m$  introduced in Eq. (2.8),

$$\alpha_{m,n_1,n_2,n_3} = 6\pi\omega_m \int_0^L dx \chi^{(3)}(x) \varphi_{m,k_0}^*(x) \varphi_{n_1,k_0}(x) \varphi_{n_2,k_0}^*(x) \varphi_{n_3,k_0}(x). \quad (3.7)$$

Notice that the matrix elements in Eq. (3.6) are those with respect to the Bloch functions at  $k_0$ . Because the orthogonality relations include the dielectric function [Eq. (2.3)], matrix elements such as  $\langle m|n \rangle$  are not known *a priori* in general. This constitutes a major difference with applying the effective-mass formalism to the Schrödinger equation. We now rewrite Eq. (3.6) symbolically as

$$\mathcal{H}A = \omega^2 A . \quad (3.8)$$

Just as in the standard derivation of the effective-mass formalism,<sup>4,11</sup> we introduce a canonical transformation to reduce the size of the off-diagonal elements in the matrix representation of the operator  $\mathcal{H}$  through

$$A = e^S B , \quad (3.9)$$

which gives rise to

$$(e^{-S} \mathcal{H} e^S) B \equiv \bar{\mathcal{H}} B = \omega^2 B . \quad (3.10)$$

Expanding the operator  $e^S$  in a power series in  $S$ , it is found that

$$\bar{\mathcal{H}} = \mathcal{H} + [\mathcal{H}, S] + \frac{1}{2} [[\mathcal{H}, S], S] + \dots , \quad (3.11)$$

where  $[C_1, C_2]$  denotes the commutator of  $C_1$  and  $C_2$ . We assume that  $S$  is small in the sense that the expansion leading to Eq. (3.11) is valid.

Now define the operators  $\mathcal{H}_0$  and  $\mathcal{H}_1$  through their representation in the basis of the functions  $\chi_{n,k}(x)$ , as follows:

$$[mk|\mathcal{H}_0|nq] = [\omega_{m,k_0}^2 + c^2(\Delta k)^2 \langle m|m \rangle] \delta_{n,m} \delta(k-q) \quad (3.12)$$

and

$$[mk|\mathcal{H}_1|nq] = [2c(\Delta k) \langle m|\Omega|n \rangle + c^2(\Delta k)^2 \langle m|n \rangle] \times \delta(k-q)(1 - \delta_{n,m}) . \quad (3.13)$$

For now we do not consider the nonlinear term in Eq. (3.5), but return to it below. As in the effective-mass formalism, the size of the off-diagonal terms (associated with the operator  $\mathcal{H}_1$ ) can be reduced in size if the operator  $S$  is chosen such that

$$\mathcal{H}_0 + [\mathcal{H}_1, S] = 0 . \quad (3.14)$$

Using this equation, and the definition of the operators  $\mathcal{H}_0$  and  $\mathcal{H}_1$ , we see that the matrix elements of  $S$  are

$$[mk|S|nq] = \frac{[2c(\Delta k) \langle m|\Omega|n \rangle + c^2(\Delta k)^2 \langle m|n \rangle]}{\omega_{n,k_0}^2 - \omega_{m,k_0}^2} \times \delta(k-q)(1 - \delta_{n,m}) . \quad (3.15)$$

In the effective-mass approximation,<sup>4,11</sup> one is usually only interested in terms of order  $(\Delta k)^2$  or lower, so that higher-order terms in  $(\Delta k)$  are dropped. We follow this convention presently. It should be mentioned, however, that this limitation is not strictly necessary.<sup>13</sup> Using Eq. (3.15) we can now find the matrix elements of the transformed Hamiltonian  $\bar{\mathcal{H}}$  to be

$$[mk|\bar{\mathcal{H}}|nq] = \left[ \omega_{m,k_0}^2 + c^2(\Delta k)^2 \left( \langle m|m \rangle + 4 \sum_{\substack{r \\ r \neq m}} \frac{|\langle r|\Omega|m \rangle|^2}{\omega_m^2 - \omega_r^2} \right) \right] \delta_{n,m} \delta(k-q) , \quad (3.16)$$

where again only terms up to  $(\Delta k)^2$  were retained. The off-diagonal terms of the transformed Hamiltonian are proportional to  $(\Delta k)^2$ , so that these contribute only to order  $(\Delta k)^4$  to the energies. But according to Eq. (A8) from I, the right-hand side of Eq. (3.16) is exactly the angular frequency at  $k$  to second order in  $(\Delta k)$ . Equation (3.10) can thus simply be rewritten as

$$\hat{\omega}^2((\Delta k)^2) B((\Delta k)) = \omega^2 B((\Delta k)) , \quad (3.17)$$

where the  $k$  dependence of the expansion coefficients  $B$  was written explicitly. It is important in Eq. (3.17) to distinguish the number  $\omega^2$ , which is the square of the angular frequency, from the function  $\hat{\omega}^2(\dots)$ , which gives the lowest-order expansion of the dispersion curve about  $k_0$ . It is now customary to transform this equation to real space by introducing the function

$$F_m(x) = \int_{\text{BZ}} e^{i(\Delta k)x} B_m((\Delta k)) d(\Delta k) , \quad (3.18)$$

which results in

$$\hat{\omega}^2 \left[ -i \frac{d}{dx} \right] F_n(x) = \omega^2 F_n(x) . \quad (3.19)$$

The notation  $\hat{\omega}^2(-id/dx)$  in this equation means that  $(\Delta k)$  should be replaced by  $-id/dx$  in the expression for  $\omega^2((\Delta k))$ , to second order in  $(\Delta k)$ .<sup>4,11</sup> This prescription turns  $\hat{\omega}^2$  into an operator, thus turning the algebraic equations in Eq. (3.17) into a set of second-order differential equations [Eq. (3.19)]. From Eq. (3.18) we see that all Fourier components which lie outside the first BZ are neglected. For this reason, the formalism is only reliable if the functions  $B_m((\Delta k))$  are negligible in the vicinity of the BZ edge.<sup>4,11</sup> In real space this means that the functions  $F_m(x)$  have to be slowly varying on the scale of the lattice constant. For this reason we refer to these functions as envelope functions.

Next we find the relation between the envelope functions and the electric field  $E(x)$ . Combining Eqs. (3.5), (3.1), and (3.9) and, again, expanding the exponential, we find that

$$E(x) = \sum_n \int_{\text{BZ}} dk (1 + S + \dots) B_n((\Delta k)) \times e^{i(\Delta k)x} \varphi_{n,k_0}(x) . \quad (3.20)$$

Using the completeness of the functions  $\chi_{n,k}(x)$  and Eq.

(3.15) for the matrix elements of the operator  $S$ , we finally find for the electric field

$$E(x) = \sum_n \left[ F_n(x) \varphi_{n,k_0}(x) + \sum_{\substack{p \\ p \neq n}} \Lambda_{p,n} d \frac{dF_p(x)}{dx} \varphi_{p,k_0}(x) \right], \quad (3.21)$$

in which the coupling coefficient  $\Lambda_{p,n}$  is defined in Eq. (2.5). Comparing to Eq. (2.11) we see that our present results are identical to those obtained using the method of multiple scales, if the summation over  $n$  in Eq. (3.21) is

$$\frac{2\omega^2}{\omega_m} \sum_{n_1} \sum_{n_2} \sum_{n_3} \alpha_{m,n_1,n_2,n_3} \int dq_1 \int dq_2 \int dq_3 B_{n_1}(q_1) B_{n_2}^*(q_2) B_{n_3}(q_3) \int_{\text{BZ}} dk e^{i(\Delta k)x} \delta(k - q_1 + q_2 - q_3) \quad (3.22)$$

in real space. Using the definition of the envelope functions in Eq. (3.18), we finally find for the contribution of the nonlinear term, in real space

$$\frac{2\omega^2}{\omega_m} \sum_{n_1} \sum_{n_2} \sum_{n_3} F_{n_1}(x) F_{n_2}^*(x) F_{n_3}(x) \alpha_{m,n_1,n_2,n_3}, \quad (3.23)$$

which is the final form for this contribution. Notice, however, that the integration over  $k$  in Eq. (3.22) was limited to the first Brillouin zone. By the  $\delta$  function in this equation, this then implies that  $|q_1| + |q_2| + |q_3|$  is limited to the first Brillouin zone as well. From this we can conclude that the contribution of the nonlinearity is correctly given by Eq. (3.23), as long as the Fourier coefficients do not extend beyond about a third of the size of the Brillouin zone. Equivalently, the envelope functions are required to vary not faster than on a scale of three lattice constants. We return to this matter below.

We now return to Eq. (3.19) to find a differential equation for the envelope functions  $F_n(x)$ . For this purpose we apply the prescription to replace  $(\Delta k)$  by  $-id/dx$  and find, after adding the nonlinear contribution,

$$\omega_m^2 F_m(x) + \omega_m'' \omega_m \left[ -i \frac{d}{dx} \right]^2 F_m(x) + 2 \frac{\omega^2}{\omega_m} \alpha_{m,m,m,m} |F_m(x)|^2 F_m(x) = \omega^2 F_m(x), \quad (3.24)$$

where only the nonlinear self-phase modulation was included. In the final step, we remember that  $\delta = \omega - \omega_m$  from Sec. II and that  $\delta$  is much smaller than both  $\omega$  and  $\omega_m$ . Under these conditions, we see that Eq. (3.24) can finally be rewritten as

$$\delta F_m(x) + \frac{1}{2} \omega_m'' \frac{d^2}{dx^2} F_m(x) + \alpha_{m,m,m,m} |F_m|^2 F_m(x) = 0. \quad (3.25)$$

After a trivial rescaling to obtain similar definitions of the effective nonlinearity, this equation is identical to Eq. (2.10). The latter equation is the result using the method

restricted to a single branch  $m$ . It now remains to be shown that the present envelope function  $F(x)$  satisfies the same equation as the envelope function  $\psi(x)$  used in I and in Sec. II. To this end we first evaluate the nonlinear term appearing in Eq. (3.6).

We know that, to lowest order, the coefficients  $A_m(q)$  and  $B_m(q)$  are equal. Since we are only interested in the lowest-order contribution, the distinction between these two sets of expansion coefficients may be neglected in evaluating the influence of the nonlinearity. Thus, taking the nonlinear term in Eq. (3.6), and determining the Fourier transform as in Eq. (3.18), it is found that it contributes

of multiple scales in the limit in which the time dependence is harmonic and the incoming radiation is tuned to a stop gap.

We thus see that the method of multiple scales and the envelope-function formalism give identical results for harmonic time dependence. The former method is more powerful, however, as it can be used for a general time dependence, a possibility not offered by the effective-mass formalism. An important conclusion from the present exercise, however, is that for our approach to be reliable, the envelope function may only contain Fourier components within the inmost one-third of the Brillouin zone. This requirement did not emerge in this explicit form from the derivation using the method of multiple scales. This limitation is intrinsic in nature, and holds, no matter how many branches of the dispersion relation are explicitly included and no matter to what order in  $(\Delta k)$  (or, equivalently, in  $\mu$  in the method of multiple scales) the expansion is followed through.

Finally, comparing Eqs. (3.21) and (2.11) we see how our formalism should be modified to include multiple branches of the dispersion relation explicitly. Such a generalization is the subject of Sec. IV.

#### IV. TWO-BRANCH APPROXIMATIONS

The ansatz in Eq. (2.2), on which the analysis in I was based, implies that a single Bloch function dominates the behavior of the fast component of the electric field. This was justified in I, since the incoming radiation was tuned to one of the stop gaps, very close to one of its edges. In such a treatment, the influence of the remaining states of the periodic structure is summarized in the group-velocity dispersion and in the companion component of the electric field.

It is clear, however, that if the incoming radiation is tuned to the middle of a stop gap, it is no longer possible to define a single dominating state; rather, both edges of the stop gap are equally important. This is just a single example of a condition under which the ansatz in Eq. (2.3) is no longer justified, and a more general approach

has to be taken. This generalization requires that the principal component of the electric field consists of two, or more, of the terms as in Eq. (2.3). In Sec. III, the effective-mass treatment gave the relevant equations for the case of harmonic time dependence [see, e.g., Eq. (3.21)]. After a derivation which is very similar to that in I, equations for a general time dependence can be obtained by application of the method of multiple scales. The final, most general results of such a treatment contain a very large number of terms and are not very informative. In this section, therefore, we concentrate on two possible experimental situations in which a two-branch treatment is required and we derive the pertinent equations for these situations. The first of these was described above and pertains to the situation in which the incoming radiation is tuned to the middle region of one of the stop gaps of the periodic structure. Since the states at both edges of the gap have to be treated on an equal footing, both appear in the principal component of the electric field. The interaction between these two bands is then considered dynamically, rather than through the group-velocity dispersion. The second of the situations in which a two-branch approach is necessary is when third-harmonic generation by the Kerr nonlinearity is included. In this case, explicit inclusion of the relevant Bloch functions at both  $\omega$  and  $3\omega$  is required. At present we only consider the very interesting situation in which both  $\omega$  and  $3\omega$  lie within stop gaps. It should be noted here that a multibranch approach as used in the present section is very common in solid-state physics. In the effective-mass treatment of superlattices, for example, it is often necessary to include two or more bands explicitly.<sup>14</sup>

We now thus start with the derivation for the equations for the slowly varying envelope when the incoming radiation falls in the middle region of a stop gap. In analogy to Eqs. (2.3) and (3.21), we start with the ansatz that the principal component of the electric field can now be

written as

$$e_1(x,t) = a_1(x,t)\varphi_{m_1}(x)e^{-i\omega_{m_1}t} + a_2(x,t)\varphi_{m_2}(x)e^{-i\omega_{m_2}t} + \text{c.c.}, \quad (4.1)$$

which satisfies the equation originating from the terms which are linear in  $\mu$  (see Sec. II and I). Notice that the states  $\varphi_{m_1}$  and  $\varphi_{m_2}$  now form the two edges of a single stop gap. Just as in I, we now continue with terms proportional to  $\mu^2$  and find first that the envelopes  $a_1$  and  $a_2$  each travel with the group velocity associated with  $\varphi_{m_1}$  and  $\varphi_{m_2}$ , respectively. Since these states border a stop gap, it is found that both envelopes are at rest in space. It is further found that the companion component can be written as the sum of two terms of the form

$$\sum_{l \neq m_1, m_2} \Lambda_{l, m_i} d \frac{\partial a_i}{\partial x} \varphi_l(x) e^{-i\omega_{m_i}t} + \text{c.c.}, \quad (4.2)$$

with  $i=1,2$ . Next the terms to third order in  $\mu$  are considered. In evaluating the expressions obtained in this way, it is important to realize that, since  $m_1$  and  $m_2$  border the same stop gap, these states have opposite parity, so that matrix elements such as  $\langle m_1 | m_2 \rangle$  vanish. Another consequence of the fact that  $m_1$  and  $m_2$  border the same stop gap, is that the angular frequency  $\Delta$ , defined as

$$\Delta = \omega_{m_2} - \omega_{m_1}, \quad (4.3)$$

which may be generated by the optical nonlinearity, is not necessarily large. This implies that terms with this time dependence have to be retained. After a calculation which is very similar to that in I, it is then found that the two envelope functions satisfy the equations

$$i \frac{\partial a_1}{\partial t} + \frac{1}{2} \bar{\omega}''_{m_1, m_1} \frac{\partial^2 a_1}{\partial x^2} + ic \frac{\langle m_1 | \Omega | m_2 \rangle}{\omega_{m_1}} \frac{\partial a_2}{\partial x} e^{-i\Delta t} + \alpha_{1111} |a_1|^2 a_1 + 2\alpha_{2211} |a_2|^2 a_1 + \left[ 1 + 2 \frac{\Delta}{\omega_{m_1}} \right]^2 \alpha_{1122} a_1^* a_2^2 e^{-2i\Delta t} = 0 \quad (4.4a)$$

and

$$i \frac{\partial a_2}{\partial t} + \frac{1}{2} \bar{\omega}''_{m_2, m_2} \frac{\partial^2 a_2}{\partial x^2} + ic \frac{\langle m_2 | \Omega | m_1 \rangle}{\omega_{m_2}} \frac{\partial a_1}{\partial x} e^{i\Delta t} + \alpha_{2222} |a_2|^2 a_2 + 2\alpha_{1122} |a_1|^2 a_2 + \left[ 1 - 2 \frac{\Delta}{\omega_{m_2}} \right]^2 \alpha_{2211} a_2^* a_1^2 e^{2i\Delta t} = 0. \quad (4.4b)$$

In these equations we used the diagonal elements of the operator defined by

$$\bar{\omega}''_{m_i, m_j} = c^2 \langle m_i | m_j \rangle + 4c^2 \sum_{l \neq m_1, m_2} \frac{\langle m_i | \Omega | l \rangle \langle l | \Omega | m_j \rangle}{\omega_{m_j}^2 - \omega_l^2}. \quad (4.5)$$

The diagonal elements bear a close resemblance to the group-velocity dispersion [see Eq. (A10) in I]. Notice, however, that the summation in Eq. (4.5) excludes both  $m_1$  and  $m_2$ . The interaction between these two states is now, however, dynamically included in the coupled differential equations through the linear cross terms. Further, for convenience the nonlinear coefficients  $\alpha_{m_i, m_j, \dots}$ , [Eq. (3.7)] were written as  $\alpha_{ij} \dots$ . Equations

very similar to Eqs. (4.4) have been derived for light propagation in birefringent nonlinear dispersive media.<sup>15</sup> Notice that since  $m_1$  and  $m_2$  define the border of a stop gap, the associated Bloch functions are real (Appendix A). The first nonlinear terms in Eq. (4.4) describe the nonlinear self-phase modulation of the two envelope functions. Such a term also appears in our earlier treatment [Eq. (2.7)]. The second and third nonlinear terms are new, and describe the cross-phase modulation between  $a_1$  and  $a_2$ . The second terms simply describe the fact that the intensities associated with each of the two envelope functions change the refractive index of the stack as seen by the other envelope function. A similar effect is also observed in optical fibers.<sup>16</sup> The third terms describe a similar interaction, except that the mutual phase of the two envelopes is now important. Such terms do not play a role in fibers because  $\Delta$  is not small in such media. We briefly return to Eqs. (4.4) in Sec. V.

We next consider the case of third-harmonic generation under the condition, considered below, that both  $\omega$  and  $3\omega$  fall within stop gaps. In the most general case, we would have to include four Bloch functions explicitly, namely the two Bloch functions bordering the gap at  $\omega$  and those at  $3\omega$ . The ensuing four coupled nonlinear partial differential equations, however, are not very informative and are not written out here. Instead, we consider the situation in which both  $\omega$  and  $3\omega$  lie close to one of the band edges, so that an ansatz as in Eq. (4.1) is justified. Notice, however, that now  $m_1$  and  $m_2$  border

different stop gaps, in contrast with the situation above. Within these restrictions, still two possibilities remain—either  $m_1$  and  $m_2$  have the same parity, or their parity is opposite. We consider the former possibility only, since, as we see later, the latter does not give rise to third-harmonic generation at all. This means that both  $m_1$  and  $m_2$  are either close to the bottom or to the top of a stop gap.

With the restrictions from the preceding paragraph and an ansatz as in Eq. (4.1), we again apply the method of multiple scales. Considering the terms proportional to  $\mu^2$ , we again conclude that both envelopes  $a_1(x,t)$  and  $a_2(x,t)$  travel with their associated group velocity. In the present situation, this again means that both are at rest in space. It is further found that  $e_2$  is given by the sum of two terms as in Eq. (4.2). Considering the terms to third order in  $\mu$ , we find significant differences with the previous case. First, we obtain different linear terms since the two states  $m_1$  and  $m_2$  have now identical parity, rather than opposite. Moreover, we obtain different nonlinear terms, since different linear combinations of the angular frequencies  $\omega_{m_1}$  and  $\omega_{m_2}$  are now small. For definiteness, we label the band edge closest to  $\omega$  by  $m_1$ , and the edge closest to  $3\omega$  by  $m_2$ .

For the case of general time dependence of the envelope functions, the following two coupled equations are found:

$$i \frac{\partial a_1}{\partial t} + \frac{1}{2} \omega''_{m_1} \frac{\partial^2 a_1}{\partial x^2} + \frac{1}{2} \tilde{\omega}''_{m_1, m_2} \frac{\partial^2 a_2}{\partial x^2} e^{-i\Delta t} + \alpha_{1111} |a_1|^2 a_1 + 2\alpha_{1122} |a_2|^2 a_1 + \left(1 - \frac{\gamma}{\omega_{m_1}}\right)^2 \alpha_{1112} a_1^* a_2^2 e^{i\gamma t} = 0, \quad (4.6a)$$

and

$$i \frac{\partial a_2}{\partial t} + \frac{1}{2} \omega''_{m_2} \frac{\partial^2 a_2}{\partial x^2} + \frac{1}{2} \tilde{\omega}''_{m_2, m_1} \frac{\partial^2 a_1}{\partial x^2} e^{i\Delta t} + \alpha_{2222} |a_2|^2 a_2 + 2\alpha_{2211} |a_1|^2 a_2 + \frac{1}{3} \left(1 + \frac{\gamma}{\omega_{m_2}}\right)^2 \alpha_{2111} a_1^3 e^{-i\gamma t} = 0, \quad (4.6b)$$

where

$$\gamma = 3\omega_{m_1} - \omega_{m_2} \quad (4.7)$$

is a low angular frequency (see Fig. 2). Equations similar to Eq. (4.6) have been derived for the mechanical problem of transverse elastic waves traveling along a uniform bar that rests on a nonlinear elastic foundation.<sup>17</sup> The element  $\tilde{\omega}''_{m_i, m_j}$  was defined through Eq. (4.5). We now demonstrate that the terms in Eqs. (4.6) which contain these matrix elements are negligibly small. First, since  $m_1$  and  $m_2$  belong to different gaps, both  $\langle m_1 | m_2 \rangle$  and either of  $\langle l | m_1 \rangle$  or  $\langle l | m_2 \rangle$  are very small, so that according to Eq. (4.5),  $\tilde{\omega}''_{m_i, m_j}$  is very small as well. Actually, in the Kogelnik approximation these small matrix elements vanish entirely. Second, the angular frequency  $\Delta$  is not small in the present situation, so that the (small) contributions of the terms under consideration average out over times longer than about  $1/\Delta$ . Finally, these off-diagonal terms are of second order in  $d/dx$ , so that, in

analogy to matrix perturbation theory, their effective contribution is only of fourth order. These terms will henceforth be neglected.

Equations (4.6) for third-harmonic generation look very similar to Eqs. (4.4), derived earlier. A most important difference, however, is that, apart from a negligibly small linear coupling term, the envelopes in Eqs. (4.6) are only coupled through the nonlinearity. In fact, there are two such nonlinear interaction terms. The fifth terms in each of Eqs. (4.6) describe the cross-phase modulation through the intensity associated with the envelope functions. Similar terms appear in Eq. (4.4). The last terms in the two equations have no counterpart in Eq. (4.4) and describe third-harmonic generation [Eq. (4.6b)] and mixing of  $3\omega$  and  $\omega$  [Eq. (4.6a)]. A further interpretation of these equations is given in Sec. V in which we show that Eqs. (4.6) can be derived from a conservative Hamiltonian.

A final point to be addressed in this section is the restriction to select  $m_1$  and  $m_2$ , such that their Bloch func-

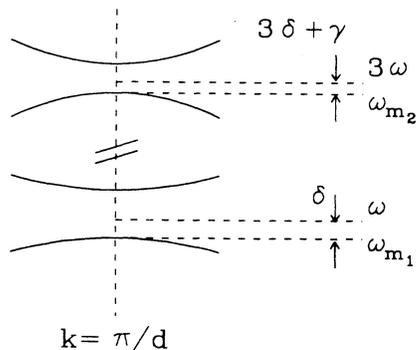


FIG. 2. Definition of the various angular frequencies for the analysis of THG. The detuning at  $3\omega$  is  $3\delta + \gamma$ , since  $\gamma \equiv 3\omega_{m_1} - \omega_{m_2}$ .

tions have the same parity. We see from Eq. (3.7) that, otherwise, the relevant nonlinear coefficients  $\alpha_{1112}$  and  $\alpha_{2111}$  for third-harmonic generation would vanish.

### V. HAMILTONIAN DESCRIPTION

The equations for the envelope functions for the situations we have studied so far [Eqs. (2.7), (4.4), and (4.6)] can, in the limit of harmonic time dependence, each formally be derived from a Lagrangian and Hamiltonian formalism. This assures us of the help of the powerful methods of theoretical mechanics<sup>18</sup> in analyzing the differential equations for the envelope functions. For example, we see that the energy flow through the system is a conserved quantity for the corresponding Hamiltonian, thus reducing the number of degrees of freedom by 1. Moreover, the study of the equivalent Hamiltonian provides insight into the origin of the various terms in the equations for the envelope function.

Equation (2.10), for example, which is obtained from Eq. (2.7) in the limit of harmonic time dependence, can be derived from the effective Lagrangian,

$$\mathcal{L} = \frac{1}{2} \frac{1}{B^2} |\psi'|^2 + \frac{1}{2} |\psi|^2 - \frac{1}{2} \frac{1}{A^2} |\psi|^4, \quad (5.1)$$

where  $A$  and  $B$  were defined in Eqs. (2.13). It is important to stress that the prime in this Lagrangian refers to a spatial, rather than to a temporal derivative, which appears in the study of mechanical systems.<sup>18</sup> It is now convenient to rewrite Eq. (5.1) in Hamiltonian form, making use of polar coordinates. The result is then

$$\mathcal{H} = \frac{1}{2} B^2 \left[ p_\psi^2 + \frac{p_\vartheta^2}{q_\psi^2} \right] - \frac{1}{2} q_\psi^2 + \frac{1}{2} \frac{1}{A^2} q_\psi^4, \quad (5.2)$$

where  $q_\psi$  and  $p_\psi$  are the conjugate variables associated with  $|\psi|$ , and  $q_\vartheta$  and  $p_\vartheta$  those associated with the polar angle  $\theta$ . We thus see that the spatial dependence of the envelope function through the stack, is formally equivalent to the temporal dependence of a particle with a mass of  $1/B^2$  in a rotationally symmetric potential of form  $-\frac{1}{2}q_\psi^2 + (1/2A^2)q_\psi^4$ . This expression assures that the particle has two turning points, which prevent escape to infinity. A sketch of a cross section of this potential is given in Fig. 3(a). In terms of our electrodynamic prob-

lem, this “potential energy” assures that the field strengths do not diverge. This should be contrasted with the case in which  $\chi^{(3)}$  has the sign such that  $A^2$  is negative [Eq. (2.13a)]. The anharmonic potential-energy term in Eq. (5.2) then changes sign so that the motion of the particle is no longer bounded. This corresponds to the instance in which, through the nonlinearity, the stack tunes itself deeper into the gap, rather than out of it. Since the potential-energy term in Eq. (5.2) is rotationally symmetric, we find in the equivalent mechanical problem that the angular momentum  $L = p_\vartheta$  is conserved.<sup>18</sup> To see to what quantity this corresponds in the original problem, we recall from I [Eq. (3.12)] that the Poynting vector through the system  $S$  can be written in terms of the envelope function as  $S = (\omega_m''/2\pi)\text{Im}(\psi^*\psi')$ . Rewriting this expression in polar coordinates and making use of the canonical variables introduced in Eq. (5.2), we simply find

$$S = \frac{\omega_m''}{2\pi} B^2 p_\vartheta = -\frac{\delta}{\pi} p_\vartheta, \quad (5.3)$$

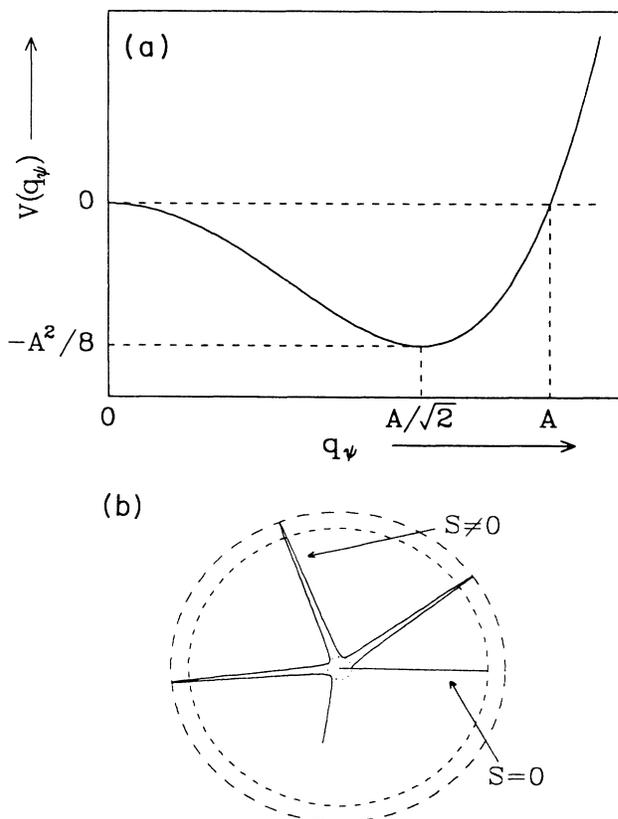


FIG. 3. (a) Radial dependence of the effective potential-energy surface of the nonlinear stack. (b) Top view of two possible orbits in this potential. The orbit labeled  $S=0$  is that of the gap soliton which transports no energy. This implies the absence of angular momentum in the equivalent mechanical system. The orbit labeled  $S \neq 0$  is that for the more general case of finite energy flow and thus finite angular momentum.

where use was made of the definition of  $B$  [Eq. (2.13)] to obtain the final expression. We thus see that the conservation of energy flow in the original problem corresponds to conservation of angular momentum in the equivalent mechanical system.

Apart from the angular momentum, the total energy is, of course, also conserved in the equivalent mechanical system. To try to find the interpretation of this conserved quantity, we first calculate the free energy of our periodic stack. Here we run into a complication as thermodynamic potentials lose their significance in externally driven systems.<sup>19</sup> For weakly nonlinear driven systems, however, one can uniquely define a *time-averaged free energy* with comparable properties.<sup>7,8,20</sup> A further complication is that electrostatics allows the definition of four such time-averaged thermodynamic potentials, which are all related through Legendre transformations.<sup>21,22</sup> As is discussed in Appendix B, the time-averaged free energy relevant to the present problem is

$$4\pi\langle \mathcal{F} - \mathcal{F}_0 \rangle = 2 \sum_i \operatorname{Re} \left[ \int_0^{B_0(\omega_i)} H_0^*(\omega_i) dB_0(\omega_i) - \int_0^{E_0(\omega_i)} D_0^*(\omega_i) dE_0(\omega_i) \right], \quad (5.4)$$

where  $E_0(\omega_i)$  is the complex amplitude of the electric field at  $\omega_i$ , defined through

$$E(x, t) = \sum_i [E_0(x, \omega_i) e^{-i\omega_i t} + \text{c.c.}], \quad (5.5)$$

and the complex amplitudes of the other fields are defined similarly. It should be mentioned that this definition can be generalized to include complex amplitudes which are slowly varying in time.<sup>20</sup>

With the field  $E(x)$  given by Eq. (2.10),  $B(x) = H(x)$  following from Maxwell's equations, and  $D(x)$  following through the relation  $D = \epsilon E + 4\pi\chi^{(3)}E^3$ , we can rewrite Eq. (5.4) in terms of the envelope function and its first derivative, as follows:

$$4\pi\langle \mathcal{F} - \mathcal{F}_0 \rangle = \frac{\omega_m \omega_m''}{\omega^2} |\psi'|^2 + \frac{\omega_m^2 - \omega^2}{\omega^2} |\psi|^2 - \frac{\alpha_m}{\omega_m} |\psi|^4 + \sum_{l \neq m} d^2 \Lambda_{l,m}^2 \frac{\omega_l^2 - \omega^2}{\omega^2} |\psi'|^2. \quad (5.6)$$

First consider the final term in this expression. Since it consists of the product of two companion contributions, it is effectively of order  $\mu^4$  (see I). However, since our derivation in I included terms up to third order in  $\mu$  only, we conclude that to be consistent, the last term in Eq. (5.6) should be dropped. Further, since  $\delta = \omega - \omega_m$  and  $\delta \ll \omega, \omega_m$ , we see that, to a good approximation,

$$-\frac{\pi\omega_m}{\delta} \langle \mathcal{F} - \mathcal{F}_0 \rangle = \frac{1}{2B^2} |\psi'|^2 + \frac{1}{2} |\psi|^2 - \frac{1}{2A^2} |\psi|^4, \quad (5.7)$$

which is identical to the Lagrangian in Eq. (5.1). The requirement that the variation of the free energy per unit area over the total length of the stack vanish,<sup>19</sup> now corresponds to the requirement that the variation of the ac-

tion integral of the equivalent mechanical system vanish.<sup>18</sup> This requirement then gives rise to the Lagrangian equations of motion in the usual way. We thus see that the free energy in our electrostatic system corresponds to the Lagrangian of the associated mechanical system, with temporal rather than spatial derivatives. Although this provides a definite interpretation of the Lagrangian, we have found no such simple interpretation of the total energy of the equivalent mechanical system. It should be stressed that the potential-energy terms in Eq. (5.7) contain contributions of both the electric and the magnetic parts of the free energy in Eq. (5.4). It would thus be incorrect to conclude that the total energy of the equivalent mechanical system would correspond to one of the thermodynamic potentials in which the contributions of the electric and magnetic fields have identical signs.

Now that we know that the electric field distribution in the nonlinear periodic stack is equivalent to the motion of a particle in a rotationally symmetric potential well, we can interpret the field profiles obtained in I in terms of the motion of the particle. As mentioned before, the form of the potential energy assures that the particle describes an orbit between two turning points [Fig. 3(a)]. In determining the outer turning point, the harmonic and the anharmonic terms in the potential energy are of equal importance. We thus come to the conclusion that the anharmonic term in the potential energy is crucial in determining the orbit of the particle. Alternatively, we can state that the nonlinearity is crucial in determining the field profile in the periodic stack.

The exact position of the two turning points can be found when the angular momentum  $L$  and the energy  $E$  of the particle are given. The electric field distribution, on the other hand, was characterized in I by two quantities, the minimum value of the square modulus of the envelope function  $I_m$  and a quantity proportional to the energy flow  $W$ . Thus  $I_m$  directly defines the inner turning point, whereas  $W$  and  $L$  must be linearly related. So  $I_m$  and  $W$ , and  $E$  and  $L$ , provide equivalent sets of information. We saw in I that the  $W$  and  $I_m$  cannot be chosen independently, and, in fact, must be linearly proportional in order to satisfy the boundary conditions at the back of the stack. Similarly, we conclude that the energy and the angular momentum of the particle cannot be chosen independently either, severely limiting the possible orbits of the particle.

Two possible orbits of the equivalent mechanical system are shown in Fig. 3(b). The gap soliton corresponds to the orbit without angular momentum, and we thus see that it originates from a delicate balancing of dispersion, which gives rise to the harmonic part of the potential energy, and the nonlinearity, which is responsible for the anharmonic part. Starting at the outer turning point, the particle will not reach the origin within a finite time. This is consistent with the asymptotic behavior of the hyperbolic secant function, which mathematically describes the gap soliton. The other orbit in Fig. 3(b) clearly has a finite period and nonzero angular momentum, and is described by Jacobi elliptic functions.

It is possible to write the coupled differential equations [Eqs. (4.4) and (4.6)] in terms of a classical mechanical

system as well. For Eq. (4.6) this is quite useful, as the corresponding Hamiltonian represents two nonlinearly coupled oscillators of the kind discussed above. Equation (4.4), however, has very strong linear coupling terms so that such a picture is not very revealing. As a consequence, considering the mechanical counterpart is not very useful in this case. We thus limit our attention in this section to a discussion of Eq. (4.6), and turn to the coupled equations describing third-harmonic generation (THG) in the stack. The arguments to show that these equations can be derived using a Lagrangian or Hamiltonian formalism are essentially identical as before, except that the expressions involved are more complicated. For this reason, only the most important results are given. The Lagrangian appropriate to Eq. (4.6) can be found to be

$$\begin{aligned} \mathcal{L} = \mathcal{L}_1 + & \left[ 3 + \frac{\gamma}{\delta} \right] \frac{\omega_{m_1}}{\omega_{m_2}} \mathcal{L}_2 - \frac{2}{A_3^2} |\psi_1|^2 |\psi_2|^2 \\ & - \frac{2}{3A_4^2} |\psi_1|^3 |\psi_2| \cos(3\vartheta_1 - \vartheta_2), \end{aligned} \quad (5.8)$$

where  $\mathcal{L}_1$  is given by Eq. (5.1) and  $\mathcal{L}_2$  by a similar expression, but with  $A$  and  $B$  replaced by  $A_2$  and  $B_2$ , respectively, where

$$A_2 = [-2(3\delta + \gamma)/\alpha_{2222}]^{1/2}, \quad (5.9a)$$

$$B_2 = [-2(3\delta + \gamma)/\omega_{m_2}'' ]^{1/2}. \quad (5.9b)$$

It should be noted that, in the present notation,  $3\delta + \gamma$  is the detuning at state  $m_2$  (see Fig. 2). Finally,  $A_3$  and  $A_4$  are defined similarly as  $A$ , but with  $\alpha_m$  replaced by  $\alpha_{1122}$  and  $\alpha_{1112}$  [Eq. (3.7)], respectively. It can be shown in the same way as before that the right-hand side of Eq. (5.8) is proportional to the time-averaged free energy [Eq. (5.4)] per unit volume, if the fields at  $\omega$  and at  $3\omega$  are both included. This is a straightforward but tedious exercise, which is not given here. It should be noted that the multiplying factor of  $\mathcal{L}_2$  in Eq. (5.8) equals unity when  $\gamma = 0$  (since under this condition  $\omega_{m_2} = 3\omega_{m_1}$ ). When this condition is not strictly satisfied, the deviation from unity is of order  $\delta/\omega \ll 1$ . For this reason, we assume henceforth for convenience that the multiplying factor equals unity exactly. It should be stressed that this restriction is by no means fundamental, and can be lifted easily if necessary.

Under this restriction we can finally write the Hamiltonian function applicable to THG, from Eq. (5.8), as

$$\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2 + \frac{2}{A_3^2} q_{\psi_1}^2 q_{\psi_2}^2 + \frac{2}{3A_4^2} q_{\psi_1}^3 q_{\psi_2} \cos(3q_{\vartheta_1} - q_{\vartheta_2}), \quad (5.10)$$

with  $\mathcal{H}_1$  given in Eq. (5.2) and  $\mathcal{H}_2$  following in analogy with  $\mathcal{L}_2$  [following Eq. (5.8)]. It can be seen from this equation that  $\mathcal{H}_1$  and  $\mathcal{H}_2$  both represent oscillators such as described earlier in this section. Although they are described by mutually different coefficients  $A$  and  $B$ , their properties are qualitatively the same. Let us denote the energy and the angular momentum of these oscillators as  $E_1$  and  $E_2$  and  $L_1$  and  $L_2$ , respectively. Of course, the

total energy of the system  $E = E_1 + E_2$  is conserved. However, because the Hamiltonian depends on the difference  $3q_{\vartheta_1} - q_{\vartheta_2}$  only, we see that in addition,  $L_1 + 3L_2$  is a conserved quantity too. Not surprisingly, it can be shown that this corresponds to the total energy flow (that at  $\omega$  and that at  $3\omega$ ) through the system. The conservation of energy flow through the system thus reduces the number of degrees of freedom of the system from 4 to 3.

We now consider the influence of the two nonlinear coupling terms in Eq. (5.10). Clearly, in the absence of these terms, the particles move independently and both  $E_1$  and  $E_2$  and both  $L_1$  and  $L_2$  are conserved separately. We now include the first coupling term in Eq. (5.10). Since this term does not have any angular dependence, it allows the exchange of energy between the two oscillators, but not that of angular momentum. In terms of the fields in the periodic stack, this term describes the cross-phase modulation of the fields at  $\omega$  and  $3\omega$  through their intensities. The second coupling term does depend on the angular variables, and so does allow the exchange of angular momentum between the particles. In terms of the fields, again, this term allows the exchange between the energy flow at  $\omega$  and  $3\omega$ , and thus describes frequency mixing, including THG. It is important to realize that the THG can be quite efficient in the periodic structure. Remember that, for each of the oscillators, the orbit of the particle was determined to a large degree by the anharmonic terms in the potential energy. Similarly, this implies that the nonlinear coupling terms influence the motion to a large degree. Conversely, we can conclude that the fields at  $\omega$  and  $3\omega$  can exchange substantial amounts of power within a characteristic length of  $1/B$ . We will contrast this behavior later with that of THG in homogeneous materials.

At this point it is good to remember the assumptions on which our analysis is based. The most important one at present is that the media were taken to be without intrinsic dispersion. When one is only interested in a monochromatic light wave (perhaps with small sidebands), this assumption is justified. When considering THG, however, this will not be true in general. Most importantly, we assumed that both  $\omega$  and  $3\omega$  fall within a stop gap, and so Eq. (5.8) is only valid if this requirement is satisfied. This requirement is quite severe, but we will see that it is far less stringent than the equivalent requirement for THG in homogeneous materials.

The relevant equations for THG in homogeneous materials are well known<sup>8</sup> and are not given here. One can show in the same way as above, that these equations can be derived from a Hamiltonian formalism too. In order to compare THG in homogeneous materials to that in the periodic structures, it is most convenient to compare the Hamiltonians for these two processes. For homogeneous materials we find that

$$\begin{aligned} \overline{\mathcal{H}} = \overline{\mathcal{H}}_1 + \overline{\mathcal{H}}_2 \\ + 4\pi\chi^{(3)}k_0^2 [3q_{E_1}^2 q_{E_2}^2 + q_{E_1}^3 q_{E_2} \cos(3q_{\vartheta_1} - q_{\vartheta_2})], \end{aligned} \quad (5.11)$$

where

$$\overline{\mathcal{H}}_1 = \frac{1}{2}(p_{E_1}^2 + p_{\vartheta_1}^2 / q_{E_1}^2) + \frac{1}{2}k_\omega^2 q_{E_1}^2 + 3\pi\chi^{(3)}k_0^2 q_{E_1}^4. \quad (5.12)$$

In these equations,  $q_{E_1}$  and  $p_{E_1}$  and  $q_{\vartheta_1}$  and  $p_{\vartheta_2}$  are the canonical variables associated with the modulus and the phase angle of the electric field at  $\omega$ , respectively. The variables with subscript 2 refer to the field at  $3\omega$ . The parameter  $k_0$  is the wave number of the radiation at  $\omega$  in vacuum, whereas  $k_\omega$  is the wave number in the material. In Eq. (5.11),  $\overline{\mathcal{H}}_2$  is defined similarly as  $\overline{\mathcal{H}}_1$ , except that  $k_\omega$  and  $k_0$  should be replaced by  $k_{3\omega}$  and  $3k_0$ , respectively. Just as before, it can be shown that the Hamiltonian in Eq. (5.11) is derivable from a free-energy argument based on the definition in Eq. (5.4). Also, again, the Hamiltonian in Eq. (5.11) implies the conservation of the total energy flow through the system.

We see from Eqs. (5.11) and (5.12) that the Hamiltonian for the homogeneous material is quite similar to that for the periodic stacks [Eqs. (5.11) and (5.2)]. The coupling terms in Eq. (5.11) are, in fact, identical in form to those in Eq. (5.10). The crucial difference, however, is that the harmonic potential-energy terms in Eq. (5.12) are positive, in contrast to those in Eq. (5.2). As a consequence, the particles oscillate around the origin where the anharmonic terms, including the last in Eq. (5.11) which gives rise to THG, are very small. We thus conclude that THG is quite inefficient, and can only be accomplished effectively when the oscillators have identical period, or, when  $k_{3\omega} = 3k_\omega$ . Of course, this is the well-known phase-matching condition.<sup>8</sup> The severity of this condition can be appreciated from the fact that the third harmonic is usually generated using two  $\chi^{(2)}$  processes.<sup>9</sup>

As mentioned before, the Hamiltonian for third-harmonic generation in the periodic structures has the property that the anharmonic terms play a much more prominent role. In addition, no phase matching is necessary in this case, except for the requirement that both  $\omega$  and  $3\omega$  fall within a stop gap. This latter requirement is not as severe as the phase-matching condition and might offer the possibility of THG in solids mediated by a Kerr nonlinearity. The dynamics of the Hamiltonian in Eq. (5.10) is quite complicated, but we hope to report on a detailed analysis in a future publication.

## VI. DISCUSSION AND CONCLUSIONS

Optical absorption and damage are the two effects most likely to hamper the experimental observation of gap solitons and their related phenomena. Optical absorption simply limits the effective length of the stack, and in order to make use of the full stack length it is thus necessary that the attenuation distance exceed the length of the stack. Optical damage, on the other hand, limits the maximum energy flow through the system [Eq. (2.18)]. Apart from these two criteria for the experimental observation of gap solitons, the wavelength of the incoming radiation should also be tuned to the appropriate location within a stop gap of the structure. The required detuning from the nearest edge of the stop gap can be found from Eq. (2.22). Particularly in the case of small

stop gaps, an actual experiment requires the use of a continuously tunable source with narrow spectral width. Another, perhaps easier way to tune the system might be to vary the angle of incidence, rather than the frequency. Our present theory does not apply to the situation of non-normal incidence, which is much harder to analyze. We expect, however, to report on such an extension in the near future.

Although our theoretical description thus far was restricted to bulk systems, it appears that waveguide and fiber geometries are better candidates for experimental observations. In fiber geometries it is possible to obtain a periodic index modulation with a considerable number of periods by the interference of two counter-propagating laser beams.<sup>23</sup> The thus obtained modulation depth is quite small, however (typically  $10^{-7}$ ).<sup>23</sup> Further, the damage threshold of silica limits the change in refractive index through the nonlinearity, to about this same amount.<sup>23</sup> This implies, through the application of Eq. (2.18), the need of a stack with about  $10^7$  periods, which, for a period of  $1 \mu\text{m}$ , corresponds to about 10 m of fiber.

In optical waveguides it is the effective guide index  $N$  (Ref. 6) for the guided modes that plays the role of a refractive index. One can obtain a modulation of the guide index by a periodic perturbation of the ideal waveguide. A well-known way to obtain this modulation (but certainly not the only one<sup>24</sup>) is to periodically vary the thickness of the waveguide.<sup>6</sup> The modulation of the guide index in these *waveguide diffraction gratings* is proportional to the amplitude of the thickness variation. This amplitude can have a value of several tens of nanometers, giving rise to a modulation in the guide index of about  $10^{-4}$  to  $10^{-3}$ . By appropriately choosing the waveguide material, it is possible to obtain similar values for the maximum change in the guide index due to the nonlinearity. According to Eq. (2.18), again, this leads to about a stack with a few thousand periods, or, alternatively, to a stack length of several millimeters. This should be contrasted to the required length in a fiber geometry.

Finally, it might also be possible to observe gap solitons in the microwave region of the electromagnetic spectrum. In this regime, it may be that large nonlinearities can be obtained without appreciable absorption.<sup>25</sup>

Another matter worth discussing at this point are the results of the Lagrangian and Hamiltonian treatment of the nonlinear periodic stack. The results for THG are particularly interesting, as the conventional phase-matching condition, which dominates the process in homogeneous materials, does not enter the discussion at all. It is easily seen that this condition is presently replaced by the condition that  $z = x - \omega't$  be the same for the fields at  $\omega$  and  $3\omega$ . In turn, this implies that the group velocity for the fields at  $\omega$  and  $3\omega$  is the same. This is a generalization of the condition in Secs. IV and V that the fields at  $\omega$  and  $3\omega$  both fall within a stop gap, so that both group velocities vanish. Note that this requirement is rather different from the usual phase-matching condition in bulk materials, which requires that the phase velocity of the two fields be the same. The present emphasis on the group velocity, rather than the phase velocity, is consistent with the notion that we are dealing with

the envelope function of the electromagnetic field, rather than with the field itself. This is a most important difference from a practical point of view, since the group velocity in the periodic structures can, to some extent, be tailored by the choice of materials and by the stack period. This thus opens the possibility of optimizing nonlinear periodic structures for THG.

In conclusion, we have considerably widened the scope of our previous investigations into the properties of nonlinear periodic structures and their associated gap solitons. These investigations focus on the envelope of the electromagnetic field, rather than on the field itself. Although our method is very general, we have restricted ourselves to envelopes which have harmonic time dependence which are at rest in space. We have shown that, in this limit, our method is equivalent to the well-known effective-mass approximation from solid-state physics. In addition, the derivation leading to this result shows that for our approach to be valid, the envelope function should vary on a slower scale than three lattice constants, a result which could not be obtained using our derivation using the method of multiple scales. We have further generalized this derivation to include more than a single Bloch function explicitly. This allows us not only to describe the situation in which the frequency of the incoming radiation corresponds to an arbitrary position in a stop gap, but also the situation in which the third harmonic of the incoming radiation is generated. The Hamiltonian and Lagrangian treatment, finally, provides a relatively simple interpretation of the coupled nonlinear differential equations which describe these situations.

#### ACKNOWLEDGMENTS

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

#### APPENDIX A

In this appendix we apply the Kogelnik approach<sup>5</sup> to the present problem. This approximation, which is exact in the limit of vanishing modulation of the (linear) refractive-index profile, allows us to estimate the size of the stop gap and the group-velocity dispersion. In addition, it gives an approximate form for the optical Bloch functions. It should be noted that the use of Bloch functions in an electromagnetic context is not very common. Rather, the fields are usually written in terms of forward- and backward-scattered waves.

The analysis starts with the time-independent wave equation for the electric field in the linear stack

$$\left[ \frac{d^2}{dx^2} + \frac{\omega^2}{c^2} n^2(x) \right] E(x) = 0, \quad (\text{A1})$$

where  $n^2$  was written rather than  $\epsilon$  as in Eq. (2.1). In the Kogelnik approximation one assumes that the variation of the refractive index,  $\Delta n$ , is small, so that second- and higher-order terms in this parameter can be ignored. Since we are only interested in the lowest stop gap of the structure, we can ignore all Fourier components of the

refractive-index distribution, except the lowest, and for the present purpose, therefore,

$$n(x) = \bar{n} + (\Delta n) \cos(2k_0 x), \quad (\text{A2})$$

so that the period of the stack is  $\pi/k_0$ . The wave numbers  $\pm k_0$  are thus located at the edges of the Brillouin zone. The electric field is now written as

$$E(x) = \varphi_+(x) e^{ik_0 x} + \varphi_-(x) e^{-ik_0 x}, \quad (\text{A3})$$

where the envelopes  $\varphi_{\pm}(x)$  are assumed to be slowly varying on the scale of the period of the refractive-index profile. This fact allows us to make two approximations in substituting Eqs. (A2) and (A3) into Eq. (A1). The first of these is the neglect of terms containing  $\varphi_{\pm}''$  in the calculation of the second derivative of the electric field. In addition, it allows us to neglect mixing between frequency components centered at  $\pm k_0$ . Using these approximations it is then found that the envelopes satisfy a set of coupled equations

$$\begin{aligned} \frac{d\varphi_+}{dx} - i\nu\varphi_+ &= +i\sigma\varphi_-, \\ \frac{d\varphi_-}{dx} + i\nu\varphi_- &= -i\sigma\varphi_+, \end{aligned} \quad (\text{A4})$$

where

$$\nu = \frac{\omega\bar{n}}{c} - k_0, \quad (\text{A5a})$$

and

$$\sigma = \frac{k_0(\Delta n)}{2\bar{n}}. \quad (\text{A5b})$$

These equations can readily be solved to give

$$\begin{aligned} \varphi_+(x) &= A_1 e^{i\mu x} + A_2 e^{-i\mu x}, \\ \varphi_-(x) &= \frac{\mu - \nu}{\sigma} A_1 e^{i\mu x} - \frac{\mu + \nu}{\sigma} A_2 e^{-i\mu x}, \end{aligned} \quad (\text{A6})$$

where  $\mu$  can be found from

$$\mu^2 = \nu^2 - \sigma^2. \quad (\text{A7})$$

We see from this equation that the functions  $\varphi_{\pm}(x)$  denote running wave solutions if  $|\nu| \geq \sigma$  only. The violation of this inequality signifies the presence of a stop gap.

Substituting Eqs. (A6) into Eq. (A3) defines two Bloch functions for the electric field. The first of these, which consists of the terms with  $A_1$  in Eqs. (A6), reads as

$$\varphi(x) = A_1 e^{i(\mu + k_0)x} \left[ 1 + \frac{\mu - \nu}{\sigma} e^{-2ik_0 x} \right], \quad (\text{A8})$$

and it thus has a crystal momentum equal to

$$\kappa = \mu + k_0. \quad (\text{A9})$$

The other Bloch function has a crystal momentum of  $-\mu + k_0$  and consists of the terms with  $A_2$  in Eqs. (A6). Since  $\pm k_0$  indicate the edges of the BZ,  $\mu$  can be interpreted as the distance (in reciprocal space) to the BZ edge. For  $\mu = 0$  the two Bloch functions simply reduce to

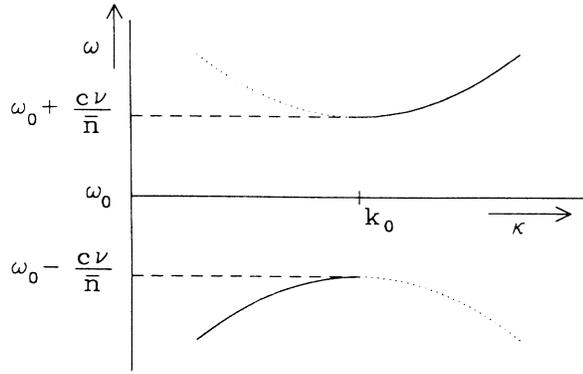


FIG. 4. Graphical interpretation of Eq. (A10), which gives the dispersion curves around BZ edge in the Kogelnik approximation. Note that  $\omega = c\nu/\bar{n}$ . The solid lines refer to the dispersion curves in an extended zone scheme.

sine and cosine functions. At the BZ edge where the Bloch functions border a stop gap, the Bloch functions can, therefore, be made real. They thus represent standing waves which do not transfer energy. These are well-known properties for Bloch functions which border a stop gap.

For the present purpose it is sufficient only to consider the Bloch function in Eq. (A8). Combining Eqs. (A9) and (A7), it is found that

$$\nu = \pm[\sigma^2 + (\kappa - k_0)^2]^{1/2}. \quad (\text{A10})$$

Figure 4 gives a graphical interpretation of this equation. The parameter  $\omega_0$  in this figure is defined as  $k_0 c/\bar{n}$ , in analogy with Eq. (A5a). It denotes the angular frequency of an electromagnetic wave with wave vector  $k_0$  traveling in a uniform medium with refractive index  $\bar{n}$ . For this reason, it corresponds approximately to the middle of the stop gap (deviations from this position are due to higher-order effects, which are explicitly excluded from the present discussion). Combining Eqs. (A10) and (A4) we can now immediately find the size of the stop gap  $\Delta\omega$ ,

$$\frac{\Delta\omega}{\omega} = \frac{\Delta n}{\bar{n}}. \quad (\text{A11})$$

We finally find the group-velocity dispersion, which is defined as the curvature of the dispersion curve. From the present analysis it is clear that, apart from the sign, this quantity is identical for the lower and upper branch of the dispersion relation around the stop gap. Deviations are, again, due to higher-order effects. We expand the radical in Eq. (A10) and, using the definitions in Eq. (A5), it is found that

$$\omega'' = \frac{2c^2}{\bar{n}(\Delta n)\omega_0}. \quad (\text{A12})$$

Equations (A11) and (A12) are both used in Sec. II to find the criteria for the experimental observation of gap solitons [Eqs. (2.18) and (2.22)] and for the validity of the envelope-function approach as a whole [Eqs. (2.20) and (2.23)].

## APPENDIX B

In this appendix we discuss the relevant free energy for the present problem. The ultimate justification for the choice in Eq. (5.4) is that it is consistent with Maxwell's equations, so that it corresponds to the effective Lagrangian of our nonlinear equations. One has, in principle, however, four possibilities to define the free energy for the electromagnetic field. These four possibilities can be generated through the total differentials<sup>21,22</sup>

$$\begin{aligned} 4\pi d\mathcal{F}_{e_1} &= E dD, & 4\pi d\mathcal{F}_{e_2} &= -D dE, \\ 4\pi d\mathcal{F}_{m_1} &= H dB, & 4\pi d\mathcal{F}_{m_2} &= -B dH, \end{aligned} \quad (\text{B1})$$

where the subscripts refer to the electric and magnetic parts of the free energy and the arguments of the fields were ignored. As mentioned in Sec. V, for externally driven systems one must use the time-averaged free energy. This distinction is ignored in the present appendix. Combining one each from the top and bottom line of Eqs. (B1) gives rise to four possible definitions. We can reduce this number by rewriting Eqs. (B1) using Maxwell's equations, the relations between the fields and the scalar and vector potentials and the fact that the medium is a dielectric. One then finds that

$$\begin{aligned} 4\pi cd\mathcal{F}_{e_1} &= -\dot{A}(D)dD, & 4\pi cd\mathcal{F}_{e_2} &= D(A)d\dot{A}, \\ 4\pi cd\mathcal{F}_{m_1} &= \dot{D}(A)dA, & 4\pi cd\mathcal{F}_{m_2} &= -A(D)d\dot{D}, \end{aligned} \quad (\text{B2})$$

where  $A$  is the vector potential and the dot denotes differentiation with respect to time. Also, the dependence  $D(A)$  and its inverse  $A(D)$  are explicitly given. The relations between these fields can be found from the constitutive relations of the dielectric. From Eqs. (B2) we immediately conclude that we should combine  $\mathcal{F}_{e_1}$  with  $\mathcal{F}_{m_2}$  and  $\mathcal{F}_{e_2}$  with  $\mathcal{F}_{m_1}$ , so that one of the fields is allowed to vary freely (within the boundary conditions), whereas the other is found using the relations  $D(A)$  or  $A(D)$ . If we had taken other combinations, we would obtain expressions for the free energy in which the variation of both  $A$  and  $D$  would be required. Note that this argument excludes the combination  $\mathcal{F}_{e_1}$  and  $\mathcal{F}_{m_1}$ , which is the energy density associated with the electromagnetic field. The two remaining choices are thus

$$4\pi d\mathcal{F}_1 = E(D,H)dD - B(D,H)dH, \quad (\text{B3a})$$

$$4\pi d\mathcal{F}_2 = H(E,B)dB - D(E,B)dE, \quad (\text{B3b})$$

where the constitutive relations were again explicitly written. Apart from the sign, the distinction between  $\mathcal{F}_1$  and  $\mathcal{F}_2$  vanishes for linear media, so that a final choice can only be made when nonlinearities are present. This final choice is given by Eq. (B3b), but we have not been able to devise an *a priori* argument to support this choice. It is, however, consistent with the notion that physically the fields  $E$  and  $B$ , and thus  $A$ , are the applied fields, whereas  $D$  and  $H$  describe the medium response. It is thus reassuring that the fundamental fields are allowed to vary freely, with the influence of materials following through the constitutive relations.

- <sup>1</sup>C. M. de Sterke and J. E. Sipe, *Phys. Rev. A* **38**, 5149 (1988).
- <sup>2</sup>Wei Chen and D. L. Mills, *Phys. Rev. Lett.* **58**, 160 (1987).
- <sup>3</sup>Wei Chen and D. L. Mills, *Phys. Rev. B* **36**, 6269 (1987).
- <sup>4</sup>J. M. Luttinger and W. Kohn, *Phys. Rev.* **97**, 869 (1955).
- <sup>5</sup>H. Kogelnik, *Bell Syst. Techn. J.* **48**, 2909 (1969).
- <sup>6</sup>H. Kogelnik, in *Integrated Optics*, edited by T. Tamir (Springer, Berlin, 1979), p. 15.
- <sup>7</sup>N. Bloembergen, *Nonlinear Optics* (Benjamin, London, 1965).
- <sup>8</sup>J. A. Armstrong, N. Bloembergen, J. Ducuing, and P. S. Pershan, *Phys. Rev.* **127**, 1918 (1962).
- <sup>9</sup>Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley, New York, 1984), Chap. 3.
- <sup>10</sup>R. K. Dodd, J. C. Eilbeck, J. D. Gibbon, and H. C. Morris, *Solitons and Nonlinear Wave Equations* (Academic, London, 1982).
- <sup>11</sup>J. Callaway, *Quantum Theory of the Solid State* (Academic, New York, 1974), p. 409.
- <sup>12</sup>Since  $L$  is evidently a finite quantity,  $k$  is not a continuous variable, so that the Dirac  $\delta$  function in Eq. (3.3) and the integral over the Brillouin zone in Eq. (3.4) are strictly inappropriate. In order to simplify the notation in Sec. III and to be consistent with existing literature (Refs. 4 and 11), the present notation was used.
- <sup>13</sup>G. L. Bir and G. E. Pikus, *Symmetry and Strain Induced Effects in Semiconductors* (Wiley, New York, 1974).
- <sup>14</sup>See, e.g., G. Bastard, in *Molecular Beam Epitaxy and Heterostructures*, edited by L. L. Chang and K. Ploog (Nijhoff, Dordrecht, 1985), Chap. 11.
- <sup>15</sup>D. N. Christodoulides and R. I. Joseph, *Opt. Lett.* **13**, 53 (1988).
- <sup>16</sup>G. P. Agrawal, *Phys. Rev. Lett.* **59**, 880 (1987).
- <sup>17</sup>A. H. Nayfeh and D. T. Mook, *Nonlinear Oscillations* (Wiley, New York, 1979), Chap. 8.
- <sup>18</sup>L. D. Landau and E. M. Lifshitz, *Mechanics*, 3rd ed. (Pergamon, London, 1976).
- <sup>19</sup>See, e.g., L. E. Reichl, *A Modern Course in Statistical Physics* (University of Texas Press, Austin, 1980), Chap. 2.
- <sup>20</sup>P. S. Pershan, *Phys. Rev.* **130**, 919 (1963).
- <sup>21</sup>L. D. Landau, E. M. Lifshitz, and L. P. Pitaevskii, *Electrodynamics of Continuous Media*, 2nd ed. (Pergamon, Oxford, 1984).
- <sup>22</sup>C. Kittel, *Thermal Physics*, 1st ed. (Wiley, New York, 1969), Chaps. 23 and 24.
- <sup>23</sup>K. O. Hill, Y. Fujii, D. C. Johnson, and B. S. Kawasaki, *Appl. Phys. Lett.* **32**, 647 (1978).
- <sup>24</sup>G. A. Reider, M. Huemer, and A. J. Schmidt, *Opt. Commun.* **68**, 149 (1988).
- <sup>25</sup>D. L. Mills (private communication).