Hamiltonian formulation of coupled-mode theory in waveguiding structures

Philip Chak,[1,2] Rajiv Iyer,[2,3] J. S. Aitchison,[2,3] and J. E. Sipe[1,2]
[1]*Department of Physics, University of Toronto, Toronto, Ontario, Canada*
[2]*Institute for Optical Sciences, University of Toronto, Toronto, Ontario, Canada*
[3]*Department of Electrical and Computer Engineering, University of Toronto, Toronto, Ontario, Canada*
(Received 13 November 2005; revised manuscript received 5 October 2006; published 18 January 2007)

We present a Hamiltonian formulation of coupled mode theory for scenarios in which the coupled modes are associated with different "parent structures," such as two nearby waveguides. The relativistic nature of the photon leads to the complication that not any set of orthonormal modes can be used as a basis if the associated amplitudes are to satisfy canonical commutation relations. This difficulty is circumvented by the introduction of "dressed parent modes," which are in fact seen to be the "coupled modes" of the system. While an exact solution of the linear problem within the restricted basis of interest formally must be found before these modes can be constructed, in practice they can be constructed directly from the modes of the parent structures. The approach can be applied to periodic parent structures, such as photonic crystal waveguiding structures, as well as to simpler waveguides. We illustrate the accuracy of the various approximations employed by studying two sample systems in detail. We derive the linear coupled mode equations, and show how the approach can be immediately generalized from the linear regime to treat problems in nonlinear quantum optics.

DOI: 10.1103/PhysRevE.75.016608 PACS number(s): 42.79.Gn, 42.25.Bs, 42.82.Et

I. INTRODUCTION

The coupling of modes in adjacent dielectric waveguides is an old subject [1]. Following its first applications in the regime of linear, classical optics, the theory of coupled modes has been extended into both nonlinear [2] and quantum regimes. Quantum linear and nonlinear properties of directional couplers have been studied using an effective Hamiltonian [3], and nonclassical behavior such as squeezed-state generation [4] and optical parametric processes [5] have been investigated. Nowadays, photonic crystals [6–8] offer a new platform for such effects, due to their miniature size and potential applications as ultracompact optical devices. If the photonic crystal possesses a band gap in the frequency range of interest, line defects can be used as waveguides for light. Coupled mode theory applicable to photonic crystal structures has been presented by a number of workers [5,9,10]. Yet a derivation of coupled mode theory for photonic crystal waveguides, applicable in the quantum and nonlinear regimes, is still lacking. In this paper we derive a general coupled mode theory for waveguides, based on a Hamiltonian formulation of Maxwell's equations, which can easily be generalized to include quantum and nonlinear effects in a systematic manner. It is applicable both to photonic crystal waveguides and to waveguides of other types.

While the analysis in this paper is focused solely on the linear regime, it is nonetheless a significant extension of earlier work on the canonical formulation of the electrodynamics of fields in artificially structured materials. To highlight this, it is useful to distinguish two types of coupled mode theories. In a single parent type of theory, one starts with the exact linear modes of one nominal medium; the coupling occurs, in a Hamiltonian language, because a perturbation Hamiltonian is introduced. A simple example is a very weak Bragg grating in an optical fiber. Here the modes can be taken to be forward- and backward-propagating waves, with the perturbation Hamiltonian describing the effect of the variation in the index of refraction responsible for the grating. Due to this perturbation, the amplitudes of the forward and backward propagating waves become coupled. This kind of approach has been extended, within a Hamiltonian framework, to treat linear and nonlinearly coupled modes in photonic crystals, even if the linear index variations are not weak [11,12]. Even there, however the defining feature is that the basis set of modes, which for photonic crystals is the set of Bloch modes, are eigenmodes of one nominal linear structure.

In contrast, in a multiple parent type of coupled mode theory one starts with the exact linear modes of different structures, and even in the linear regime the modes are coupled because the actual structure is neither of the parent structures. We refer to these modes, which are typically nonorthogonal, as the parent modes. The classic example here is the directional coupler, where one begins with modes that would be exact for isolated channels, but in fact are coupled because the channels are close to each other. The generalization to photonic crystal waveguides, shown in Fig. 1, involves the coupling of waveguide modes of two nominal structures [Figs. 1(a) and 1(b)], which arises because the two channels are close to each other in the actual structure of interest [Figs. 1(c)]. The complication arising in a canonical formulation of a multiple parent theory is that there is no single, physical unperturbed Hamiltonian that governs the uncoupled evolution of all of the parent modes employed. We provide a framework for addressing the optical dynamics of such structures by introducing a formal "unperturbed Hamiltonian" governing what we will call "dressed parent modes," or simply "dressed modes," and a formal perturbation Hamiltonian that then couples them in a description of the full problem.

A related situation can arise in calculations in quantum chemistry, where one often starts with a basis of nonorthogonal orbitals involving eigenstates of different atomic Hamiltonians. The problem in quantum optics is more complicated,

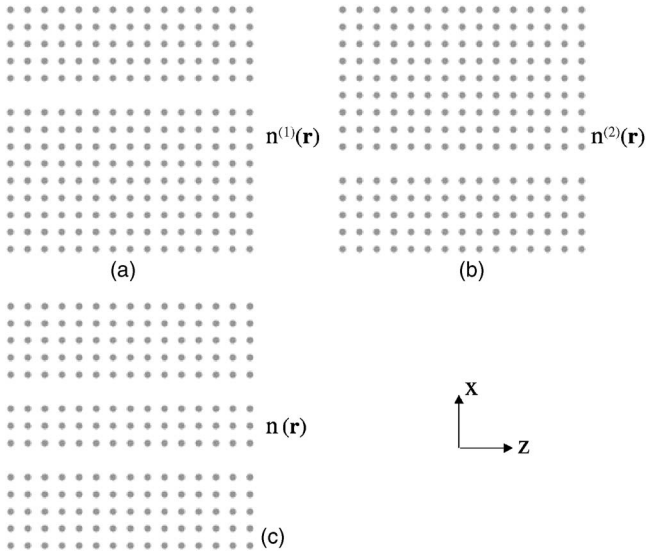

FIG. 1. (Color online) (a),(b) Dielectric profiles of individual photonic crystal waveguide. (c) Dielectric profile of the coupled photonic crystal waveguide system.

however, because of the inherently relativistic nature of the photon. In a general expansion of electromagnetic fields there appear products of raising and lowering operators with normalized modes *and* factors of $\sqrt{\hbar\omega}$, where ω is the frequency of the mode. And so an expansion involving raising and lowering operators that satisfy canonical commutation relations cannot immediately be constructed in terms of *any* basis set, but only by using the exact modes of the system, and only when their eigenfrequencies are known. By a canonical transformation it is possible to construct the dressed modes mentioned above, which are associated with different, but still canonical, raising and lowering operators. However, this can only be done once the exact eigenfunctions and eigenfrequencies are found.

It might then appear that there is no advantage to identifying the dressed modes and their coupling constants. For the exact eigenmodes must be found first, and if they are already found why should one bother to construct approximate eigenmodes? There are two answers to this. The first is that it is the dressed modes that actually appear as the "coupled modes" when the linear coupled mode equations are constructed from a Hamiltonian formalism. Thus their identification helps establish the physics of the problem. Also, when any nonlinearity introduced in the problem is weak, the dressed modes form a natural basis for expanding the nonlinear part of the Hamiltonian.

A second answer is that the difficulty in identifying the dressed modes and their coupling constants is not as serious in practice as it is in principle. This is because any frequency shifts due to mode coupling in the optical regime are typically much smaller than the frequencies of the modes themselves. As a result, although the exact modes must be formally constructed, and the dressed modes indeed built from them, in fact a series expansion of the dressed modes in terms of the parent modes can be extracted from the formal construction. This can be done in such a way that practitioners need never actually determine the exact modes; they can move directly from the parent modes to the dressed modes, which are the natural basis for the coupled mode equations and their extension into the nonlinear regime.

In the rest of this paper we put mathematical flesh on the bones of the argument given here, and establish the accuracy of the approximations necessary in moving along this route. In Sec. II we review the canonical formulation of Maxwell's equations, in the presence of material media, in the form most useful for our purposes [11]. In Sec. III we introduce the parent modes, and the restricted space they form. We confirm in two typical examples that, as is often the case, the exact modes of interest can be well described using this restricted space. We then introduce a set of dressed modes in Sec. IV, and in Sec. V introduce an expansion in terms of which the dressed modes can be written directly from the parent modes; we illustrate how well this expansion recovers the actual dressed modes for our example systems. In Sec. VI we show how the coupled mode equations follow from the Hamiltonian written in terms of the dressed mode raising and lowering operators, which we call the "dressed mode Hamiltonian," and examine for our example systems the validity of the kind of approximations that must be made. We conclude in Sec. VII.

II. CANONICAL FORMULATION OF MAXWELL'S EQUATIONS

To set our notation, in this section we briefly review a canonical formulation of Maxwell's equations that was used earlier [11,17] to study pulse dynamics in higher dimension photonic crystals. We focus on classical optics here, but because its easy generalization to quantum optics is one of the strengths of this approach, we adopt a quantum notation and, for the classical Poisson bracket $\{\ldots,\ldots\}$, we write $(i\hbar)^{-1}[\ldots,\ldots]$; we also use a dagger to indicate complex conjugation. We will also often speak of operators rather than variables, especially when it makes the physics more clear. Beginning with the familiar Maxwell's equations

$$\frac{\partial \mathbf{D}}{\partial t} = \nabla \times \mathbf{H}, \quad \frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E}, \tag{1}$$

$$\nabla \cdot \mathbf{D} = 0, \quad \nabla \cdot \mathbf{B} = 0, \tag{2}$$

in what follows we treat Eq. (2) as initial conditions for fields that satisfy Eq. (1); if Eqs. (2) are satisfied at any one particular time, then the dynamical equations (1) guarantee that they are satisfied at all later (and earlier) times. Neglecting magnetic effects in our structures we have

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

$$\mathbf{D}(\mathbf{r},t) = \varepsilon_0 \mathbf{E}(\mathbf{r},t) + \mathbf{P}(\mathbf{r},t), \tag{3}$$

where $\mathbf{P}(\mathbf{r},t)$ is the polarization of the material. In this paper we restrict ourselves to linear material response, for which $\mathbf{P}$ is typically written in terms of $\mathbf{E}$ through a constitutive relation of the form

$$P^i(\mathbf{r},t) = \varepsilon_0 \chi_1^{ij}(\mathbf{r}) E^j(\mathbf{r},t), \tag{4}$$

where the superscripts represent Cartesian components and are to be summed over if repeated. For example, if we have an inhomogeneous medium that is nonetheless isotropic in its linear response, we can write

$$\chi_1^{ij}(\mathbf{r}) = \delta^{ij}[n^2(\mathbf{r}) - 1], \tag{5}$$

where the real $n(\mathbf{r})$ is the local index of refraction. However, since the divergence Maxwell equations (2) take the form of initial conditions, it is convenient to treat $\mathbf{D}$ and $\mathbf{B}$ as our fundamental fields, and we here write $\mathbf{P}$ as a function of $\mathbf{D}$,

$$P^i(\mathbf{r},t) = \Gamma_1^{ij}(\mathbf{r}) D^j(\mathbf{r},t). \tag{6}$$

This introduces no loss of generality, because for any specified $\chi_1^{ij}(\mathbf{r})$ appearing in Eq. (4) we can easily find the $\Gamma_1^{ij}(\mathbf{r})$ appearing in Eq. (6) by using the second of (3); in the special case of Eq. (5), for example, we have

$$\Gamma_1^{ij}(\mathbf{r}) = \delta^{ij}\left(1 - \frac{1}{n^2(\mathbf{r})}\right). \tag{7}$$

The total energy of the electromagnetic field, in the most general model of the form (6), can be written as [13]

$$H = \frac{1}{2\mu_0}\int d\mathbf{r} B^i(\mathbf{r}) B^i(\mathbf{r}) + \frac{1}{2\varepsilon_0}\int d\mathbf{r} D^i(\mathbf{r}) D^i(\mathbf{r}) - \frac{1}{2\varepsilon_0}\int d\mathbf{r} D^i(\mathbf{r}) \Gamma_1^{ij}(\mathbf{r}) D^j(\mathbf{r}). \tag{8}$$

In writing Eqs. (6) and (8) we have implicitly assumed that the carrier frequency is far from any material resonance, so $\mathbf{P}(\mathbf{r},t)$ is a function only of $\mathbf{D}(\mathbf{r},t)$ [or alternatively of $\mathbf{E}(\mathbf{r},t)$, using Eq. (4)] at the same time; hence we neglect any material dispersion.

In order to construct a canonical formulation of electromagnetic field we require a set of Poisson bracket relations that yield the dynamical equation (1) as a consequence of Hamilton's equation using the total energy as the Hamiltonian. The appropriate equal time commutation relations are

$$[D^i(\mathbf{r}), D^j(\mathbf{r}')] = [B^i(\mathbf{r}), B^j(\mathbf{r}')] = 0,$$

$$[D^i(\mathbf{r}), B^j(\mathbf{r}')] = i\hbar \varepsilon^{ijl} \frac{\partial}{\partial r^l}[\delta(\mathbf{r} - \mathbf{r}')], \tag{9}$$

where ε^{ijl} is the Levi-Civita symbol. Taking these as our fundamental commutation relations, it follows that we recover Eq. (1) using Hamilton's equations

$$i\hbar \frac{\partial \mathbf{D}}{\partial t} = [\mathbf{D}, H],$$

$$i\hbar \frac{\partial \mathbf{B}}{\partial t} = [\mathbf{B}, H]. \tag{10}$$

This approach for constructing a canonical formulation of the electromagnetic field has its roots in the early work of Born and Infeld [14].

In the special case of a periodic structure, the $\mathbf{D}$ and $\mathbf{B}$ fields can be expanded in terms of the Bloch modes of the structure, as illustrated earlier [11]. If there is one dimensional periodicity in the structure, such as a photonic crystal slab missing one or a few rows of rods (holes), the modes of interest are labeled by a (positive or negative) crystal wave number k, $-\pi/\Lambda < k \leqslant \pi/\Lambda$, where Λ is the periodicity of the structure; we can write

$$\mathbf{B}(\mathbf{r},t) = \sum_m \int dk \sqrt{\frac{\hbar\omega_{mk}}{2}} a_{mk}(t) \mathbf{B}_{mk}(\mathbf{r}) + \text{c.c.},$$

$$\mathbf{D}(\mathbf{r},t) = \sum_m \int dk \sqrt{\frac{\hbar\omega_{mk}}{2}} a_{mk}(t) \mathbf{D}_{mk}(\mathbf{r}) + \text{c.c.}, \tag{11}$$

where m is the band index, a_{mk} is the mode amplitude for the Bloch modes $\mathbf{B}_{mk}(\mathbf{r})$, obtained as the solution to the so-called "master equation"

$$\boldsymbol{\nabla} \times \left[\frac{\boldsymbol{\nabla} \times \mathbf{B}_{mk}(\mathbf{r})}{n^2(\mathbf{r})}\right] = \left(\frac{\omega_{mk}}{c}\right)^2 \mathbf{B}_{mk}(\mathbf{r}). \tag{12}$$

Here we have specialized again to the particular model (5) and (7), but an extension to a more general $\Gamma^{ij}(\mathbf{r})$ can be easily constructed. It is well-known that Eq. (12) is a Hermitian eigenvalue problem, and its numerical implementation has been discussed extensively in the literature [15]. The corresponding displacement field $\mathbf{D}_{mk}$ can then be obtained from $\mathbf{B}_{mk}$ using Maxwell's equations

$$\mathbf{D}_{mk}(\mathbf{r}) = \frac{i}{\mu_0 \omega_{mk}} \boldsymbol{\nabla} \times \mathbf{B}_{mk}(\mathbf{r}). \tag{13}$$

Without loss of generality we assume the structure is oriented such that it is periodic in the $\hat{\mathbf{z}}$ direction. From Bloch's theorem, $\mathbf{B}_{mk}(\mathbf{r})$ takes the form

$$\mathbf{B}_{mk}(\mathbf{r}) = \sqrt{\frac{\Lambda}{2\pi}} \mathbf{b}_{mk}(\mathbf{r}) e^{ikz}, \tag{14}$$

where $\mathbf{b}_{mk}(\mathbf{r}) = \mathbf{b}_{mk}(\mathbf{r} + \Lambda\hat{\mathbf{z}})$, and the factor $\sqrt{\Lambda/2\pi}$ is introduced to simplify the formulas below; similar relations hold for $\mathbf{D}_{mk}(\mathbf{r})$. With our choice [16] of the constants appearing in Eq. (11), the mode functions $\mathbf{D}_{mk}(\mathbf{r})$ and $\mathbf{B}_{mk}(\mathbf{r})$ satisfy the normalization conditions

$$\int d\mathbf{r} \frac{\mathbf{B}^*_{m'k'}(\mathbf{r}) \cdot \mathbf{B}_{mk}(\mathbf{r})}{\mu_0} = \delta_{mm'} \delta(k - k'),$$

$$\int d\mathbf{r} \frac{\mathbf{D}^*_{m'k'}(\mathbf{r}) \cdot \mathbf{D}_{mk}(\mathbf{r})}{\varepsilon_0 n^2(\mathbf{r})} = \delta_{mm'} \delta(k - k'). \tag{15}$$

These result from first constructing the modes in a large volume using periodic boundary conditions, and then passing to the limit of an infinite normalization volume [11]; the integrals in Eq. (15) hence range over all space. From these follow the normalization conditions

$$\int_{\text{cell}} d\mathbf{r} \frac{\mathbf{b}^*_{m'k}(\mathbf{r}) \cdot \mathbf{b}_{mk}(\mathbf{r})}{\mu_0} = \delta_{mm'},$$

$$\int_{\text{cell}} d\mathbf{r} \frac{\mathbf{d}^*_{m'k}(\mathbf{r}) \cdot \mathbf{d}_{mk}(\mathbf{r})}{\varepsilon_0 n^2(\mathbf{r})} = \delta_{mm'}, \tag{16}$$

where in Eq. (16) the integrals range over all x and y, but with z ranging over one periodicity length Λ; we restrict ourselves to modes that are bounded in the x and y directions, although generalizations have been presented [11,12] earlier. With Eqs. (11)–(16), the operators a_{mk} and $a^\dagger_{m'k'}$ satisfy the commutation relations

$$[a_{mk}, a_{m'k'}] = 0,$$

$$[a_{mk}, a^\dagger_{m'k'}] = \delta_{mm'}\delta(k-k'), \tag{17}$$

and the Hamiltonian of the electromagnetic fields can be written as

$$H = \sum_m \int dk \hbar \omega_{mk} a^\dagger_{mk} a_{mk}. \tag{18}$$

The Hamiltonian approach can easily be generalized to include nonlinearity of arbitrary order in a systematic way. Earlier [11,12] it was used with an effective field formalism to describe pulse propagation in three-dimensional (3D) photonic crystals, taking into account diffraction and mode dispersion inside the structure.

III. EIGENMODES IN A RESTRICTED BASIS

While, in general, the Bloch modes of any periodic structure can be obtained by numerically solving the Hermitian eigenvalue problem (12), the approach we consider here relies on the formal construction of the exact waveguide modes of the overall structure using the approximation that the space of the modes of interest is spanned, to good approximation, by what we call "parent modes," which are exact solutions of Maxwell's equations for simpler, parent structures. Indeed, any type of coupled mode theory necesssarily relies on the restriction of the space of fields to a subset of those spanned by a full basis set. So before entering into any of the details of the coupled mode theory, it is necessary to argue or establish that such a restricted basis provides a good description, as is of course often the case. In this section we do that for two simple structures where the true modes can easily be found and compared with those found within the restricted space. These two structures will be used as examples throughout the following sections.

The first structure is the 2D photonic crystal waveguide depicted in Fig. 1, where the parent structures are formed by removing a line of rods in a square lattice. The radius of the rod is given by $r=0.2a$, where a is the lattice constant. Figure 1(a) [1(b)] shows the first (second) parent structure and Fig. 1(c) shows the coupled waveguide system. The rods are taken to have an index $n=3.4$, and the background medium an index $n=1$. The dispersion relation for the parent waveguide modes is depicted in Fig. 2. While a perfect 2D square lattice (without defects) exhibits a 2D photonic bandgap, our particular structure [Figs. 1(a) and 1(b)] supports guided modes for $\omega a/2\pi c$ within the range [0.30387, 0.42059]. In

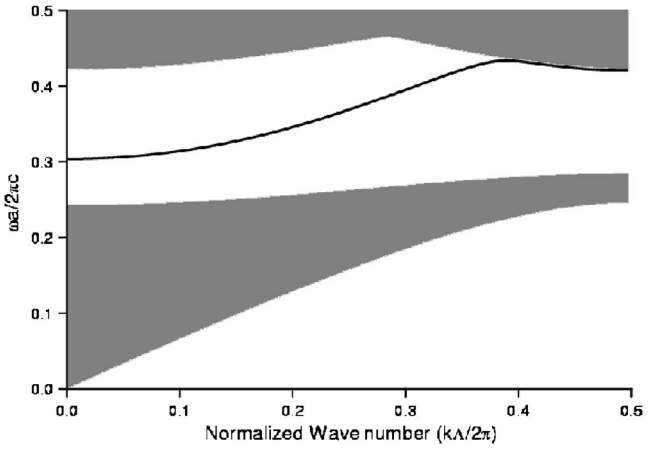

FIG. 2. Dispersion relation of the photonic crystal waveguide in Figs. 1(a) and 1(b). The waveguide supports guided mode for $\omega c/2\pi a \in [0.30387, 0.42059]$.

the analysis that follows we will restrict ourselves to this frequency regime. While for the overall structure shown in Fig. 1(c) we have only two parent structures, namely those shown in Figs. 1(a) and 1(b), more generally one could imagine more parent structures; we might, for example, be dealing with three or more missing rows in the overall structure. In general, we let N label the number of parent structures. For simplicity we assume throughout that there is only one parent mode at each k for each parent structure, and so there is no need of a band index for the parent modes; our notation could be easily extended to relax this restriction.

Our second example is a simpler structure of two neighboring planar waveguides shown in Fig. 3(c), with the parent modes being those of the structures shown in Figs. 3(a) and 3(b); again we have $N=2$. The index of the waveguides is taken to be $n=3.16$, and that of the background $n=1$. The thickness of the waveguides are taken as $d=200$ nm; our focus will be on modes at a frequency corresponding to a vacuum wavelength of $\lambda=1.54$ μm, and we consider the modes at that frequency as a function of the gap θ between the waveguides. At the wavelength $\lambda=1.54$ μm there is only one bound s-polarized (TE) mode for each parent structure, and we consider only modes with $k_y=0$; so again there is no need for a band index for the parent modes.

In either of the structures, or more generally for the whole family of structures of this sort, the parent waveguide modes $\mathbf{B}^{(i)}_k$ of the ith parent structure satisfy

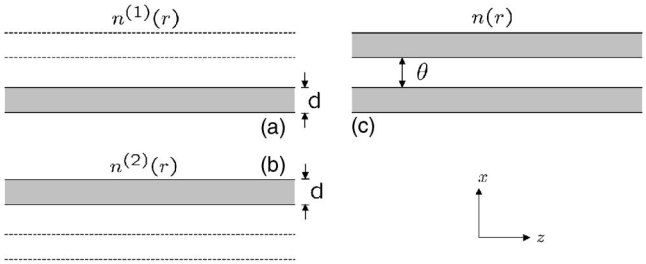

FIG. 3. (a),(b) Dielectric profiles of individual planar waveguide. (c) Dielectric profile of the coupled planar waveguide system.

$$\nabla \times \left[\frac{\nabla \times \mathbf{B}_k^{(i)}(\mathbf{r})}{[n^{(i)}(\mathbf{r})]^2} \right] = \left(\frac{\omega_k^{(i)}}{c} \right)^2 \mathbf{B}_k^{(i)}(\mathbf{r}), \tag{19}$$

where $n^{(i)}(\mathbf{r})$ represents the refractive index profile of the ith parent structure [Figs. 1(a) and 1(b); Figs. 3(a) and 3(b)]. Here we are simply applying the general results of Sec. II to the parent structures. Continuing along these lines, for a given $\mathbf{B}_k^{(i)}(\mathbf{r})$, the associated displacement amplitude $\mathbf{D}_k^{(i)}(\mathbf{r})$ follows from Eq. (1),

$$\mathbf{D}_k^{(i)}(\mathbf{r}) = \frac{i}{\mu_0 \omega_k^{(i)}} \nabla \times \mathbf{B}_k^{(i)}(\mathbf{r}). \tag{20}$$

We normalize the modes according to

$$\int d\mathbf{r} \frac{\mathbf{B}_{k'}^{*(i)}(\mathbf{r}) \cdot \mathbf{B}_k^{(i)}(\mathbf{r})}{\mu_0} = \delta(k-k'),$$

$$\int d\mathbf{r} \frac{\mathbf{D}_{k'}^{*(i)}(\mathbf{r}) \cdot \mathbf{D}_k^{(i)}(\mathbf{r})}{\varepsilon_0 [n^{(i)}(\mathbf{r})]^2} = \delta(k-k'), \tag{21}$$

[see Eq. (15)], and introducing periodic functions $\mathbf{b}_k^{(i)}(\mathbf{r})$ and $\mathbf{d}_k^{(i)}(\mathbf{r})$ as done above [see Eq. (14)], we have

$$\int_{\text{cell}} d\mathbf{r} \frac{\mathbf{b}_k^{*(i)}(\mathbf{r}) \cdot \mathbf{b}_k^{(i)}(\mathbf{r})}{\mu_0} = 1,$$

$$\int_{\text{cell}} d\mathbf{r} \frac{\mathbf{d}_k^{*(i)}(\mathbf{r}) \cdot \mathbf{d}_k^{(i)}(\mathbf{r})}{\varepsilon_0 [n^{(i)}(\mathbf{r})]^2} = 1. \tag{22}$$

The normalization conditions (21) and (22) are appropriate for truly three-dimensional structures where the bound modes are coupled in both the x and y directions. This is not the case in our simple examples of Figs. 1 and 3, where, strictly speaking, one should have normalization conditions that also restrict the y components of the wave vectors of two modes to be equal. Such a generalization is straightforward [11,12] but we do not do it here since it does not alter our results for the determination of the frequencies of the modes considered in this section, or for the mode dynamics considered later in Sec. VI.

We assume that parent modes and their frequencies are known from a previous calculation or model, and now we seek waveguide modes of the full system as the solution to the master equation

$$\nabla \times \left[\frac{\nabla \times \mathbf{B}_k(\mathbf{r})}{n^2(\mathbf{r})} \right] = \left(\frac{\omega_k}{c} \right)^2 \mathbf{B}_k(\mathbf{r}), \tag{23}$$

where $n(\mathbf{r})$ is the index profile of the structure of interest [e.g., Fig. 1(c) or 3(c)], and ω_k is the frequency of the approximate mode. With the assumption that the parent modes span the space of the full waveguide modes of interest to good approximation, we write

$$\mathbf{B}_k(\mathbf{r}) = \sum_j \eta^{(j)}(k) \mathbf{B}_k^{(j)}(\mathbf{r}), \tag{24}$$

where $\eta^{(j)}(k)$ are expansion coefficients to be determined. Inserting Eq. (24) into the master equation (23), multiplying that equation by $[\mathbf{B}_k^{(i)}(\mathbf{r})]^*$, integrating over all space, and finally using Eq. (20) together with Eq. (21) and the second of Eq. (22), we find the equation

$$\sum_j M^{ij}(k) \eta^{(j)}(k) = \left(\frac{\omega_k}{c} \right)^2 \sum_j S^{ij}(k) \eta^{(j)}(k), \tag{25}$$

with the matrix elements $M^{ij}(k)$ and $S^{ij}(k)$ given by

$$M^{ij}(k) = \frac{\omega_k^{(i)} \omega_k^{(j)}}{c^2} \int_{\text{cell}} d\mathbf{r} \frac{\mathbf{d}_k^{(i)*}(\mathbf{r}) \cdot \mathbf{d}_k^{(j)}(\mathbf{r})}{\varepsilon_0 n^2(\mathbf{r})},$$

$$S^{ij}(k) = \int_{\text{cell}} d\mathbf{r} \frac{\mathbf{b}_k^{(i)*}(\mathbf{r}) \cdot \mathbf{b}_k^{(j)}(\mathbf{r})}{\mu_0}. \tag{26}$$

In matrix form, we have

$$\mathsf{M}(k) \cdot \eta(k) = \left(\frac{\omega_k}{c} \right)^2 \mathsf{S}(k) \cdot \eta(k), \tag{27}$$

where $\mathsf{M}(k)$ indicates the (in general) $N \times N$ Hermitian matrix with elements $M^{ij}(k)$, $\mathsf{S}(k)$ indicates the $N \times N$ Hermitian matrix with elements $S^{ij}(k)$, and $\eta(k)$ the N element column vector with elements $\eta^{(i)}(k)$; in our particular examples (Figs. 1 and 3), recall $N=2$. Next we define the column vector $\alpha(k) \equiv \mathsf{S}^{1/2}(k) \cdot \eta(k)$, so the equation above becomes

$$\tilde{\mathsf{M}}(k) \cdot \alpha(k) = \left(\frac{\omega_k}{c} \right)^2 \alpha(k), \tag{28}$$

with the matrix $\tilde{\mathsf{M}}(k) \equiv \mathsf{S}^{-1/2}(k) \cdot \mathsf{M}(k) \cdot \mathsf{S}^{-1/2}(k)$. It immediately follows that $\tilde{\mathsf{M}}(k)$ is Hermitian, and therefore Eq. (28) is a standard Hermitian eigenvalue problem. Labelling the mth eigenvectors of $\tilde{M}(k)$ as $\alpha_m(k)$, and its associated frequency ω_{mk},

$$\tilde{\mathsf{M}}(k) \cdot \alpha_m(k) = \left(\frac{\omega_{mk}}{c} \right)^2 \alpha_m(k) \tag{29}$$

the mth mode of the overall system with wave number k then follows from Eq. (24); since $\eta_m(k) = \mathsf{S}^{-1/2}(k) \cdot \alpha_m(k)$ it can be written as

$$\mathbf{B}_{mk}(\mathbf{r}) = \sum_j \eta_m^{(j)}(k) \mathbf{B}_k^{(j)}(\mathbf{r}) \tag{30}$$

with

$$\eta_m^{(j)}(k) = \sum_l [\mathsf{S}^{-1/2}(k)]^{jl} \alpha_m^{(l)}, \tag{31}$$

where $\alpha_m^{(l)}(k)$ is the lth element of $\alpha_m(k)$.

Since the eigenvectors $\alpha_m(k)$ of the Hermitian matrix $\tilde{M}(k)$ can be chosen to satisfy the orthonormality conditions

$$\sum_m \alpha_m^{(n)}(k) [\alpha_m^{(n')}(k)]^* = \delta_{nn'},$$

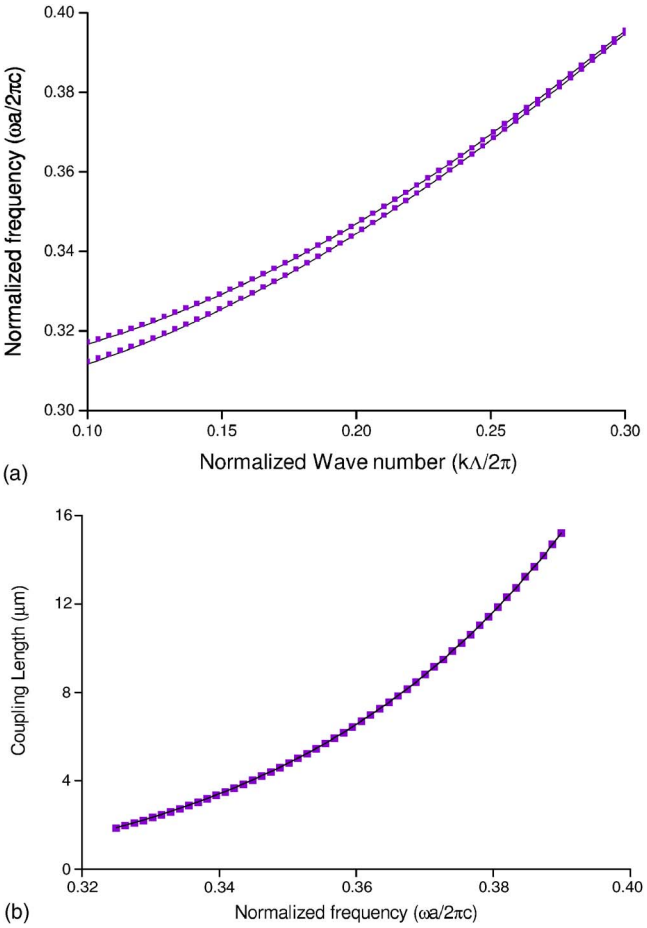

FIG. 4. (Color online) (a) Dispersion relation for $\omega a/2\pi c \in [0.30, 0.40]$ for the structure depicted in Fig. 1(c). Squares indicate the dispersion relation obtained by diagonalizing Eq. (28) and solid line indicates the dispersion relation obtained by a freely available software package (Ref. [17]). (b) Coupling length vs normalized frequency for light propagating in the structure depicted in Fig. 1(c). Coupling lengths calculated using Eq. (28) are indicated by squares; coupling lengths calculated numerically are indicated by the solid line.

$$\sum_n \alpha_m^{(n)}(k)[\alpha_{m'}^{(n)}(k)]^* = \delta_{mm'}, \tag{32}$$

we can easily establish that the $\mathbf{B}_{mk}(\mathbf{r})$ [Eq. (30)] satisfies the orthonormality condition (15), as expected. In terms of the raising and lowering operators associated with the modes $\mathbf{B}_{mk}(\mathbf{r})$, and their frequencies ω_{mk} following from Eq. (29), the Hamiltonian of the overall structure is given by Eq. (18).

We emphasize that this is all within the assumption that, to good approximation, the parent modes span the space of the exact modes of interest. We can confirm that this is indeed reasonable for our sample structures by examining the results of such an approximate calculation and comparing with exact numerical solutions. In Fig. 4(a) we plot the dispersion relations of the two bound modes of the overall structure of Fig. 1(c); one mode is symmetric and one antisymmetric in x about the middle of the central three rows of rods in Fig. 1(c). The squares indicate the dispersion

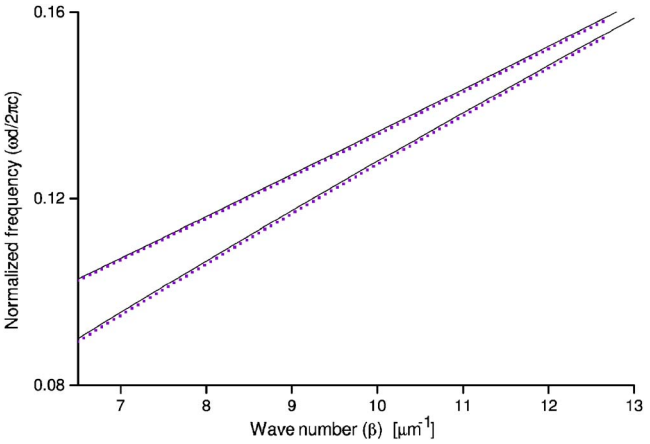

FIG. 5. (Color online) Dispersion relation for $\omega d/2\pi c \in [0.08, 0.16]$ for the structure depicted in Fig. 3(c). In this specific example, we have $d=200$ nm and $\theta=200$ nm. Squares indicate the dispersion relation obtained by diagonalizing Eq. (28) and solid line indicates the dispersion relation obtained by a freely available software package.

relations obtained by diagonalizing Eq. (28), and the solid line the exact dispersion relations of the structure obtained numerically using a freely available software package [17]. At a given frequency it is often the coupling length between the symmetric and antisymmetric modes that is of physical interest; it is given by $\pi/(\Delta k)$, where Δk is the difference in wave numbers of the two modes. This is plotted in Fig. 4(b), again with the squares indicating the results from Eq. (28), and the solid line that from the exact numerical solution. In Fig. 5 we plot the dispersion relations of the two bound modes of the overall structure of Fig. 3(c), for a gap of $\theta=200$ nm. Again there is a symmetric and an antisymmetric mode, the squares are the results of diagonalizing Eq. (28), and the solid lines the result of an exact numerical solution. In the second two rows of Table I we compare the coupling length at $\lambda=1.54$ μm obtained from the restricted basis solution (28) and the exact numerical solutions for different gap sizes.

Naturally there is more relative error in the coupling length than in the dispersion relations of the individual modes, since the former involves the subtraction of two approximate quantities that are close in magnitude. Nonetheless, these results indicate that for our two sample structures, as indeed for many structures of interest, the use of the

TABLE I. Comparison of the exact coupling length, for the structure in Fig. 3(c), with that predicted by two approximate schemes discussed in this paper. The exact result is calculated using a freely available numerical package (Ref. [17]). The gap size Θ is defined as the distance between the edge of one waveguide to the edge of the other.

gap size	100 nm	150 nm	200 nm	250 nm
Exact	1.88 μm	2.98 μm	4.71 μm	7.43 μm
Restricted basis	1.81 μm	2.93 μm	4.67 μm	7.41 μm
CMT	1.84 μm	2.97 μm	4.71 μm	7.44 μm

restricted basis of parent modes is a good one for the determination of the exact modes. Henceforth we refer to the modes of overall structures identified by Eq. (30), found by solving the eigenvalue equation (28), as the "nominal modes," and treat them as the exact modes. The fields $\mathbf{D}_{mk}(\mathbf{r})$ associated with the $\mathbf{B}_{mk}(\mathbf{r})$ follow from the condition (13); using the corresponding relations (20) that the parent modes satisfy, we can write

$$\mathbf{D}_{mk}(\mathbf{r}) = \sum_j \nu_m^{(j)}(k)\mathbf{D}_k^{(j)}(\mathbf{r}), \tag{33}$$

where

$$\nu_m^{(j)}(k) = \frac{\omega_k^{(j)}}{\omega_{mk}}\eta_m^{(j)}(k). \tag{34}$$

IV. DRESSED MODES

For the derivation of our coupled mode equations we will require neither the nominal modes $\mathbf{B}_{mk}(\mathbf{r})$ (30), nor the parent modes $\mathbf{B}_k^{(i)}(\mathbf{r})$ (19), but rather a set of "dressed modes" that are superpositions of the modes $\mathbf{B}_{mk}(\mathbf{r})$ resembling the parent modes. We define dressed mode operators according to

$$b_{nk} = \sum_m \alpha_m^{(n)}(k)a_{mk}. \tag{35}$$

From this it follows that dressed mode operators satisfy the canonical commutation relation

$$[b_{nk}, b_{n'k'}^\dagger] = \delta_{nn'}\delta(k-k'),$$

$$[b_{nk}, b_{n'k'}] = 0, \tag{36}$$

where we have used the orthonormality relations (32) and the fact that the a_{mk} and $a_{m'k'}^\dagger$ satisfy the canonical commutation relations (17). The inverse relation of Eq. (35) is easily found to be

$$a_{mk} = \sum_n [\alpha_m^{(n)}(k)]^* b_{nk}, \tag{37}$$

and using it we can write the Hamiltonian (18) as

$$H = \hbar \sum_{n,n'} \int dk \gamma^{nn'}(k) b_{nk}^\dagger b_{n'k}, \tag{38}$$

with

$$\gamma^{nn'}(k) = \sum_m \omega_{mk}\alpha_m^{(n)}(k)[\alpha_m^{(n')}(k)]^*. \tag{39}$$

Using relation (37) for the nominal mode amplitudes in terms of the dressed mode amplitudes, and Eq. (30) relating the $\mathbf{B}_{mk}(\mathbf{r})$ to the parent modes, we write our field as

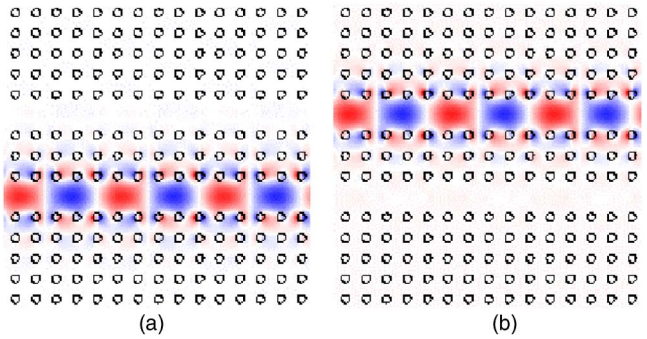

FIG. 6. (Color online) Real part of the x component of the magnetic field (color online) of the two dressed modes for the structure in Fig. 1(c).

$$\mathbf{B}(\mathbf{r},t) = \sum_m \int dk \sqrt{\frac{\hbar\omega_{mk}}{2}} a_{mk}\mathbf{B}_{mk}(\mathbf{r}) + \text{c.c.}$$

$$= \sum_n \int dk \sqrt{\frac{\hbar\omega_k^{\text{ref}}}{2}} b_{nk}\bar{\mathbf{B}}_{nk}(\mathbf{r}) + \text{c.c.}, \tag{40}$$

where we have introduced a reference frequency ω_k^{ref}, which may be chosen to depend on k; we will turn to its choice in the section below. The dressed mode $\bar{\mathbf{B}}_{nk}(\mathbf{r})$ is then identified as

$$\bar{\mathbf{B}}_{nk}(\mathbf{r}) = \sum_j \bar{\eta}_n^{(j)}(k)\mathbf{B}_k^{(j)}(\mathbf{r}), \tag{41}$$

where

$$\bar{\eta}_n^{(j)}(k) = \sum_{l,m} [\mathsf{S}^{-1/2}(k)]^{jl} \sqrt{\frac{\omega_{mk}}{\omega_k^{\text{ref}}}} \alpha_m^{(l)}(k)[\alpha_m^{(n)}(k)]^*. \tag{42}$$

Similarly, we have

$$\mathbf{D}(\mathbf{r},t) = \sum_m \int dk \sqrt{\frac{\hbar\omega_{mk}}{2}} a_{mk}\mathbf{D}_{mk}(\mathbf{r}) + \text{c.c.}$$

$$= \sum_n \int dk \sqrt{\frac{\hbar\omega_k^{\text{ref}}}{2}} b_{nk}\bar{\mathbf{D}}_{nk}(\mathbf{r}) + \text{c.c.}, \tag{43}$$

where

$$\bar{\mathbf{D}}_{nk}(\mathbf{r}) = \sum_j \bar{\nu}_n^{(j)}(k)\mathbf{D}_k^{(j)}(\mathbf{r}), \tag{44}$$

with

$$\bar{\nu}_n^{(j)}(k) = \sum_{l,m} [\mathsf{S}^{-1/2}(k)]^{jl} \frac{\omega_k^{(j)}}{\sqrt{\omega_{mk}\omega_k^{\text{ref}}}} \alpha_m^{(l)}(k)[\alpha_m^{(n)}(k)]^*. \tag{45}$$

While the raising and lowering operators $b_{nk}^\dagger$ and b_{nk} associated with the dressed states clearly satisfy the canonical commutation relations (36), the mode functions $\bar{\mathbf{B}}_{nk}(\mathbf{r})$ and $\bar{\mathbf{D}}_{nk}(\mathbf{r})$ do not satisfy any of the orthonormalization conditions (15). In Figs. 6 and 7 the two dressed states at a given frequency for the structures depicted in Figs. 1(c) and 3(c) are shown; in each case it is evident that the first dressed state has its energy localized in the first waveguide, and the

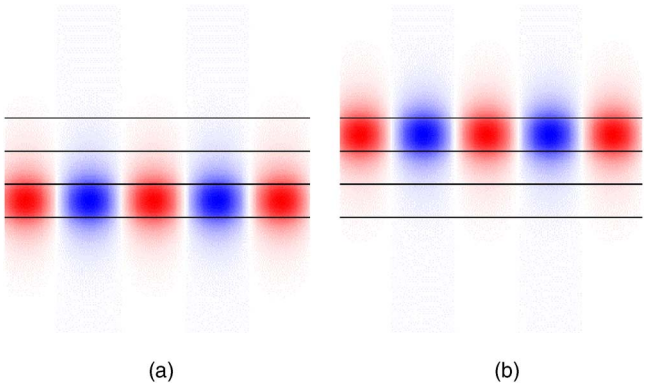

FIG. 7. (Color online) Real part of the x component of the magnetic field (color online) of the two dressed modes for the structure in Fig. 3(c).

second dressed state has its energy localized in the second. Thus they resemble the parent modes. In fact, the dressed modes $\bar{\mathbf{B}}_{nk}(\mathbf{r})$ differ from the parent modes $\mathbf{B}_k^{(n)}(\mathbf{r})$ only because of two effects: (a) the overlap matrix $\mathsf{S}(k)$ is not diagonal and (b) the mode frequencies ω_{mk} differ from the parent mode frequencies $\omega_k^{(i)}$. Indeed, if we set $S^{ij}(k)=\delta_{ij}$ and put all the ω_{mk} equal in Eq. (42), we see using the orthogonality relations (32) that Eq. (41) reduces to $\bar{\mathbf{B}}_{nk}(\mathbf{r})=\mathbf{B}_k^{(n)}(\mathbf{r})$.

The first of these two effects is familiar from problems in quantum chemistry, where often uses a basis of nonorthogonal orbitals and similar overlap matrix element effects arise. However, the appearance of the term involving $\sqrt{\omega_{mk}}$ in Eq. (42) is particular to problems in quantum optics; its appearance can be traced back to the presence of frequency dependent factors appearing in the expressions for the fundamental fields (11) in terms of canonical raising and lowering operators, which arise because of the inherently relativistic nature of the photon.

As alluded to in the Introduction, these frequency factors make our problem more complicated than the simpler situation that arises in quantum chemistry. We can now illustrate that explicitly. In quantum chemistry one typically starts with a Hamiltonian of the form

$$H = H_s + H_{e\text{-}e}, \tag{46}$$

where H_{e-e} describes the electron-electron interaction, which is the analog of nonlinearity in the optical case, and H_s describes the single particle Hamiltonian, which is the analog of the linear Hamiltonian (8) in the optical problem; here

$$H_s = \int \psi^\dagger(\mathbf{r})\mathcal{H}(\mathbf{r},\mathbf{r}')\psi(\mathbf{r}')d\mathbf{r}d\mathbf{r}', \tag{47}$$

where

$$\int \mathcal{H}(\mathbf{r},\mathbf{r}')\psi(\mathbf{r}')d\mathbf{r}' = -\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r}) + V(\mathbf{r})\psi(\mathbf{r}), \tag{48}$$

with $V(\mathbf{r})$ characterizing the single particle potential to which the electrons are subject. In the electron problem the field operators $\psi(\mathbf{r})$ satisfy anticommutation relations

$$\psi(\mathbf{r})\psi^\dagger(\mathbf{r}') + \psi^\dagger(\mathbf{r}')\psi(\mathbf{r}) = \delta(\mathbf{r}-\mathbf{r}'), \tag{49}$$

rather that the commutation relations that are important for bosons, but that difference is not relevant to the issue at hand. The point is that we can begin with *any* set of "parent orbitals" $\phi^\alpha(\mathbf{r})$, whether or not they are orthogonal, and introducing an overlap matrix with elements

$$S^{\alpha\beta} = \int [\phi^\alpha(\mathbf{r})]^* \phi^\beta(\mathbf{r})d\mathbf{r}, \tag{50}$$

we can immediately construct dressed orbitals according to

$$\bar{\phi}_\alpha(\mathbf{r}) = \sum_\beta (S^{-1/2})^{\beta\alpha}\phi^\beta(\mathbf{r}). \tag{51}$$

By construction these dressed orbitals satisfy

$$\int \bar{\phi}_\alpha^*(\mathbf{r})\bar{\phi}_\beta(\mathbf{r})d\mathbf{r} = \delta_{\alpha\beta}, \tag{52}$$

and since they are orthonormal we can use as an expansion basis for our field operators

$$\psi(\mathbf{r}) = \sum_\alpha a_\alpha \bar{\phi}_\alpha(\mathbf{r}), \tag{53}$$

where a_α is a destruction operator associated with the orbital $\bar{\phi}_\alpha(\mathbf{r})$,

$$a_\alpha a_\beta^\dagger + a_\beta^\dagger a_\alpha = \delta_{\alpha\beta}. \tag{54}$$

Then the single-particle part (47) of the Hamiltonian can be written

$$H_s = \sum_{\alpha,\beta}\left(\int \bar{\phi}_\alpha^*(\mathbf{r})\mathcal{H}(\mathbf{r},\mathbf{r}')\bar{\phi}_\beta(\mathbf{r}')d\mathbf{r}d\mathbf{r}'\right)a_\alpha^\dagger a_\beta, \tag{55}$$

similar to Eq. (38). Note that this can be done without solving for the exact eigenstates of $\mathcal{H}(\mathbf{r},\mathbf{r}')$ because any orthonormal basis (52) can be used to expand the field operator $\psi(\mathbf{r})$ in the form (53), where a_α satisfies the canonical anticommutation relations (54). In the optics problem, on the other hand, the expansion (11) cannot be written down for *any* normalized $\mathbf{B}_{mk}(\mathbf{r})$ and any ω_{mk} if the a_{mk} are to satisfy the canonical commutation relations; those modes must be the true modes and the frequencies the true frequencies. Thus here the dressed modes can be constructed only after the true modes are identified.

V. SIMPLIFYING THE DRESSED MODE HAMILTONIAN

A Hamiltonian of the form (38) is simpler than the most general Hamiltonian bilinear in the raising and lowering operators because it contains no "counter-rotating" terms involving the product of two raising or two lowering operators; while no approximations have been made, beyond taking the nominal modes to be exact, the Hamiltonian is of a "rotating-wave" form. Hence it is easier to diagonalize, for example, than the most general bilinear Hamiltonian. Associated with this, we will see below, is the ability to construct coupled mode equations in an easy way. But since the nominal modes

themselves must be found before the dressed modes can be constructed and Eq. (38) identified, a skeptic might still question the benefit of identifying these dressed modes and working in terms of them rather than simply working with the nominal modes. It turns out, however, that to good approximation the dressed mode Hamiltonian (38) can essentially be written down immediately once the parent modes are given. Thus in practice it is not necessary to find the nominal eigenmodes and eigenfrequencies to identify the dressed modes and their coupling constants $\gamma^{nn'}(k)$.

This is possible because, in a typical mode coupling problem at optical frequencies, the true eigenmode frequencies ω_{mk} of interest lie much closer to each other than does any of them to zero frequency. That is, we can choose a reference frequency ω_k^{ref} such that it is reasonable to assume that

$$\Delta_{mk} \equiv \frac{\omega_{mk} - \omega_k^{\mathrm{ref}}}{\omega_k^{\mathrm{ref}}} \tag{56}$$

has a norm much less than unity for all the true eigenmodes m of interest. Let us suppose such a choice has been made; in our structures of Figs. 1(c) and 3(c) the obvious choice of ω_k^{ref} is the frequency of the parent modes at the specified k. We can then write our expression (39) for $\gamma^{nn'}(k)$ in terms of the Δ_{mk} and the $\alpha_m^{(n)}(k)$,

$$\gamma^{nn'}(k) = \omega_k^{\mathrm{ref}}[\delta_{nn'} + \Gamma_1^{nn'}(k)], \tag{57}$$

where we have used the first term of Eq. (32) and defined the components of matrices $\Gamma_p(k)$ according to

$$\Gamma_p^{nn'}(k) \equiv \sum_m (\Delta_{mk})^p \alpha_m^{(n)}(k)[\alpha_m^{(n')}(k)]^*. \tag{58}$$

To come up with an approximation for $\Gamma_1(k)$ and thus $\gamma^{nn'}(k)$, we rewrite Eq. (29) in an exactly equivalent form

$$\sum_j \tilde{N}^{ij}(k)\alpha_m^{(j)}(k) = \left(\Delta_{mk} + \frac{1}{2}\Delta_{mk}^2\right)\alpha_m^{(i)}(k), \tag{59}$$

where we have defined the components of a matrix $\tilde{\mathsf{N}}(k)$ according to

$$\tilde{N}^{ij}(k) \equiv \frac{c^2}{2(\omega_k^{\mathrm{ref}})^2}\left[\tilde{M}^{ij}(k) - \left(\frac{\omega_k^{\mathrm{ref}}}{c}\right)^2 \delta_{ij}\right]. \tag{60}$$

Multiplying the eigenvalue equation (59) by $[\alpha_m^{(l)}(k)]^*$ and summing over m we find

$$\tilde{N}^{il}(k) = \Gamma_1^{il}(k) + \frac{1}{2}\Gamma_2^{il}(k), \tag{61}$$

where we have used the first of Eq. (32) and the definition (58) of the $\Gamma_p(k)$. Since $\Gamma_p(k)$ involves $(\Delta_{mk})^p$, and $|\Delta_{mk}| \ll 1$, it is clear that to lowest order we can here neglect $\Gamma_2(k)$ compared to $\Gamma_1(k)$ and write

$$\Gamma_1^{il}(k) = \tilde{N}^{il}(k) + \cdots. \tag{62}$$

In fact, we show in the Appendix that formal series expansions (A10)–(A12) for the $\Gamma_p(k)$ in terms of powers of $\tilde{\mathsf{N}}(k)$ can be constructed to find the better approximations that we employ below.

To then identify an appropriate expansion for $\gamma^{nn'}(k)$, we write

$$\mathsf{S}(k) = U + \mathsf{s}(k), \tag{63}$$

and

$$\mathsf{N}(k) \equiv \frac{c^2}{2(\omega_k^{\mathrm{ref}})^2}\left[\mathsf{M}(k) - \left(\frac{\omega_k^{\mathrm{ref}}}{c}\right)^2 \mathsf{S}(k)\right], \tag{64}$$

so $\tilde{\mathsf{N}}(k) = \mathsf{S}^{-1/2}(k)\cdot\mathsf{N}(k)\cdot\mathsf{S}^{-1/2}(k)$. Note that $\mathsf{N}(k)$ and $\mathsf{S}(k)$ are matrices with small components that can be constructed immediately, without any diagonalization, once $\mathsf{M}(k)$ and the overlap matrix $\mathsf{S}(k)$ are calculated. From Eqs. (57) and (A10) we then find the expansion

$$\frac{\gamma^{nn'}(k)}{\omega_k^{\mathrm{ref}}} = \delta^{nn'} + Y_{\gamma 1}^{nn'}(k) + Y_{\gamma 2}^{nn'}(k) + Y_{\gamma 3}^{nn'}(k) + \cdots, \tag{65}$$

where

$$\mathsf{Y}_{\gamma 1} = \mathsf{N},$$

$$\mathsf{Y}_{\gamma 2} = -\frac{1}{2}\mathsf{N}^2 - \frac{1}{2}(\mathsf{Ns} + \mathsf{sN}),$$

$$\mathsf{Y}_{\gamma 3} = \frac{1}{2}\mathsf{N}^3 + \frac{1}{4}(\mathsf{sN}^2 + \mathsf{N}^2\mathsf{s}) + \frac{1}{2}\mathsf{NsN} + \frac{3}{8}(\mathsf{Ns}^2 + \mathsf{s}^2\mathsf{N}) + \frac{1}{4}\mathsf{sNs}.$$

Here we have used the expansions (A10)–(A12) $\Gamma_p(k)$ in terms of powers $\tilde{\mathsf{N}}(k)$, and expanding $\tilde{\mathsf{N}}(k)$ in terms of powers of $\mathsf{N}(k)$ and $\mathsf{s}(k)$.

We can also use this approach to find approximate expressions for $\bar{\mathbf{B}}_{nk}(\mathbf{r})$ and $\bar{\mathbf{D}}_{nk}(\mathbf{r})$ without having to solve the eigenvalue equation for ω_{mk}. Returning to Eqs. (42) and (45) and writing

$$\left(\frac{\omega_{mk}}{\omega_k^{\mathrm{ref}}}\right)^{\pm 1/2} = (1 + \Delta_{mk})^{\pm 1/2}, \tag{66}$$

from Eq. (56), we can use the expansion of these terms in powers of Δ_{mk} to generate approximate expressions for $\bar{\eta}_n^{(j)}(k)$ and $\bar{\nu}_n^{(j)}(k)$ in terms of the $\Gamma_p(k)$ and $\mathsf{S}(k)$. Then again using the expansions (A10)–(A12) of $\Gamma_p(k)$ in terms of powers $\tilde{\mathsf{N}}(k)$, and expanding $\tilde{\mathsf{N}}(k)$ in terms of powers of $\mathsf{N}(k)$ and $\mathsf{S}(k)$, we find

$$\bar{\eta}_n^{(j)}(k) = \delta^{jn} + Y_{\eta 1}^{nn'}(k) + Y_{\eta 2}^{nn'}(k) + Y_{\eta 3}^{nn'}(k) + \cdots, \tag{67}$$

where

$$\mathsf{Y}_{\eta 1} = \frac{1}{2}\mathsf{N} - \frac{1}{2}\mathsf{s},$$

$$\mathsf{Y}_{\eta 2} = -\frac{3}{8}\mathsf{N}^2 - \frac{1}{4}\mathsf{Ns} - \frac{1}{2}\mathsf{sN} + \frac{3}{8}\mathsf{s}^2,$$

TABLE II. Comparison of quantities $\mu^{nn'}$, $\tilde{\eta}_n^{(j)}$, and $\tilde{\nu}_n^{(j)}$ associated with the structure depicted in Fig. 1(c) using exact and approximate expressions.

	$\mu^{11}=\mu^{22}$	$\mu^{12}=\mu^{21}$	$\tilde{\eta}_1^1=\tilde{\eta}_2^2$	$\tilde{\eta}_1^2=\tilde{\eta}_2^1$	$\tilde{\nu}_1^1=\tilde{\nu}_2^2$	$\tilde{\nu}_1^2=\tilde{\nu}_2^1$
1st order	3.9775×10^{-4}	1.2698×10^{-4}	1.9887×10^{-4}	-7.6527×10^{-3}	-1.9887×10^{-4}	-1.1335×10^{-2}
2nd order	3.2098×10^{-4}	1.2667×10^{-3}	2.7649×10^{-4}	-7.6594×10^{-3}	-2.6780×10^{-6}	-1.1327×10^{-2}
3rd order	3.2119×10^{-4}	1.2672×10^{-3}	2.7669×10^{-4}	-7.6601×10^{-3}	-2.9145×10^{-6}	-1.1331×10^{-2}
exact	3.2115×10^{-4}	1.2672×10^{-3}	2.7668×10^{-4}	-7.6601×10^{-3}	-2.8346×10^{-6}	-1.1331×10^{-2}

$$\mathsf{Y}_{\eta 3}=\frac{7}{16}\mathsf{N}^3+\frac{3}{16}\mathsf{N}^2\mathsf{s}+\frac{3}{8}\mathsf{N}\mathsf{s}\mathsf{N}+\frac{3}{8}\mathsf{s}\mathsf{N}^2+\frac{3}{16}\mathsf{N}\mathsf{s}^2+\frac{1}{4}\mathsf{s}\mathsf{N}\mathsf{s}+\frac{1}{2}\mathsf{s}^2\mathsf{N}-\frac{5}{16}\mathsf{s}^3$$

and

$$\frac{\omega_k^{\mathrm{ref}}}{\omega_k^{(j)}}\bar{\nu}_n^{(j)}(k)=\delta^{jn}+Y_{\nu 1}^{nn'}(k)+Y_{\nu 2}^{nn'}(k)+Y_{\nu 3}^{nn'}(k)+\cdots, \tag{68}$$

where

$$\mathsf{Y}_{\nu 1}=-\frac{1}{2}\mathsf{N}-\frac{1}{2}\mathsf{s},$$

$$\mathsf{Y}_{\nu 2}=\frac{5}{8}\mathsf{N}^2+\frac{1}{4}\mathsf{N}\mathsf{s}+\frac{1}{2}\mathsf{s}\mathsf{N}+\frac{3}{8}\mathsf{s}^2,$$

$$\mathsf{Y}_{\nu 3}=-\frac{15}{16}\mathsf{N}^3-\frac{5}{16}\mathsf{N}^2\mathsf{s}-\frac{5}{8}\mathsf{N}\mathsf{s}\mathsf{N}-\frac{5}{8}\mathsf{s}\mathsf{N}^2-\frac{3}{16}\mathsf{N}\mathsf{s}^2-\frac{1}{4}\mathsf{s}\mathsf{N}\mathsf{s}-\frac{1}{2}\mathsf{s}^2\mathsf{N}-\frac{5}{16}\mathsf{s}^3.$$

To investigate the accuracy of this approach we introduce quantities

$$\mu^{nn'}(k)\equiv\frac{\gamma^{nn'}(k)}{\omega_k^{\mathrm{ref}}}-\delta^{nn'}, \tag{69}$$

$$\tilde{\eta}_n^{(j)}(k)\equiv\bar{\eta}_n^{(j)}(k)-\delta^{jn},$$

$$\tilde{\nu}_n^{(j)}(k)\equiv\frac{\omega_k^{\mathrm{ref}}}{\omega_k^{(j)}}\bar{\nu}_n^{(j)}(k)-\delta^{jn}.$$

In Table II we consider the structure of Fig. 1(c) and evaluate these quantities at a k that, for the parent modes, corresponds to a vacuum wavelength of $\lambda=1.54$ μm. We do this both by using the exact expressions (39), (42), and (45), and by using the corresponding approximate expressions (65), (67), and (68) to different orders, where by an mth order calculation we mean including the $Y_{is}^{nn'}(k)$ up to $s=m$, for $i=\gamma$, η, and ν. In Table III we present the corresponding data for the structure of Fig. 3(c) at $\lambda=1.54$ μm. We see that doing even a low order calculation of the quantities $\mu^{nn'}(k)$, $\tilde{\eta}_n^{(j)}(k)$, and $\tilde{\nu}_n^{(j)}(k)$ yields a good approximation to the exact values.

VI. EFFECTIVE FIELDS AND THE COUPLED MODE EQUATIONS

The dressed mode Hamiltonian (38) is precisely of the form required for a straightforward derivation of coupled mode equations to describe the propagation of light. We begin by writing that Hamiltonian in a form that explicitly separates out the small terms involving the coupling between the parent modes

$$H=\int dk\hbar\omega^{\mathrm{self}}(k)\sum_n b_{nk}^\dagger b_{nk}+\int dk\hbar\omega^{\mathrm{cross}}(k)\sum_{n,n'}{}' b_{nk}^\dagger b_{n'k}, \tag{70}$$

where the prime on the second sum indicates that the terms with $n'=n$ are to be excluded; we have also specialized to our particular sample structures and put

$$\hbar\omega^{\mathrm{self}}(k)\equiv\gamma^{11}(k)=\gamma^{22}(k),$$

$$\hbar\omega^{\mathrm{cross}}(k)\equiv\gamma^{12}(k)=\gamma^{21}(k).$$

For pulses centered about a frequency ω equal to $\omega_{\bar{k}}^{\mathrm{ref}}$, for some $\bar{k}$, we follow earlier work [11] and introduce effective fields according to

TABLE III. Comparison of quantities $\mu^{nn'}$, $\tilde{\eta}_n^{(j)}$, and $\tilde{\nu}_n^{(j)}$ associated with the structure depicted in Fig. 3(c) using exact and approximate expressions.

	$\mu^{11}=\mu^{22}$	$\mu^{12}=\mu^{21}$	$\tilde{\eta}_1^1=\tilde{\eta}_2^2$	$\tilde{\eta}_1^2=\tilde{\eta}_2^1$	$\tilde{\nu}_1^1=\tilde{\nu}_2^2$	$\tilde{\nu}_1^2=\tilde{\nu}_2^1$
1st order	-1.8576×10^{-4}	-8.0656×10^{-4}	-9.2878×10^{-5}	-4.7962×10^{-2}	9.2878×10^{-4}	-2.3068×10^{-2}
2nd order	1.2726×10^{-3}	-8.0628×10^{-4}	2.8929×10^{-3}	-4.7956×10^{-2}	1.0461×10^{-3}	-2.3073×10^{-2}
3rd order	1.2721×10^{-3}	-8.0918×10^{-4}	2.8925×10^{-3}	-4.8151×10^{-2}	1.0462×10^{-3}	-2.3121×10^{-2}
exact	1.2778×10^{-3}	-8.0919×10^{-4}	2.9056×10^{-3}	-4.8152×10^{-2}	1.0490×10^{-3}	-2.3121×10^{-2}

$$g_n(z,t) = \frac{1}{\sqrt{2\pi}} \int_{\mathrm{BZ}} dk b_{nk} e^{i(k-\bar{k})z}, \tag{71}$$

where the integral ranges over the first Brillouin zone. From the fact that the b_{nk} and $b_{nk}^{\dagger}$ satisfy the canonical commutation relations follows the commutation relation for the effective fields,

$$[g_{n'}(z',t), g_n^{\dagger}(z,t)] = \delta_{n,n'} \hat{\delta}(z-z'), \tag{72}$$

where $\hat{\delta}(z-z')$ represents an effective delta function, in the sense that

$$f(z,t) = \int dz' f(z',t) \hat{\delta}(z'-z) \tag{73}$$

when the function $f(z,t)$ contains only wave number (Fourier components) in the first Brillouin zone. The idea is that the effective fields $g_n(z,t)$ are assumed to be slowly varying functions of position compared with the length scale $2\pi/\bar{k}$, and so the expression (71) contains significant contributions only from k close to $\bar{k}$. We then expand the ω_k^{ref} appearing in the first term of Eq. (70) about $\bar{k}$, and write

$$H \approx H_{cme} \equiv \int dk [\hbar \omega^{\mathrm{self}}(\bar{k}) + \hbar v_g (k-\bar{k})] \sum_n b_{nk}^{\dagger} b_{nk} + \int dk \sum_{n,n'}{}' \hbar \omega^{\mathrm{cross}}(\bar{k}) b_{nk}^{\dagger} b_{n'k} \tag{74}$$

in place of Eq. (70), where

$$v_g = \left. \frac{\partial \omega^{\mathrm{self}}(\bar{k})}{\partial k} \right|_{k=\bar{k}} \tag{75}$$

is the group velocity associated with the parent modes. In the second term in Eq. (74) we have evaluated $\hbar \omega^{\mathrm{cross}}(k)$ at $k=\bar{k}$; since we assume that $\mu^{12}(k) = \mu^{21}(k) \ll 1$ the effect of this term is small, and we neglect its dispersion. This approach can be made more rigorous within the context of a formal multiple scales analysis, which we do not pursue here. Using (71) in the approximate Hamiltonian (74) we find we can write H_{cme} as

$$H_{cme} = \hbar \omega^{\mathrm{self}}(\bar{k}) \sum_n \int g_n^{\dagger}(z) g_n(z) dz + \frac{i}{2} \hbar v_g \sum_n \int \left(\frac{\partial g_n^{\dagger}(z,t)}{\partial z} g_n(z,t) - g_n^{\dagger}(z,t) \frac{\partial g_n(z,t)}{\partial z} \right) dz + \hbar \omega^{\mathrm{cross}}(\bar{k}) \sum_{n,n'}{}' \int g_n^{\dagger}(z) g_{n'}(z) dz.$$

Applying this Hamiltonian to our examples, from the dynamical equations for the effective fields

$$\frac{\partial}{\partial t} g_n(z,t) = -\frac{i}{\hbar} [g_n(z,t), H_{cme}], \tag{76}$$

for $n=1,2$ labeling the dressed modes of our examples (see Figs. 6 and 7), we find

$$\frac{\partial g_{1,2}}{\partial t} + v_g \frac{\partial g_{1,2}}{\partial z} = -i\omega^{\mathrm{self}}(\bar{k}) g_{1,2}(z,t) - i\omega^{\mathrm{cross}}(\bar{k}) g_{2,1}(z,t). \tag{77}$$

Equations (77) are the (linear) coupled mode equations for the examples we consider here. We identify the coupling length they predict in the usual way [18] by looking for a stationary solution

$$g_{1,2}(z,t) = \tilde{g}_{1,2}(z) \exp[-i\omega^{\mathrm{self}}(\bar{k}) t], \tag{78}$$

and finding that for $\tilde{g}_2(0)=0$ we have

$$|\tilde{g}_1(z)|^2 = |\tilde{g}_1(0)|^2 \cos^2 \kappa z,$$

$$|\tilde{g}_2(z)|^2 = |\tilde{g}_1(0)|^2 \sin^2 \kappa z,$$

where

$$\kappa \equiv \omega^{\mathrm{cross}}(\bar{k})/v_g, \tag{79}$$

and so we find the coupling length predicted to be

$$\Delta z = \frac{\pi}{2|\kappa|} = \frac{\pi v_g}{2|\omega^{\mathrm{cross}}(\bar{k})|}. \tag{80}$$

For the second of our sample structures [see Fig. 3(c)], we give in the third row of Table I the predicted value Δz of the coupling length that follows from this formula for different gap sizes. The small differences from the "restricted basis" results are due solely to the approximations made in moving from the Hamiltonian (70) to the Hamiltonian (74), since the dressed mode Hamiltonian (38) is equivalent to the exact Hamiltonian (18) within the approximation that the nominal modes found within the restricted basis are equal to the exact modes. Less drastic approximations in moving from Eq. (70)–(74), such as keeping higher terms in the expansion of $\omega^{\mathrm{self}}(k)$ and keeping the lowest order dispersion in $\omega^{\mathrm{cross}}(k)$, can easily be constructed. They would lead to more complicated coupled mode equations, but a more accurate description of the coupling within the restricted basis assumption. However, since at least for the structures considered here the error in the coupling length associated with the approximations in this section is comparable to the error associated with using the restricted basis, such an extension does not seem justified.

VII. DISCUSSION

We have presented a Hamiltonian formulation of linear coupled mode theory applicable to photonic crystal structures, and other waveguiding structures. The modes are associated with parent structures that are related in a natural way to the actual structure of interest, as the parent structures of Figs. 1(a) and 1(b) are related to the actual structure of

Fig. 1(c), and the parent structures of Figs. 3(a) and 3(b) are related to the actual structure of Fig. 3(c). Yet the "coupled modes" in the theory are not modes of the parent structures, but rather dressed modes that resemble them. Even within the approximation that a restricted set of modes of the parent structures form a basis for expanding the exact modes of the actual structure, these exact modes must, strictly speaking, be found before the dressed modes can be constructed. However, if as usual the frequency shifts due to the coupling of the dressed modes are small, in practice the dressed modes can be found immediately as series expansions involving modes of the parent structures. While the use of such an expansion is arguably not needed for the simple structures we considered as our examples in this paper, it will present a real advantage when many coupled modes are involved, such as in a SCISSOR structure where modes associated with a number of resonators are coupled with travelling modes in nearby waveguide channels [7]. We plan to turn to this in a future communication.

The resulting linear dressed mode Hamiltonian, which is simply the Hamiltonian of the linear system written in terms of canonical raising and lowering operators for the dressed modes, contains no "counter-rotating" terms, and forms the natural starting point for deriving the coupled mode equations. Our sample calculations for two typical systems illustrate the accuracy of the approximations inherent in this kind of approach. Better approximations to the exact modes can naturally be achieved by using more parent modes in the expansion, and the accuracy of the coupled mode equations can be improved by including terms involving spatial derivatives higher than the first.

One of the strengths of this approach is its easy generalization to nonlinear quantum optics. A (nonresonant) nonlinear response is typically described by an additional, nonlinear polarization of the form

$$P^i_{\mathrm{NL}}(\mathbf{r},t) = \varepsilon_0 \chi_2^{ijk}(\mathbf{r}) E^j(\mathbf{r},t) E^k(\mathbf{r},t) + \varepsilon_0 \chi_3^{ijkl}(\mathbf{r}) E^j(\mathbf{r},t) E^k(\mathbf{r},t) E^l(\mathbf{r},t) + \cdots, \quad (81)$$

where $\chi_2^{ijk}(\mathbf{r})$ and $\chi_3^{ijkl}(\mathbf{r})$ are, respectively, the usual second- and third-order susceptibility tensors of nonlinear optics, which we here allow to depend on position. From our point of view it is more convenient to construct this kind of power series expansion in terms of our fundamental field $\mathbf{D}(\mathbf{r},t)$, writing

$$P^i_{\mathrm{NL}}(\mathbf{r},t) = \Gamma_2^{ijk}(\mathbf{r}) D^j(\mathbf{r},t) D^k(\mathbf{r},t) + \Gamma_3^{ijkl}(\mathbf{r}) D^j(\mathbf{r},t) D^k(\mathbf{r},t) D^l(\mathbf{r},t) + \cdots. \quad (82)$$

As in our discussion of the linear response in Sec. II, there is no loss of generality here, because the $\Gamma's$ can easily be constructed in terms of the $\chi's$ using the second of Eq. (3). Then the full Hamiltonian is obtained by replacing the H of Eq. (8) by

$$H_{\mathrm{NL}} = H - \frac{1}{3\varepsilon_0} \int d\mathbf{r} D^i(\mathbf{r}) \Gamma_2^{ijk}(\mathbf{r}) D^j(\mathbf{r}) D^k(\mathbf{r}) - \frac{1}{4\varepsilon_0} \int d\mathbf{r} D^i(\mathbf{r}) \Gamma_3^{ijkl}(\mathbf{r}) D^j(\mathbf{r}) D^k(\mathbf{r}) D^l(\mathbf{r}) + \cdots,$$

as discussed earlier [11]. The route to nonlinear coupled mode equations is now straightforward. As in the linear limit we take $H \to H_{cme}$ [see Eq. (74)] and then in the nonlinear contributions to H_{NL} above we substitute our expression (43) for $\mathbf{D}(\mathbf{r})$ in terms of the dressed modes. The derivation then proceeds in precisely the same way as was earlier shown [11,12] for the kind of single parent type of theory mentioned in the Introduction; the expression for $H_{\mathrm{NL}} - H$ can be written, within the usual approximations, in terms of powers of $g_n(z,t)$ and $g_n^\dagger(z,t)$. Hence the inclusion of nonlinear effects in the canonical description of a multiple parent structure is now no longer more complicated than it is for single parent structures. We plan to turn to this in detail, and its applications to artificially structured materials, in later publications. Particularly interesting is the extension to describe coupling between N coupled nonlinear waveguides and the description of discrete spatial solitons with a $\chi^{(3)}$ or cascaded $\chi^{(2)}$ nonlinearity. To date, coupled mode equations governing the evolution of the electromagnetic field in discrete coupled systems have been derived phenomenologically, with the nonlinear coefficients determined by numerical means (e.g., by finite difference time domain simulations). The results presented here should facilitate the derivation of nonlinear coefficients in these phenomenological coupled equations for discrete systems, and in particular aid in understanding the physics of their size and behavior at different frequencies. Such an easy generalization to nonlinear structures within a canonical framework suitable for quantization is not available for other coupled mode theory presented in the literature [2]. As well, since the Hamiltonian we use is identified with the energy of the electromagnetic field, generalizations to include in the Hamiltonian the interaction of modes with nearby atoms, molecules, or quantum dots is straightforward [19–22].

ACKNOWLEDGMENTS

The authors acknowledge financial support from Photonics Research Ontario (PRO), the Natural Science and Engineering Research Council of Canada, and the DARPA Slow-Light program. P.C. acknowledges an Ontario Graduate Scholarship.

APPENDIX

Here we detail a formal series expansion for the matrix $\Gamma_1(k)$ introduced in Sec. V. For simplicity, we keep the k label on all quantities implicit in this appendix, writing Γ_1 in place of $\Gamma_1(k)$, for example. From Eq. (59) we can, by applying the matrix $\tilde{\mathsf{N}}$ to both sides of the equation $p-1$ times, find

$$\sum_j \Lambda_p^{ij} \alpha_m^{(j)} = \left(\Delta_m + \frac{1}{2}\Delta_m^2\right)^p \alpha_m^{(i)}, \tag{A1}$$

where we have defined the matrices

$$\Lambda_p \equiv \tilde{\mathsf{N}}^p. \tag{A2}$$

Proceeding as in Sec. V, we multiply both sides of the equation by $(\alpha_m^{(l)})^*$ and sum over m to find

$$\begin{aligned}\Lambda_p^{il} &= \sum_m \left(\Delta_m + \frac{1}{2}\Delta_m^2\right)^p \alpha_m^{(i)} (\alpha_m^{(l)})^* \\ &= \sum_m \sum_{q=0}^{p} \binom{p}{q} (\Delta_m)^{p-q} \left(\frac{1}{2}\Delta_m^2\right)^q \alpha_m^{(i)} (\alpha_m^{(l)})^* \\ &= \sum_m \left[(\Delta_m)^p + \sum_{q=1}^{p} \binom{p}{q} \frac{\Delta_m^{p+q}}{2^q}\right] \alpha_m^{(i)} (\alpha_m^{(l)})^*,\end{aligned}$$

and so using our definitions (58) of the Γ_p we can write this as

$$\Lambda_p = \Gamma_p - \sum_{s=1}^{\infty} c_{ps} \Gamma_s, \tag{A3}$$

where we have put

$$\begin{aligned} c_{p(p+q)} &\equiv -\frac{1}{2^q}\binom{p}{q} \quad \text{if } q = 1,2,\ldots,p \\ &\equiv 0 \quad \text{otherwise}. \end{aligned}$$

That is c_{ps} only exists for $s > p$, and in particular only for $s = p+1, p+2, \ldots, 2p$. We can then write Eq. (A3) as

$$\sum_s K_{ps} \Gamma_s = \Lambda_p, \tag{A4}$$

where we have set

$$K_{ps} \equiv \delta_{ps} - c_{ps}. \tag{A5}$$

Then we have

$$\Gamma_p = \sum_s (K^{-1})_{ps} \Lambda_s. \tag{A6}$$

Requiring

$$\sum_s (K^{-1})_{ps} K_{sr} = \delta_{pr} \tag{A7}$$

and using the definition (A5) for K_{sr}, we find the equation

$$(K^{-1})_{ps} = \delta_{ps} + \sum_r (K^{-1})_{pr} c_{rs}. \tag{A8}$$

Iterating this equation, we find

$$(K^{-1})_{ps} = \delta_{ps} + c_{ps} + \sum_q c_{pq} c_{qs} + \sum_{q,r} c_{pq} c_{qr} c_{rs} + \cdots. \tag{A9}$$

Recalling the range over indices over which the c_{ps} are nonzero, we can easily construct the first few terms of $(K^{-1})_{1s}$:

$$(K^{-1})_{11} = 1,$$

$$(K^{-1})_{12} = c_{12} = -\frac{1}{2},$$

$$(K^{-1})_{13} = c_{12}c_{23} = \frac{1}{2},$$

$$(K^{-1})_{14} = c_{12}c_{24} + c_{12}c_{23}c_{24} = -\frac{5}{8},$$

$$(K^{-1})_{15} = c_{12}c_{23}c_{35} + c_{12}c_{24}c_{45} + c_{12}c_{23}c_{34}c_{45} = \frac{7}{8},$$

and so we have

$$\Gamma_1 = \tilde{\mathsf{N}} - \frac{1}{2}\tilde{\mathsf{N}}^2 + \frac{1}{2}\tilde{\mathsf{N}}^3 - \frac{5}{8}\tilde{\mathsf{N}}^4 + \frac{7}{8}\tilde{\mathsf{N}}^5 + \cdots, \tag{A10}$$

where we have used the definition (A2) of the Λ_p. Similarly we find

$$(K^{-1})_{22} = 1,$$

$$(K^{-1})_{23} = c_{23} = -1,$$

$$(K^{-1})_{24} = c_{24} + c_{23}c_{34} = \frac{5}{4},$$

$$(K^{-1})_{25} = c_{24}c_{45} + c_{23}c_{35} + c_{23}c_{34}c_{45} = -\frac{7}{4}$$

and

$$(K^{-1})_{33} = 1,$$

$$(K^{-1})_{34} = c_{34} = -\frac{3}{2},$$

$$(K^{-1})_{35} = c_{35} + c_{34}c_{45} = \frac{9}{4},$$

which yields

$$\Gamma_2 = \tilde{\mathsf{N}}^2 - \tilde{\mathsf{N}}^3 + \frac{5}{4}\tilde{\mathsf{N}}^4 - \frac{7}{4}\tilde{\mathsf{N}}^5 + \cdots, \tag{A11}$$

$$\Gamma_3 = \tilde{\mathsf{N}}^3 - \frac{3}{2}\tilde{\mathsf{N}}^4 + \frac{9}{4}\tilde{\mathsf{N}}^5 + \cdots. \tag{A12}$$

[1] For the standard textbook treatment see, e. g., A. Yariv and P. Yeh, *Optical Waves in Crystals* (Wiley, New York, 1984).
[2] An early reference is the work of S. M. Jensen, IEEE J. Quantum Electron. **18**, 1580 (1982).
[3] J. Perina Jr. and J. Perina, Prog. Opt. **41**, 361 (2000), and reference therein.
[4] A. Berzanskia, K. H. Feller, and A. Stabinis, Opt. Commun. **118**, 438 (1995).
[5] J. Janszky, A. Petak, C. Sibilia, M. Bertolotti, and P. Adam, Quantum Semiclassic. Opt. **7**, 145 (1995).
[6] E. Yablonovitch, Phys. Rev. Lett. **58**, 2059 (1987).
[7] John E. Heebner, Robert W. Boyd, and Q-Han Park, J. Opt. Soc. Am. B **19**, 722 (2002).
[8] S. John, Phys. Rev. Lett. **58**, 2486 (1987).
[9] S. Kuchinsky, V. Y. Golyatin, A. Y. Kutikov, T. P. Pearsall, and D. Nedeljkovic, IEEE J. Quantum Electron. **38**, 1349 (2002).
[10] D. Michaelis, U. Peschel, C. Wachter, and A. Brauer, Phys. Rev. E **68**, 065601(R) (2003).
[11] J. E. Sipe, Navin A. R. Bhat, Philip Chak, and Suresh Pereira, Phys. Rev. E **69**, 016604 (2004).
[12] S. N. Volkov and J. E. Sipe, Phys. Rev. E **70**, 066621 (2004).
[13] J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1998).
[14] M. Born, Proc. R. Soc. London, Ser. A **143**, 410 (1934); M. Born and L. Infeld, *ibid.* **144**, 425 (1934).
[15] J. D. Joannopoulos, R. D. Meade, and J. N. Winn, *Photonic Crystals* (Princeton University Press, Princeton, 1995).
[16] For convenience, we use a slightly different normalization condition than in earlier work (Ref. [11,12]).
[17] Steven G. Johnson and J. D. Joannopoulos, Opt. Express **8**, 173 (2001).
[18] A. Yariv, *Quantum Electronics* (Wiley, New York, 1984).
[19] D. N. Christodoulildes and R. I. Joseph, Opt. Lett. **13**, 794 (1988).
[20] D. J. Griffiths, *Introduction to Quantum Mechanics* (Prentice Hall, Upper Saddle River, NJ, 1994).
[21] Paul E. Barclay, Kartik Srinivasan, and Oskar Painter, J. Opt. Soc. Am. B **20**, 2274 (2003).
[22] A. Taflove and S. Hagness, *Computational Electrodynamics* (Artech House, Norwood, MA, 2000).