Perturbation theory for Maxwell's equations with shifting material boundaries

Steven G. Johnson, M. Ibanescu, M. A. Skorobogatiy,* O. Weisberg, J. D. Joannopoulos,
and Y. Fink

OmniGuide Communications, One Kendall Square, Building 100, No. 3, Cambridge, Massachusetts 02139

Perturbation theory permits the analytic study of small changes on known solutions, and is especially useful in electromagnetism for understanding weak interactions and imperfections. Standard perturbation-theory techniques, however, have difficulties when applied to Maxwell's equations for small shifts in dielectric interfaces (especially in high-index-contrast, three-dimensional systems) due to the discontinous field boundary conditions—in fact, the usual methods fail even to predict the lowest-order behavior. By considering a sharp boundary as a limit of anisotropically smoothed systems, we are able to derive a correct first-order perturbation theory and mode-coupling constants, involving only surface integrals of the unperturbed fields over the perturbed interface. In addition, we discuss further considerations that arise for higher-order perturbative methods in electromagnetism.

I. INTRODUCTION

Perturbation theory, a class of techniques to find the effect of *small* changes on *known* solutions to a set of equations, is as important a tool for classical electromagnetism as it is for quantum mechanics and other fields. Not only does it allow one to apply the computational efficiency of idealized systems to more realistic problems, or to study effects too small and weak to easily characterize numerically, but it also provides a window of analytical insight into complex systems otherwise accessible only via opaque numerical experiments. Surprisingly, the standard forms of perturbation theory for electromagnetism [1–3] have serious limitations, rarely remarked upon, when handling material-boundary perturbations [4–6] as depicted in Fig. 1. Since Maxwell's equations for lossless media can be cast in the form of a generalized Hermitian eigenproblem in the frequency ω [6,7] or, for a waveguide, the axial wave number β [5,6], it might, at first, seem that the general algebraic machinery of perturbation theory developed in quantum mechanics [8] could apply directly. We show, however, that the vectorial nature of the electromagnetic field and its peculiar boundary conditions at discontinuous material interfaces require that special care be taken in applying perturbative methods to the common problem of slightly shifted interfaces (e.g., from fabrication disorder). In fact, ordinary perturbation theory produces expressions that are *ill defined*, and we demonstrate how correct general expressions can be derived using a limit of anisotropically smoothed systems. (We do not consider shifting metallic boundaries, for which accurate methods are already available [9].) Finally, we remark on further considerations that arise in evaluating perturbation theory for Maxwell's equations beyond the first order.

A clear statement of the problem with the usual perturbation theory when it is applied to shifting dielectric boundaries can be found in Ref. [4], similar to our discussion in Sec. II A. They constructed a Green's function formulation for rough-surface scattering, limited to one-dimensional unperturbed systems $\varepsilon(z)$. Here, in contrast, we treat completely arbitrary, three-dimensional unperturbed systems, and derive the perturbed eigenvalue and the eigenmode coupling coefficients rather than an integral equation for scattered light. In the special case of a shifted flat boundary in a waveguide of uniform cross section, our result for the eigenvalue is equivalent to a variational approach that was postulated, without proof, in Ref. [10]. For another particular case, that of a waveguide with a slowly z-varying cross section, our mode-coupling coefficients are the same as previous expressions derived by matching boundary conditions [1,11]. Moreover, our result reduces to the conventional perturbative expressions in the limit of low index contrast, where those expressions have been most commonly employed (e.g., as in Ref. [12]).

II. PERTURBATION THEORY IN ELECTROMAGNETISM

We first review the application of standard perturbation-theory techniques to electromagnetism, employing an explicit analogy with the algebraic eigenproblem framework
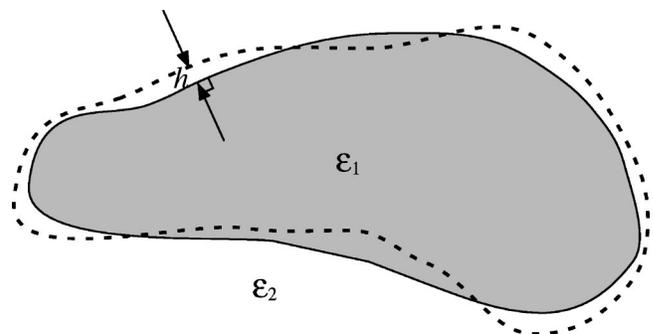

FIG. 1. Schematic of a perturbation due to a shifting interface between two dielectrics ε_1 and ε_2, where the shift (to the dashed boundary) is by some small amount h (measured perpendicular to the unperturbed boundary), which may vary across the surface. For Large index contrasts, a modified perturbation theory is required to obtain the correct change in the eigensolutions.

*Electronic address: maksim@omni-guide.com

developed for quantum mechanics. There are a number of ways in which Maxwell's equations can be written as an eigenproblem, but it suffices to focus on one: the generalized eigenproblem in the electric field $|E\rangle$ with time dependence $e^{-i\omega t}$ in a source-free linear dielectric $\varepsilon(\mathbf{x})$:

$$\mathbf{\nabla}\times\mathbf{\nabla}\times|E\rangle=\left(\frac{\omega}{c}\right)^2\varepsilon|E\rangle, \tag{1}$$

where we use the Dirac notation of basis-independent state kets $|E\rangle$ and inner products $\langle E|E'\rangle\equiv\int\mathbf{E}^*\cdot\mathbf{E}'\,dV$ [8,13]. Assuming that ε is purely real (lossless) and positive, then this eigenproblem is Hermitian and positive semidefinite, leading to real-ω solutions. Since it is a generalized eigenproblem, the eigenstates are orthogonal under the inner product $\langle E|\varepsilon|E'\rangle$. Typically, one is concerned with bound modes (e.g., in a cavity or waveguide) and/or periodic systems where Bloch's theorem applies [7], and so the integrals are effectively of finite spatial extent and the eigenvalues are discrete.

To apply perturbation theory, one must have some small parameter $\Delta\alpha$ characterizing the perturbation—for example, $\Delta\alpha$ could be $\sim\Delta\varepsilon$ for a *small* change in ε, or the volume $\Delta\alpha\sim\Delta V$ of the change for a small boundary shift of a piecewise constant ε. Then, in the standard method [8], the new eigensolutions $|E\rangle$ and ω are expanded in powers n of $\Delta\alpha$: $|E\rangle=\Sigma_{n=0}^{\infty}|E^{(n)}\rangle$ and $\omega=\Sigma_{n=0}^{\infty}\omega^{(n)}$, with $|E^{(0)}\rangle$ and $\omega^{(0)}$ being the unperturbed eigensolution and where the nth term is proportional to $(\Delta\alpha)^n$. Corrections $|E^{(n>0)}\rangle$ are defined such that $\langle E^{(0)}|\varepsilon|E^{(n>0)}\rangle=0$, and the series are substituted into Eq. (1) and solved order by order. [In the case of degenerate (equal-ω) unperturbed modes, the well-known modification of degenerate perturbation theory must be applied: linear combinations are chosen to diagonalize the first-order correction [8].] The first-order correction $\omega^{(1)}$ from a perturbation $\Delta\varepsilon$ is then easily found to be

$$\omega^{(1)}=-\frac{\omega^{(0)}}{2}\frac{\langle E^{(0)}|\Delta\varepsilon|E^{(0)}\rangle}{\langle E^{(0)}|\varepsilon|E^{(0)}\rangle}. \tag{2}$$

This can be thought of as either an approximate expression for the change in ω due to the perturbation (accurate as long as $\langle E^{(0)}|\Delta\varepsilon|E^{(0)}\rangle$ is small), or an *exact* expression for the derivative of ω in the limit of infinitesimal $\Delta\alpha$,

$$\frac{d\omega}{d\alpha}=-\frac{\omega^{(0)}}{2}\frac{\left\langle E^{(0)}\left|\frac{d\varepsilon}{d\alpha}\right|E^{(0)}\right\rangle}{\langle E^{(0)}|\varepsilon|E^{(0)}\rangle}, \tag{3}$$

which is simply the Hellman-Feynman theorem [8]. Similarly, higher-order perturbation theory can be recast as exact expressions for higher-order derivatives of the eigenvalue or the eigenfields.

Such an electromagnetic perturbation theory, and equivalent formulations (sometimes derived via the variational principle instead of the explicit eigenproblem), has seen widespread use [1–3,5], e.g., to determine the effect of material losses (small imaginary $\Delta\varepsilon$) or nonlinearities ($\Delta\varepsilon$

$\sim|\mathbf{E}|^2$). Often, for uniform waveguide fields with (z,t) dependence $e^{i(\beta z-\omega t)}$, the eigenproblem is cast instead in terms of the wave number β [5,6], in which case the shift can be either derived independently by similar methods or be inferred via the group velocity, $\beta^{(1)}=-\omega^{(1)}/(d\omega/d\beta)$. For convenience, we will focus mainly on the differential form of Eq. (3).

A. The problem of shifting boundaries

Consider the numerator of Eq. (3) in the case of an interface between two dielectrics, ε_1 and ε_2, that shifts a distance $h(\alpha,u,v)$ towards ε_2, where (u,v) parametrize the interfacial surface, as depicted in Fig. 1. Since ε is a step function, its derivative is a Dirac δ function that produces a surface integral over the interface:

$$\left\langle E^{(0)}\left|\frac{d\varepsilon}{d\alpha}\right|E^{(0)}\right\rangle=\int dA\frac{dh}{d\alpha}(\varepsilon_1-\varepsilon_2)|\mathbf{E}^{(0)}|^2. \tag{4}$$

This integral, however, is manifestly undefined, since the normal component $E_\perp$ to the interface is discontinuous at the boundary (only $D_\perp=\varepsilon E_\perp$ and the parallel components $\mathbf{E}_\parallel$ are continuous) [9]. Alternatively, naively employing the finite-perturbation form of Eq. (2) corresponds to simply picking one side of the interface on which to evaluate $E_\perp$, which has been shown to yield incorrect results [5,6]; the error worsens as the dielectric contrast (and thus the magnitude of the field discontinuity) increases, but it has been a popular method for low-contrast systems [12]. (Of course, for TM fields in two dimensions (2D) that are everywhere parallel to the boundaries [7], there is no problem.) Why has perturbation theory failed? The source of the error was the assumption that the lowest-order correction $|E^{(1)}\rangle$ is of first order in $\Delta\alpha$—here, because of the discontinuous boundary conditions, there are points where the correction $E_\perp^{(1)}$ is finite even for infinitesimal $\Delta\alpha$, foiling the order-by-order solution of perturbation theory.

It might seem that one could simply recast the eigenproblem in terms of the magnetic field [7], which is everywhere continuous, but a similar difficulty arises—not only the field correction, but *also* the eigen-operators applied to this correction must be of first order in $\Delta\alpha$, and $\mathbf{\nabla}\times\mathbf{H}$ is discontinuous. In fact, comparing with the $\mathbf{H}$ eigenproblem is another way to see that there is a problem in perturbation theory for Maxwell's equations with large $\Delta\varepsilon$. Applying the same procedure as for Eq. (2) to the $\mathbf{H}$ eigenequation, $\mathbf{\nabla}\times 1/\varepsilon\mathbf{\nabla}\times\mathbf{H}=(\omega/c)^2\mathbf{H}$, and *then* rewriting in terms of $\mathbf{E}$, one obtains a *different* result:

$$\omega^{(1)}=\frac{\omega^{(0)}}{2}\frac{\left\langle E^{(0)}\left|\varepsilon^2\Delta\left(\frac{1}{\varepsilon}\right)\right|E^{(0)}\right\rangle}{\langle E^{(0)}|\varepsilon|E^{(0)}\rangle}, \tag{5}$$

which is only equivalent to Eq. (2) if $\Delta\varepsilon$ is the small perturbation parameter [19]. Both formulations are ill defined for shifting boundaries [with Eq. (5), the problem is the discon-
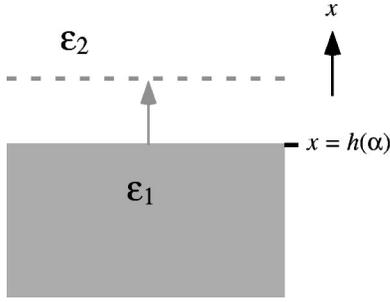
FIG. 2. Schematic of a small, locally flat region of the dielectric interface from ε_1 to ε_2 at $x=h(\alpha)$, where x designates the direction perpendicular to the surface. We consider a perturbation consisting of a small shift of the interface in the direction of ε_2.

tinuity of $\mathbf{D}_\parallel$] and their inconsistency has further unfortunate implications for higher-order perturbation theory, as we discuss in Sec. IV.

One way of solving this problem in certain cases is to express the perturbation not as a $\Delta\varepsilon$, but as a transformation of the coordinate system that moves the boundaries—in this way, the field boundary conditions can be preserved, and the perturbation is expressed via a distorted $\nabla\times$ operation. Such a perturbation theory was developed and successfully applied to the problem of uniformly scaled or stretched waveguides: unlike the conventional method, it yields correct results even in high-contrast systems (e.g., for fiber birefringence) [5,6]. As discussed in Sec. IV, the coordinate-transform method may have additional advantages for computing higher-order corrections. Coordinate transformations, however, are cumbersome to apply for arbitrary interface perturbations, and also result in integrals that are not conveniently localized to the perturbed surface. We circumvent both of these shortcomings by instead deriving a perturbation theory from a *limit* of systems with *smoothed* boundaries.

B. A solution for shifting boundaries

If, instead of a discontinuous transition from ε_1 to ε_2, the dielectric function changes smoothly, then *all* field components are continuous, $\Delta\varepsilon$ is small for a small boundary shift, and we can apply Eq. (3) without difficulty. The answer for the discontinous system should then be the limit as the transition becomes sharper and sharper—this limit must be unique, so *it does not matter how we do the smoothing so long as the limit is well defined*. In order to consider smoothing explicitly, we focus on a small area dA where the interface is locally flat (deferring until later the question of kinks or corners in the boundary), and define x as the coordinate perpendicular to the boundary at $x=h(\alpha)$, depicted in Fig. 2. The local dielectric function is then

$$\varepsilon(x)=\varepsilon_1+(\varepsilon_2-\varepsilon_1)\Theta(x-h), \tag{6}$$

where $\Theta(x)$ is the unit step function at $x=0$.

To start with, let us consider a simple isotropic smoothing, replacing ε with $\bar{\varepsilon}$ given by

$$\bar{\varepsilon}(x)\equiv\int g_s(x-x')\varepsilon(x')dx', \tag{7}$$

where $g_s(x)$ is some smoothing function: a localized function (distribution) around $x=0$, of unit integral, that goes to a Dirac delta function $\delta(x)$ in the limit as $s\rightarrow0$. Thus, $d\bar{\varepsilon}/dh=(\varepsilon_1-\varepsilon_2)g_s(x-h)$, and the contribution to $\langle E^{(0)}|d\bar{\varepsilon}/d\alpha|E^{(0)}\rangle$ from dA in this smoothed system is $(\varepsilon_1-\varepsilon_2)dh/d\alpha dA\int|\mathbf{E}^{(0)}|^2 g_s(x-h)dx$. If we now take $s\rightarrow0$, however, we merely recover the original problem: we have the integral of a step function ($E_\perp^2$) times a δ function, and the limit is undefined.

If we can stumble across *any* smoothing method that circumvents this problem, yielding a well-defined limit, then the uniqueness theorem for Maxwell's equations will mean that we are done—there is no need to otherwise prove that a given smoothing is "correct" (and such well-defined smoothings are certainly not unique, if any exist). At this point, we take a hint from effective medium theory, and realize that the most appropriate boundary smoothing in electromagnetism is *anisotropic*—different field components should "see" different average dielectric constants [14–16]. Specifically, there is an effective *tensor*

$$\varepsilon_s(x)\equiv\begin{pmatrix} \widetilde{\varepsilon}(x) & & \\ & \bar{\varepsilon}(x) & \\ & & \bar{\varepsilon}(x) \end{pmatrix}, \tag{8}$$

so that $\mathbf{E}_\parallel$ ($\mathbf{E}_{yz}$) sees $\bar{\varepsilon}$ from Eq. (7), while $E_\perp$ (E_x) sees instead a *harmonic* mean $\widetilde{\varepsilon}$,

$$\widetilde{\varepsilon}(x)^{-1}\equiv\int g_s(x-x')\varepsilon(x')^{-1}dx'. \tag{9}$$

(Precisely such an anisotropic smoothing has been employed to greatly speed convergence, compared to unsmoothed or isotropically smoothed boundaries, in numerical simulations with finite spatial resolution [15,16].) From Eq. (9), one finds that $d\widetilde{\varepsilon}/dh=-\widetilde{\varepsilon}(x)^2(\varepsilon_1^{-1}-\varepsilon_2^{-1})g_s(x-h)$, and thus the contribution to $\langle E^{(0)}|d\varepsilon_s/d\alpha|E^{(0)}\rangle$ from dA is

$$dA\frac{dh}{d\alpha}\int dx[\Delta\varepsilon_{12}|\mathbf{E}_\parallel^{(0)}|^2-\Delta(\varepsilon_{12}^{-1})|\widetilde{\varepsilon}E_\perp^{(0)}|^2]g_s(x-h), \tag{10}$$

where $\Delta\varepsilon_{12}\equiv\varepsilon_1-\varepsilon_2$ and $\Delta(\varepsilon_{12}^{-1})\equiv\varepsilon_1^{-1}-\varepsilon_2^{-1}$. Note that $D_\perp=\widetilde{\varepsilon}E_\perp$ is continuous, so when we take the $s\rightarrow0$ limit and $g_s(x-h)$ becomes $\delta(x-h)$, the result is well defined, giving

$$\left\langle E^{(0)}\left|\frac{d\varepsilon}{d\alpha}\right|E^{(0)}\right\rangle=\int dA\frac{dh}{d\alpha}[\Delta\varepsilon_{12}|\mathbf{E}_\parallel^{(0)}|^2-\Delta(\varepsilon_{12}^{-1})|D_\perp^{(0)}|^2]. \tag{11}$$

This expression, combined with Eq. (3), yields a correct first-order perturbation theory for arbitrary boundary deformations and arbitrary index contrasts. Reassuringly, if we
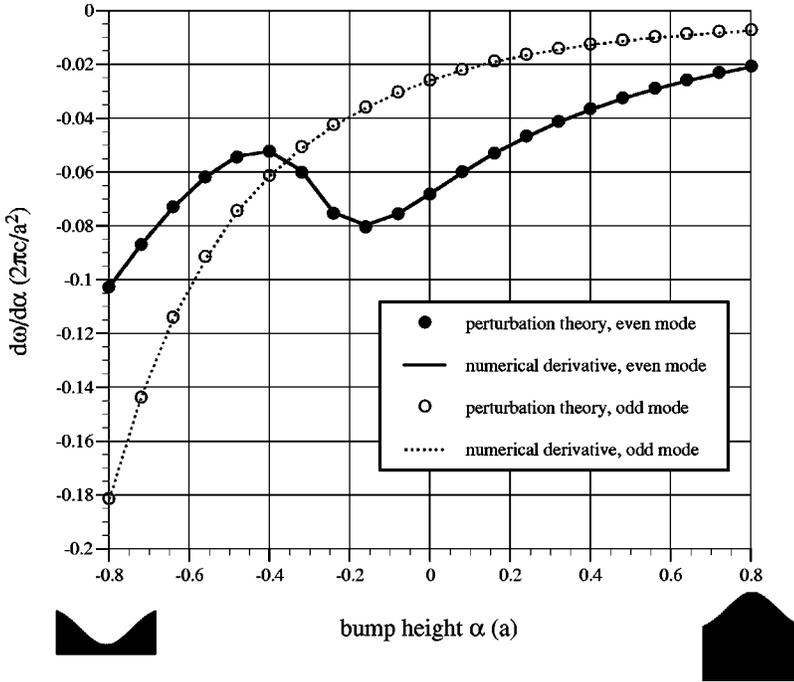
FIG. 3. Comparison of perturbation theory (symbols) and numerical differentiation (lines) for $d\omega/d\alpha$ as a function of α, applied to a Gaussian "bump" of height α on top of a $2a \times a$ rectangular dielectric waveguide (illustrated at the two extremes of the α axis). The comparison is shown for the lowest-order modes of even (filled symbols) and odd (hollow symbols) parity with respect to the mirror symmetry plane of the waveguide.

apply the same limit process to the alternate first-order perturbation theory of Eq. (5), we get the same result. (We also check it numerically below.)

There are two points that we glossed over in the derivation above: the effect of kinks in the surface and of changes in the interface orientation, both of which contributions turn out to be of measure zero. A kink or corner can be thought of as the limit of a tighter and tighter bend in the interface—but in this limit, the area of the kink region goes to zero and the field remains finite, so its contribution to Eq. (10) vanishes. Kinks yield discontinuous changes in the interface orientation, whereas any continuous change can be expressed as a rotation $d\theta$ of the surface in addition to the shift dh, and enters the theory as a rotation matrix transforming the dielectric tensor of Eq. (8). This results in a term proportional to $(\bar{\varepsilon} - \tilde{\varepsilon})d\theta/d\alpha$, which integrates to zero in the $s \to 0$ limit (it is everywhere finite and is zero away from the interface).

The same method can be used to determine the coupling coefficients between modes, e.g., for time-dependent perturbation theory (or z dependent, in waveguides), also known as coupled-mode theory [1,11], as well as for higher-order perturbation theory. Such coupling coefficients, derived in the usual way [8], stumble over the same problem with discontinuities as in first-order perturbation theory. The correct coupling coefficient between the unperturbed modes $|E\rangle$ and $|E'\rangle$ for a shift $h(\alpha, u, v)$ in the interface, using the notation from above, involves [20] the surface integral

$$\left\langle E \left| \frac{d\varepsilon}{d\alpha} \right| E' \right\rangle = \int dA \frac{dh}{d\alpha} [\Delta \varepsilon_{12} (\mathbf{E}_\parallel^* \cdot \mathbf{E}_\parallel')$$

$$- \Delta(\varepsilon_{12}^{-1})(D_\perp^* \cdot D_\perp')], \quad (12)$$

which is again defined purely in terms of the field components that are continuous across the interface. In the case of a waveguide with slowly z-varying cross section, for

coupled-mode equations expanded in the instantaneous eigenmodes at each z, one uses $\alpha = z$ in Eq. (12). If one expands in the eigenmodes of a fixed cross section rather than those of the instantaneous cross section, dh/dz is replaced (to first order) by Δh in Eq. (12). This equation can also be used for higher-order perturbation theory (higher-order derivatives in α), given that the additional considerations discussed in Sec. IV are addressed. (The perturbative expansion for finite $\Delta \alpha$ then follows from the Taylor series.)

The generalization to multiple shifted interfaces, and/or to a $\Delta \varepsilon$ that varies over an interface, is obvious. In the case of shifting boundaries for magnetic materials (nonuniform $\mu \neq 1$), the perturbation theory's difficulty and its solution are precisely analogous, with μ, $\mathbf{H}$, and $\mathbf{B}$ substituting for ε, $\mathbf{E}$, and $\mathbf{D}$, respectively.

III. A TEST CASE

In order to numerically verify that our perturbation theory for shifting boundaries yields the correct result, we consider an arbitrary test case that exhibits a nonuniform shift of a curved boundary, changing boundary orientations, kinks, and large index contrast, and also lacks any special relationship between the field direction and the surface normal. In particular, we start with a uniform rectangular waveguide (along z) of index $n = 3$ and dimensions $2a \times a$ (where a is an arbitrary length scale), surrounded by air ($n = 1$). Then, to one of the $2a$ edges (lying along the $x = 0$ plane), we add a Gaussian "bump" of height $h(y) = \alpha e^{-y^2/2w^2}$, where the bump width is $w = a/2$, as depicted in the insets of Fig. 3. Note that $\alpha < 0$ corresponds to an indentation, and the bump is abruptly terminated at $y = \pm a$, yielding a shifting surface kink. Given this "bumped" waveguide, we consider the
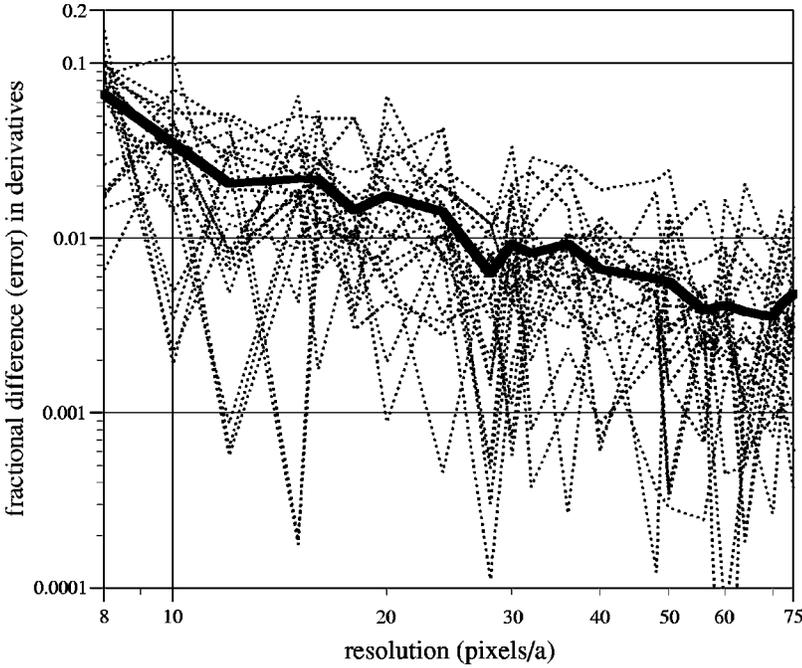
FIG. 4. Absolute fractional difference in $d\omega/d\alpha$ between perturbation theory and numerical differentiation as a function of computational resolution (in pixels/a) for the even mode of Fig. 3. Each dotted line is the difference at a given α for the lowest-order even mode (i.e., one line per filled circle of Fig. 3), and the thick black line is the mean value. The large oscillations stem from both discretization noise and the fact that the sign of the difference is not fixed (it sometimes passes almost through zero), but the overall trend is the expected power-law decline.

lowest-order guided modes, with z dependence $e^{i\beta z}$, at a wave number $\beta = \pi/a$, and compute $d\omega/d\alpha$ at various values of α. Since the waveguide is symmetric around $y=0$, we examine the lowest-order modes that are even or odd with respect to this mirror plane (corresponding to modes mostly polarized along the x or y direction, respectively).

The fully vectorial eigenfields and frequencies of this structure are computed in a $5a \times 5a$ supercell with periodic boundaries by preconditioned conjugate-gradient minimization of the block Rayleigh quotient in a plane-wave basis, using a freely available software package [16]. (The size of the supercell is actually irrelevant for our purposes; a small supercell would merely produce the modes of coupled periodic waveguides, and would be just as stringent a test of the perturbation theory.) Given these fields for a particular α, we compute the line integral of Eq. (11) over the bump surface (bilinearly interpolating the fields from the discrete computational grid) and then employ Eq. (3) for $d\omega/d\alpha$. For comparison, we evaluate the explicit numerical derivative $d\omega/d\alpha$ by computing the eigenfrequency at various nearby values of α and applying Ridder's method of polynomial extrapolation (which maximizes the resulting precision) [17]. The two answers are compared, as a function of α, in Fig. 3 (at a computational resolution of 75 pixels/a), and demonstrate that the perturbation method yields accurate results.

Of course, there are errors in both the perturbation theory and the numerical derivative due to the finite computational resolution, so the two results do not match precisely. Such differences, however, should decline in roughly a power-law relationship with the resolution [21], and we display this decline in Fig. 4 for each value of α evaluated above. The absolute fractional difference between the two derivatives oscillates widely around this decline, of course, due to both the discretization noise (the boundary shape does not change in a continuous fashion) and the fact that the difference is not of a fixed sign (so it sometimes passes almost through zero,

yielding sharp accidental dips). Averaging over α to smooth these oscillations results in a clearer picture of the decaying error (shown as a thick black line in Fig. 4).

It is also instructive to compare our perturbation theory of Eq. (11) with the incorrect Eq. (4) that ignores the boundary discontinuity—in the latter case, we simply evaluate $|\mathbf{E}|^2$ on the low-index side of the boundary [corresponding to the naive application of the standard Eq. (2) for positive $\Delta\alpha$]. Then, at a resolution of 75 pixels/a, we compute the absolute fractional difference in $d\omega/d\alpha$ with the numerical derivative, as a function of α, and plot the results in Fig. 5. For the incorrect perturbation theory, as in Ref. [6], systematic errors are revealed: significant differences (exceeding 100%) that do not decrease with increasing resolution. The errors here are considerably worse for the even mode than for the odd mode because the latter is mostly polarized parallel to the interface (minimizing the $E_\perp$ discontinuity)—although the errors for the odd mode increase as the bump becomes larger and thereby less parallel to $\mathbf{E}$. The error in the incorrect theory is proportional to the surface integral $[\Delta\varepsilon_{12}/\varepsilon_x^2 + \Delta(\varepsilon_{12}^{-1})]\int dA |D_\perp|^2$, where ε_x here is the dielectric on the low-index side of the boundary. If we instead evaluate the incorrect theory on the high-index side, only this ε_x is affected, and so the error changes merely by a constant factor (of 1/9, in this case).

IV. REMAINING QUESTIONS

Perturbation theory, in principle, provides not only a first-order correction to an eigenvalue, but also a systematic way to find the series of higher-order corrections to both the eigenvalue and eigenfield. Similarly, coupled-mode theory need not be used in the small-perturbation limit—it can be thought of as an exact set of coupled linear differential equa-
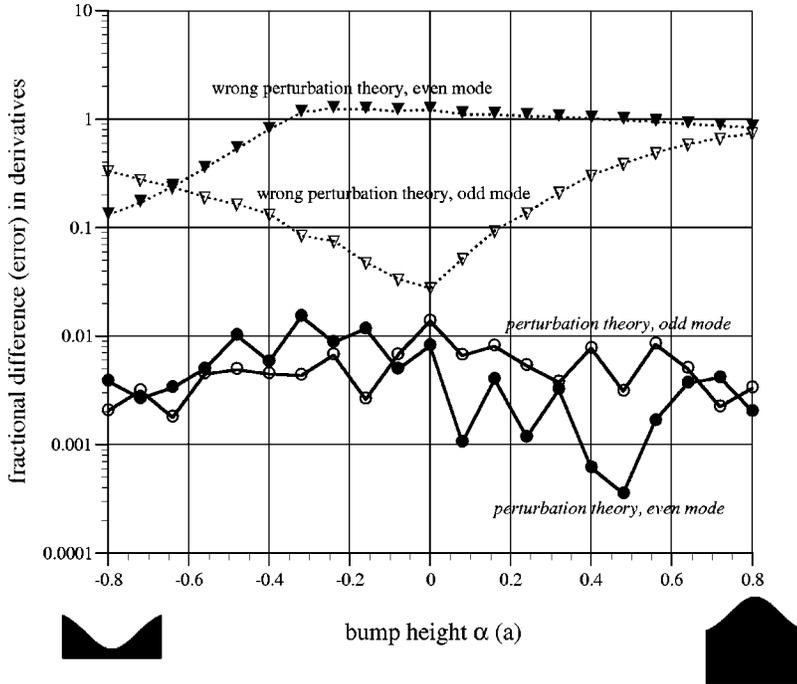
FIG. 5. Absolute fractional difference in $d\omega/d\alpha$ between perturbation theory and numerical differentiation as a function of bump height α, for a resolution of 75 pixels/a. We plot both our corrected perturbation theory (solid lines, circles) and the ill-defined perturbation theory (dashed lines, triangles) of Eq. (4) in which the discontinuous field is simply evaluated on the low-index side of the interface. Results for the lowest-order even mode are shown as filled symbols and for the odd mode as hollow symbols.

tions, expressing the mode in terms of the unperturbed eigenmodes with varying coefficients. For both of these techniques, however, the discrepancy between the perturbative expressions of Eqs. (2) and (5) from the **E** and **H** eigenproblems is symptomatic of an underlying problem—the two formulations are only equivalent to first order, and they cannot *both* be correct to higher orders. How could this be? Both coupled-mode theory and the standard perturbation-theory method [8] for first-order eigenfield or higher-order eigenvalue corrections rely on one key assumption that is false for Maxwell's equations: they assume that the basis of the unperturbed eigenstates is *complete*.

In fact, when the **E** eigenproblem is solved, one normally imposes the additional constraint of zero free charge, or $\nabla \cdot \varepsilon \mathbf{E} = 0$ (whence the discontinuous boundary condition on $E_{\perp}$). The perturbed field $\mathbf{E}'$, however, satisfies $\nabla \cdot (\varepsilon + \Delta \varepsilon) \mathbf{E}' = 0$, so in general, $\nabla \cdot \varepsilon \mathbf{E}' \neq 0$ and $\mathbf{E}'$ cannot be expanded in the basis of the unperturbed fields. Relaxing this divergence-free constraint would mean the inclusion of infinitely many static-field solutions at $\omega = 0$, as seen by taking the divergence of both sides of Eq. (1), which would be (at the least) computationally inconvenient. In contrast, the **H** eigenproblem involves the constraint $\nabla \cdot \mu \mathbf{H} = 0$, which is not altered by $\Delta \varepsilon$, so it appears that the **H** formulation is correct to higher orders. Alternatively, one could formulate the eigenproblem **D** or **B** (handling $\Delta \mu$ as well as $\Delta \varepsilon$). For the β eigenproblem (in waveguides), these completeness issues do not seem to arise—there, one works at a fixed $\omega \neq 0$ frequency, so nondivergenceless fields are always excluded. Even when a complete basis is employed, however, one is likely to encounter convergence difficulties due to Gibb's phenomena [18] that will arise from the shifted field discontinuities. Thus, it may be desirable to find some corrections to the unperturbed modes in order to make a complete, fast-converging basis for the perturbed system,

preferably without solving any new differential equations; perhaps such corrections can themselves be a perturbative expansion.

Fortunately, in most cases where perturbation theory is of interest, first-order accuracy is sufficient, and there a poor basis is not a problem. There are situations in which this is not enough, however; for example, when the first-order correction is zero by some symmetry, or if one wishes to explore intentional mode conversion in a strong-coupling limit. If the perturbation is due to shifting boundaries, one possible solution is the coordinate-transformation method alluded to earlier [5,6]—since it can preserve the boundary conditions, there are indications that it is able to efficiently compute, e.g., second-order eigenvalues.

V. CONCLUDING REMARKS

In this paper, we explained and solved a difficulty that arises on applying perturbation theory to Maxwell's equations for small shifts in dielectric interfaces, especially in three-dimensional, high-index-contrast systems. The resulting expression, Eqs. (11) and (12), is a simple surface integral over the perturbed interface(s), and we have also numerically illustrated its correctness for an arbitrary curved boundary distortion. Such a perturbative method is useful for a wide variety of applications, from fiber birefringence [6,12], to waveguide tapering and adiabatic coupling, to surface roughness, to tuning of cavity modes and photonic band gaps by geometric alterations (e.g., strain induced). Open questions persist, however, in the computation of higher-order perturbative corrections. We hope to further expand the reach of perturbative techniques in future work.

[1] D. Marcuse, *Theory of Dielectric Optical Waveguides*, 2nd ed. (Academic Press, San Diego, 1991).

[2] A. W. Snyder and J. D. Love, *Optical Waveguide Theory* (Chapman and Hall, London, 1983).

[3] C. Vassallo, *Optical Waveguide Concepts* (Elsevier, Amsterdam, 1991).

[4] N. R. Hill, Phys. Rev. B **24**, 7112 (1981).

[5] S. G. Johnson, M. Ibanescu, M. Skorobogatiy, O. Weisberg, T. D. Engeness, M. Soljačić, S. A. Jacobs, J. D. Joannopoulos, and Y. Fink, Opt. Express **9**, 748 (2001).

[6] M. Skorobogatiy, M. Ibanescu, S. G. Johnson, O. Weisberg, T. D. Engeness, M. Soljačić, S. A. Jacobs, and T. Fink, J. Opt. Soc. Am. B (to be published).

[7] J. D. Joannopoulos, R. D. Meade, and J. N. Winn, *Photonic Crystals: Molding the Flow of Light* (Princeton University, Princeton, NJ, 1995).

[8] C. Cohen-Tannoudji, B. Din, and F. Laloë, *Quantum Mechanics* (Hermann, Paris, 1977), Vol. 1, Chap. 2; Vol. 2, Chaps. 11 and 13.

[9] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1998).

[10] M. Lohmeyer, N. Bahlmann, and P. Hertel, Opt. Commun. **163**, 86 (1999).

[11] B. Z. Katsenelenbaum, L. Mercader del Río, M. Pereyaslavets, M. Sorolla Ayza, and M. Thumm, *Theory of Nonuniform Waveguides: The Cross-Section Method* (Inst. of Electrical Engineers, London, 1998).

[12] D. Q. Chowdhury and D. A. Nolan, Opt. Lett. **20**, 1973 (1995).

[13] P. A. M. Dirac, *Principles of Quantum Mechanics* (Clarendon, Oxford, 1982).

[14] D. E. Aspnes, Am. J. Phys. **50**, 704 (1982).

[15] R. D. Meade, A. M. Rappe, K. D. Brommer, J. D. Joannopoulos, and O. L. Alerhand, Phys. Rev. B **48**, 8434 (1993); S. G. Johnson, *ibid.* **55**, 15 942(E) (1997).

[16] S. G. Johnson and J. D. Joannopoulos, Opt. Express **8**, 173 (2001).

[17] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, 1992).

[18] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, 5th ed. (Harcourt, San Diego, 2001).

[19] As yet another alternative, the waveguide-mode eigenproblem in terms of β yields a first-order correction that is a mixture between Eqs. (2) and (5) for different components of $\mathbf{E}$ [1,5,6].

[20] Depending upon whether one is doing time- or z-(in)dependent perturbation/coupled-mode theory, there are additional (well known and easily derived) normalization factors multiplying the coupling integral [8] (see, e.g., Ref. [5]).

[21] The exact power law would be at best $\sim \Delta x^2$, since this is the convergence rate of the eigenfrequencies [16], but is actually closer to Δx because of errors inherent in the numerical differentiation and the interpolated line integrals on a discrete grid.