

Analysis of general geometric scaling perturbations in a transmitting waveguide: fundamental connection between polarization-mode dispersion and group-velocity dispersion

Maksim Skorobogatiy, Mihai Ibanescu, Steven G. Johnson, Ori Weisberg, Torkel D. Engeness, Marin Soljačić, Steven A. Jacobs, and Yoel Fink

OmniGuide Communications, One Kendall Square, Building 100, Cambridge, Massachusetts 02139

Received November 28, 2001; revised manuscript received May 17, 2002

We develop a novel perturbation theory formulation to evaluate polarization-mode dispersion (PMD) for a general class of scaling perturbations of a waveguide profile based on generalized Hermitian Hamiltonian formulation of Maxwell's equations. Such perturbations include ellipticity and uniform scaling of a fiber cross section, as well as changes in the horizontal or vertical sizes of a planar waveguide. Our theory is valid even for discontinuous high-index contrast variations of the refractive index across a waveguide cross section. We establish that, if at some frequencies a particular mode behaves like pure TE or TM polarized mode (polarization is judged by the relative amounts of the electric and magnetic longitudinal energies in the waveguide cross section), then at such frequencies for fibers under elliptical deformation its PMD as defined by an intermode dispersion parameter τ becomes proportional to group-velocity dispersion D such that $\tau = \lambda \delta |D|$, where δ is a measure of the fiber ellipticity and λ is a wavelength of operation. As an example, we investigate a relation between PMD and group-velocity dispersion of a multiple-core step-index fiber as a function of the core-clad index contrast. We establish that in this case the positions of the maximum PMD and maximum absolute value of group-velocity dispersion are strongly correlated, with the ratio of PMD to group-velocity dispersion being proportional to the core-clad dielectric contrast. © 2002 Optical Society of America

OCIS codes: 060.2310, 060.2400, 060.2280, 060.0060.

1. INTRODUCTION

Recently, much effort has been placed on understanding the detrimental effects of polarization-mode dispersion (PMD) because of the fiber imperfections in high-bit-rate transmission systems (see Ref. 1 and references therein). The task of quantifying PMD involves two problems. The first problem is to understand signal propagation along a fiber with a stochastically varying birefringence. The second problem is to quantify the local birefringence by assuming a particular type of perturbation and employing some method to find the new perturbed modes. In this paper we analyze the PMD induced by geometric imperfections that fall into a general class of scaling perturbations, including ellipticity or a uniform scaling of a fiber cross section, as well as changes in the horizontal or vertical size of a planar waveguide.

In general, PMD arises when an imperfection splits an originally degenerate mode (such as the doubly degenerate linearly polarized fundamental mode of a silica fiber or a square planar buried waveguide) into two closely spaced modes, each carrying substantial signal power and traveling at slightly different group velocities. A standard way to quantify local PMD is to define an intermode dispersion parameter τ , which equals the difference between the inverse of the group velocities of the newly split modes.

There has been a significant amount of research on the estimation of such quantities as the local birefringence induced by perturbations in the fiber profile. Most of that

study is currently a part of several standard textbooks in the field.²⁻⁵ However, most of these treatments were geared toward understanding low-index contrast, weakly guiding systems such as silica optical fiber, and are not directly applicable to high contrast systems such as Bragg fibers, photonic crystal fibers, and integrated-optics waveguides that are emerging as an integral part of state-of-the-art optical communication systems. Although several approaches have been suggested for the perturbative treatment of arbitrary fiber profiles,^{3,6-9} they were either developed for specific index profile geometries or their validity deteriorated substantially with an increase in the index contrast.

Here we derive a generalized Hermitian Hamiltonian formulation of Maxwell's equations in waveguides, as well as develop a perturbation theory for the general class of scaling perturbations that for fibers include ellipticity and uniform scaling of an arbitrary index profile, whereas for planar waveguides it treats the changes in their horizontal or vertical sizes. Because of the Hermitian nature of the formulation most of the results from the well-developed perturbation theory of quantum mechanical systems can be applied directly to light propagation in the waveguides. Such a formulation provides an intuitive way to understand PMD and birefringence in elliptically perturbed fibers. In particular, we show that the regions in which the mode is predominantly TE or TM polarized, as judged by the magnitudes of the longitudinal electric and magnetic energies in the fiber cross section, corre-

spond to the regions of high PMD. Moreover, within the same limit, we show that the PMD of such a polarized mode is proportional to the group-velocity dispersion, which might become of concern in the case of dispersion-compensating fibers. Finally, we demonstrate these findings for the example system of a dual-core, step-index fiber. We find that the point of maximum absolute value of PMD correlates strongly with the point of maximum absolute value of dispersion in the 1–200% range of core–clad dielectric contrasts. Moreover, for small core–clad dielectric contrasts the ratio of PMD to mode dispersion increases proportional to such a contrast. Finally, the results of our perturbative formulation are compared to numerical calculations by use of a general plane-wave expansion method¹⁰ and excellent correspondence between theoretical and numerical results is found. Predictions of a standard perturbation theory are shown to match our results in a region of small core–clad dielectric contrast whereas for large core–clad dielectric contrast substantial deviations from the correct results are found.

2. GENERALIZED HERMITIAN HAMILTONIAN FORMULATION OF THE PROBLEM OF FINDING EIGENMODES OF A WAVEGUIDE

A. Hamiltonian Formulation in β

First we derive a Hamiltonian formulation, also present in Ref. 11, for the guided eigenfields of a generic waveguide that exhibits translational symmetry in the longitudinal \hat{z} direction. We start with a well-known expres-

sion of Maxwell's equations in terms of the transverse and longitudinal fields.⁴ Assuming that the form of the fields is

$$\begin{pmatrix} \mathbf{E}(x, y, z, t) \\ \mathbf{H}(x, y, z, t) \end{pmatrix} = \begin{pmatrix} \mathbf{E}(x, y) \\ \mathbf{H}(x, y) \end{pmatrix} \exp(i\beta z - i\omega t) \quad (1)$$

and introducing transverse and longitudinal components of the fields,

$$\mathbf{E} = \mathbf{E}_t + \mathbf{E}_z, \quad \mathbf{E}_z = \hat{z}E_z, \quad \mathbf{E}_t = (\hat{z} \times \mathbf{E}) \times \hat{z}, \quad (2)$$

Maxwell's equations,

$$\begin{aligned} \nabla \times \mathbf{E} &= i \frac{\omega}{c} \mu \mathbf{H}, & \nabla \times \mathbf{H} &= -i \frac{\omega}{c} \epsilon \mathbf{E}, \\ \nabla \cdot \mu \mathbf{H} &= 0, & \nabla \cdot \epsilon \mathbf{E} &= 0, \end{aligned} \quad (3)$$

take the form

$$\frac{\partial \mathbf{E}_t}{\partial z} + i \frac{\omega}{c} \mu \hat{z} \times \mathbf{H}_t = \nabla_t \mathbf{E}_z, \quad \frac{\partial \mathbf{H}_t}{\partial z} - i \frac{\omega}{c} \epsilon \hat{z} \times \mathbf{E}_t = \nabla_t H_z, \quad (4)$$

$$\hat{z}(\nabla_t \times \mathbf{E}_t) = i \frac{\omega}{c} \mu H_z, \quad \hat{z}(\nabla_t \times \mathbf{H}_t) = -i \frac{\omega}{c} \epsilon E_z, \quad (5)$$

$$\nabla_t \cdot \epsilon \mathbf{E}_t = -\frac{\partial \epsilon E_z}{\partial z}, \quad \nabla_t \cdot \mu \mathbf{H}_t = -\frac{\partial \mu H_z}{\partial z}. \quad (6)$$

Eliminating the E_z and H_z components from Eqs. (4) by using Eqs. (5) and following the derivation in Refs. 11 and 12 we arrive at

$$-i \frac{\partial}{\partial z} \begin{pmatrix} \mathbf{E}_t \\ \mathbf{H}_t \end{pmatrix} = \begin{pmatrix} 0 & -\frac{\omega}{c} \mu (\hat{z} \times) + \frac{c}{\omega} \nabla_t \left[\frac{1}{\epsilon} \hat{z} \cdot (\nabla_t \times) \right] \\ \frac{\omega}{c} \epsilon (\hat{z} \times) - \frac{c}{\omega} \nabla_t \left[\frac{1}{\mu} \hat{z} \cdot (\nabla_t \times) \right] & 0 \end{pmatrix} \begin{pmatrix} \mathbf{E}_t \\ \mathbf{H}_t \end{pmatrix}. \quad (7)$$

Substituting Eq. (1) into Eq. (7), and after some rearrangement employing the identity $\hat{z} \times (\nabla_t \cdot) = -\nabla_t \times (\hat{z} \cdot)$, we arrive at the following eigenproblem:

$$\beta \begin{pmatrix} 0 & -\hat{z} \times \\ \hat{z} \times & 0 \end{pmatrix} \begin{pmatrix} \mathbf{E}_t \\ \mathbf{H}_t \end{pmatrix} = \begin{pmatrix} \frac{\omega}{c} \epsilon - \frac{c}{\omega} \nabla_t \times \left\{ \hat{z} \left[\frac{1}{\mu} \hat{z} \cdot (\nabla_t \times) \right] \right\} & 0 \\ 0 & \frac{\omega}{c} \mu - \frac{c}{\omega} \nabla_t \times \left\{ \hat{z} \left[\frac{1}{\epsilon} \hat{z} \cdot (\nabla_t \times) \right] \right\} \end{pmatrix} \begin{pmatrix} \mathbf{E}_t \\ \mathbf{H}_t \end{pmatrix}. \quad (8)$$

In this form the operators at left and at right are Hermitian, thus defining a generalized Hermitian eigenproblem¹³ and allowing for all the convenient properties that pertain to such a form, including real eigenvalues β for guided modes as well as orthogonality of the modes that correspond to the different β 's (for more discussion see Refs. 11 and 12). Defining operators \hat{A} and \hat{B} as

$$\begin{aligned} \hat{B} &= \begin{pmatrix} 0 & -\hat{z} \times \\ \hat{z} \times & 0 \end{pmatrix}, \\ \hat{A} &= \begin{pmatrix} \frac{\omega}{c} \epsilon - \frac{c}{\omega} \nabla_t \times \left\{ \hat{z} \left[\frac{1}{\mu} \hat{z} \cdot (\nabla_t \times) \right] \right\} & 0 \\ 0 & \frac{\omega}{c} \mu - \frac{c}{\omega} \nabla_t \times \left\{ \hat{z} \left[\frac{1}{\epsilon} \hat{z} \cdot (\nabla_t \times) \right] \right\} \end{pmatrix}, \end{aligned} \quad (9)$$

and introducing Dirac notation

$$|\psi\rangle = \begin{pmatrix} \mathbf{E}_t \\ \mathbf{H}_t \end{pmatrix},$$

we introduce more compact notation for the generalized eigenproblem of Eq. (8):

$$\hat{A}|\psi_\beta\rangle = \beta\hat{B}|\psi_\beta\rangle, \quad (10)$$

with an orthogonality condition between guided modes β and β' having the form

$$\langle\psi_\beta|\hat{B}|\psi_{\beta'}\rangle = \frac{\beta}{|\beta|}\delta_{\beta,\beta'}. \quad (11)$$

In the following we analyze perturbations that are uniform along the \hat{z} axis. If a perturbation such as a small uniform rescaling (a uniform increase in the sizes of all transverse dimensions of a structure by a small multiplicative parameter) is introduced into the system, it will modify operator \hat{A} . Denoting the correction to an original operator by $\Delta\hat{A}$, the new eigenvalues $\tilde{\beta}$ (up to the first order) can be expressed with the standard perturbation theory¹⁴:

$$\tilde{\beta} = \beta + \frac{\langle\psi_\beta|\Delta\hat{A}|\psi_\beta\rangle}{\langle\psi_\beta|\hat{B}|\psi_\beta\rangle} \quad (12)$$

In general, at a particular frequency, knowledge of β alone is not enough to characterize an eigenmode of a fiber uniquely. Additional labels are needed, such as angular index m , which indicates the type of angular dependence. In the case of an elliptical perturbation (equivalent to the rescaling of the axes such that one of them increases while the other decreases by some small multiplicative factor), a split occurs in the doubly degenerate eigenmodes characterized by the same β but having opposite angular indices m and $-m$. For the case of a doubly degenerate mode, Eq. (12) is not directly applicable. Instead, the split is obtained from degenerate perturbation theory.¹⁴ Then, new linearly polarized nondegenerate eigenmodes, which we denote $|\psi_\beta^+\rangle$ and $|\psi_\beta^-\rangle$, up to the phase are found to be to the first order

$$|\psi_\beta^\pm\rangle = \frac{1}{\sqrt{2}}(|\psi_{\beta,m}\rangle \pm |\psi_{\beta,-m}\rangle), \quad (13)$$

and the perturbed eigenvalues are

$$\beta^\pm = \beta + \frac{\langle\psi_{\beta,m}|\Delta\hat{A}|\psi_{\beta,m}\rangle}{\langle\psi_{\beta,m}|\hat{B}|\psi_{\beta,m}\rangle} \pm \frac{\langle\psi_{\beta,m}|\Delta\hat{A}|\psi_{\beta,-m}\rangle}{\langle\psi_{\beta,m}|\hat{B}|\psi_{\beta,m}\rangle}. \quad (14)$$

The intermode dispersion parameter τ is defined to be the mismatch of the inverse group velocities of the perturbed modes that can be expressed in terms of the frequency derivative

$$\tau = \frac{1}{v_g^+} - \frac{1}{v_g^-} = \frac{\partial(\beta^+ - \beta^-)}{\partial\omega} = \frac{\partial\Delta\beta_e}{\partial\omega}, \quad (15)$$

where $\Delta\beta_e = (\beta^+ - \beta^-)$. The PMD of a doubly degenerate mode that exhibits splitting caused by a perturbation is defined to be proportional to τ .¹⁵ The desirable condition of zero PMD at a particular frequency ω then implies a zero value of the frequency derivative of the degeneracy split $\Delta\beta_e$, or equivalently $\Delta\beta_e$ must be stationary at such a frequency.

Finally, Eq. (7) permits perturbations that are, in general, nonuniform along the \hat{z} axis. In this case one has to resort to an analog of the time-dependent perturbation theories of quantum mechanics.¹⁴

B. Hamiltonian Formulation in ω

An alternative formulation of a generalized Hermitian Hamiltonian eigenproblem in terms of ω rather than β is also possible. Following the same derivation as in Subsection 2.A, we can also eliminate components E_z and H_z from Eqs. (4) by using Eqs. (6). Following an explicit substitution of Eq. (1) into Eqs. (6) and after some manipulations, we arrive at

$$\frac{\omega}{c} \begin{pmatrix} 0 & -\hat{z} \times \\ \hat{z} \times & 0 \end{pmatrix} \begin{pmatrix} \mathbf{D}_t \\ \mathbf{B}_t \end{pmatrix} = \begin{pmatrix} \frac{\beta}{\epsilon} - \frac{1}{\beta} \frac{1}{\epsilon} \nabla_t \cdot \nabla_t & 0 \\ 0 & \frac{\beta}{\mu} - \frac{1}{\beta} \frac{1}{\mu} \nabla_t \cdot \nabla_t \end{pmatrix} \begin{pmatrix} \mathbf{D}_t \\ \mathbf{B}_t \end{pmatrix}, \quad (16)$$

where $\mathbf{D}_t = \epsilon\mathbf{E}_t$, $\mathbf{B}_t = \mu\mathbf{H}_t$, and operators at left and at right are Hermitian. Unfortunately, the formulation in ω for the transverse fields does not offer perturbative treatment of \hat{z} dependent perturbations as there is no analog of Eq. (7) for Eq. (16). However, \hat{z} independent perturbations can be successfully treated by the same approach as described in Subsection 2.A.

3. GENERIC PERTURBATION OF THE DIELECTRIC PROFILE (WEAK PERTURBATION THEORY)

We now derive an expression for the matrix elements of the perturbation corresponding to a generic change in the dielectric profile of a fiber. In all that follows we assume nonmagnetic materials with $\mu = 1$. Starting with the definitions in Eqs. (9) for operators \hat{A} and \hat{B} and assuming perturbations of the form $\tilde{\epsilon} = \epsilon + \Delta\epsilon$, we arrive at a modified generalized Hamiltonian eigenproblem:

$$(\hat{A}_o + \Delta\hat{A})|\psi_\beta^-\rangle = \tilde{\beta}\hat{B}|\psi_\beta^-\rangle, \quad (17)$$

where

$$\Delta\hat{A} = \begin{pmatrix} \frac{\omega}{c} \Delta\epsilon & 0 \\ 0 & \frac{c}{\omega} \nabla_t \times \left\{ \hat{z} \left[\frac{\Delta\epsilon}{\epsilon + \Delta\epsilon} \hat{z} \cdot (\nabla_t \times) \right] \right\} \end{pmatrix}. \quad (18)$$

Expanding Dirac notation into an integral form we proceed with an evaluation of the matrix elements:

$$\begin{aligned} \langle \psi_\beta | \hat{B} | \psi_{\beta'} \rangle &= \int_S ds \begin{pmatrix} \mathbf{E}_t \\ \mathbf{H}_t \end{pmatrix}_\beta^\dagger \begin{pmatrix} 0 & -\hat{z} \times \\ \hat{z} \times & 0 \end{pmatrix} \begin{pmatrix} \mathbf{E}_t \\ \mathbf{H}_t \end{pmatrix}_{\beta'} \\ &= \int_S ds \hat{z} \cdot (\mathbf{E}_{t,\beta}^* \times \mathbf{H}_{t,\beta'} + \mathbf{E}_{t,\beta'} \times \mathbf{H}_{t,\beta}^*), \end{aligned} \tag{19}$$

where the integration is performed over the waveguide cross section. Using the vector identity

$$\mathbf{a}(\nabla \times \mathbf{b}) = \nabla(\mathbf{b} \times \mathbf{a}) + \mathbf{b}(\nabla \times \mathbf{a}), \tag{20}$$

the divergence theorem, the square integrability of the fields, and Eqs. (4) and (5), after some manipulation involving integration by parts we derived

$$\begin{aligned} \langle \psi_\beta | \hat{A}_o | \psi_{\beta'} \rangle &= \int_S ds \begin{pmatrix} \mathbf{E}_t \\ \mathbf{H}_t \end{pmatrix}_\beta^\dagger \begin{pmatrix} \frac{\omega}{c} \epsilon - \frac{c}{\omega} \nabla_t \times \{ \hat{z} [\hat{z} \cdot (\nabla_t \times)] \} & 0 \\ 0 & \frac{\omega}{c} - \frac{c}{\omega} \nabla_t \times \left\{ \hat{z} \left[\frac{1}{\epsilon} \hat{z} \cdot (\nabla_t \times) \right] \right\} \end{pmatrix} \begin{pmatrix} \mathbf{E}_t \\ \mathbf{H}_t \end{pmatrix}_{\beta'} \\ &= \frac{\omega}{c} \int_S ds \begin{pmatrix} E_z \\ \mathbf{E}_t \\ H_z \\ \mathbf{H}_t \end{pmatrix}_\beta^\dagger \begin{pmatrix} -\epsilon & 0 & 0 & 0 \\ 0 & \epsilon & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} E_z \\ \mathbf{E}_t \\ H_z \\ \mathbf{H}_t \end{pmatrix}_{\beta'}, \end{aligned} \tag{21}$$

where

$$\beta = \frac{\langle \psi_\beta | \hat{A}_o | \psi_\beta \rangle}{\langle \psi_\beta | \hat{B} | \psi_\beta \rangle}, \tag{22}$$

$$\begin{aligned} \langle \psi_\beta | \Delta \hat{A} | \psi_{\beta'} \rangle &= \int_S ds \begin{pmatrix} \mathbf{E}_t \\ \mathbf{H}_t \end{pmatrix}_\beta^\dagger \begin{pmatrix} \frac{\omega}{c} \Delta \epsilon & 0 \\ 0 & \frac{c}{\omega} \nabla_t \times \left\{ \hat{z} \left[\frac{\Delta \epsilon}{\epsilon + \Delta \epsilon} \hat{z} \cdot (\nabla_t \times) \right] \right\} \end{pmatrix} \begin{pmatrix} \mathbf{E}_t \\ \mathbf{H}_t \end{pmatrix}_{\beta'} \\ &= \frac{\omega}{c} \int_S ds \begin{pmatrix} E_z \\ \mathbf{E}_t \\ H_z \\ \mathbf{H}_t \end{pmatrix}_\beta^\dagger \begin{pmatrix} \frac{\epsilon \Delta \epsilon}{\epsilon + \Delta \epsilon} & 0 & 0 & 0 \\ 0 & \Delta \epsilon & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} E_z \\ \mathbf{E}_t \\ H_z \\ \mathbf{H}_t \end{pmatrix}_{\beta'}. \end{aligned} \tag{23}$$

Given the eigenmodes of a structure, one can use Eqs. (19), (21), and (23) to evaluate the new eigenvalues of the eigenmodes given by Eqs. (12) and (14).

As we demonstrate later in this paper, when applied to the problem of elliptical perturbations of a fiber, the theory derived above gives a correct result only in the limit of small ellipticity as well as small index changes over the cross section of a waveguide. We, therefore, refer to this form of the perturbation theory as a weak perturbation theory. Its form and region of applicability is

similar to the vector perturbation theories developed in Refs. 2, 5, 7, and 16. The problem of validity of such a form of perturbation theory in the case of high-index contrast systems stems from its inadequacy in treating the field discontinuities at the perturbed interfaces of substantially dissimilar dielectrics. By literally moving the high-index contrast interfaces the change in the dielectric profile $\Delta\epsilon$ becomes large in absolute value near the high-index contrast interfaces. The eigenmodes of an unperturbed fiber do not constitute an adequate basis to approximate large changes of the fields in the perturbed region and a standard perturbation theory breaks down.

With the development of integrated optics, as well as microstructured and photonic crystal fibers, there is also

a need for a perturbation theory that is valid for arbitrary index contrast. Some perturbation theory formulations for arbitrary contrast are known, but most of them are particular to some analytic geometries with sharp boundaries between layers.^{6,8,9} Although current development of a vectorial perturbation theory^{3,16} might also be applicable to an arbitrary index contrast, there is no detailed discussion of that issue by the authors. In the following, we describe a vectorial perturbation theory based on the generalized Hermitian Hamiltonian formulation of Eqs.

(9) and (18) for a general scaling of a structure. Our theory is valid for an arbitrary index profile as long as the perturbation is small.

4. GENERIC SCALING PERTURBATION OF THE DIELECTRIC PROFILE (STRONG PERTURBATION THEORY)

Here we develop a vector perturbation theory for the general scaling of a waveguide refractive-index profile. We classify a perturbation as general scaling, as depicted in Fig. 1 if the coordinates (x, y) of the unperturbed waveguide can be related to the coordinates of the perturbed waveguide $(x_{\text{scaled}}, y_{\text{scaled}})$ by

$$x_{\text{scaled}} = x(1 + \delta_x), \quad y_{\text{scaled}} = y(1 + \delta_y). \quad (24)$$

We start with a perturbed structure defined in a regular Euclidian coordinate system. To formulate a converging perturbation theory we perform a coordinate transformation so that in the new (generally curvilinear) coordinate system the initially perturbed geometric structure becomes unperturbed. Then we use the modes of an unperturbed structure to form an adequate basis set for further expansions. Proceeding in such a way we effectively substitute the problem of evaluating new fields in a structure with shifted dielectric interfaces by a problem with unchanged geometry but a slightly changed form of underlying electromagnetic equations.

We start with a perturbed waveguide and a generalized Hermitian Hamiltonian formulation [Eqs. (9), (19), and (21)], where the derivatives in Eq. (21) should be understood as the derivatives over the coordinates x_{scaled} and y_{scaled} . We then transform into the coordinate system in which a scaled waveguide becomes unperturbed using the transformations in Eqs. (24). Thus, starting with Eq. (21) written in scaled coordinates $(x_{\text{scaled}}, y_{\text{scaled}})$, we re-express the same operator in terms of the unperturbed coordinates (x, y) . First, we note that if \mathbf{F} is a vector field then¹⁷

$$\begin{aligned} \nabla_{t,\text{scaled}} \times \mathbf{F} &= \det \begin{pmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x_{\text{scaled}}} & \frac{\partial}{\partial y_{\text{scaled}}} & 0 \\ F_{x,\text{scaled}} & F_{y,\text{scaled}} & F_z \end{pmatrix} \\ &= \det \begin{bmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{(1 + \delta_x)\partial x} & \frac{\partial}{(1 + \delta_y)\partial y} & 0 \\ F_x & F_y & F_z \end{bmatrix} \\ &= \nabla_t \times \mathbf{F} - \hat{\mathcal{O}}, \end{aligned} \quad (25)$$

where the operator $\hat{\mathcal{O}}$ is defined in terms of $\eta_x = \delta_x/(1 + \delta_x)$ and $\eta_y = \delta_y/(1 + \delta_y)$ by

$$\Delta \hat{\mathcal{A}} = \frac{c}{\omega} \left\{ \begin{array}{l} \nabla_t \times [\hat{z}(\hat{z} \cdot \hat{\mathcal{O}})] + \hat{\mathcal{O}}[\hat{z}(\hat{z} \cdot \nabla_t \times)] \\ 0 \\ \nabla_t \times \left[\hat{z} \left(\frac{1}{\epsilon} \hat{z} \cdot \hat{\mathcal{O}} \right) \right] + \hat{\mathcal{O}} \left[\hat{z} \left(\frac{1}{\epsilon} \hat{z} \cdot \nabla_t \times \right) \right] \end{array} \right\}. \quad (28)$$

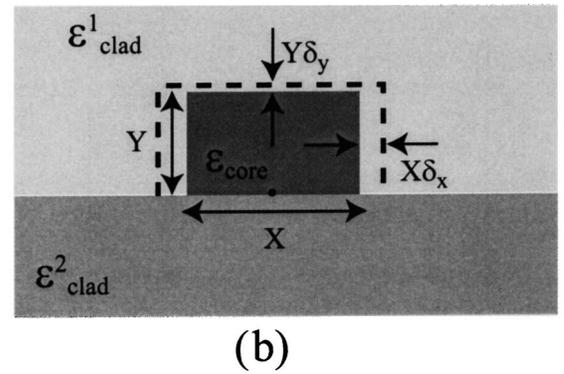
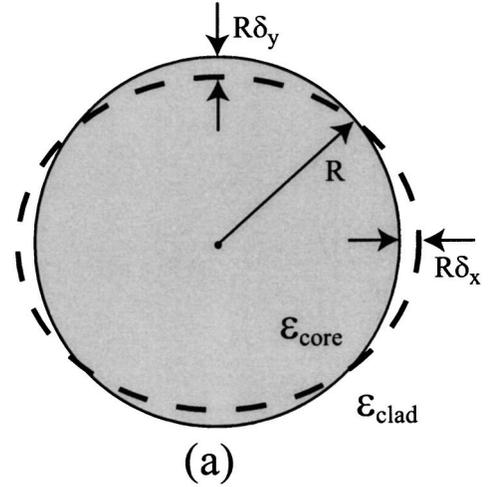


Fig. 1. General scaling perturbation defined by a scaling of coordinates $x_{\text{scaled}} = x(1 + \delta_x)$ and $y_{\text{scaled}} = y(1 + \delta_y)$: (a) the particular case of $\delta_x = -\delta_y$ corresponds to uniform elliptical perturbation of a fiber, and (b) scaling perturbations can also be used to analyze size variations of planar waveguides.

$$\hat{\mathcal{O}} = \det \begin{pmatrix} \hat{x} & \hat{y} & \hat{z} \\ \eta_x \frac{\partial}{\partial x} & \eta_y \frac{\partial}{\partial y} & 0 \\ F_x & F_y & F_z \end{pmatrix}. \quad (26)$$

Substituting Eqs. (25) and (26) into Eq. (21), we find

$$\hat{\mathcal{A}}_{\text{scaled}} = \hat{\mathcal{A}}_0 + \Delta \hat{\mathcal{A}}, \quad (27)$$

where $\hat{\mathcal{A}}_0$ has the same form as in Eq. (21) and is written for a waveguide with an unperturbed index profile, whereas the perturbation operator $\Delta \hat{\mathcal{A}}$ is

We assume that the eigenmodes of the unperturbed waveguide are known, and we will use them to calculate the matrix elements in Eqs. (12) and (14) by using Eq. (28) as the perturbation.

In the following we evaluate the matrix element $\langle \psi_{\beta} | \Delta \hat{A} | \psi_{\beta'} \rangle$ for a generic scaling perturbation characterized by η_x and η_y . Using the vector identity in Eq. (20) as well as Maxwell's equations in Eq. (8) after some manipulation involving integration by parts, we found that the matrix elements of Eq. (28) take the form

$$\langle \psi_{\beta,m} | \Delta \hat{A} | \psi_{\beta',m'} \rangle = \int_S ds \begin{pmatrix} E_x \\ E_y \\ H_x \\ H_y \end{pmatrix}_{\beta}^{\dagger} \begin{pmatrix} 2 \frac{\omega}{c} \epsilon \eta_y & 0 & 0 & -(\beta \eta_x + \beta' \eta_y) \\ 0 & 2 \frac{\omega}{c} \epsilon \eta_x & \beta \eta_y + \beta' \eta_x & 0 \\ 0 & \beta \eta_x + \beta' \eta_y & 2 \frac{\omega}{c} \eta_y & 0 \\ -(\beta \eta_y + \beta' \eta_x) & 0 & 0 & 2 \frac{\omega}{c} \eta_x \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ H_x \\ H_y \end{pmatrix}_{\beta'} \quad (29)$$

which is valid for arbitrary geometries, including cylindrical fibers and planar rectangular waveguides.

For a cylindrically symmetric fiber the eigenmodes of the fiber are characterized by their wave vector β and a corresponding angular index m so that, in cylindrical coordinates (ρ, θ, z) , the fields take the form

$$\begin{pmatrix} \mathbf{E}(\mathbf{r}, t) \\ \mathbf{H}(\mathbf{r}, t) \end{pmatrix}_{\beta,m} = \begin{pmatrix} \mathbf{E}(\rho) \\ \mathbf{H}(\rho) \end{pmatrix}_{\beta,m} \exp(i\beta z - i\omega t + im\theta). \quad (30)$$

Because of this symmetry of fibers it is advantageous to operate in such cylindrical coordinates. After a change of coordinates and explicit integration over angle θ , Eq. (29) transforms into

$$\begin{aligned} \langle \psi_{\beta,m} | \Delta \hat{A} | \psi_{\beta',m'} \rangle = \int dr \begin{pmatrix} E_r \\ E_{\theta} \\ H_r \\ H_{\theta} \end{pmatrix}_{\beta,m}^{\dagger} & \left\{ \frac{(\eta_x + \eta_y)}{2} \delta_{m,m'} \begin{bmatrix} 2 \frac{\omega}{c} \epsilon & 0 & 0 & -(\beta + \beta') \\ 0 & 2 \frac{\omega}{c} \epsilon & (\beta + \beta') & 0 \\ 0 & (\beta + \beta') & 2 \frac{\omega}{c} & 0 \\ -(\beta + \beta') & 0 & 0 & 2 \frac{\omega}{c} \end{bmatrix} \right. \\ & + \left. \frac{(\eta_x - \eta_y)}{2} \delta_{m,m' \mp 2} \begin{bmatrix} -2 \frac{\omega}{c} \epsilon & \pm i \frac{\omega}{c} \epsilon & \mp i \frac{(\beta - \beta')}{2} & -\frac{(\beta - \beta')}{2} \\ \pm i \frac{\omega}{c} \epsilon & 2 \frac{\omega}{c} \epsilon & -\frac{(\beta - \beta')}{2} & \pm i \frac{(\beta - \beta')}{2} \\ \pm i \frac{(\beta - \beta')}{2} & \frac{(\beta - \beta')}{2} & -2 \frac{\omega}{c} & \pm i \frac{\omega}{c} \\ \frac{(\beta - \beta')}{2} & \mp i \frac{(\beta - \beta')}{2} & \pm i \frac{\omega}{c} & 2 \frac{\omega}{c} \end{bmatrix} \right\} \\ & \times \begin{pmatrix} E_r \\ E_{\theta} \\ H_r \\ H_{\theta} \end{pmatrix}_{\beta',m'} \quad (31) \end{aligned}$$

The form of the matrix elements suggests that generic scaling can be decomposed into uniform scaling with a scaling parameter $\eta_s = (\eta_x + \eta_y)/2$ and uniform ellipticity with ellipticity parameter $\eta_e = (\eta_x - \eta_y)/2$. We now address these cases of uniform scaling and uniform ellipticity in greater detail.

A. Uniform Scaling Perturbation

For uniform scaling, $\eta_x = \eta_y = \eta$, the matrix element corresponds to the first term in Eq. (31), which implies that uniform scaling directly couples only modes with the same angular index, $m = m'$. In particular, the first-order correction to the propagation constant of a mode (β, m) that is due to uniform scaling is, by means of Eq. (12),

$$\begin{aligned} \Delta\beta_s &= \langle \psi_{\beta,m} | \Delta\hat{A} | \psi_{\beta,m} \rangle \\ &= 2\eta \int dr \begin{pmatrix} E_r \\ E_\theta \\ H_r \\ H_\theta \end{pmatrix}_{\beta,m}^\dagger \\ &\quad \times \begin{pmatrix} \frac{\omega}{c}\epsilon & 0 & 0 & -\beta \\ 0 & \frac{\omega}{c}\epsilon & \beta & 0 \\ 0 & \beta & \frac{\omega}{c} & 0 \\ -\beta & 0 & 0 & \frac{\omega}{c} \end{pmatrix} \begin{pmatrix} E_r \\ E_\theta \\ H_r \\ H_\theta \end{pmatrix}_{\beta,m} \\ &= 2\eta \frac{\omega}{c} \int dr (\epsilon |E_z|^2 + |H_z|^2), \end{aligned} \quad (32)$$

where the final expression was derived with the help of the mode orthogonality condition in Eq. (11), representations in Eqs. (19), (21), and (22) of operators A , B and eigenvalue β . From the expression in Eq. (32) we see that the shift in the propagation constant of the mode that is due to uniform scaling is proportional to the time-average longitudinal (magnetic and electric) energy in the cross section of a structure.

Another interesting and independent result about the change in the propagation constant of a mode under uniform scaling is that its frequency derivative is proportional to the group-velocity dispersion of the mode. To derive this result, we consider a dispersion relation for some mode of the waveguide, $\beta = f(\omega)$. From the form of Eq. (8), it is obvious that, if we uniformly rescale all the transverse dimensions in a system by a factor of $(1 + \delta)$, the new $\tilde{\beta} = \beta + \Delta\beta_s$ for the same ω will satisfy $\tilde{\beta} = \{f[\omega(1 + \delta)]/(1 + \delta)\}$. Expanding this expression in a Taylor series and collecting terms of the same order in δ , we derive expressions for $\Delta\beta_s$ valid up to second order in δ^2 :

$$\Delta\beta_s = \delta \left[\omega \frac{\partial\beta}{\partial\omega} - \beta \right]. \quad (33)$$

Taking the derivative of Eq. (33) with respect to ω , we obtain up to second order in δ^2 :

$$\frac{\partial\Delta\beta_s}{\partial\omega} = \delta\omega \frac{\partial^2\beta}{\partial\omega^2} = -\lambda \delta D(\omega), \quad (34)$$

where $D(\omega) = (\omega^2/2\pi c)(\partial^2\beta/\partial\omega^2)$ is the group-velocity dispersion of the mode.

B. Uniform Ellipticity Perturbation

For uniform ellipticity, $\eta_x = -\eta_y = \eta$, the matrix element corresponds to the second term in Eq. (31). It shows that uniform ellipticity directly couples only modes with angular momentum states different by 2, i.e., $m = m' \pm 2$. In particular, for the case of an $m = 1$ doubly degenerate mode, PMD arises because of the ellipticity-induced split between the originally degenerate $m = 1$ and $m' = -1$ modes. The first-order correction of Eq. (14) to the split in the propagation constants of modes $(\beta, 1)$ and $(\beta, -1)$ because of uniform ellipticity becomes

$$\begin{aligned} \Delta\beta_e &= 2|\langle \psi_{\beta,1} | \Delta\hat{A} | \psi_{\beta,-1} \rangle| \\ &= 2\eta \frac{\omega}{c} \left| \int dr \begin{pmatrix} E_r \\ E_\theta \\ H_r \\ H_\theta \end{pmatrix}_{\beta,1}^\dagger \right. \\ &\quad \times \begin{pmatrix} -\epsilon & -i\epsilon & 0 & 0 \\ -i\epsilon & \epsilon & 0 & 0 \\ 0 & 0 & -1 & -i \\ 0 & 0 & -i & 1 \end{pmatrix} \begin{pmatrix} E_r \\ E_\theta \\ H_r \\ H_\theta \end{pmatrix}_{\beta,-1} \left. \right| \\ &= 2\eta \frac{\omega}{c} \left| \int dr [(-\epsilon |E_z|^2 + |H_z|^2) \right. \\ &\quad \left. + 2Im(\epsilon E_r^* E_\theta - H_r^* H_\theta)] \right|, \end{aligned} \quad (35)$$

where the E and the H are those of the $(\beta, 1)$ mode. In derivation we used the fact that the fields of the degenerate modes (β, m) and $(\beta, -m)$ can be related by

$$\begin{aligned} E_z^{-m} &= -E_z^m, & E_r^{-m} &= -E_r^m, & E_\theta^{-m} &= E_\theta^m, \\ H_z^{-m} &= H_z^m, & H_r^{-m} &= H_r^m, & H_\theta^{-m} &= -H_\theta^m, \end{aligned} \quad (36)$$

and the total electric and magnetic energies are equal

$$\int dr \epsilon (|E_z|^2 + |E_t|^2) = \int dr (|H_z|^2 + |\mathbf{H}_t|^2).$$

We find that, for high-index contrast profiles, $\Delta\beta_e$ is generally dominated by the diagonal term $\sim |\int dr [(-\epsilon |E_z|^2 + |H_z|^2)]|$, whereas for low-index contrast profiles the cross terms of Eq. (35) also become important. For any index contrast, an important conclusion about the PMD of a structure can be drawn when either the electric or the magnetic longitudinal energy dominates substantially over the other. If a mode behaves like a pure TE ($\int dr \epsilon |E_z|^2 \ll \int dr |H_z|^2$) or TM ($\int dr |H_z|^2 \ll \int dr \epsilon |E_z|^2$) mode, the split that is due to uniform scaling [see Eq. (32)] becomes almost identical to the split in the degeneracy of the modes that is due to uniform ellipticity perturbation [see Eq. (35)] and thus $\Delta\beta_s$

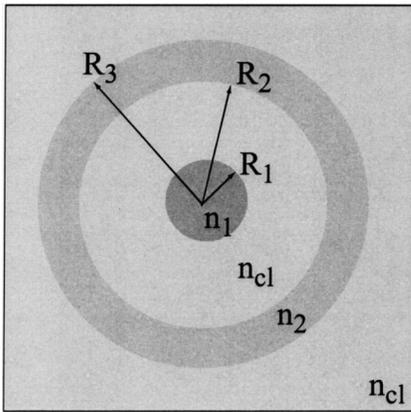


Fig. 2. Dual-core dielectric waveguide. The inner core is a dielectric cylinder of radius R_1 and index n_1 , whereas the outer core is a ring of index n_2 with inner and outer radii R_2 and R_3 , respectively. Two cores are separated by cladding with an index n_{cl} .

$\approx \Delta\beta_e$. As PMD is proportional to $\tau = \partial\Delta\beta_e/\partial\omega$ and taking into account Eq. (34) for the frequency derivatives of $\Delta\beta_s$, we arrive at the conclusion that, for such modes, PMD is proportional to the dispersion of a mode:

$$\tau = \left| \frac{\partial\Delta\beta_e}{\partial\omega} \right| \approx \left| \frac{\partial\Delta\beta_s}{\partial\omega} \right| = \lambda \delta |D|. \quad (37)$$

5. PMD AND DISPERSION OF A DUAL-CORE STEP-INDEX FIBER

In the following we consider a relationship between PMD and dispersion in a dual-core step-index fiber (Fig. 2). In particular, we are interested in how PMD behaves, compared with group-velocity dispersion D , as the core-clad index contrast increases.

An inner core of a dual-core dielectric waveguide is a dielectric cylinder of radius $R_1 = a$ and a variable index n_1 , whereas an outer core is a ring of index $n_2 = (n_1 + n_{cl})/2$, with inner and outer radii $R_2 = 5a$ and $R_3 = 6a$, respectively. Two cores are separated by cladding with an index $n_{cl} = 1.5$. For a given core index of refraction n_1 , we chose the radius of an inner core a in such a way that the absolute value of group-velocity dispersion of a fundamental $m = 1$ mode reaches its maximum at $\lambda = 1.55 \mu\text{m}$.

In Fig. 3 we plot a split $|\Delta\beta_e| = |\beta_{m=1} - \beta_{m=-1}|$ in a propagation constant of a doubly degenerate $m = 1$ fundamental mode of a dual-core fiber with $n_1 = 2.5$ that is due to a uniform elliptical perturbation of $\delta = 2\%$. $|\Delta\beta_e|$ is measured in units of $2\pi/a$, where $a = 0.2046 \mu\text{m}$. The solid curve corresponds to $|\Delta\beta_e|$ as calculated by the first-order perturbation theory in Eq. (35). For comparison, we computed the dispersion relations of an elliptically perturbed fiber by using iterative solutions of a general plane-wave expansion implemented in a frequency-domain code, i.e., MIT photonic bands (MPB).¹⁰ The crosses correspond to $|\Delta\beta_e|$ as calculated with the MPB. Excellent agreement was observed between our first-order perturbation theory and the MPB code across a wide frequency range.

Defining the PMD parameter $D_e = \tau/(\lambda\delta)$, we know from the discussion in previous sections and Eq. (37), that D_e is loosely bound from the above by the value of the group-velocity dispersion D . Thus, for modes in which electric and magnetic longitudinal energy densities are comparable, $D_e \ll D$, whereas for pure TE and TM modes $D_e \approx D$. In Fig. 4 we plot the dispersion curves $D(\omega)$ and $D_e(\omega)$ for different values of the index of refraction of the core $n_1 = [2.5, 2.0, 1.6]$, keeping the index of refraction of the cladding fixed and equal to $n_{cl} = 1.5$. When the core-clad index contrast is small, the PMD parameter

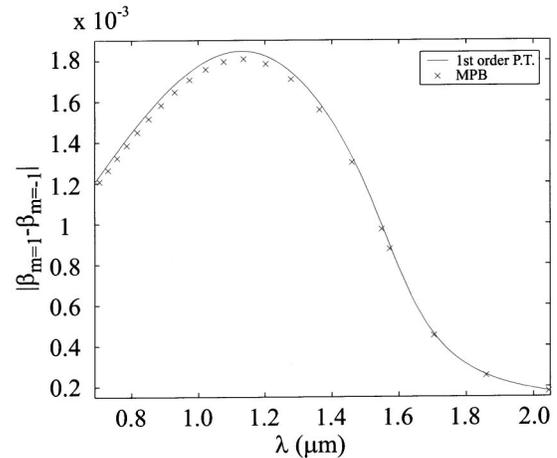


Fig. 3. Value of a split $|\Delta\beta_e| = |\beta_{m=1} - \beta_{m=-1}|$ in a propagation constant of a doubly degenerate $m = 1$ fundamental mode of a dual core ($n_1 = 2.5$) fiber (see Fig. 2) that is due to a uniform elliptical perturbation of $\delta = 2\%$. $|\Delta\beta_e|$ is measured in units of $2\pi/a$, where $a = 0.2046 \mu\text{m}$. The solid curve corresponds to $|\Delta\beta_e|$ as calculated by the first-order perturbation theory in Eq. (35). The crosses correspond to $|\Delta\beta_e|$ as calculated by the frequency-domain plane-wave expansion code MPB. Excellent agreement was observed between the first-order perturbation theory and the frequency-domain code across a wide frequency range.

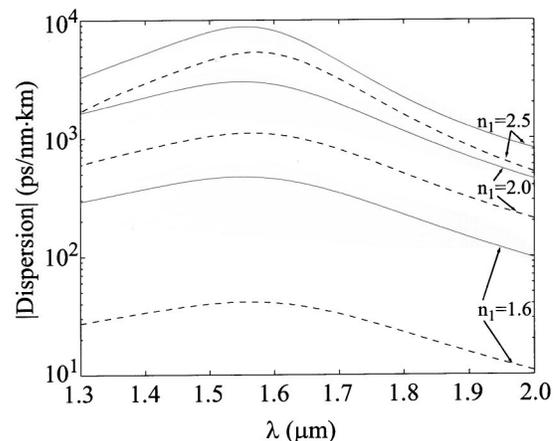


Fig. 4. PMD parameter D_e (dotted curves) and group-velocity dispersion D (solid curves) plotted for the $m = 1$ fundamental mode as a function of frequency for different core-clad index contrasts. The value of scaling factor a was chosen to maximize the absolute value of the negative dispersion of the waveguide for the $m = 1$ fundamental mode at $\lambda = 1.55 \mu\text{m}$. For a fixed $n_{clad} = 1.5$, varying the dielectric constant of the core leads to the D_e curves that approach the D curves for the high values of index contrast.

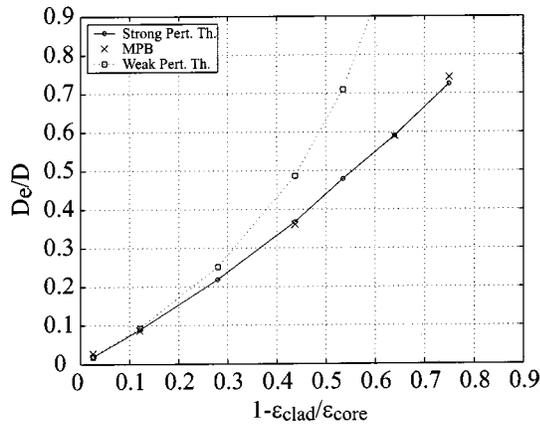


Fig. 5. Ratio of the PMD parameter and group-velocity dispersion D_e/D plotted as a function of the core-clad dielectric contrast relative to the dielectric constant of the core. The solid curve corresponds to the results of a strong perturbation theory (circles). The crosses correspond to a numerical simulation that was performed with the frequency-domain plane-wave expansion code MPB, and they closely match a curve that is due to a strong perturbation theory. Weak perturbation theory (squares) predicts correct results only for the relative dielectric contrasts that are less than $\sim 10\%$.

D_e is much smaller than the group-velocity dispersion D over the whole single/mode frequency range, with the negative minimum of D approximately coinciding with the negative minimum of the PMD parameter. When the index contrast is increased, the PMD parameter starts to approach the group-velocity dispersion, and for large index contrasts the curves of the PMD parameter and the group-velocity dispersion almost overlap.

In Fig. 5, the ratio of the PMD parameter to group-velocity dispersion at $\lambda = 1.55 \mu\text{m}$, D_e/D , is plotted versus the ratio of the core-clad dielectric contrast to the core dielectric constant $(\epsilon_{\text{core}} - \epsilon_{\text{clad}})/\epsilon_{\text{core}}$. The same curve computed by the frequency-domain MPB code is plotted in Fig. 5 as crosses. The results of our (strong) perturbation theory are in excellent agreement with the MPB calculations for small as well as large dielectric contrasts. We also note that for small index contrast, the ratio of the PMD parameter to group-velocity dispersion is directly proportional to the relative dielectric contrast. For high-index contrasts, the difference between the PMD parameter and dispersion approaches zero linearly with the ratio of clad-to-core dielectric constants.

To demonstrate inadequacies of the standard (weak) perturbation theory derived in Section 3, we performed calculations of the same D_e/D curve using Eq. (23). As shown in Fig. 3 as squares, this curve matches the correct result only for the relative dielectric contrast smaller than 10% whereas it gives erroneous results for higher dielectric contrasts.

6. CONCLUSION

We have presented a novel perturbation theory approach to treat general scaling perturbations in a waveguide with an arbitrary dielectric profile. This formulation, unlike much previous research, holds equally well for dis-

continuous high-index contrast profiles. We have demonstrated that the PMD of a fiber that is due to an elliptical perturbation can be conveniently characterized by a PMD parameter D_e that is proportional to intermode dispersion parameter τ . In the case of low-index contrast or a mixed polarization mode (judged by the electric and magnetic energies in the longitudinal direction), this PMD parameter is much smaller than dispersion ($D_e \ll D$), whereas for a high-index contrast or a purely polarized mode $D_e \approx D$ in general.

The e-mail address for M. Skorobogatiy is maksim@omni-guide.com.

REFERENCES AND NOTES

1. S. J. Savori, "Pulse propagation in fibers with polarization-mode dispersion," *J. Lightwave Technol.* **19**, 350–357 (2001).
2. A. W. Snyder and J. D. Love, *Optical Waveguide Theory* (Chapman & Hall, London, 1983).
3. C. Vassallo, *Optical Waveguide Concepts* (Elsevier, Amsterdam, 1991).
4. J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1998), p. 357.
5. D. Marcuse, *Theory of Dielectric Optical Waveguides* (Academic, New York, 1974).
6. M. Lohmeyer, N. Bahlmann, and P. Hertel, "Geometry tolerance estimation for rectangular dielectric waveguide devices by means of perturbation theory," *Opt. Commun.* **163**, 86–94 (1999).
7. D. Chowdhury, "Comparison between optical fiber birefringence induced by stress anisotropy and geometric deformation," *IEEE J. Sel. Top. Quantum Electron.* **6**, 227–232 (2000).
8. V. P. Kalosha and A. P. Khapalyuk, "Mode birefringence in a single-mode elliptic optical fiber," *Sov. J. Quantum Electron.* **13**, 109–111 (1983).
9. V. P. Kalosha and A. P. Khapalyuk, "Mode birefringence of a three-layer elliptic single-mode fiber waveguide," *Sov. J. Quantum Electron.* **14**, 427–430 (1984).
10. S. G. Johnson and J. D. Joannopoulos, "Block-iterative frequency-domain methods for Maxwell's equations in a planewave basis," *Opt. Express* **8**, 173–190 (2001), <http://www.opticsexpress.org>.
11. S. G. Johnson, M. Ibanescu, M. A. Skorobogatiy, O. Weisberg, J. D. Joannopoulos, and Y. Fink, "Perturbation theory for Maxwell's equations with shifting material boundaries," *Phys. Rev. E* **65**, 066611 (June 2002).
12. S. G. Johnson, M. Ibanescu, M. Skorobogatiy, O. Weisberg, T. D. Engeness, M. Soljačić, S. A. Jacobs, J. D. Joannopoulos, and Y. Fink, "Low-loss asymptotically single-mode propagation in large-core OmniGuide fibers," *Opt. Express* **9**, 748–779 (2001), <http://www.opticsexpress.org>.
13. Note, multiplying from the left by \hat{B}^{-1} on both sides of Eq. (8) and, thus, trying to make it look as a standard eigenvalue problem is erroneous as, in general, a resulting matrix on the right would not be Hermitian.
14. L. D. Landau and E. M. Lifshitz, *Quantum Mechanics (Non-Relativistic Theory)* (Butterworth-Heinemann, Stoneham, Mass., 2000), p. 140.
15. F. Curti, B. Daino, G. De Marchis, and F. Matera, "Statistical treatment of the evolution of the principal states of polarization in single-mode fibers," *J. Lightwave Technol.* **8**, 1162–1166 (1990).
16. D. Q. Chowdhury and D. A. Nolan, "Perturbation model for computing optical fiber birefringence from a two-dimensional refractive-index profile," *Opt. Lett.* **20**, 1973–1975 (1995).
17. G. B. Arfken and H. J. Webber, *Mathematical Methods for Physicists* (Academic, New York, 1995).