Analysis of the Effective-Index Method for the Vector Modes of Rectangular-Core Dielectric Waveguides

Kin Seng Chiang, *Member, IEEE*

***Abstract*—The approximations involved in the effective-index method for analyzing the vector modes of rectangular-core dielectric waveguides are examined in detail. It is shown that the effective-index method does not solve the full vector-wave equation that governs the modes. Instead, it solves the reduced vector-wave equation, which is accurate only for approximately linearly polarized waves. Furthermore, in solving the reduced vector-wave equation, the method of separation of variables is used, which leads to additional errors as the waveguide being analyzed is a mathematically nonseparable structure. To characterize the performance of the effective-index method, asymptotic expressions are derived for the errors in the calculation of propagation constants. Apart from separating the effects of different approximations involved, these expressions show explicitly how the accuracy of the method depends on the mode type, the normalized frequency, the mode orders, the dimensions of the waveguide, and the relative refractive-index differences between the core and the surrounding media. With the help of these expressions, it is demonstrated that more accurate results can be obtained by combining various effective-index solutions.**

I. Introduction

THE effective-index method was first proposed by Knox and Toulios in 1970 [1] with a view to improving Marcatili's results [2] for the fundamental mode of a simple rectangular-core waveguide. The basic idea of the method is to replace the rectangular structure by an equivalent slab with an effective refractive index obtained from another slab. The generic feature of the concept has since been recognized and the method has been generalized and applied to more complicated structures, such as composite rectangular waveguides [3]–[9], diffused channel waveguides [10], various types of optical fibers and fiber devices [9], [11]–[15], and nonlinear waveguides [16]. In spite of the fact that the method has become a popular waveguide design tool for a long time [17], the theoretical basis of the method has not been firmly established for many applications.

The first theoretical study of the method was given by Peng and Oliner [6], [18], who pointed out that, in the case of analyzing composite rectangular waveguides, the effective-index method was in fact the lowest-order version of the mode-matching method, which they developed. The approximate nature of the effective-index method can thus be understood with reference to a rigorous numerical method, namely, the mode-matching method. In the study of the polarization properties of optical fibers with the effective-index method, Chiang [12] presented the reduced vector-wave equation (in contrast with the full vector-wave equation) that the method actually solved, and pointed out which terms in the equation were ignored in the application of the method. This study reveals, in the most fundamental way, the basic assumptions made in the method. Considering only the scalar modes, Kumar *et al.* [19] derived the rectangular structure that was exact for the effective-index method. By treating this structure as a perturbation of the original structure, Chiang developed an asymptotic theory for the effective-index method for rectangular waveguides [20] and waveguide arrays [9]. This theory fully characterizes the performance of the method for the scalar modes in the far-from-cutoff region.

In this paper, a detailed analysis of the effective-index method for the vector modes of rectangular-core waveguides is presented. It is shown that the solution obtained from the effective-index method in general contains errors arising from the use of the reduced vector-wave equation to approximate the full vector-wave equation, and from the use of a different structure to approximate the original structure. Analytical asymptotic expressions for these errors are derived to describe explicitly the performance of the method. Based on these expressions, a procedure to obtain more accurate solutions from combining various effective-index solutions is proposed. The results presented in this paper should provide a much clearer picture of the nature of the effective-index method for the analysis of rectangular-core waveguides.

II. Wave Equations and Vector Modes

It is well known that the transverse electric field $\mathbf{E}_t$ in an isotropic waveguide with a refractive-index profile $n(x, y)$ satisfies the full vector-wave equation [17]

$$\nabla_t^2 \mathbf{E}_t + (n^2 k^2 - \beta^2)\mathbf{E}_t + \nabla_t \left[\mathbf{E}_t \frac{\nabla_t (n^2)}{n^2} \right] = 0 \qquad (1)$$

where

$$\nabla_t = \frac{\partial}{\partial x}\hat{\mathbf{x}} + \frac{\partial}{\partial y}\hat{\mathbf{y}}$$

with $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ the unit vectors in the x and y directions, respectively, $k = 2\pi/\lambda$ is the free-space wavenumber with λ the free-space wavelength, and β is the propagation constant. Equation (1) actually consists of two equations coupling the x and y components of the electric field. The solution obtained from (1) in general contains both components (except for some special geometries, such as slab waveguides) and is referred to

Manuscript received June 11, 1995; revised January 17, 1996. This work was supported by UGC Research Grant of the Hong Kong Government.

The author is with the Department of Electronic Engineering, City University of Hong Kong, Hong Kong.

Publisher Item Identifier S 0018-9480(96)03038-4.

0018–9480/96$05.00 © 1996 IEEE

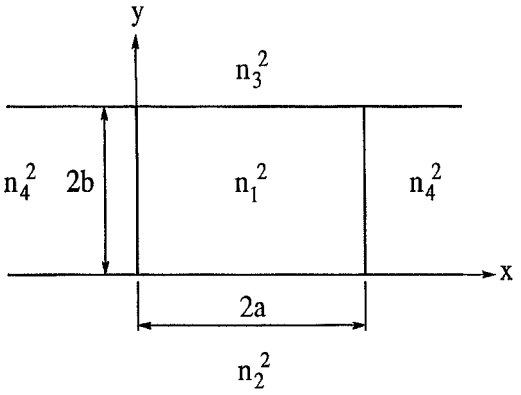


Fig. 1. Dielectric waveguide of rectangular cross section.

as a "hybrid" mode or a vector mode. If the coupling terms in (1) are ignored, the following reduced vector-wave equations are obtained [12]

$$\nabla_t^2 E_x + (n^2k^2 - \beta^2)E_x + \frac{\partial}{\partial x}\left(\frac{E_x}{n^2}\frac{\partial n^2}{\partial x}\right) = 0, \quad (2)$$

$$\nabla_t^2 E_y + (n^2k^2 - \beta^2)E_y + \frac{\partial}{\partial y}\left(\frac{E_y}{n^2}\frac{\partial n^2}{\partial y}\right) = 0 \quad (3)$$

where E_x and E_y represent the x and y components of the electric field, respectively. The propagation constants calculated from (2) and (3), which are separate equations, are in general different and represent two different modes, which are sometimes referred to as the quasi-TE and quasi-TM modes or the E^x and E^y modes. Obviously, these modes are linearly polarized modes as each of these contains only one transverse field component.

It has been shown [12] that the effective-index method solves the reduced vector-wave equations, instead of the full vector-wave equation. This means that, in the effective-index method, the linearly polarized modes are used to approximate the hybrid modes. It is known that the guided mode of a rectangular-core waveguide, though hybrid in nature, does contain a predominant transverse electric field component [2], and, therefore, can be well approximated by the linearly polarized mode, as long as the mode is not close to cutoff. This explains why the effective-index method can usually give good results for the vector modes of rectangular waveguides (except for cases where cross-polarization coupling is involved [6]). In the weak-guidance limit (i.e., when the relative index difference between the core and a surrounding medium is small), however, the modes are always linearly polarized, regardless of the geometrical shape of the waveguide, and the effective-index method can be accurate for a wide range of geometrical shapes [9]–[15].

III. The Effective-Index Method

The cross-section of a rectangular-core waveguide is shown in Fig. 1, where a and b are the half-thickness and the half-width of the core, respectively, and n_1, n_2, n_3, and n_4 are the refractive indices of the core and the surrounding claddings with $n_1 > n_2 \geq n_3, n_4$. The refractive-index profile of this waveguide can be characterized by three relative index steps:

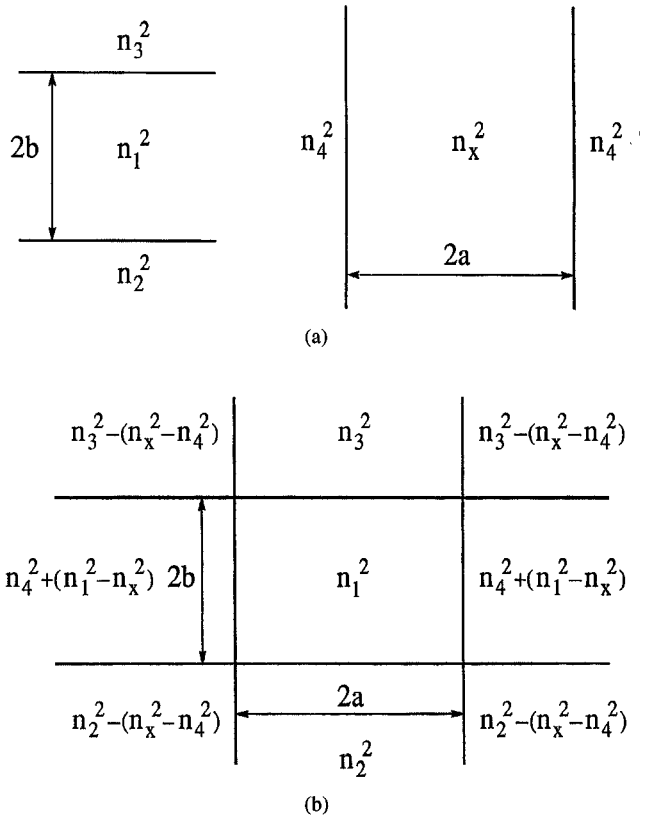


Fig. 2. (a) Effective-index method that results in an x-dependent profile (the x-method). (b) Mathematically separable profile that the x-method actually analyzes.

$\Delta_i = (n_1^2 - n_i^2)/2n_1^2$ for $i = 2, 3, 4$. This structure is general enough to represent several important classes of waveguides including fully embedded waveguides ($\Delta_2 = \Delta_3 = \Delta_4$), channel waveguides ($\Delta_3 > \Delta_2 = \Delta_4$), and strip waveguides ($\Delta_3 = \Delta_4 > \Delta_2$). The guided mode of the waveguide is denoted by the E_{mn}^i mode, where $i = x$ or y indicates the direction of the predominant transverse electric field, and m and n are the numbers of the peaks in the field in the x and y directions, respectively. Two different ways of applying the effective-index method to the waveguide are possible, giving different solutions for the same mode [7].

A. The x-Method

The effective-index method that results in an x-dependent refractive-index profile (the x-method) is shown in Fig. 2(a). The mode index (the propagation constant divided by the free-space wavenumber) for the $\mathrm{TE}_{n-1}(\mathrm{TM}_{n-1})$ mode of the slab waveguide of thickness $2b$ is used as the effective refractive index n_x of a second slab of thickness $2a$. The propagation constant for the $\mathrm{TM}_{m-1}(\mathrm{TE}_{m-1})$ mode of the second slab waveguide is then regarded as the approximate propagation constant for the $E_{mn}^x(E_{mn}^y)$ mode of the rectangular waveguide [7]. The fact that the effective-index method solves two slab waveguides in a sequence suggests that it is a method of separation of variables. There must exist a separable refractive-index profile that the method actually analyzes. The extent to which this profile differs from the original profile is therefore

a measure of the errors in the method. To find this profile, the method of Kumar *et al.* [19] is used.

According to the separation of variables, the mode field and the profile can be expressed by

$$\psi(x,y) = \psi_a(x)\psi_b(y), \tag{4}$$

$$n^2(x,y) = n_a^2(x) + n_b^2(y). \tag{5}$$

For the E_{mn}^x mode with $E_x = \psi$, the two steps of the x-method just described can be represented mathematically by

$$\frac{d^2\psi_b}{dy^2} + \left[n_b^2(y)k^2 - n_x^2k^2\right]\psi_b = 0, \tag{6}$$

$$\frac{d^2\psi_a}{dx^2} + \frac{d}{dx}\left[\frac{\psi_a}{(n_a^2+n_x^2)}\frac{d(n_a^2+n_x^2)}{dx}\right] + \left[n_a^2(x)k^2 + n_x^2k^2 - \beta_{x1}^2\right]\psi_a = 0 \tag{7}$$

with

$$n_b^2(y) = \begin{cases} n_3^2, & 2b < y < +\infty, \\ n_1^2, & 0 \le y \le 2b, \\ n_2^2, & -\infty < y < 0, \end{cases} \tag{8}$$

$$n_a^2(x) + n_x^2 = \begin{cases} n_r^2, & 0 \le x \le 2a, \\ n_4^2, & \text{otherwise} \end{cases} \tag{9}$$

where β_{x1} denotes the propagation constant calculated from (7). It is noted that (6) and (7) are the TE and TM wave equations for the two slab waveguides shown in Fig. 2(a). From (9), $n_a^2(x)$ can be solved

$$n_a^2(x) = \begin{cases} 0, & x \le 2a, \\ n_4^2 - n_x^2, & \text{otherwise.} \end{cases} \tag{10}$$

The profile that the x-method analyzes for the E_{mn}^x mode is simply given by $n_a^2 + n_b^2$, i.e., the sum of (8) and (10). This profile, which is shown in Fig. 2(b), differs from the original profile in the left and right cladding regions.

For the E_{mn}^y mode with $E_y = \psi$, the effective-index equations are

$$\frac{d^2\psi_b}{dy^2} + \frac{d}{dy}\left(\frac{\psi_b}{n_b^2}\frac{dn_b^2}{dy}\right)\left[n_b^2(y)k^2 - n_x^2k^2\right]\psi_b = 0, \tag{11}$$

$$\frac{d^2\psi_a}{dx^2} + \left[n_a^2(x)k^2 + n_x^2k^2 - \beta_{y1}^2\right]\psi_a = 0 \tag{12}$$

where $n_b^2(y)$ and $n_a^2(x) + n_x^2$ are given by (8) and (9), respectively, and β_{y1} is the corresponding propagation constant. Obviously, $n_a^2(x)$ is also given by (10). The values of n_x for the E_{mn}^x and E_{mn}^y modes are, of course, different, as they are calculated from different equations, i.e., (6) and (11), respectively. The form of the profile that the method analyzes, shown in Fig. 2(b), is, however, independent of the mode type and the mode orders. In fact, it is identical to that already obtained for the scalar modes [19].

B. The y-Method

The effective-index method can also result in a y-dependent refractive-index profile (the y-method), as shown in Fig. 3(a). In this case, the mode index for the $\mathrm{TM}_{m-1}(\mathrm{TE}_{m-1})$ mode of the slab waveguide of thickness $2a$ is used as the effective refractive index n_y of a second slab of thickness $2b$. The propagation constant for the $\mathrm{TE}_{n-1}(\mathrm{TM}_{n-1})$ mode of the second slab waveguide is then used to approximate the propagation constant for the $E_{mn}^x(E_{mn}^y)$ mode of the rectangular waveguide [7].

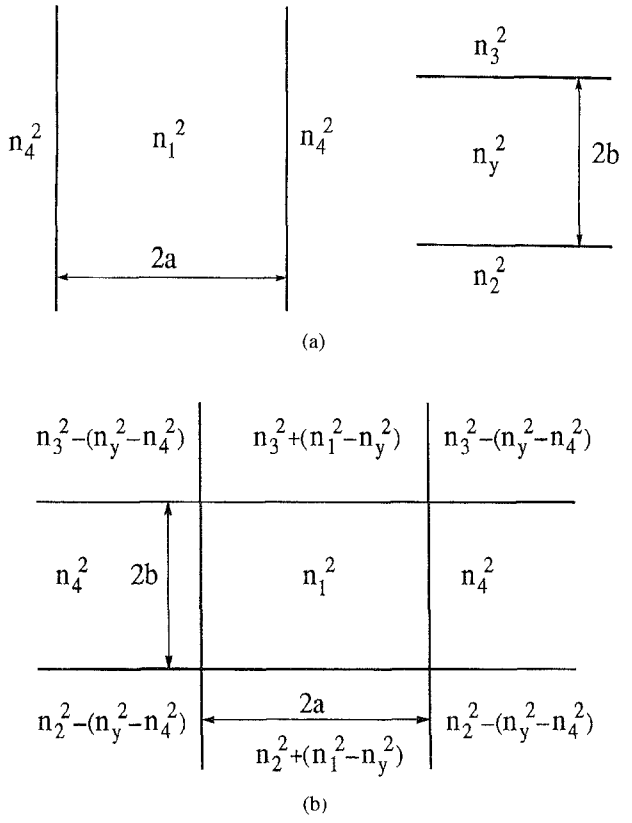


Fig. 3. (a) Effective-index method that results in a y-dependent profile (the y-method). (b) Mathematically separable profile that the y-method actually analyzes.

The E_{mn}^x mode is described by the following two separate equations

$$\frac{d^2\psi_a}{dx^2} + \frac{d}{dx}\left(\frac{\psi_a}{n_a^2}\frac{dn_a^2}{dx}\right)\left[n_a^2(x)k^2 - n_y^2k^2\right]\psi_a = 0, \tag{13}$$

$$\frac{d^2\psi_b}{dy^2} + \left[n_b^2(y)k^2 + n_y^2k^2 - \beta_{x2}^2\right]\psi_b = 0 \tag{14}$$

with

$$n_a^2(x) = \begin{cases} n_1^2, & 0 \le x \le 2a, \\ n_4^2, & \text{otherwise,} \end{cases} \tag{15}$$

$$n_b^2(y) + n_y^2 = \begin{cases} n_y^2, & 0 \le y \le 2b, \\ n_3^2, & 2b < y < +\infty, \\ n_2^2, & -\infty < y < 0 \end{cases} \tag{16}$$

where β_{x2} is the propagation constant calculated by the method. From (16), $n_b^2(y)$ can be solved

$$n_b^2(y) = \begin{cases} 0, & 0 \le y \le 2b, \\ n_3^2 - n_y^2, & 2b < y < +\infty, \\ n_2^2 - n_y^2, & -\infty < y < 0. \end{cases} \tag{17}$$

The profile for the y-method for the E_{mn}^x mode is given by the sum of (15) and (17), as shown in Fig. 3(b), which differs from the original profile in the upper and lower cladding regions.

For the sake of completeness, the equations for the E_{mn}^{y} mode are also given

$$\frac{d^2\psi_a}{dx^2} + \left[n_a^2(x)k^2 - n_y^2k^2\right]\psi_a = 0, \tag{18}$$

$$\frac{d^2\psi_b}{dy^2} + \frac{d}{dy}\left[\frac{\psi_b}{(n_b^2+n_y^2)}\frac{d(n_b^2+n_y^2)}{dy}\right] + \left[n_b^2(y)k^2 + n_y^2k^2 - \beta_{y2}^2\right]\psi_b = 0 \tag{19}$$

where $n_a^2(x), n_b^2(y) + n_y^2$, and $n_b(y)$ are given by (15), (16), and (17), respectively, and β_{y2} is the corresponding propagation constant. The form of the profile for the E_{mn}^{y} mode is just the same as that for the E_{mn}^{x} mode, and, in fact, the same as that for the scalar mode [19].

IV. Error Analysis

From the results obtained in the previous sections, the errors in the effective-index method can be understood in the following way. First, errors arise from the use of the solution of the reduced vector-wave equation for a separable profile to approximate the full vector-wave solution for the original profile. Second, errors arise from solving the reduced vector-wave equation for the separable profile. Unlike the scalar-wave equation, the reduced vector-wave equation is nonseparable even for a separable profile. For example, the pair (6) and (7) for the x-method for the E_{mn}^{x} mode is not identical to the reduced vector-wave equation (2) for the separable profile. There exist errors in approximating (2) with (6) and (7). These two types of errors can be separated by using perturbation theory. This forms the basis of quantifying the accuracy of the effective-index method.

A. Perturbation Formulas

The errors in the effective-index solution can be written as

$$\beta_{ij}^2 = \beta_i^2 + \epsilon_{ij} + \eta_{ij} \tag{20}$$

where β_i is the exact propagation constant, ϵ_{ij} represents the error arising from the use of the separable linearly polarized solution to approximate the exact vector solution, and η_{ij} represents the error arising from the use of the effective-index equations to approximate the reduced vector-wave equation for the separable profile, with $i = x$ for the E_{mn}^{x} mode and $i = y$ for the E_{mn}^{y} mode, and $j = 1$ for the x-method and $j = 2$ for the y-method.

The exact expression for ϵ_{ij} can be readily found from perturbation theory [21]

$$\epsilon_{ij} = 2n_1k^2\left(\frac{\varepsilon_0}{\mu_0}\right)^{1/2} \times \frac{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[(n_a^2(x)+n_b^2(y)-n^2(x,y)\right]\mathbf{E}\cdot\mathbf{E}_e\,dxdy}{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(\mathbf{E}\times\mathbf{H}_e^* + \mathbf{E}_e^*\times\mathbf{H})\cdot\hat{\mathbf{z}}\,dxdy} \tag{21}$$

where ε_0 and μ_0 are the free-space permittivity and permeability, respectively, $n^2(x,y)$ is the original profile shown in Fig. 1, $n_a^2(x) + n_b^2(y)$ is the separable profile, which depends on the method used (i.e., the value of j), $\mathbf{E}$ and $\mathbf{H}$ are the electric and magnetic fields for the separable profile, respectively, $\mathbf{E}_e$ and $\mathbf{H}_e$ are the electric and magnetic fields for the original profile, respectively, and $\hat{\mathbf{z}}$ is the unit vector along the z direction. The lowest-order approximation of ϵ_{ij} can be obtained by replacing the fields in (21) with those calculated by the effective-index method

$$\epsilon_{ij} \simeq k^2\frac{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[n_a^2(x)+n_b^2(y)-n^2(x,y)](E_i^2+E_z^2)\,dxdy}{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}E_i^2\,dxdy} \tag{22}$$

with

$$E_z^2 = -\frac{1}{n_1^2k^2}\left(\frac{dE_i}{di}\right)^2 \tag{23}$$

where $E_i = \psi_a(x)\psi_b(y)$ is the effective-index solution with $i = x$ for the E_{mn}^{x} mode and $i = y$ for the E_{mn}^{y} mode. The relation $\mathbf{H}_t = (\varepsilon_0/\mu_0)^{1/2}n_1\hat{\mathbf{z}}\times\mathbf{E}_t$, where $\mathbf{H}_t$ and $\mathbf{E}_t$ are the transverse magnetic and electric fields, respectively, has been used in deriving (22).

To find an expression for η_{ij}, the reduced vector-wave equation is multiplied by the effective-index solution $\psi_a\psi_b$ and integrated over the entire two-dimensional space. The appropriate effective-index equations, in integral forms, are then used to simplify the result. In this way, the effective-index equations are effectively treated as a perturbation of the corresponding reduced vector-wave equation for the separable profile. The lowest-order approximation of η_{ij} can then be obtained by putting the field calculated by the effective-index method into the final expression. This procedure leads to the following results, as in (24)–(27), shown at the bottom of the next page. With the help of these expressions, the errors in the calculated propagation constants can be estimated with the solutions obtained from the effective-index method. This provides a way of assessing the accuracy of the method without the need to refer to exact numerical data. In fact, these formulas could be evaluated numerically to provide perturbation corrections to the solutions calculated by the effective-index method to produce more accurate results.

B. Asymptotic Results

In this section, an attempt is made to derive analytical expressions from the perturbation formulas to characterize explicitly the performance of the effective-index method. The experience with the scalar modes [9], [20] shows that an explicit error analysis appears to be possible only in the case when the normalized frequency V defined by

$$V \equiv bkn_1(2\Delta_2)^{1/2} \tag{28}$$

tends to infinity. In such a case, the mode is far from cutoff and the mode field is mainly confined in the core area. According to an asymptotic theory [22], in the limit $V \to +\infty$, the fields

TABLE I
ASYMPTOTIC ERRORS IN THE PROPAGATION CONSTANTS (SQUARED) CALCULATED BY THE TWO VERSIONS OF THE EFFECTIVE-INDEX METHOD FOR THE VECTOR MODES OF RECTANGULAR-CORE WAVEGUIDES

Mode	Method	Asymptotic Error in β^2
E^x_{mn}	x-Method	$\frac{m^2n^2\pi^4}{16R^3b^2V^3}\left(\frac{\Delta_2}{\Delta_4}\right)^{3/2}(1-2\Delta_4)(1+4\Delta_4)$
	y-Method	$\frac{m^2n^2\pi^4}{32R^2b^2V^3}[1+\left(\frac{\Delta_2}{\Delta_3}\right)^{3/2}]$
E^y_{mn}	x-Method	$\frac{m^2n^2\pi^4}{16R^3b^2V^3}\left(\frac{\Delta_2}{\Delta_4}\right)^{3/2}$
	y-Method	$\frac{m^2n^2\pi^4}{32R^2b^2V^3}[(1-2\Delta_2)(1+4\Delta_2)+\left(\frac{\Delta_2}{\Delta_3}\right)^{3/2}(1-2\Delta_3)(1+4\Delta_3)]$

$\psi_a(x)$ and $\psi_b(y)$ can be written as

$$\psi_a(x)\simeq\begin{cases}\sin\frac{m\pi x}{2a}, & 0<x<2a,\\ \frac{m\pi b}{2aV}\left(\frac{\Delta_2}{\Delta_4}\right)^{1/2}\exp\left[\frac{V}{b}\left(\frac{\Delta_4}{\Delta_2}\right)^{1/2}x\right], & -\infty<x\le 0,\\ \frac{(-1)^{m+1}m\pi b}{2aV}\left(\frac{\Delta_2}{\Delta_4}\right)^{1/2}\exp\left[-\frac{V}{b}\left(\frac{\Delta_4}{\Delta_2}\right)^{1/2}(x-2a)\right], & 2a\le x<+\infty,\end{cases}\tag{29}$$

$$\psi_b(y)\simeq\begin{cases}\sin\frac{n\pi y}{2b}, & 0<y<2b,\\ \frac{(-1)^{n+1}n\pi}{2V}\exp\left[\frac{V}{b}y\right], & -\infty<y\le 0,\\ \frac{n\pi}{2V}\left(\frac{\Delta_2}{\Delta_3}\right)^{1/2}\exp\left[-\frac{V}{b}\left(\frac{\Delta_3}{\Delta_2}\right)^{1/2}(y-2b)\right], & 2b\le y<+\infty.\end{cases}\tag{30}$$

It should be noted that the fields are not continuous across the boundaries. The fields at the boundaries given in (29) and (30) are assumed to be on the cladding sides of the boundaries. The fields on the core sides of the boundaries can be found from the condition that $n^2(x,y)\psi(x,y)$ is continuous everywhere.

The substitution of (29) and (30) into the perturbation formulas (22)–(27) leads to the following simple asymptotic results (after a significant amount of analytical work has been done)

$$\epsilon_{x1}\simeq\frac{m^2n^2\pi^4}{16R^3b^2V^3}\left(\frac{\Delta_2}{\Delta_4}\right)^{3/2}(1-2\Delta_4),\tag{31}$$

$$\eta_{x1}\simeq\frac{m^2n^2\pi^4\Delta_2}{4R^3b^2V^3}\left(\frac{\Delta_2}{\Delta_4}\right)^{1/2}(1-2\Delta_4),\tag{32}$$

$$\epsilon_{x2}\simeq\frac{m^2n^2\pi^4}{32R^2b^2V^3}\left[1+\left(\frac{\Delta_2}{\Delta_3}\right)^{3/2}\right],\tag{33}$$

$$\eta_{x2}\sim O\left(\frac{1}{V^4}\right),\tag{34}$$

$$\epsilon_{y1}\simeq\frac{m^2n^2\pi^4}{16R^3b^2V^3}\left(\frac{\Delta_2}{\Delta_4}\right)^{3/2},\tag{35}$$

$$\eta_{y1}\sim O\left(\frac{1}{V^4}\right),\tag{36}$$

$$\epsilon_{y2}\simeq\frac{m^2n^2\pi^4}{32R^2b^2V^3}\left[(1-2\Delta_2)+\left(\frac{\Delta_2}{\Delta_3}\right)^{3/2}(1-2\Delta_3)\right],\tag{37}$$

$$\eta_{y2}\simeq\frac{m^2n^2\pi^4\Delta_2}{8R^2b^2V^3}\left[(1-2\Delta_2)+\left(\frac{\Delta_2}{\Delta_3}\right)^{1/2}(1-2\Delta_3)\right]\tag{38}$$

where $R=a/b$ is the aspect ratio of the core, and $O(1/V^4)$ means "of the order of $1/V^4$." In writing (31)–(38), only the most significant terms are retained, which are of the order of $1/V^3$. η_{x2} and η_{y1}, which are of the higher order, are therefore neglected. The asymptotic errors in various effective-index solutions are summarized in Table I. The expressions given in Table I show explicitly how the accuracy of the method is affected by the way of applying the method, the mode type, the normalized frequency, the mode orders, the dimensions of the structure, and the relative index differences. It is noted that the errors in the x-method decrease with increasing R faster than those in the y-method ($1/R^3$ versus $1/R^2$). The x-method is particularly accurate for waveguides with $\Delta_2\ll\Delta_4$ (such as strip waveguides). In the weak-guidance case when Δ_2 tends to zero, the results reduce to those already obtained for the scalar modes [20].

$$\eta_{x1}\simeq\frac{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\psi_b^2\psi_a\left\{\frac{d}{dx}\left[\frac{\psi_a}{(n_a^2+n_x^2)}\frac{d(n_a^2+n_x^2)}{dx}\right]-\frac{d}{dx}\left[\frac{\psi_a}{(n_a^2+n_b^2)}\frac{d(n_a^2+n_b^2)}{dx}\right]\right\}dxdy}{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\psi_a^2\psi_b^2\,dxdy}\tag{24}$$

$$\eta_{x2}\simeq\frac{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\psi_b^2\psi_a\left\{\frac{d}{dx}\left[\frac{\psi_a}{n_a^2}\frac{dn_a^2}{dx}\right]-\frac{d}{dx}\left[\frac{\psi_a}{(n_a^2+n_b^2)}\frac{d(n_a^2+n_b^2)}{dx}\right]\right\}dxdy}{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\psi_a^2\psi_b^2\,dxdy}\tag{25}$$

$$\eta_{y1}\simeq\frac{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\psi_a^2\psi_b\left\{\frac{d}{dy}\left[\frac{\psi_b}{n_b^2}\frac{dn_b^2}{dy}\right]-\frac{d}{dy}\left[\frac{\psi_b}{(n_a^2+n_b^2)}\frac{d(n_a^2+n_b^2)}{dy}\right]\right\}dxdy}{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\psi_a^2\psi_b^2\,dxdy}\tag{26}$$

$$\eta_{y2}\simeq\frac{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\psi_a^2\psi_b\left\{\frac{d}{dy}\left[\frac{\psi_b}{(n_b^2+n_y^2)}\frac{d(n_b^2+n_y^2)}{dy}\right]-\frac{d}{dy}\left[\frac{\psi_b}{(n_a^2+n_b^2)}\frac{d(n_a^2+n_b^2)}{dy}\right]\right\}dxdy}{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\psi_a^2\psi_b^2\,dxdy}\tag{27}$$

C. Numerical Verification

The validity of the asymptotic expressions with $\Delta_2 \to 0$ for the scalar modes has been verified numerically [20]. Here the effects of a finite relative index difference are examined.

For the sake of simplicity, a fully embedded waveguide with $\Delta_2 = \Delta_3 = \Delta_4$ is considered. From Table I, the asymptotic errors for this structure are given by

$$\beta_{x1}^2 \simeq \beta_x^2 + \frac{m^2n^2\pi^4}{16R^3b^2V^3}(1-2\Delta_2)(1+4\Delta_2), \quad (39)$$

$$\beta_{x2}^2 \simeq \beta_x^2 + \frac{m^2n^2\pi^4}{16R^2b^2V^3}, \quad (40)$$

$$\beta_{y1}^2 \simeq \beta_y^2 + \frac{m^2n^2\pi^4}{16R^3b^2V^3}, \quad (41)$$

$$\beta_{y2}^2 \simeq \beta_y^2 + \frac{m^2n^2\pi^4}{16R^2b^2V^3}(1-2\Delta_2)(1+4\Delta_2). \quad (42)$$

These equations predict that the x-method and the y-method can produce equal results only for a special value of R, which is given by

$$R = (1-2\Delta_2)(1+4\Delta_2) \quad (43)$$

for the E_{mn}^x mode, and by

$$R = \frac{1}{(1-2\Delta_2)(1+4\Delta_2)} \quad (44)$$

for the E_{mn}^y mode. To check these predictions numerically, the x-method and the y-method are applied to waveguides that cover a range of aspect ratios and the particular aspect ratio for which the results calculated from both methods are equal is determined. This particular aspect ratio as a function of Δ_2 is shown in Fig. 4(a) for the E_{11}^x mode, and in Fig. 4(b) for the E_{11}^y mode, for $V = 10$ and 20. The agreement between the numerical results and the asymptotic values calculated from (43) and (44) is good. As expected, the agreement is better for a larger value of V. To check roughly the range of V values over which the asymptotic theory is valid, the numerically calculated aspect ratio is plotted in Fig. 5 as a function of V for the case $\Delta_2 = 0.25$. The value of the aspect ratio predicted by (43) and (44) is equal to unity for both the E_{mn}^x and E_{mn}^y modes. The numerical results, as shown in Fig. 5, are indeed close to unity except at small values of V. Significant deviation from unity occurs when the value of V is smaller than about 4 for both the E_{11}^x and E_{11}^y modes. It appears that the asymptotic results remain accurate even at moderate values of V.

To test the validity of the asymptotic expressions more thoroughly, the calculation of the birefringence in the waveguide is considered. The birefringences calculated by the x-method and the y-method, denoted by $\delta\beta_1^2$ and $\delta\beta_2^2$, respectively, are given by

$$\delta\beta_1^2 \equiv \beta_{x1}^2 - \beta_{y1}^2 \simeq \delta\beta^2 + \frac{m^2n^2\pi^4}{8R^3b^2V^3}\Delta_2(1-4\Delta_2), \quad (45)$$

$$\delta\beta_2^2 \equiv \beta_{x2}^2 - \beta_{y2}^2 \simeq \delta\beta^2 - \frac{m^2n^2\pi^4}{8R^2b^2V^3}\Delta_2(1-4\Delta_2) \quad (46)$$

where $\delta\beta^2 \equiv \beta_x^2 - \beta_y^2$ is the exact birefringence. Birefringence does not exist in the scalar approximation (since $\delta\beta_1^2 = \delta\beta_2^2 = 0$ when $\Delta_2 = 0$); its presence is a characteristic

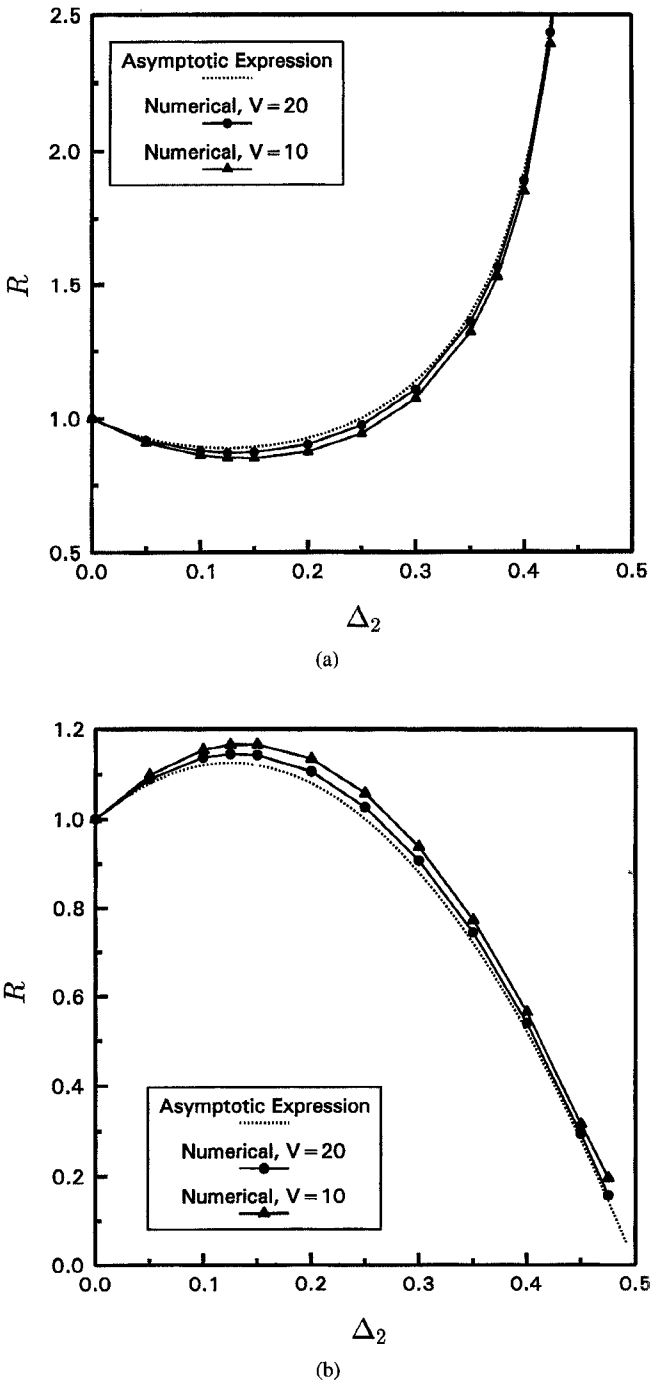


Fig. 4. Aspect ratio R of the rectangular core for which the x-method and the y-method give the same results as a function of the relative index difference Δ_2 calculated for (a) the E_{11}^x mode, and (b) the E_{11}^y mode.

of the vector modes. Obviously, whether the effective-index method overestimates or underestimates the exact birefringence depends on the value of Δ_2. The difference between the normalized birefringences calculated by the x-method and the y-method for a fully embedded waveguide with an aspect ratio of $R = 2$ is plotted as a function of V in Fig. 6(a) for $\Delta_2 = 0.1$ and 0.2, and in Fig. 6(b) for $\Delta_2 = 0.3$ and 0.4. The normalized birefringence is equal to the difference between the normalized propagation constants for the E_{11}^x and

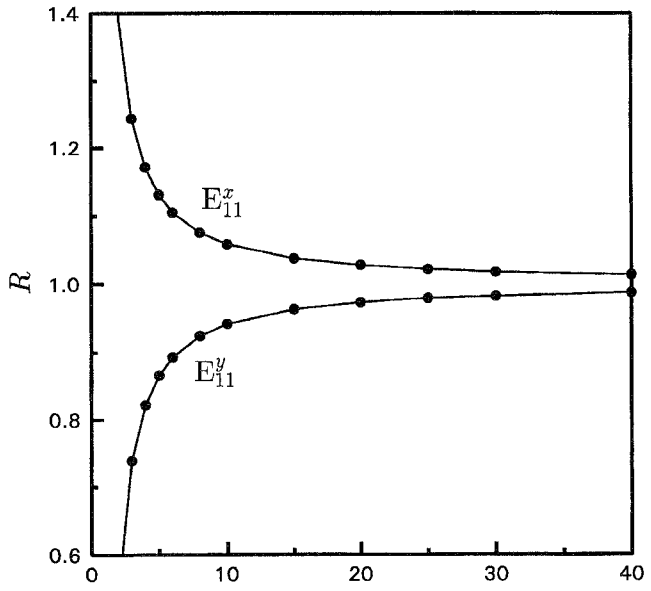


Fig. 5. Aspect ratio R of the rectangular core for which the x-method and the y-method give the same results as a function of V for $\Delta_2 = 0.25$.

E_{11}^y modes, i.e., $\delta P_j \equiv P_{xj} - P_{yj}$, where the normalized propagation constant $P_{ij}(i = x, y$ and $j = 1, 2)$ is defined by

$$P_{ij} = \frac{(\beta_{ij}/k)^2 - n_2^2}{n_1^2 - n_2^2}. \tag{47}$$

It is clear from (45) and (46) that, for $\Delta_2 = 0.1$ and $0.2, \delta P_1 > \delta P_2$, whereas for $\Delta_2 = 0.3$ and $0.4, \delta P_2 > \delta P_1$. As shown in Fig. 6, the asymptotic results agree well with the exact numerical results and are therefore verified. The agreement is less favorable for a large value of Δ_2 because of the omission of the higher-order terms in the asymptotic expressions, which could become significant when the value of Δ_2 is large.

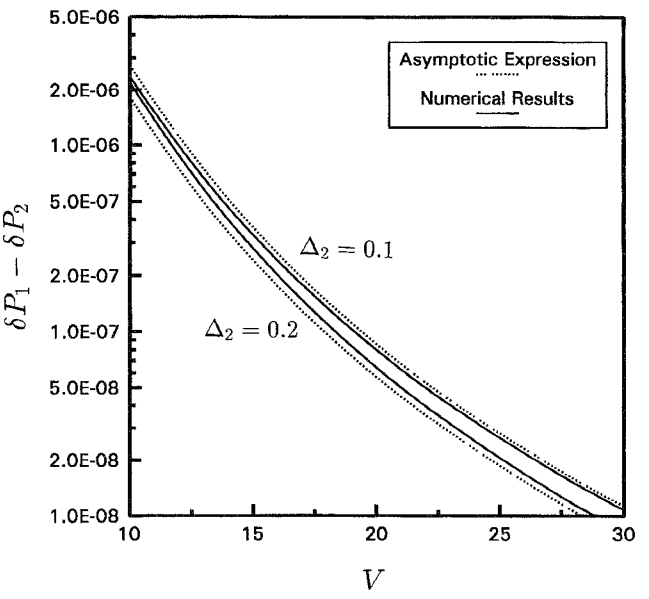
V. The Dual Effective-Index Method

The perturbation formulas (22)–(27) could be evaluated to provide perturbation corrections to the effective-index solution. This could lead to improved results but the evaluation of the integrals involved is tedious. A simpler method to obtain more accurate results is based on cancellation of the errors in various effective-index solutions. As an example, improved results can be obtained for fully embedded waveguides by eliminating the errors that appear in (39)–(42). It is noted from (39)–(42) that the errors in β_{y2}^2 and β_{x2}^2 are R times of those in β_{x1}^2 and β_{y1}^2, respectively. These errors can be eliminated by multiplying the appropriate equations by R or R^2 and subtracting the resultant equations. These procedures lead to

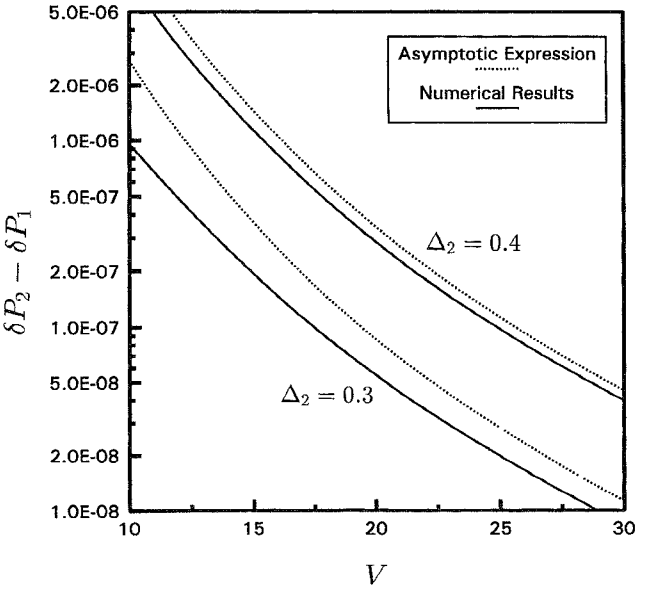
$$\beta_x^2 \simeq \frac{(R^2\beta_{x1}^2 - \beta_{x2}^2) + R(\beta_{y1}^2 - \beta_{y2}^2)}{R^2 - 1}, \tag{48}$$

$$\beta_y^2 \simeq \frac{(R^2\beta_{y1}^2 - \beta_{y2}^2) + R(\beta_{x1}^2 - \beta_{x2}^2)}{R^2 - 1}. \tag{49}$$

More accurate results are therefore expected by combining the four effective-index solutions according to (48) and (49).



(a)



(b)

Fig. 6. The difference between the normalized birefringences δP_1 and δP_2 calculated by the x-method and the y-method as a function of V for a fully embedded rectangular waveguide: (a) $\Delta_2 = 0.1$ and 0.2; (b) $\Delta_2 = 0.3$ and 0 4.

The normalized propagation constants calculated from (48) and (49) for the E_{11}^x and E_{11}^y modes, together with the results obtained from the x-method and the y-method, are presented in Table II for $R = 2$, and in Table III for $R = 1$, where the point-matching results of Goell [23] (from graph reading) are shown as references. In principle, (48) and (49) are not applicable for $R = 1$, but in practice, a waveguide with R slightly larger than unity can be used to approximate the $R = 1$ waveguide. The effective-index results shown in Table III are in fact calculated for a waveguide with $R = 1.000\,01$ to ensure that the errors due to the slight difference in R can be neglected, yet (48) and (49) remains applicable. As shown in Tables II and III,

TABLE II
NORMALIZED PROPAGATION CONSTANTS CALCULATED BY VARIOUS METHODS FOR THE E_{11}^x AND E_{11}^y MODES OF A FULLY EMBEDDED RECTANGULAR WAVEGUIDE WITH $n_1 = 1.5, n_2 = 1.0$, AND $R = 2$

	x-Method		y-Method		Dual Method		Point Matching	
V	P_{x1}	P_{y1}	P_{x2}	P_{y2}	P_x	P_y	P_x	P_y
0.9425	0.1848	0.0756	0.1850	0.1272	0.1504	0.0583	0.122	0.049
1.2566	0.3471	0.2124	0.3539	0.2515	0.3187	0.1948	0.309	0.183
1.5708	0.4851	0.3635	0.4911	0.3852	0.4686	0.3522	0.469	0.344
1.8850	0.5896	0.4915	0.5937	0.5029	0.5807	0.4850	0.580	0.473
2.1991	0.6673	0.5906	0.6700	0.5967	0.6624	0.5868	0.662	0.583
2.5133	0.7256	0.6658	0.7274	0.6693	0.7228	0.6635	0.721	0.659
3.1416	0.8048	0.7675	0.8056	0.7687	0.8037	0.7665	0.803	0.764
3.7699	0.8543	0.8298	0.8547	0.8304	0.8538	0.8293	0.854	0.829

TABLE III
NORMALIZED PROPAGATION CONSTANTS CALCULATED BY VARIOUS METHODS FOR THE E_{11}^x AND E_{11}^y MODES OF A FULLY EMBEDDED RECTANGULAR WAVEGUIDE WITH $n_1 = 1.5, n_2 = 1.0$, AND $R = 1$

	x-Method		y-Method		Dual Method	Point Matching
V	P_{x1}	P_{y1}	P_{x2}	P_{y2}	$P_x = P_y$	$P_x = P_y$
1.2566	0.1493	0.1108	0.1108	0.1493	0.0882	0.062
1.5708	0.2596	0.2339	0.2339	0.2596	0.2012	0.184
1.8850	0.3726	0.3586	0.3586	0.3726	0.3315	0.323
2.1991	0.4731	0.4659	0.4659	0.4731	0.4472	0.440
2.5133	0.5568	0.5531	0.5531	0.5568	0.5406	0.533
3.1416	0.6792	0.6781	0.6781	0.6792	0.6724	0.669
3.7699	0.7593	0.7589	0.7589	0.7593	0.7561	0.755
4.7124	0.8341	0.8340	0.8340	0.8341	0.8328	0.833

the results obtained from (48) and (49) are significantly more accurate than the original effective-index solutions. It can be observed from the results in Tables II and III that, in the case of $R = 2$, the x-method is more accurate than the y-method, whereas in the case of $R = 1, P_{y1} = P_{x2}$ is more accurate than $P_{x1} = P_{y2}$, as predicted by the asymptotic equations (39)–(42).

There are other ways of eliminating the errors in the effective-index solutions and such error-elimination procedures have been named the dual effective-index method [7], [11], and [12], as results from both the x-method and the y-method are required. Equations (48) and (49) can therefore be regarded as a version of the dual effective-index method. To eliminate the errors, it is only necessary to know the relations of the errors in different solutions. In the early work [7], such relations were found heuristically without knowing the detailed expressions for the errors. The expressions derived in the present paper (Table I) actually provide a more rigorous proof of those relations that were used preivously [7]. All the numerical examples given so far [7], [11], [12], [20], including those presented in this paper, indicate that, although the error-elimination procedure (the dual effective-index method) is exact only at $V \to +\infty$, it is effective even down to the close-to-cutoff region.

VI. CONCLUSION

A detailed and, in many ways, quantitative analysis of the effective-index method applied to the vector modes of rectangular-core waveguides has been presented. The contributions of this study can be summarized as follows:

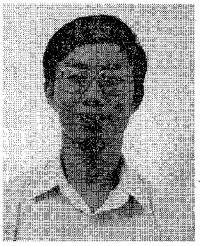
1) The one-dimensional wave equations that describe the effective-index method for analyzing the vector modes of rectangular-core waveguides are presented (Section III). It is found that solving these equations is equivalent to solving the reduced vector-wave equation approximately for a mathematically separable structure, which is different from the original structure in certain cladding regions. The approximations involved in the method are clarified at the most fundamental level—the equation level.

2) With the findings outlined in 1), general perturbation formulas (22)–(27) are presented to describe the errors in the method.
3) The perturbation formulas are evaluated for the special case $V \rightarrow +\infty$. The resultant asymptotic expressions for the errors, summarized in Table I, can describe explicitly the performance of the effective-index method in great detail. These expressions quantify the errors in the method to a high accuracy at large values of V. At moderate values of V, they still provide useful qualitative information about the performance of the method.
4) A simple procedure (the dual effective-index method) is proposed to obtain accurate results by combining various effective-index solutions. This procedure is found to be effective even if the mode is not far from cutoff.

The approach presented in this paper should be readily extended to more complicated rectangular structures, such as directional couplers and arrays, in a way similar to that reported in [9] for the scalar modes.

REFERENCES

[1] R. M. Knox and P. P. Toulios, "Integrated circuits for the millimeter through optical frequency range," in *Proceedings, Symposium on Submillimeter Waves*. Brooklyn: Polytechnic Press, 1970, pp. 497–516.
[2] E. A. J. Marcatili, "Dielectric rectangular waveguide and directional coupler for integrated optics," *Bell Syst. Tech. J.*, vol. 48, pp. 2071–2102, 1969.
[3] H. Furuta, H. Noda, and A. Ihaya, "Novel optical waveguide for integrated optics," *Appl. Opt.*, vol. 13, pp. 322–326, 1974.
[4] W. V. McLevige, T. Itoh, and R. Mittra, "New waveguide structures for millimeter-wave and optical integrated circuits," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-23, pp. 788–794, 1975.
[5] T. Itoh, "Inverted strip dielectric waveguide for millimeter-wave integrated circuits," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-24, pp. 821–828, 1976.
[6] A. A. Oliner, S. T. Peng, T. I. Hsu, and A. Sanchez, "Guidance and leakage properties of a class of open dielectric waveguides: Part II—New physical effects," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-29, pp. 855–870, 1981.
[7] K. S. Chiang, "Dual effective-index method for the analysis of rectangular dielectric waveguides," *Appl. Opt.*, vol. 25, pp. 2169–2174, 1986.
[8] ——, "Effects of cores in fused tapered single-mode fiber couplers," *Opt. Lett.*, vol. 12, pp. 431–433, 1987.
[9] ——, "Effective-index method for the analysis of optical waveguide couplers and arrays: An asymptotic theory," *J. Lightwave Technol.*, vol. 9, pp. 62–72, 1991.
[10] G. B. Hocker and W. K. Burns, "Mode dispersion in diffused channel waveguides by the effective index method," *Appl. Opt.*, vol. 16, pp. 113–118, 1977.
[11] K. S. Chiang, "Analysis of optical fibers by the effective-index method," *Appl. Opt.*, vol. 25, pp. 348–354, 1986.
[12] ——, "Geometric birefringence in a class of step-index fiber," *J. Lightwave Technol.*, vol. LT-5, pp. 737–744, 1987.
[13] K. Van de Velde, H. Thienpont, and R. Van Geen, "Extending the effective index method for arbitrarily shaped inhomogeneous optical waveguides," *J. Lightwave Technol.*, vol. 6, pp. 1153–1159, 1988.
[14] K. S. Chiang, "Analysis of fused couplers by the effective-index method," *Electron. Lett.*, vol. 22, pp. 1221–1222, 1986.
[15] ——, "Stress-induced birefringence fibers designed for single-polarization single-mode operation," *J. Lightwave Technol.*, vol. 7, pp. 436–441, 1989.
[16] K. S. Chiang and R. A. Sammut, "Effective-index method for spatial solitons in planar waveguides with Kerr-type nonlinearity," *J. Opt. Soc. Am. B*, vol. 10, pp. 704–708, 1993.
[17] K. S. Chiang, "Review of numerical and approximate methods for the modal analysis of general optical dielectric waveguides," *Opt. Quantum Electron.*, vol. 26, pp. S113–S134, 1994.
[18] S. T. Peng and A. A. Oliner, "Guidance and leakage properties of a class of open dielectric waveguides: Part I—Mathematical formulations," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-29, pp. 845–855, 1981.
[19] A. Kumur, D. F. Clark, and B. Culshaw, "Explanation of errors inherent in the effective-index method for analyzing rectangular-core waveguides," *Opt. Lett.*, vol. 13, pp. 1129–1131, 1988.
[20] K. S. Chiang, "Performance of the effective-index method for the analysis of dielectric waveguides," *Opt. Lett.*, vol. 16, pp. 714–716, 1991.
[21] A. W. Snyder and J. D. Love, *Optical Waveguide Theory*. London: Chapman & Hall, 1983, p. 610.
[22] A. W. Snyder and X.-H. Zheng, "Optical fibers of arbitrary cross sections," *J. Opt. Soc. Am. A*, vol. 3, pp. 600–609, 1986.
[23] J. E. Goell, "A circular-harmonic computer analysis of rectangular dielectric waveguides," *Bell Syst. Tech. J.*, vol. 48, pp. 2133–2160, 1969.



Kin Seng Chiang (M'94) was born in Guangdong, China, in 1957. He received the B.E. (Hon. I) and the Ph.D. degrees in electrical engineering from the University of New South Wales, Sydney, Australia, in 1982 and 1986, respectively.

In 1986, he was with the Department of Mathematics of the Australian Defence Force Academy in Canberra. From 1986 to 1993, he was with the Division of Applied Physics of the Commonwealth Scientific and Industrial Research Organization (CSIRO), Sydney. From 1987 to 1988, he received a Japanese Government Research Award and was with the Electrotechnical Laboratory in Tsukuba City, Japan. From 1992 to 1993, he was on secondment to the Optical Fiber Technology Centre (OFTC) of the University of Sydney. In August 1993, he joined the Department of Electronic Engineering of the City University of Hong Kong (previously the City Polytechnic of Hong Kong), where he is currently Associate Professor. He has published over 80 technical papers on optical waveguide theory and characterization, numerical methods, fiber devices and sensors, and nonlinear guided-wave optics.

Dr. Chiang is a member of the Optical Society of America, the International Society for Optical Engineering (SPIE), and the Australian Optical Society.