An Improved Formulation of an Optimizing Rayleigh–Ritz Technique for Closed Dielectric Waveguide Analysis

Brian Young, *Member, IEEE*

Abstract—**Previous work in the application of the Rayleigh–Ritz method to the analysis of closed dielectric waveguides has shown that if the modes of the homogeneous rectangular waveguide are used to model the modes of an inhomogeneous rectangular waveguide, then it is numerically advantageous to use an optimized value for the permittivity of the homogenous waveguide's dielectric filling. The paper reformulates this work to use a complete set of basis functions. It is shown that use of the E_z/H_z formulation to describe the modes of the homogeneous rectangular waveguide leads to a relative convergence phenomena as well as to incorrect loss calculations. The paper reformulates the method using an $E_x/E_y/H_x/H_y$ description of the homogeneous modes. The new formulation is validated for step-index waveguides through non-perturbational calculations of the propagation and attenuation constants of the round step-index dielectric waveguide. Comparison with the direct eigenvalue solution shows excellent agreement for the dominant and three higher-order modes. The new formulation is validated for graded-index waveguides through calculation of the dispersion curves for three modes of a Gaussian–Gaussian graded-index channel waveguide. Comparison of the results with two other methods shows excellent agreement.**

I. Introduction

In a recent paper [1], this author detailed a modified variational method designed to analyze arbitrary closed dielectric waveguides. The method uses the modes of the homogeneously-filled rectangular waveguide (HRW) to model the modes of the inhomogeneously-filled rectangular waveguide (IRW) in a variational expression for the complex propagation constant. Following the classic Rayleigh–Ritz technique, the variational expression is then minimized with respect to the amplitude of each HRW mode. The innovation described in the paper is to allow the dielectric constant of the HRW filling to remain variable so that the variational expression can also be minimized with respect to it. This additional optimization eliminates the difficulties with the Rayleigh–Ritz technique that had been previously observed [2]–[7], namely slow convergence.

Manuscript received December 20, 1990; revised May 28, 1991

The author is with Hughes Aircraft Company, Radar Systems Group, 2000 East Imperial Highway, El Segundo, CA 90245.

IEEE Log Number 9102817.

The method of [1] uses the E_z/H_z formulation to derive the HRW modes. This selection results in the use of a basis set that is not complete, as discussed in [8]. It is the purpose of this paper to reformulate the method to use a complete set of basis functions and to numerically prove the new formulation. The formulation based on the incomplete basis set is also discussed as it exhibits interesting numerical behavior that supports the new formulation.

II. Discussion on Modal Expansions

As it seems that there is some confusion on the completeness of the modes of the HRW, it is worth a short discussion even though the theory has already been detailed [8]. This discussion will concentrate on an intuitive explanation.

First, consider the physical scattering problem shown in Fig. 1, where an HWR sees a discontinuity into an IRW. If a single mode or multiple modes are incident onto the interface from either waveguide, the interface scatters energy into all higher-order modes in both waveguides so that the total fields in each waveguide, given by the superposition of all modes in that waveguide, satisfy the boundary conditions at the interface. Therefore, the HRW modes are complete in that they can exactly match any field configuration produced by any IRW. In short, an infinite sum of HRW modes can exactly match an infinite sum of IRW modes.

The method presented in [1] attempts to match a *single* IRW mode by an infinite sum of HRW modes. This situation is not a physical problem as was the one above; therefore, there is no guarantee that the infinite sum of HRW modes can in fact match the desired fields. In general for this case, the HRW modes are not complete, so any formulation using HRW modes to model IRW modes will result in some error.

The incompleteness of the E_z/H_z formulation can be more explicitly explored with an alternative, more specific, argument. Following [9], the E_x, E_y, H_x, and H_y field components can be written in terms of E_z and H_z

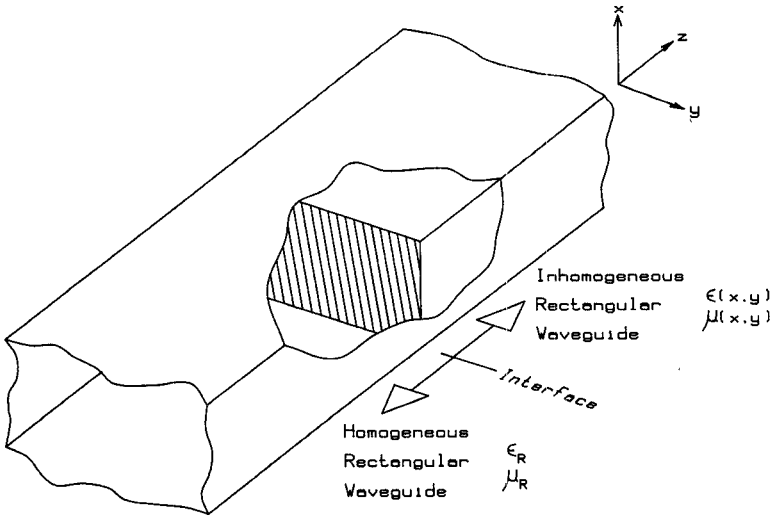


Fig. 1. On the problem of scattering in a rectangular waveguide on transition from a homogeneously-filled section to an inhomogeneously-filled section.

according to the following expressions:

$$E_x = -\frac{j}{k^2-\beta^2}\left(\beta\frac{\partial E_z}{\partial x}+\omega\mu\frac{\partial H_z}{\partial y}\right) \tag{1}$$

$$E_y = \frac{j}{k^2-\beta^2}\left(-\beta\frac{\partial E_z}{\partial y}+\omega\mu\frac{\partial H_z}{\partial x}\right) \tag{2}$$

$$H_x = \frac{j}{k^2-\beta^2}\left(\omega\epsilon\frac{\partial E_z}{\partial y}-\beta\frac{\partial H_z}{\partial x}\right) \tag{3}$$

$$H_y = -\frac{j}{k^2-\beta^2}\left(\omega\epsilon\frac{\partial E_z}{\partial x}+\beta\frac{\partial H_z}{\partial y}\right). \tag{4}$$

Concentrating on E_x as all components have similar behavior, it can be seen from (1) that if β is real, as it is for a propagating mode in a lossless HRW, then E_x and E_z are 90° out of phase. If β is imaginary, as it is for a cutoff mode in lossless HRW, then E_x and E_z are in phase (or 180° out). The phase relationship for E_x and H_z is fixed at 90° out of phase. With these fixed relationships, E_x cannot take on an arbitrary phase with respect to E_z and H_z as would be required to model an arbitrary IRW. Therefore, the modes of the HRW are not complete and cannot be used to exactly model a *single* mode of the IRW. The fixed phase relationships can be alleviated somewhat by using a lossy HRW, for which β is complex. Later numerical examples demonstrate this effect.

III. Improved Formulation

The discussion in Section II demonstrates that the E_z/H_z formulation is not general enough to model an arbitrary IRW due to the fixed phase relationships between field components. However, it is still desirable to use the fields derived from a homogeneous structure to model the IRW. To alleviate this problem, a formulation is needed that allows for arbitrary phase relationships between components. As indicated in [8], it is not necessary for each field component to have a separate eigenfunction expansion. From Maxwell's equations with no sources we have

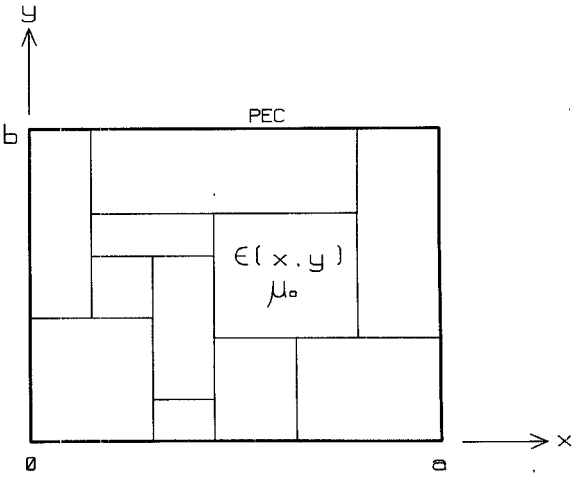


Fig. 2. Generic inhomogeneously-filled rectangular waveguide (IRW) with rectangular subdivision in ϵ and homogeneous in μ.

$$E_z = \frac{1}{j\omega\epsilon}\left(\frac{\partial H_y}{\partial x}-\frac{\partial H_x}{\partial y}\right) \tag{5}$$

$$H_z = -\frac{1}{j\omega\mu}\left(\frac{\partial E_y}{\partial x}-\frac{\partial E_x}{\partial y}\right), \tag{6}$$

so a solution with E_x, E_y, H_x, and H_y as variables would allow arbitrary phase relationships to exist between components as ϵ and μ are free to modify H_z and E_z. Another aspect of this approach is that the Maxwell equation $\nabla\cdot\boldsymbol{H}=0$ is explicitly satisfied so that spurious modes will not be generated.

For the homogeneous structure, all field components must satisfy the Helmholtz equations

$$\nabla^2\boldsymbol{E}+k_R^2\boldsymbol{E}=0 \tag{7}$$

or

$$\nabla^2 \boldsymbol{H} + k_R^2 \boldsymbol{H} = 0, \tag{8}$$

where $k_R = \omega\sqrt{\mu_R \epsilon_R}$ and μ_R and ϵ_R are complex constants. The generic IRW structure for which a solution is sought is shown in Fig. 2, where an inhomogeneous dielectric with $\mu = \mu_o$ fills a rectangular perfectly conducting box. The HRW with the same outer box is chosen to form the basis set, and with the PEC boundary conditions, the general solutions to (7)–(8) for all but the z-components are

$$E_x = \sum_{m=0}^{\infty}\sum_{n=1}^{\infty} E_{xmn} \cos\left(\frac{m\pi x}{a}\right)\sin\left(\frac{n\pi y}{b}\right) \tag{9}$$

$$E_y = \sum_{m=1}^{\infty}\sum_{n=0}^{\infty} E_{ymn} \sin\left(\frac{m\pi x}{a}\right)\cos\left(\frac{n\pi y}{b}\right) \tag{10}$$

$$H_x = \sum_{m=1}^{\infty}\sum_{n=0}^{\infty} H_{xmn} \sin\left(\frac{m\pi x}{a}\right)\cos\left(\frac{n\pi y}{b}\right) \tag{11}$$

$$H_y = \sum_{m=0}^{\infty}\sum_{n=1}^{\infty} H_{ymn} \cos\left(\frac{m\pi x}{a}\right)\sin\left(\frac{n\pi y}{b}\right), \tag{12}$$

where E_{xmn}, E_{ymn}, H_{xmn}, and H_{ymn} are the unknown expansion coefficients. Substituting (9)–(12) into (5)–(6) yields

$$E_z = \frac{1}{j\omega\epsilon_R}\left(\sum_{m=1}^{\infty}\sum_{n=0}^{\infty} H_{xmn} \sin\left(\frac{m\pi x}{a}\right)\frac{n\pi}{b}\sin\left(\frac{n\pi y}{b}\right) - \sum_{m=0}^{\infty}\sum_{n=1}^{\infty} H_{ymn}\frac{m\pi}{a}\sin\left(\frac{m\pi x}{a}\right)\sin\left(\frac{n\pi y}{b}\right)\right) \tag{13}$$

and

$$H_z = \frac{1}{j\omega\mu_R}\left(\sum_{m=0}^{\infty}\sum_{n=1}^{\infty} E_{xmn} \cos\left(\frac{m\pi x}{a}\right)\frac{n\pi}{b}\cos\left(\frac{n\pi y}{b}\right) - \sum_{m=1}^{\infty}\sum_{n=0}^{\infty} E_{ymn}\frac{m\pi}{a}\cos\left(\frac{m\pi x}{a}\right)\cos\left(\frac{n\pi y}{b}\right)\right). \tag{14}$$

A suitable variational expression for the propagation constant is given by [10] as

$$\beta = \frac{\iint_S \left(\omega\epsilon\hat{\boldsymbol{E}}^+\cdot\hat{\boldsymbol{E}}^- - \omega\mu\hat{\boldsymbol{H}}^+\cdot\hat{\boldsymbol{H}}^- + j\hat{\boldsymbol{H}}^-\cdot\nabla\times\hat{\boldsymbol{E}}^+ + j\hat{\boldsymbol{E}}^+\cdot\nabla\times\hat{\boldsymbol{H}}^-\right) ds}{2\iint_S \hat{\boldsymbol{E}}_t\times\hat{\boldsymbol{H}}_t\cdot\boldsymbol{u}_z\, ds} \tag{15}$$

where

$$\boldsymbol{E}^+ = \hat{\boldsymbol{E}}^+(x,y)e^{-j\beta z} = \left(\hat{\boldsymbol{E}}_t + \boldsymbol{u}_z\hat{E}_z\right)e^{-j\beta z}$$

$$\boldsymbol{H}^+ = \hat{\boldsymbol{H}}^+(x,y)e^{-j\beta z} = \left(\hat{\boldsymbol{H}}_t + \boldsymbol{u}_z\hat{H}_z\right)e^{-j\beta z}$$

$$\boldsymbol{E}^- = \hat{\boldsymbol{E}}^-(x,y)e^{j\beta z} = \left(\hat{\boldsymbol{E}}_t - \boldsymbol{u}_z\hat{E}_z\right)e^{j\beta z}$$

$$\boldsymbol{H}^- = \hat{\boldsymbol{H}}^-(x,y)e^{j\beta z} = \left(-\hat{\boldsymbol{H}}_t + \boldsymbol{u}_z\hat{H}_z\right)e^{j\beta z}.$$

This mixed-field formula requires only that $\boldsymbol{n}\times\boldsymbol{E} = 0$ on the contour of the waveguide, a requirement that the waveguide cross section, S, be bound by a perfect electric conductor (PEC). The formula remains stationary for the lossy case, for which $j\beta$ is replaced by $\gamma = \alpha + j\beta$. The structure is defined through inhomogeneous ϵ and μ, both of which can be complex to include loss.

The general field expressions in (9)–(14) satisfy the requirements for use in (15), so the unknown fields in (15) can be approximated by truncated expansions based on (9)–(14). A matrix equation is formed by directly substituting (9)–(14) into (15) and applying the Rayleigh–Ritz technique, where we set $\partial\beta/\partial A_{mn} = 0\ \forall mn$ for A_{mn} equal to E_{xmn}, E_{ymn}, H_{xmn}, and H_{ymn}, a process that entails considerable algebra. To alleviate the effort and to minimize the chance of error, the derivation was programmed and solved using a symbolic manipulator [11]. As the intermediate results are not even available for inspection, only the final results are presented, and they appear in the Appendix as (20)–(23). As indicated in Fig. 2, ϵ for the IRW may be restricted to homogeneous rectangular regions. This restriction allows all integrations to be performed in closed form.

It is convenient for discussion to recast the equations in matrix form. Defining the vectors:

$$\overline{E}_x = \{E_{x01}, E_{x02}, \ldots, E_{xmn}\},\quad m = 0,1,\ldots,L_x - 1,\quad n = 1,2,\ldots,L_y$$

$$\overline{E}_y = \{E_{y10}, E_{y11}, \ldots, E_{ymn}\},\quad m = 1,2,\ldots,L_x,\quad n = 0,1,\ldots,L_y - 1$$

$$\overline{H}_x = \{H_{x10}, H_{x11}, \ldots, H_{xmn}\},\quad m = 1,2,\ldots,L_x,\quad n = 0,1,\ldots,L_y - 1$$

$$\overline{H}_y = \{H_{y01}, H_{y02}, \ldots, H_{ymn}\},\quad m = 0,1,\ldots,L_x - 1,\quad n = 1,2,\ldots,L_y$$

and then

$$\overline{E} = \left\{\overline{E}_x, \overline{E}_y\right\}$$

$$\overline{H} = \left\{\overline{H}_y, \overline{H}_x\right\},$$

then (20) and (21) can be recast as

$$A\overline{E} = \beta\overline{H} \tag{16}$$

with (23) and (22) taking the form

$$B\overline{H} = \beta\overline{E}. \tag{17}$$

Note that either $\overline{E}$ or $\overline{H}$ can be eliminated to cut the order of the matrix solution in half. Eliminating $\overline{H}$ yields

$$BA\overline{E} = \beta^2\overline{E}. \tag{18}$$

Equation (18) is in the form of the standard eigenvalue equation, an advantageous form since each solution represents one mode of the IRW (so all modes are available with one calculation) and since numerically stable algorithms exist for its solution.

It is interesting to compare the result here given by (18) with the E_z/H_z formulation from [1]. Both result in a fully-dense standard eigenvalue equation in two sets of

variables, even though the formulation here is originally based on four sets, so the basic computational speed is identical. However, in [1], the eigenvalues of the matrix equation are given by β, whereas in the formulation given by (18), the eigenvalues are given by β^2. So the computation time spent computing eigenvalues for waves traveling in the negative z-direction, a direction not supported by the underlying variational expression (15), is avoided in the current formulation.

At this point, the dielectric filling of the HRW is still undetermined. As discussed in [1], it is necessary to choose ϵ_R and μ_R, the material parameters of the HRW, appropriately so that the variational expression is minimized, as was also done for E_{xmn}, E_{ymn}, H_{xmn}, and H_{ymn}. Hence, it is required that $\partial\beta/\partial\epsilon_R = 0$ and $\partial\beta/\partial\mu_R = 0$. In this work, $\mu(x, y) = \mu_o$ for the IRW so it is sufficient to set $\mu_R = \mu_o$, leaving just one additional constraint. The required derivative was also programmed and solved using a symbolic manipulator, and the result appears in the Appendix as (24).

The final algorithm is then as follows. An initial guess is made for ϵ_R and (18) is solved. One eigenvalue is chosen, usually the one with the largest real part to get the dominant mode, and its corresponding eigenvector is set equal to $\bar{E}$. Equation (16) is used to compute $\bar{H}$. With these $\bar{E}$ and $\bar{H}$, (24) is used to compute the derivative $\partial\beta/\partial\epsilon_R$. Muller's method is used to calculate an updated ϵ_R and the process is repeated until $\partial\beta/\partial\epsilon_R = 0$. It is important to note that because of the variational nature of the problem, only one root exists.

IV. Some Implementation Notes

Various techniques are used to speed execution without affecting the computational accuracy. Those discussed below are the utilization of symmetry, selective mode inclusion, and precomputation. Symmetry is easily implemented by including only even or odd terms in the modal basis set. For symmetry in one direction, the number of modes can be reduced by two resulting in a four times speed improvement. Symmetry in both the x- and the y-directions reduces the number of modes by four resulting in a sixteen times speed improvement.

Considerable speed improvement can also be obtained by emphasizing lower- over higher-order modes. The specific implementation used here is to limit the modes to those with indices that satisfy $m + n \leqslant L + 1$, where $L = \max(L_x, L_y)$, resulting in a "triangular" set as opposed to the "rectangular" set implied in the formulation given in the Appendix.

There are two opportunities for precomputation. First, it is noted that the matrix A in (18) does not depend on ϵ_R, so it does not need to be recomputed for each iteration in optimizing ϵ_R. The second opportunity lies with the fact that all of the integrations required are dependent only on the physical layout of the structure. These can be precomputed for dramatic improvement in speed.

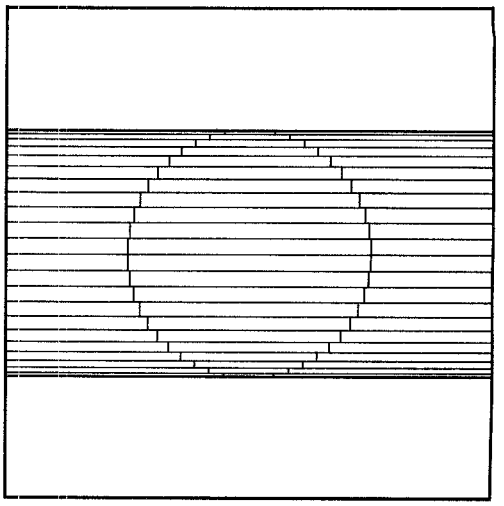


Fig. 3. Staircase approximation for the round step-index dielectric waveguide based on a rectangular subdivision using 74 rectangles. Numerical results use 122 rectangles for the staircase approximation.

V. Results

For accurate numerical validation, the open round step-index dielectric waveguide is used as the basis for comparison due to its relatively simple and unambiguous numerical analysis. Analysis of the waveguide by the variational methods described here and in [1] involves two approximations. The first is the presence of the PEC box enclosing the structure, and the second is that the round waveguide must be modeled with a staircase approximation. The perturbation by the box is minimized by placing it a sufficient distance from the dielectric waveguide so that its effect is negligible. The box compresses the fields slightly, resulting in more field residing in the dielectric, and causes an increase in the propagation constant. Larger increases occur at lower frequencies and for higher-order modes as the exponential tail extends over a greater distance for these cases resulting in more compression.

A sample staircase approximation for the round waveguide is shown in Fig. 3, where 74 rectangles are used. For the numerical results below, 122 rectangles are used (making a plot unreadable) resulting in a 0.0497% error in the modeled dielectric area compared to the actual area. In the numerical analysis, the basis set of modes do not include terms with sufficiently high spatial-frequency content to resolve the jagged boundary, so the close match in areas results in negligible error for the propagation and attenuation constants. It is worth noting that the matrix size is independent of the number of rectangles used, so the computation time is only weakly linked to the number of rectangles.

Shown in the inset in Fig. 4 is the structure analyzed to obtain the following results. For the variational methods, the staircase approximation discussed above is used. For comparison, the open structure is analyzed using the exact eigenvalue equation as given, for example, in [12]. Since the loss tangent is included in each analysis, both perform non-perturbational loss calculations.

TABLE I
TOTAL NUMBER OF MODES FOR CONVERGENCE ANALYSIS

L_x or L_y	#HRW Modes
4	6
6	12
8	20
10	30
12	42
14	56
16	72
18	90
20	110

Propagation and attenuation constant convergence plots for the HE_{11} mode at 130 GHz for both the E_z/H_z and the $E_x/E_y/H_x/H_y$ formulations are shown in Fig. 4. Since the structure is square, equal numbers of basis functions are included for each direction. The total number of modes, hence the matrix size, utilized in the solutions are listed in Table I since they cannot be easily calculated from the mode limits due to selective mode inclusion. Symmetry in each direction is fully exploited.

The propagation constant results in Fig. 4(a) show rapid convergence to a stable result for both techniques. Comparison with the open structure shows the expected slightly greater value for the propagation constant for the closed structure. The E_z/H_z formulation incurs about twice the error as the $E_x/E_y/H_x/H_y$ formulation, so although the actual change due to the PEC box is unknown, it appears that the E_z/H_z formulation incurs an offset in excess of the actual amount. This excess leads to the conclusion that the E_z/H_z formulation as implemented in [1] suffers from relative convergence [13].

The attenuation constant results in Fig. 4(b) demonstrate the underlying behavior of the E_z/H_z formulation. In contrast to the $E_x/E_y/H_x/H_y$ formulation, which achieves good convergence to an accurate and stable result, the E_z/H_z formulation demonstrates good convergence up to $L_x = L_y = 10$ where the solution takes a radical turn. With $L_x = L_y = 12$, cutoff modes of the HRW are introduced into the basis set for the first time. As discussed in Section II, cutoff HRW modes have different phase relationships between components than do the propagating modes. Numerically, the E_z/H_z formulation adjusts for the poor phase match by shifting ϵ_R to describe a very lossy dielectric. This effect is shown in Fig. 5, where the real part is not much affected by the inclusion of cutoff HRW modes whereas the imaginary part jumps by several orders of magnitude (note the scale change in Fig. 5(b)). The complex propagation constant for the HRW modes resulting from the lossy material induces phase relationships to compensate for the incompleteness of the basis functions. In contrast, the $E_x/E_y/H_x/H_y$ formulation demonstrates smooth convergence in ϵ_R to a stable result.

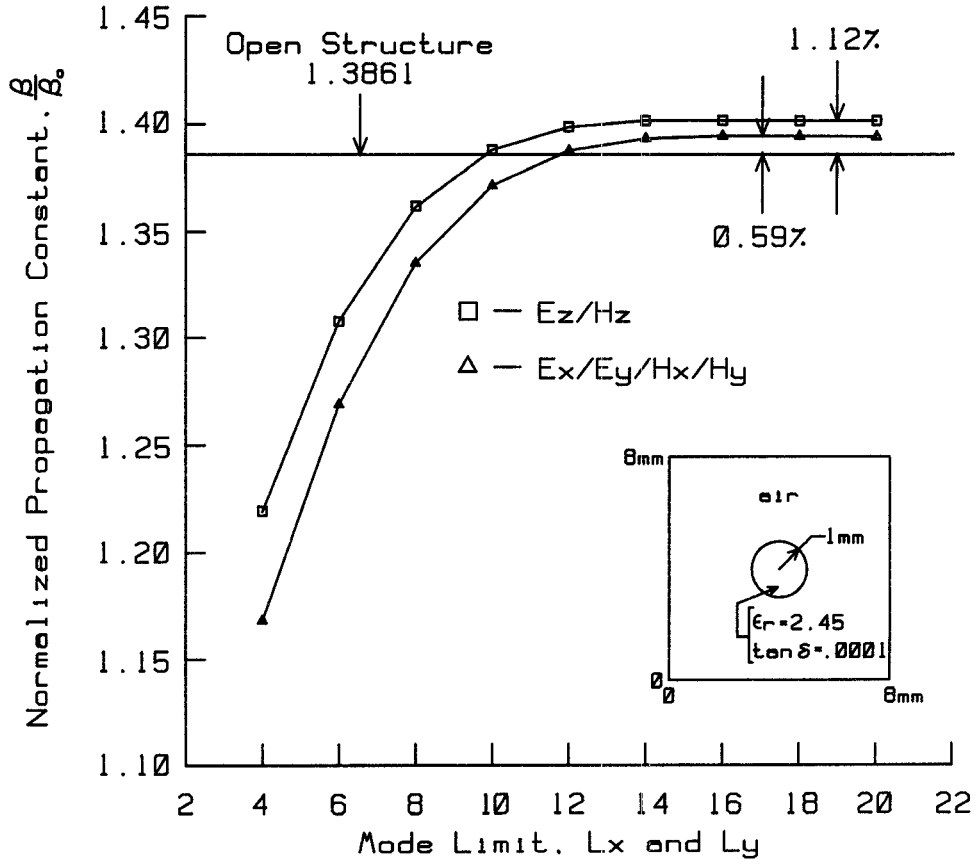
The problems demonstrated in Fig. 4(b) in calculating the attenuation constant using the E_z/H_z formulation contrasts sharply with the success in attenuation constant calculation in [1]. There, the attenuation of an image line with $\epsilon_r = 1$ and variable $\tan\delta$ is used to validate the attenuation calculation with very good success. The success is due to the fact that this particular IRW is very similar to the HRW for the range of $\tan\delta$'s employed, so the error introduced by the incomplete basis set is very small. This over-simplified test case leads to a false validation of the method. The use of a realistic waveguide for all validations below for the $E_x/E_y/H_x/H_y$ formulation yields a solid verification of the method.

The $E_x/E_y/H_x/H_y$ formulation is fully verified with the results in Fig. 6. The convergence curves in Fig. 4 show that convergence is reached for $L_x = L_y = 14$ at 130 GHz for the HE_{11} mode. Convergence depends on the ability of the basis functions to model the spatial distribution of the field structure. So convergence is also achieved for frequencies and modes with similar spatial distributions. For the results in Fig. 6, all results are calculated with $L_x = L_y = 20$, with one exception discussed below. This selection ensures convergent results for all frequencies covered and for all modes with a similar field structure. The exception mentioned is for the HE_{11} mode for 80–110 GHz. The lower frequencies induce a greater spread in the fields, resulting in less need for high spatial-frequency content to model them. These results were computed with $L_x = L_y = 18$ to avoid the numerical difficulties induced by including the unnecessary high spatial-frequency terms. This effect is evident in the numerical data used to make Fig. 4, but the effect is not large enough by $L_x = L_y = 20$ to be visible in the plot.

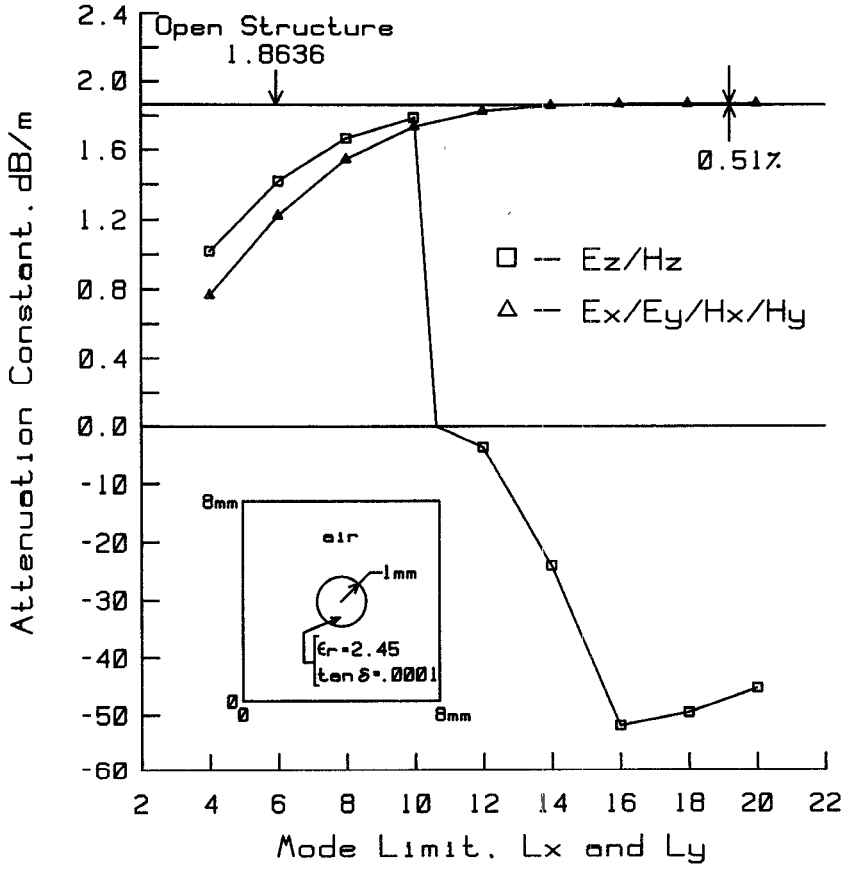
The propagation constant results of Fig. 6(a) show accurate results for the HE_{11} mode with diminishing agreement with reduced frequency as the PEC box has a greater effect. The TE_{01} mode shows good agreement regardless of frequency. Accuracy similar to the HE_{11} mode should be achieved due to the similar spatial extent. This is observed at higher frequencies as expected, but the decrease in accuracy due to the PEC box is not observed as expected. The TM_{01} mode is not shown as it is nearly degenerate with the TE_{01} mode and would unnecessarily confuse the figure. Accuracy and trends are similar to the HE_{21} mode. The HE_{21} mode shows reduced agreement, but this is expected as it has roughly twice the spatial-frequency content as the HE_{11} mode. To achieve results for the HE_{21} mode that are convergent to the same level as the HE_{11} mode, it would be necessary to require at least $L_x = L_y = 28$ (twice the HE_{11} value of 14).

The attenuation constant results shown in Fig. 6(b) show excellent agreement at all frequencies. Note that the attenuations are similar for each mode, so an offset scale is used to separate the curves. Again, similar results were obtained for the TM_{01} mode.

The method demonstrated excellent numerical stability. In no cases were spurious modes observed. The complex root search for ϵ_R typically required six to eight iterations, of which three were required to begin the root search. In all cases, over ten orders of magnitude reduction in the derivative $\partial\beta/\partial\epsilon_R$ were achieved in the root search. In no

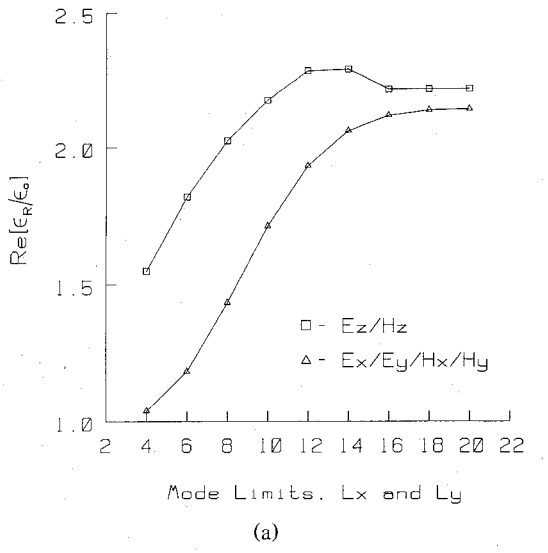


(a)

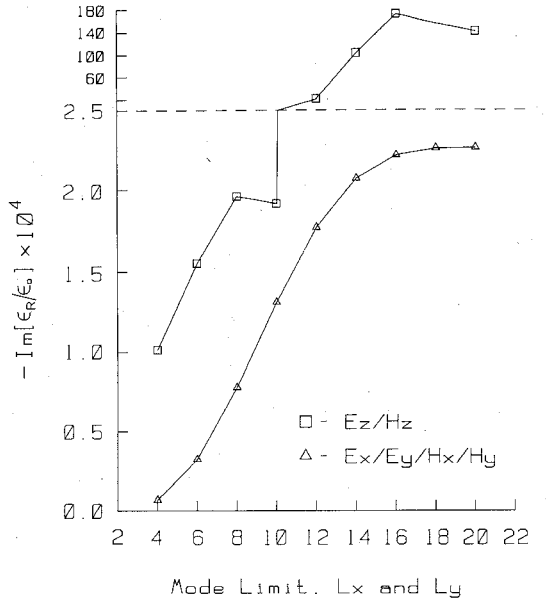


(b)

Fig. 4. Propagation and attenuation constant convergence curves for the HE_{11} mode at 130 GHz. (a) Normalized propagation constant. (b) Attenuation constant in dB/m. Note the scale change.

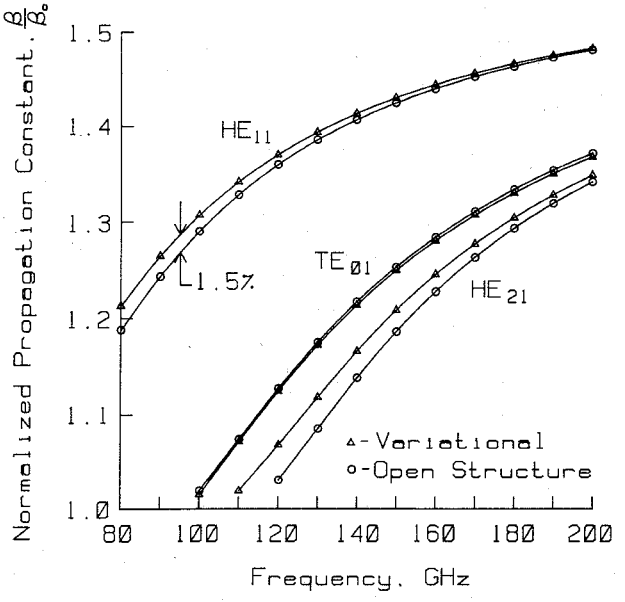


(a)

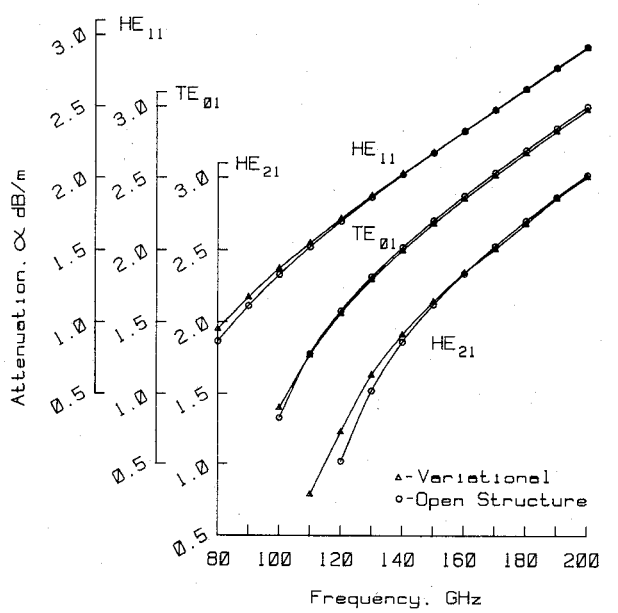


(b)

Fig. 5. Optimized values for ϵ_R required to achieve $\partial\beta/\partial\epsilon_R = 0$. (a) Real part. (b) Imaginary part. Note the scale change.



(a)



(b)

Fig. 6. Propagation and attenuation constant comparison between the $E_x/E_y/H_x/H_y$ formulation and the open structure eigenvalue results. (a) Normalized propagation constant. (b) Attenuation constant in dB/m. Note the offset scales.

cases were secondary roots observed. As the number of HRW modes included in the solution is increased, non-physical solutions do appear; however, these solutions are easily identified in the list of eigenvalues since they have gain instead of attenuation. Since the eigenvalues must be sorted to find the one with the largest real part, the non-physical solutions are easily removed in this process.

As an additional confirmation of the method and to demonstrate its versitility, a graded-index waveguide is analyzed. The structure that is analyzed is shown in inset in Fig. 7. The dielectric constant of the substrate follows a Gaussian–Gaussian distribution according to

$$\sqrt{\epsilon} = \sqrt{\epsilon_r} + \sqrt{\Delta\epsilon_r}\, e^{-(x-x_o)^2} e^{-(y-y_o)^2}, \qquad (19)$$

where (x_o, y_o) is the location of the peak value. For the inset figure, $x_o = 2.25$ mm, $y_o = 3$ mm, and the dielectric constants are as labeled. To model the channel waveguide, 200 rectangles are used as shown in Fig. 8. The dielectric constant of each rectangle is set to the value of (19) calculated at its center.

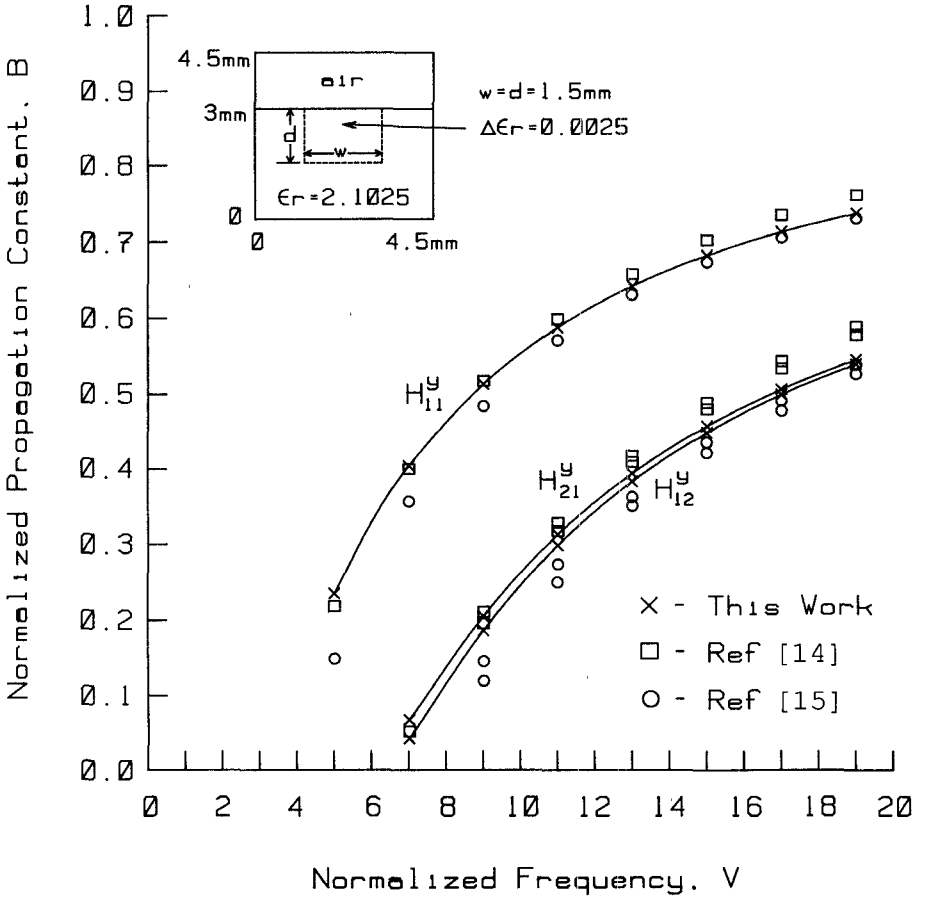


Fig. 7. Dispersion curves for the Gaussian–Gaussian graded-index channel waveguide shown in inset.

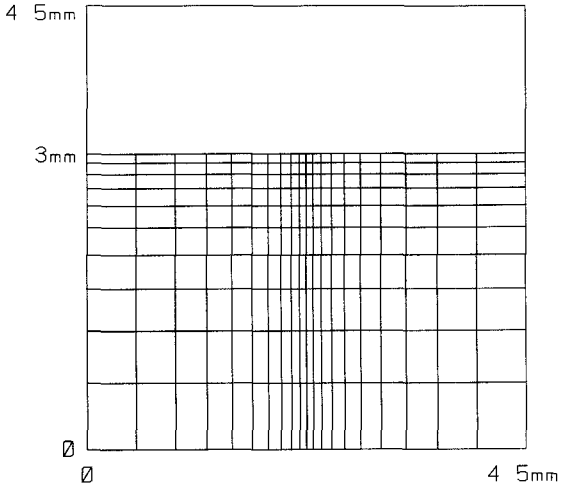


Fig. 8. Reduction of the graded-index profile into homogeneous rectangular regions.

To determine the number of HRW modes needed to model the IRW, several cases of increasing accuracy were computed. Table II shows the results at $V = 19$ for the effective dielectric constant, where V is the normalized frequency. Note that $L_y > L_x$ since the y-direction includes a step discontinuity. The total number of HRW modes used, and hence the matrix size, is given in the third column. Symmetry is exploited to reduce the number of modes used. Note that for all practical purposes, all of the results are equally accurate and the smallest matrix size could be used for generating data. However, for comparison with previous work, a normalized frequency is used that requires extreme accuracy to achieve good agreement, so for the computation of all following data, $L_x = 16$ and $L_y = 20$.

TABLE II
EFFECTIVE DIELECTRIC CONSTANT RESULTS AT $V = 19$ FOR VARYING BASIS SET SIZE

L_x	L_y	#HRW Modes	H_{11}^y	H_{21}^y	H_{12}^y
10	14	95	2.205	2.173	2.172
12	16	126	2.208	2.178	2.177
14	18	161	2.210	2.181	2.180
16	20	200	2.211	2.183	2.182

The dispersion curves of Fig. 7 show a comparison between the results computed with the optimizing Rayleigh–Ritz technique and with two other methods.

The normalized propagation constant is given by

$$B = \frac{\left(\frac{\beta}{k_o}\right)^2 - \epsilon_r}{\epsilon_{\max} - \epsilon_r},$$

where $\epsilon_{\max} = 2.25$ for this work. Due to the subtraction, this definition emphasizes the less significant digits. The normalized frequency is given as $V = dk_o$, $k_o = \omega\sqrt{\epsilon_o \mu_o}$. Note that all results required 6 iterations in the complex root search, including the three required to start the search, and that in no instances were multiple roots or spurious modes observed. The agreement between this method and previous work is quite good. This method tends towards better agreement with [14] at lower frequencies and with [15] at higher frequencies.

VI. Conclusion

The paper details a variational formulation that utilizes the modes of an optimized homogeneous rectangular waveguide to model the modes of an inhomogeneous rectangular waveguide. The specific implementation discussed is shown to make accurate non-perturbational propagation and attenuation constant calculations for closed dielectric waveguides for the dominant and higher-order modes. This work represents a new formulation of a previously described method that improves on the previous work by using a complete set of basis functions. It is shown that an E_z/H_z formulation of the homogeneous waveguide modes leads to a relative convergence phenomena and to incorrect loss calculations. The improved formulation uses an $E_x/E_y/H_x/H_y$ formulation that shows no difficulties.

The accuracy of the method is verified by analyzing the round step-index dielectric waveguide and comparing the results to a direct solution of the eigenvalue problem for the open case. Very good agreement is demonstrated for non-perturbational calculations of the propagation and attenuation constants for the dominant and three higher-order modes. The results also show that the staircase approximation required by the rectangular subdivision of the waveguide does not introduce any numerical difficulties or observable error. The computation time is only weakly linked to the number of rectangular subregions. An additional example demonstrates the ability to analyze graded-index waveguides.

Numerical stability is excellent. Formulation as a standard eigenvalue problem allows the use of stable numerical algorithms. Although a complex root search is required, the function for which the root is needed is available analytically. In addition, there is only one root–required theoretically and verified numerically. Finally, the rectangular subdivision of the inhomogeneous waveguide enhances numerical stability by allowing all integrations to be performed in closed form.

The method is intended to compete with the finite-element method for the analysis of closed dielectric waveguides. The variational technique described here is shown to possess excellent accuracy, stability, and generality. In addition, its use of entire-domain basis functions eliminates the need for any sort of mesh or grid, resulting in an implementation that is compact and very easy to use. The maximum set of required information consists of the structure itself, frequency, symmetry, desired solution (usually the dominant mode, but not required), x- and y-mode limits, and an initial guess. The latter two are strong candidates for automation, where the theoretical existence of only one root eliminates the uncertainty associated with the initial guess, and the mode limits can be set from physical considerations.

Appendix

The following four equations result from setting $\partial\beta/\partial A_{ij} = 0$ for A_{ij} equal to E_{xij}, E_{yij}, H_{xij}, and H_{yij} respectively. These results cover the general case of an inhomogeneous permeability:

$$\sum_{m=0}^{L_x-1}\sum_{n=1}^{L_y} E_{xmn}\frac{\pi^2}{b^2\omega\mu_R}\int_0^b\int_0^a\left(2jng_{xim}g_{yjn}\left(\frac{\mu(x,y)}{\mu_R}-1\right)+g_{xim}h_{yjn}\left(2\omega^2\mu_R\epsilon(x,y)b^2/\pi^2-j^2-n^2\right)\right)dx\,dy$$
$$+\sum_{m=1}^{L_x}\sum_{n=0}^{L_y-1} E_{ymn}\frac{\pi^2}{ab\omega\mu_R}\int_0^b\int_0^a\left(ijh_{xim}g_{yjn}+2mjg_{xim}g_{yjn}\left(1-\frac{\mu(x,y)}{\mu_R}\right)+mng_{xim}h_{yjn}\right)dx\,dy$$
$$=2\beta H_{yij}\int_0^b\int_0^a g_{xii}h_{yjj}\,dx\,dy,$$
$$i=0,1,\ldots,L_x-1;\quad j=1,2,\ldots,L_y \tag{20}$$

$$\sum_{m=0}^{L_x-1}\sum_{n=1}^{L_y} E_{xmn}\frac{\pi^2}{ab\omega\mu_R}\int_0^b\int_0^a\left(ijg_{xim}h_{yjn}+2nig_{xim}g_{yjn}\left(1-\frac{\mu(x,y)}{\mu_R}\right)+mnh_{xim}g_{yjn}\right)dx\,dy$$
$$+\sum_{m=1}^{L_x}\sum_{n=0}^{L_y-1} E_{ymn}\frac{\pi^2}{a^2\omega\mu_R}\int_0^b\int_0^a\left(2img_{xim}g_{yjn}\cdot\left(\frac{\mu(x,y)}{\mu_R}-1\right)+h_{xim}g_{yjn}\cdot\left(2\omega^2\mu_R\epsilon(x,y)a^2/\pi^2-i^2-m^2\right)\right)dx\,dy$$
$$=-2\beta H_{xij}\int_0^b\int_0^a h_{xii}g_{yjj}\,dx\,dy,$$
$$i=1,2,\ldots,L_x;\quad j=0,1,\ldots,L_y-1 \tag{21}$$

$$\sum_{m=1}^{L_x}\sum_{n=0}^{L_y-1} H_{xmn}\frac{\pi^2}{b^2\omega\epsilon_R}\int_0^b\int_0^a\Bigg(2jnh_{xim}h_{yjn}\cdot\left(\frac{\epsilon(x,y)}{\epsilon_R}-1\right)+h_{xim}g_{yjn}\cdot\left(2\omega^2\epsilon_R\mu(x,y)b^2/\pi^2-j^2-n^2\right)\Bigg)dx\,dy$$
$$+\sum_{m=0}^{L_x-1}\sum_{n=1}^{L_y} H_{ymn}\frac{\pi^2}{ab\omega\epsilon_R}\int_0^b\int_0^a\Bigg(ijg_{xim}h_{yjn}+2mjh_{xim}h_{yjn}\cdot\left(1-\frac{\epsilon(x,y)}{\epsilon_R}\right)+mnh_{xim}g_{yjn}\Bigg)dx\,dy$$
$$=-2\beta E_{yij}\int_0^b\int_0^a h_{xii}g_{yjj}\,dx\,dy,\qquad i=1,2,\ldots,L_x;\quad j=0,1,\ldots,L_y-1 \quad (22)$$

$$\sum_{m=1}^{L_x}\sum_{n=0}^{L_y-1} H_{xmn}\frac{\pi^2}{ab\omega\epsilon_R}\cdot\int_0^b\int_0^a\Bigg(ijh_{xim}g_{yjn}+2nih_{xim}h_{yjn}\left(1-\frac{\epsilon(x,y)}{\epsilon_R}\right)+mng_{xim}h_{yjn}\Bigg)dx\,dy+\sum_{m=0}^{L_x-1}\sum_{n=1}^{L_y} H_{ymn}\frac{\pi^2}{a^2\omega\epsilon_R}\cdot\int_0^b\int_0^a\Bigg(2imh_{xim}h_{yjn}\left(\frac{\epsilon(x,y)}{\epsilon_R}-1\right)+g_{xim}h_{yjn}\left(2\omega^2\epsilon_R\mu(x,y)a^2/\pi^2-i^2-m^2\right)\Bigg)dx\,dy$$
$$=2\beta E_{xij}\int_0^b\int_0^a g_{xii}h_{yjj}\,dx\,dy,\qquad i=0,1,\ldots,L_x-1;\quad j=1,2,\ldots,L_y \quad (23)$$

where

$$g_{xim}=\cos\left(\frac{i\pi x}{a}\right)\cos\left(\frac{m\pi x}{a}\right)$$
$$h_{xim}=\sin\left(\frac{i\pi x}{a}\right)\sin\left(\frac{m\pi x}{a}\right)$$
$$g_{yim}=\cos\left(\frac{i\pi y}{b}\right)\cos\left(\frac{m\pi y}{b}\right)$$
$$h_{yim}=\sin\left(\frac{i\pi y}{b}\right)\sin\left(\frac{m\pi y}{b}\right)$$

and ϵ_R and μ_R are the permittivity and permeability of the HRW. Note that the complex number $j=\sqrt{-1}$ appears nowhere in these results.

The required derivative $\partial\beta/\partial\epsilon_R$ can be found as

$$\frac{\partial\beta}{\partial\epsilon_R}=\frac{f_1}{f_2} \quad (24)$$

where, using a compact double-sum notation,

$$f_1=\sum_{mn}\sum_{ij}\frac{m\pi}{a\omega\epsilon_R^2}\int_0^b\int_0^a g_{xim}h_{yjn}\,dx\,dy\cdot\left(H_{ymn}H_{yij}\frac{m\pi}{a}-H_{xmn}H_{yij}\frac{n\pi}{b}\right)+\sum_{mn}\sum_{ij}\frac{n\pi}{b\omega\epsilon_R^2}\int_0^b\int_0^a h_{xim}g_{yjn}\,dx\,dy\cdot\left(H_{xmn}H_{xij}\frac{n\pi}{b}-H_{ymn}H_{xij}\frac{m\pi}{a}\right)+\sum_{mn}\sum_{ij}\frac{1}{\omega\epsilon_R^2}\int_0^b\int_0^a h_{xim}h_{yjn}\cdot\Bigg(H_{ymn}H_{yij}\frac{i\pi}{a}\frac{m\pi}{a}\left(1-2\frac{\epsilon(x,y)}{\epsilon_R}\right)+H_{ymn}H_{xij}\frac{m\pi}{a}\frac{j\pi}{b}\left(2\frac{\epsilon(x,y)}{\epsilon_R}-1\right)+H_{xmn}H_{yij}\frac{i\pi}{a}\frac{n\pi}{b}\left(2\frac{\epsilon(x,y)}{\epsilon_R}-1\right)+H_{xmn}H_{xij}\frac{j\pi}{b}\frac{n\pi}{b}\left(1-2\frac{\epsilon(x,y)}{\epsilon_R}\right)\Bigg)dx\,dy$$

and

$$f_2=2\sum_{mn}\sum_{ij}E_{xmn}H_{yij}\int_0^b\int_0^a g_{xim}h_{yjn}\,dx\,dy-2\sum_{mn}\sum_{ij}E_{ymn}H_{xij}\int_0^b\int_0^a h_{xim}g_{yjn}\,dx\,dy.$$

References

[1] B. Young, "Analysis of closed arbitrary dielectric waveguides using a modified Rayleigh-Ritz technique," *IEEE Trans. Microwave Theory Tech.*, vol. 39, no. 3, pp. 431–437, Mar. 1991.

[2] R. H. Sheikh and M. W. Gunn, "Wave propagation in a rectangular waveguide inhomogeneously filled with semiconductors," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-16, pp. 117–121, Feb. 1968.

[3] W. E. Hord and F. J. Rosenbaum, "Approximation technique for dielectric loaded waveguides," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-16, pp. 228–233, Apr. 1968.

[4] A. Vander Vorst and R. Govaerts, "On the accuracy obtained when using variational techniques for asymmetrically loaded waveguides," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 51–52, Jan. 1969.

[5] A. S. Vander Vorst, A. A. Laloux, and R. J. M. Govaerts, "A computer optimization of the Rayleigh-Ritz method," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 454–460, Aug. 1969.

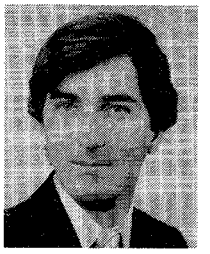
[6] S. Halevy, S. Raz, and H. Cory, "Bandwidth optimization by dielectric loading," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-26, pp. 406–412, June 1978.

[7] C.-T. Liu and C. H. Chen, "A variational theory for wave propagation in inhomogeneous dielectric slab loaded waveguides," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-29, pp. 805–812, Aug. 1981.

[8] J. B. Davies, "Propagation in rectangular waveguide filled with skew uniaxial dielectric," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-15, pp. 372–376, June 1967.

[9] S. Ramo, J. Whinnery, and T. Van Duzer, *Fields and Waves in*

Communication Electronics, 2nd ed. New York: Wiley, 1984, pp. 394–395.
[10] R. F. Harrington, *Time Harmonic Electromagnetic Fields*. New York: McGraw-Hill, 1961, p. 347.
[11] S. Wolfram, *Mathematica™ – A System for Doing Mathematics by Computer*. New York: Addison-Wesley, 1988.
[12] D. Marcuse, *Light Transmission Optics*, 2nd ed. New York: Van Nostrand Reinhold, 1982, p. 296.
[13] R. Mittra and S. W. Lee, *Analytical Techniques in the Theory of Guided Waves*. New York: Macmillan, 1971, p. 39.
[14] N. Schulz, K.-H. Bierwirth, and F. Arndt, "Finite difference analysis of integrated optical channel waveguides with arbitrarily graded index profile," in *1988 IEEE MTT-S Digest, Rec.*, pp. 731–734.
[15] R. Lagu and R. Ramaswamy, "A variational finite-difference method for analyzing channel waveguides with arbitrary index profiles," *IEEE J. Quantum Electron.*, vol. QE-22, no. 6, pp. 968–976, June 1986.



Brian Young (S'85–M'88) was born in Dallas, TX, on November 21, 1961. He received the B.S. degree from Texas A&M University in 1984, the M.S. degree from the University of Illinois at Champaign/Urbana in 1985, and the Ph.D. degree from the University of Texas at Austin in 1987, all in electrical engineering.

In 1988, he joined Texas A&M University as an Assistant Professor. In 1991 he joined the Solid-State Microwave Department of the Radar Systems Group, Hughes Aircraft Company, El Segundo, CA. His research interests are in numerical waveguide analysis, microwave/millimeter-wave/optical waveguides, electromagnetic theory, and computer-aided design.