A Circular-Harmonic Computer Analysis of Rectangular Dielectric Waveguides

By J. E. GOELL

(Manuscript received April 8, 1969)

This paper describes a computer analysis of the propagating modes of a rectangular dielectric waveguide. The analysis is based on an expansion of the electromagnetic field in terms of a series of circular harmonics, that is, Bessel and modified Bessel functions multiplied by trigonometric functions. The electric and magnetic fields inside the waveguide core are matched to those outside the core at appropriate points on the boundary to yield equations which are then solved on a computer for the propagation constants and field configurations of the various modes.

The paper presents the results of the computations in the form of curves of the propagation constants and as computer generated mode patterns. The propagation curves are presented in a form which makes them refractive-index independent as long as the difference of the index of the core and the surrounding medium is small, the case which applies to integrated optics. In addition to those for small index difference, it also gives results for larger index differences such as might be encountered for microwave applications.

I. INTRODUCTION

It is anticipated that dielectric waveguides will be used as the fundamental building blocks of integrated optical circuits. These waveguides can serve not only as a transmission medium to confine and direct optical signals, but also as the basis for circuits such as filters and directional couplers.[1] Thus, it is important to have a thorough knowledge of the properties of their modes.

Circular dielectric waveguides have received considerable attention because circular geometry is commonly used in fiber optics.[2–5] In many integrated optics applications it is expected that waveguides will consist of a rectangular, or near rectangular, dielectric core embedded in a dielectric medium of slightly lower refractive index. The modes

for this geometry are more difficult to analyze than those of the metallic rectangular waveguide because of the nature of the boundary.

Marcatili, using approximations based on the assumption that most of the power flow is confined to the waveguide core, has derived in closed form the properties of a rectangular dielectric waveguide.[6] In his solution, fields with sinusoidal variation in the core are matched to exponentially decaying fields in the external medium. In each region only a single mode is used. The results of this method are obtained in a relatively simple form for numerical evaluation.

The properties of the principal mode of the rectangular dielectric waveguide have been studied by Schlosser and Unger using a high-speed digital computer.[7] In their approach the transverse plane was divided into regions, as shown in Fig. 1, and rectangular coordinate solutions assumed in each of the regions. The longitudinal propagation constant was then adjusted so that a field match could be achieved at discrete points along the boundary. This method gives results which, theoretically, are valid over a wider range than Marcatili's, but with a significant increase in computational difficulty. One shortcoming of the method is that for a given mode, as the wavelength increases the field extent increases, so, in the limit it becomes increasingly difficult to match the fields along the boundaries between regions [1] and [2] and between regions [2] and [3].

A variational approach has been undertaken by Shaw and others.[8] They assume a test solution with two or three variable parameters in the core. From this test solution, the fields outside the core are then derived and the parameters are varied to achieve a consistent

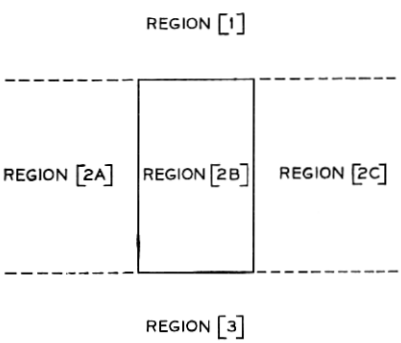

Fig. 1 — Matching boundaries for rectangular mode analysis.

solution. This approach, like that of Schlosser, requires involved computations. Also, it has the disadvantage that the test function must be assumed in advance. In addition, some of his preliminary results do not show the proper behavior for the limiting cases (waveguide dimensions which are very large or very small compared with the wavelength).

In the present analysis the radial variation of the longitudinal electric and magnetic fields of the modes are represented by a sum of Bessel functions inside the waveguide core and by a sum of modified Bessel functions outside the waveguide core. Solutions are found by matching the fields along the perimeter of the core. Thus, the matching boundary is not a function of the waveguide parameters, so the computational complexity does not increase with wavelength.

Section II discusses the underlying theory of the circular-harmonic analysis of rectangular dielectric waveguides. This is followed by a description of computational techniques and special graphical methods of presentation used. Section III is divided into three parts, the first describing the accuracy of the computations, the second describing field patterns, and the third presenting propagation curves.

II. DERIVATION OF EQUATIONS

The waveguide considered here consists of a rectangular core of dielectric constant, ϵ_1, surrounded by an infinite medium of dielectric constant, ϵ_0. Both media are assumed to be isotropic, and have the permeability of free space, μ_0. Figure 2 shows the coordinate systems (rectangular and cylindrical) and rod dimension used in this paper. The direction of propagation is in the $+z$ direction (towards the observer).

In cylindrical coordinates the field solutions of Maxwell's equations take the form of Bessel functions and modified Bessel functions multiplied by trigometric functions, and their derivatives. In order for propagation to take place in the z direction, the field solutions must be Bessel functions in the core and modified Bessel functions outside. Since Bessel functions of the second kind have a pole at the origin and modified Bessel functions of the first kind a pole at infinity, the radial variation of the fields is assumed to be a sum of Bessel functions of the first kind and their derivatives inside the core and a sum of modified Bessel functions and their derivatives outside the core.

In cylindrical coordinates, the z components of the electric and magnetic fields are given by

$$E_{z1} = \sum_{n=0}^{\infty} a_n J_n(hr) \sin(n\theta + \varphi_n) \exp[i(k_z z - \omega t)] \tag{1a}$$

and

$$H_{z1} = \sum_{n=0}^{\infty} b_n J_n(hr) \sin(n\theta + \psi_n) \exp[i(k_z z - \omega t)] \tag{1b}$$

inside the core, and by

$$E_{z0} = \sum_{n=0}^{\infty} c_n K_n(pr) \sin(n\theta + \varphi_n) \exp[i(k_z z - \omega t)] \tag{1c}$$

and

$$H_{z0} = \sum_{n=0}^{\infty} d_n K_n(pr) \sin(n\theta + \psi_n) \exp[i(k_z z - \omega t)] \tag{1d}$$

outside the core, where ω is the radian frequency and k_z the longitudinal propagation constant. The transverse propagation constants are given by

$$h = (k_1^2 - k_z^2)^{\frac{1}{2}} \tag{2a}$$

and

$$p = (k_z^2 - k_0^2)^{\frac{1}{2}} \tag{2b}$$

where $k_1 = \omega(\mu_0 \epsilon_1)^{\frac{1}{2}}$ and $k_0 = \omega(\mu_0 \epsilon_0)^{\frac{1}{2}}$. The terms J_n and K_n are the nth order Bessel functions and modified Bessel functions, respectively, and ψ_n and φ_n are arbitrary phase angles.

The transverse components of the fields are given by[9]

$$E_r = \frac{ik_z}{k^2 - k_z^2}\left[\frac{\partial E_z}{\partial r} + \left(\frac{\mu_0 \omega}{k_z r}\right)\frac{\partial H_z}{\partial \theta}\right] \tag{3a}$$

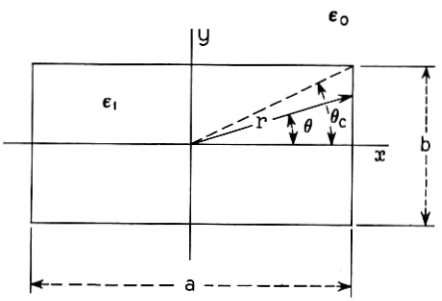

Fig. 2 — Dimensions and coordinate system.

$$E_\theta = \frac{ik_z}{k^2 - k_z^2}\left[\frac{1}{r}\frac{\partial E_z}{\partial \theta} - \left(\frac{\mu_0\omega}{k_z}\right)\frac{\partial H_z}{\partial r}\right] \tag{3b}$$

$$H_r = \frac{ik_z}{k^2 - k_z^2}\left[-\left(\frac{k^2}{\mu_0\omega k_z r}\right)\frac{\partial E_z}{\partial \theta} + \frac{\partial H_z}{\partial r}\right] \tag{3c}$$

$$H_\theta = \frac{ik_z}{k^2 - k_z^2}\left[\left(\frac{k^2}{\mu_0\omega k_z}\right)\frac{\partial E_z}{\partial r} + \frac{1}{r}\frac{\partial H_z}{\partial \theta}\right], \tag{3d}$$

where k can be either k_1 or k_0.

Finally, the component of the electric field tangent to the rectangular core is given by

$$E_t = \pm(E_r \sin\theta + E_\theta \cos\theta) \qquad \begin{matrix} -\theta_c < \theta < \theta_c \\ \pi - \theta_c < \theta < \pi + \theta_c \end{matrix} \tag{4a}$$

or

$$E_t = \pm(-E_r \cos\theta + E_\theta \sin\theta), \qquad \begin{matrix} \theta_c < \theta < \pi - \theta_c \\ \pi + \theta_c < \theta < -\theta_c \end{matrix}, \tag{4b}$$

where θ_c is the angle which a radial line to the corner in the first quadrant makes with the x axis. Similar expressions exist for the tangential magnetic field.

2.1 *Effects of Symmetry*

Since the waveguide is symmetrical about the x axis the fields must be either symmetric or antisymmetric about this axis. This is true because the structure is invarient under 180° rotations and therefore the field patterns must be invarient under a 180° rotation, except for sign. From this and the fact that $\partial/\partial\theta$ appears in each of equations (3), it is evident that two types of modes must exist, the first type with $\varphi_n = 0$ and $\psi_n = \pi/2$ and the second type with $\varphi_n = \pi/2$ and $\psi_n = \pi$.

Similarly, the field functions must also be symmetric or antisymmetric about the y axis. Suppose, for example, E_{z0} exhibits a sinusoidal angular dependence about $\theta =$ (E_{z0} is odd about the x axis). Then, letting $\alpha = \theta - \pi/2$, equation (1c) can be put in the form

$$E_{z0} = \sum_{n=0}^{\infty} c_n K_n(pr)(\sin n\alpha \cos n\pi/2 + \cos n\alpha \sin n\pi/2). \tag{5}$$

For E_{z0} to be purely symmetric about $\alpha = 0$ (the y axis), all n must be odd; for E_{z0} to be antisymmetric about $\alpha = 0$ all n must be even.

Since similar results apply for cosinusoidal variation of E_{z0} about $\theta = 0$, and all other field functions as well, any given mode must consist of either even harmonics or odd harmonics.

From the preceeding analysis it is evident that if the matching points are selected symmetrically about both the x and y axes, then, except possibly for sign, every point will have an equivalent point in each quadrant. Therefore, the field matching need only be performed in one quadrant. Thus, the use of the symmetry of the structure not only reduces the number of constants required to calculate the properties of a given mode by a factor of four, it also decreases the number of points to achieve a given degree of accuracy by the same factor.

2.2 *Selection of Matching Points*

As mentioned in Section 2.1, the matching point locations should be symmetrical about the x and y axes. For the odd harmonic cases, the points used to compute the results to be presented in Section III were $\theta_m = (m - 1/2)\pi/2N$; $m = 1, \cdots, N$, where N was the number of space harmonics.

The choice of points for the even harmonic cases was more complicated since simultaneous existence of an $n = 0$ harmonic for both the TE and TM circular modes is inconsistent with the waveguide symmetries. Thus, if the maximum n for both the TE and TM solutions are equal, the total number of coefficients to be found will be $4N - 2$ rather than $4N$ as in the previous case.

The method of choosing points for the even harmonic modes used for the computation of the results of Section III was to pick the points for the field components with even symmetry about $\theta = 0$ to be $\theta_m = (m - 1/2)\pi/2N$; $m = 1, 2, \cdots, N$, and for the field components with odd symmetry about $\theta = 0$ to be $\theta_m = (m - N - 1/2)\pi/2(N - 1)$; $m = N + 1, N + 2, \cdots, (2N - 1)$ for cases with unity aspect ratio, $(a/b = 1)$. For aspect ratios other than unity, all points were chosen according to the first formula, except that the first and last points for the odd z component were omitted.

2.3 *Formulation of Matrix Elements*

The coefficients of equation (1) were found by matching the tangential electric and magnetic fields along the boundary of the waveguide core. Since each type of field consists of both longitudinal and transverse components, four types of matching equations exist.

To facilitate computer analysis the matching equations were put in

matrix form. The matching equations in matrix form for the longitudinal field components are

$$E^{LA}A = E^{LC}C \tag{6a}$$

for the electric field and

$$H^{LB}B = H^{LD}D \tag{6b}$$

for the magnetic field. For the transverse fields the matrix matching equations are given by

$$E^{TA}A + E^{TB}B = E^{TC}C + E^{TD}D \tag{6c}$$

for the electric field and

$$H^{TA}A + H^{TB}B = H^{TC}C + H^{TD}D \tag{6d}$$

for the magnetic field. The A, B, C, and D matrices are N element column matrices of the a_n, b_n, c_n, and d_n mode coefficients, respectively. The elements of the $m \times n$ matrices E^{LA}, E^{LC}, H^{LB}, H^{LD}, E^{TA}, E^{TB}, E^{TC}, E^{TD}, H^{TA}, H^{TB}, H^{TC}, and H^{TD} are given by

$$e_{mn}^{LA} = JS, \tag{7a}$$

$$e_{mn}^{LC} = KS, \tag{7b}$$

$$h_{mn}^{LB} = JC, \tag{7c}$$

$$h_{mn}^{LD} = KC, \tag{7d}$$

$$e_{mn}^{TA} = -k_z(\mathbf{J}'SR + \mathbf{J}CT), \tag{7e}$$

$$e_{mn}^{TB} = k_0 Z_0(\mathbf{J}SR + \mathbf{J}'CT), \tag{7f}$$

$$e_{mn}^{TC} = k_z(\mathbf{K}'SR + \mathbf{K}CT), \tag{7g}$$

$$e_{mn}^{TD} = -k_0 Z_0(\mathbf{K}SR + \mathbf{K}'CT), \tag{7h}$$

$$h_{mn}^{TA} = \epsilon_r k_0(\mathbf{J}CR - J'ST)/Z_0, \tag{7i}$$

$$h_{mn}^{TB} = -k_z(\mathbf{J}'CR - \mathbf{J}ST), \tag{7j}$$

$$h_{mn}^{TC} = -k_0(\mathbf{K}CR - \mathbf{K}'ST)/Z_0, \tag{7k}$$

$$h_{mn}^{TD} = k_z(\mathbf{K}'CR - \mathbf{K}ST), \tag{7l}$$

where

$$Z_0 = (\mu_0/\epsilon_0)^{\frac{1}{2}},$$

$$\epsilon_r = \epsilon_1/\epsilon_0,$$

$$\left.\begin{aligned} S &= \sin(n\theta_m + \varphi) \\ C &= \cos(n\theta_m + \varphi) \end{aligned}\right\} \quad \begin{aligned} &\text{or} \\ &\end{aligned} \quad \begin{aligned} \varphi &= 0 \\ \varphi &= \pi/2 \end{aligned} ,$$

$$J = J_n(hr_m), \qquad K = K_n(pr_m),$$

$$J' = J_n'(hr_m), \qquad K' = K_n'(pr_m),$$

$$\mathbf{J} = \frac{nJ_n(hr_m)}{h^2 r_m}, \qquad \mathbf{K} = \frac{nK_n(pr_m)}{p^2 r_m},$$

$$\mathbf{J}' = \frac{J_n'(hr_m)}{h}, \qquad \mathbf{K}' = \frac{K_n'(pr_m)}{p},$$

and

$$\left.\begin{aligned} R &= \sin\theta_m \\ T &= \cos\theta_m \\ r_m &= (a/2)\cos\theta_m \end{aligned}\right\} \quad \theta < \theta_c , \qquad \left.\begin{aligned} R &= -\cos\theta_m \\ T &= \sin\theta_m \\ r_m &= (b/2)\sin\theta_m \end{aligned}\right\} \quad \theta > \theta_c .$$

For $\theta = \theta_c$, the boundary at the corner was assumed to be perpendicular to the radial line connecting it to the origin, so for this case $R = \cos(\theta_m + \pi/4)$, $T = \cos(\theta_m - \pi/4)$, and $r_m = (a^2 + b^2)^{\frac{1}{2}}/4$.

2.4 *Mode Designation*

Unlike metallic waveguides, the field patterns of dielectric waveguides are sensitive to refractive index difference, wavelength, and aspect ratio. This complicates the problem of finding a reasonably descriptive mode designation scheme.

For rectangular metallic waveguides, the accepted approach is to designate the modes as TE (or H) and TM (or E), and to specify the number of field maxima in the x and y directions with a double subscript. When there is no variation the subscript 0 is used.

Since the rectangular dielectric waveguide modes are neither pure TE nor pure TM, a different scheme must be used. The scheme adopted is based on the fact that in the limit, for large aspect ratio, short wavelength, and small refractive index difference, the transverse electric field is primarily parallel to one of the transverse axes. Modes are designated as E^y_{mn} if in the limit their electric field is parallel to the y axis and as E^x_{mn} if in the limit their electric field is parallel to the x axis. The m and n subscript are used to designate the number of maxima in the x and y directions, respectively.†

† This scheme agrees with that used by Marcatili in Ref. 6.

2.5 *Electric and Magnetic Field Function Differences*

For a hollow metallic waveguide where pure TE and TM modes can exist, it is evident from equation (3) that E_r and H_θ have similar transverse variations as do E_θ and H_r, so that the impedance is independent of position. Furthermore, the transverse electric and magnetic fields are perpendicular and the power flow, Re $\{E \times H^*\}$, does not change sign anywhere across the waveguide.

By examination of equation (3), it is clear that for the mixed modes of the dielectric waveguide, the field functions are not similar and the impedance is a function of position. In order for the transverse fields E_t and H_t to be perpendicular,

$$E_t \cdot H_t = E_r H_r + E_\theta H_\theta = 0. \tag{8}$$

Now, from equation (3)

$$E_t \cdot H_t = \frac{k_z^2 - k^2}{k_z^2} \left(\frac{\partial H_z}{\partial r} \frac{\partial E_z}{\partial r} + \frac{1}{r^2} \frac{\partial H_z}{\partial \theta} \frac{\partial E_z}{\partial \theta} \right). \tag{9}$$

Thus, E_t and H_t are not necessarily perpendicular. Finally, since the transverse variations of E_t and H_t are not the same, the electric field and magnetic field can change sign at different points, which results in negative power flow.†

Three special cases exist where the electric and magnetic fields, and the impedance, have the same positional dependence, and where the power flow does not change sign across the waveguide:

(*i*) in one of the regions if the propagation constant is approximately equal to the bulk propagation constant of that region, that is, if $k \approx k_1$ or $k \approx k_0$,

(*ii*) everywhere in the limit for small refractive index difference, since case *i* will then hold in both regions, and

(*iii*) everywhere for circular symmetry of both the structure and the modes.

2.6 *Normalization*

The arguments of the Bessel and modified Bessel functions are given by $hr = (k_1^2 - k_z^2)^{\frac{1}{2}} r$ and $pr = (k_z^2 - k_0^2)^{\frac{1}{2}} r$, respectively. The first argument can be put in the form

$$hr = [k_1^2 - k_0^2 - p^2]^{\frac{1}{2}} r. \tag{10}$$

† This unusual property has also been observed for helices.[10] Presumably, if loss were included there would be a radial component of power to feed the reverse flow, and the lossless case can be thought of as the limit of the lossy case.

Letting

$$\mathcal{P}^2 = \frac{(k_z/k_0)^2 - 1}{n_r^2 - 1}, \tag{11}$$

and

$$\mathcal{R} = rk_0(n_r^2 - 1)^{\frac{1}{2}}, \tag{12}$$

where

$$n_r = (k_1/k_0)^{\frac{1}{2}} \tag{13}$$

is the index of refraction of the core relative to the outer medium, gives

$$pr = \mathcal{P}\mathcal{R} \tag{14}$$

and

$$hr = \mathcal{R}(1 - \mathcal{P}^2)^{\frac{1}{2}}. \tag{15}$$

The curves of the propagation constant given in Section III are drawn in terms of $\mathcal{P}^2$ and $\mathcal{B}$, where

$$\mathcal{B} = \frac{2b}{\lambda_0}(n_r^2 - 1)^{\frac{1}{2}} \tag{16}$$

and $\lambda_0 = 2\pi/k_0$. Since $\mathcal{R}$ is proportional to $1/(n_r^2 - 1)^{\frac{1}{2}}$ and $\mathcal{P}$ and $\mathcal{B}$ are proportional to $(n_r^2 - 1)^{\frac{1}{2}}$, the use of $\mathcal{P}^2$ and $\mathcal{B}$ as plotting variables eliminates the explicit dependence of the Bessel and modified Bessel function arguments on the refractive indices of the media.

Examination of the matching equations, equations (6), reveals that ϵ_r appears in the H^{TA} term. However, since ϵ_r appears as a multiplicative factor in H^{TA}, for sufficiently small values the normalized propagation constant, $\mathcal{P}^2$, is independent of ϵ_r.

The normalized propagation constant, $\mathcal{P}^2$, has two additional properties which make its use convenient. First, its range of variation is on the interval (0, 1). Second, for $n_r \approx 1$,

$$\mathcal{P}^2 \approx \frac{k_z/k_0 - 1}{\Delta n_r}, \tag{17}$$

where $\Delta n_r = n_r - 1$; so for small n_r, $\mathcal{P}^2$ is proportional to $k_z - k_0$. The latter property is the reason that $\mathcal{P}^2$ rather than $\mathcal{P}$ was used as a plotting variable.

2.7 *Method of Computation*

2.7.1 *Propagation Constant*

Equation (6) yields $4N$ simultaneous homogeneous linear equations for the a_n, b_n, c_n, and d_n for the odd modes and $4N$-2 equations for

the even modes, using the matching points previously described. The equations can be combined to form a single matrix equation

$$[Q][T] = 0, \tag{18}$$

where

$$Q = \begin{bmatrix} E^{LA} & 0 & -E^{LC} & 0 \\ 0 & H^{LB} & 0 & -H^{LD} \\ E^{TA} & E^{TB} & -E^{TC} & -E^{TD} \\ H^{TA} & H^{TB} & -H^{TC} & -H^{TD} \end{bmatrix}$$

and the column matrix

$$[T] = \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix}.$$

All of the quantities in the matrices $[Q]$ and $[T]$ are themselves matrices as defined by equations (1), (6), and (7).

In order for a nontrivial solution to equation (18) to exist

$$\text{Det } [Q] = 0. \tag{19}$$

The normalized propagation constant, Φ^2, was found by substituting test values into equation (19). First, values of Φ^2 evenly distributed in the interval (0, 1) were substituted to crudely locate the roots. Then, Newton's method was used to find the roots to the desired accuracy.[11] Generally, one Newton approximation was used to find Φ^2 for the propagation curves and about ten Newton's approximations when Φ^2 was to be used to calculate field plots.

Both the simple method of triangulation[12] and the more complicated Gauss pivotal condensation method[13] were used to evaluate the determinant, the former for almost all cases and the latter for a few cases when roundoff error was apparent because the value of the determinant was not a smooth function of Φ^2. In all cases double precision arithmetic was used. For five space harmonics, about 0.1 second of IBM 360/65 computing time was required for each value of Φ^2 to evaluate the determinant using the triangulation method.

Due to the wide dynamic range of the coefficients, steps had to be taken to prevent underflow and overflow of the computer and to re-

duce the effects of roundoff. Multiplying a row or column of the matrix by a finite constant is equivalent to multiplying the determinant by that constant. Thus, any row or column of the determinant can be multiplied by a positive function without shifting its zeroes.

A detailed theory giving the "best functions" can be derived. However, since a "brute force" method was used, the more sophisticated method, which was not used because it would have required a substantial increase in the complexity of the program logic, is not discussed. It was found that multiplying the Bessel function terms by $h^2d/|J_n(hb)|$ and the modified Bessel function terms by $p^2d/k_n(pb)$, where d is the average of the waveguide dimensions, kept the variation of the terms "under control." A further simplification was made by setting Z_0 to unity, which does not shift the zeroes of the determinant because if the H_t rows are multiplied by Z_0, then if Z_0 appears in a column, it will appear in a similar manner in every element of the column.

2.7.2 *Mode Configurations*

The electric and magnetic fields were calculated for representative cases from equation (3). To find the a_n, b_n, c_n, and d_n coefficients, k_z was first found from equation (19). Its value was then substituted into equation (18). By setting one of the elements of the T column matrix to unity, all of the other elements were then found by standard matrix techniques.[13]

Several approaches were used to obtain information that could be used to derive the field patterns. These included computation of the field components along radial cuts of the waveguide cross section, computer generated isoclines giving the direction of the electric field, and computer generated mode pictures.

The isoclines and pictures were drawn using a simulated Stromberg Carlson SC-4020 cathode ray tube plotter, which is capable of generating points and lines on a 1024 × 1024 grid.† A single quadrant was used for the isoclines and intensity picture since the results for all quadrants are identical except for orientation. In general, the dimensions were scaled so that the long dimension of the rectangular waveguide core extended over 80 percent of the displayed width. All figures were plotted at the points $(20m, 20n)$, where m and n take on all integer values from 0 to 49.

Isocline drawings were made by drawing a line at each of the coordinate points parallel to the electric field at that point (all lines

† An SC-4060 plotter was used to simulate the SC-4020 plotter to take advantage of previously existing programs.

had the same length). The isocline drawings were used as working tools to derive the field line drawings in Section III.

In order to draw pictures of mode patterns, the power density was calculated at each of the points to be plotted. The square root of the power density was then normalized to the square root of the peak power density and quantized into 21 levels. About each point in the picture, a portion of the figure shown in Fig. 3 was then plotted, starting at 1 and going to the point corresponding to the appropriate quantized level (except at the points where the quantized power was zero where no plotting was done). Since the size of the cathode ray tube spot is approximately equal to the line spacing in the figure, the plotted figures are filled in. Therefore, the light passed by these figures is approximately equal to the power density to be represented. For small index difference, since the power density is proportional to the square of the transverse electric field, the dynamic range of the pictures (in terms of the electric field) is 400.

Starting with the single quadrant pictures, complete pictures were generated by making quadruple exposures of the microfilm. In general, about 30 to 60 seconds of IBM 360/65 computing time were required for each picture.

III. RESULTS OF COMPUTATION

This section gives the computed results. Section 3.1 discusses accuracy. This is followed by a discussion of field plots and mode

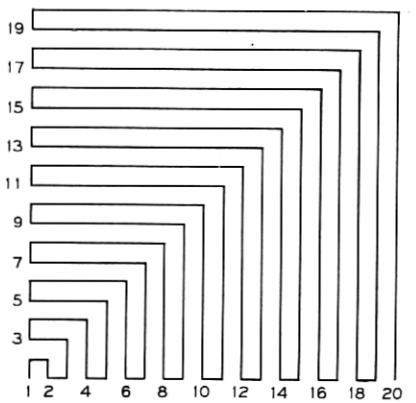

Fig. 3 — Intensity picture figure.

TABLE I—SAMPLE ACCURACY RESULTS

Number of Harmonics Used	$\mathcal{P}^2$			
	$a/b=1$	$a/b=2$	$a/b=3$	$a/b=4$
3	0.714	0.811	0.820	0.828
4	0.713	0.811	0.820	0.819
5	0.715	0.808	0.819	0.813
6	0.714	0.808	0.822	0.820
7	0.715	0.808	0.820	0.813
8	0.715	0.807	0.820	0.814
9	0.715	0.807	0.823	0.815
Variation	0.2%	0.4%	0.4%	1.5%

pictures in Section 3.2. Finally, curves of the propagation constant for a variety of conditions are presented in Section 3.3.

3.1 *Accuracy*

Numerous test runs were made in order to obtain an estimate of the accuracy of the computed results. The results of several of these runs are given in Table I for the first mode with $\mathcal{B} = 2$. The numbers at the bottom of the table represent the total variation for a given aspect ratio taken as a percentage of the full range possible (one).

For small aspect ratios, it is clear that the convergence is very rapid. However, for larger aspect ratios the convergence is not as good. For example, the variation for an aspect ratio of four is 1.5 percent (taken as a percentage of the full range of variation). For this case, from the table and from the limit for infinite aspect ratio[14] which is an upper bound for $\mathcal{P}^2$, it appears the error is about 3 percent. This error is achieved with a relatively small number of harmonics and can only be improved by using a prohibitively large number of harmonics on a computer which carries more significant digits than the one which was available for this study. However, since solutions exist for an infinite aspect ratio, the decrease in accuracy for the large aspect ratio of the circular-harmonic method is not a serious problem.

Computations similar to those for Table I were performed to obtain an estimate of the upper bound of the accuracy of the cases presented in Section 3.3. From these calculations, it is believed that all of the data to be presented in the following sections is accurate to 1 percent, except for the results of calculations using even harmonics for aspect ratios other than unity which are believed to be accurate to better than 2 percent. In general, accuracy decreases as the mode order increases, although not monotonically.

The results of the circular-harmonic analysis and of Marcatili's analysis agree.[6] In the regions where his method and the circular-harmonic method are both theoretically valid, the agreement is well within the tolerances given above. To avoid duplication, the reader is directed to his curves for a comparison.

The effect of the number of harmonics used in the field patterns is of some interest. This question has not been explored in great detail; however, a few comparisons of intensity pictures for different numbers of circular harmonics were made. In general, it was found that five harmonics were sufficient to give a good representation of the modes that this paper presents. An example of this is given in Fig. 4, comparing the E_{11}^{y} mode intensity patterns for five and nine harmonics. For the results which follow, five circular harmonics were used.

3.2 *Mode Configurations*

Figure 5 shows intensity pictures for the first six modes for unity aspect ratio, $\mathcal{B} = 3$, and an index difference of 0.01. Figure 6 gives similar data for an aspect ratio of two and $\mathcal{B} = 2$. For both, the plots are arranged in ascending order of cutoff frequency. All of the pictures are for E_{mn}^{y} modes. These pictures are virtually indistinguishable from the corresponding E_{mn}^{x} modes so both sets are not presented. In general, for small index differences the E_{mn}^{y} and E_{mn}^{x} can be considered to be near duals, that is, to have identical field patterns except that the electric and magnetic fields are interchanged.

The field distribution patterns for the modes of Figs. 5 and 6 are more complicated than those for the rectangular metallic waveguide

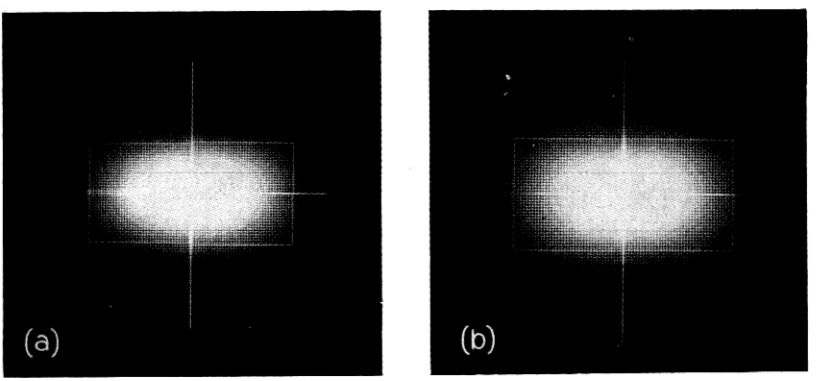

Fig. 4 — Intensity for the E_{11}^{y} mode for $a/b = 2$, $\mathcal{B} = 2$, and $\Delta n_r = .01$: (a) for five harmonics and (b) for nine harmonics.

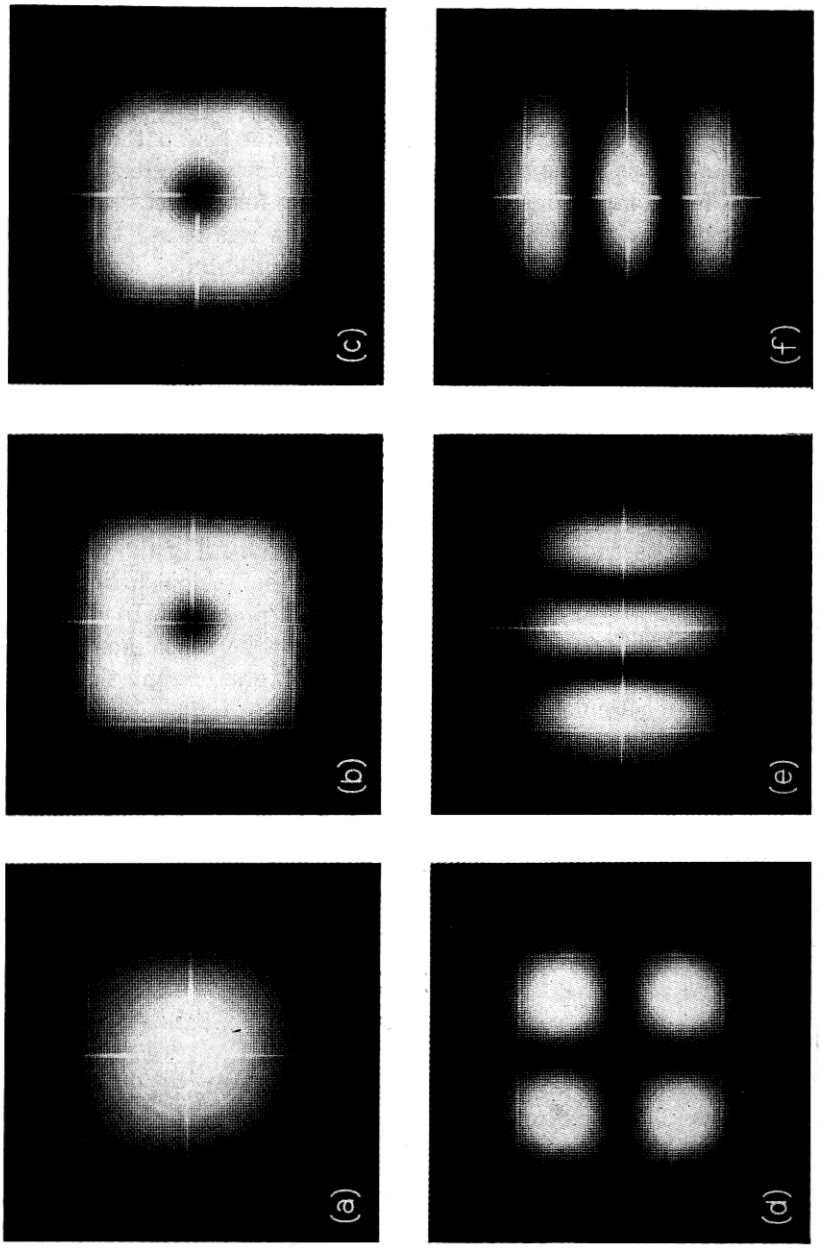

Fig. 5 — Intensity for some E^y_{mn} modes with unity aspect ratio, $\mathcal{B} = 3$, and $\Delta n_r = 0.01$: (a) E^y_{11}, (b) E^y_{21}, (c) E^y_{12}, (d) E^y_{22}, (e) E^y_{31}, and (f) E^y_{13}.

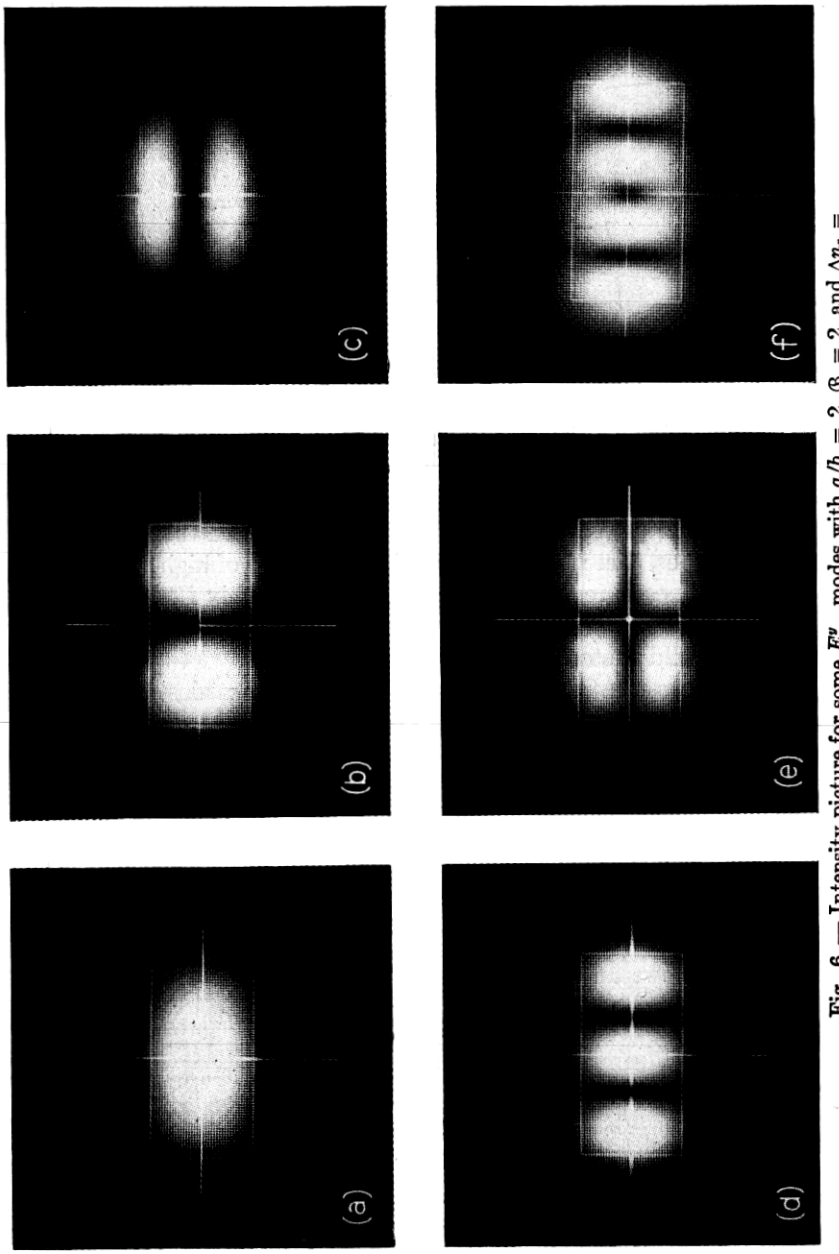

Fig. 6 — Intensity picture for some E^y_{mn} modes with $a/b = 2$, $\mathcal{B} = 2$, and $\Delta n_r = 0.01$: (a) E^y_{11}, (b) E^y_{21}, (c) E^y_{12}, (d) E^y_{31}, (e) E^y_{22}, and (f) E^y_{41}.

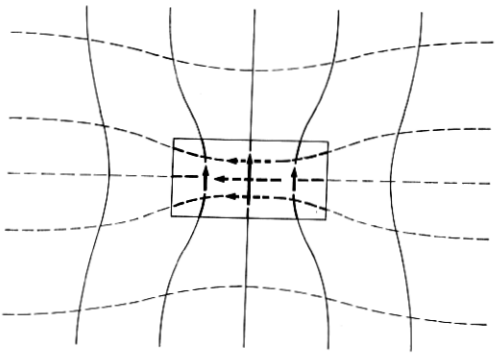

Fig. 7— Field configuration of the E^y_{11} mode.

since they extend beyond the waveguide boundary and, in general, their shape is dependent on waveguide parameters other than shape. The E^x_{11} and E^y_{11} modes have the simplest field patterns. Figure 7 shows the electric and magnetic field orientations for the E^y_{11} mode. In this figure and the following ones, there are heavy lines in the regions of high field intensity and light lines in regions of low field intensity. Only E^y_{mn} modes are shown since the E^x_{mn} modes can be obtained by interchanging the electric and magnetic field vectors.

Figure 8 shows the field lines for the E^y_{21} and E^y_{12} modes for a large aspect ratio. (For $a/b \to \infty$ the fields have the appearance of rectangular metallic waveguide modes.) However, as the aspect ratio approaches unity, the E^y_{12} and E^x_{21} modes and the E^y_{21} and E^x_{12} modes couple and shift to the patterns shown in Fig. 9. Most of the change takes place with the aspect ratio close to unity.

Figures 10, 11, and 12 show the field configurations for the E^y_{22} mode, the E^y_{31} mode, and the E^y_{13} mode, respectively. The field patterns of these modes do not change drastically with the aspect ratios.

Figure 13a shows an intensity picture of the E^y_{32} mode and Figure

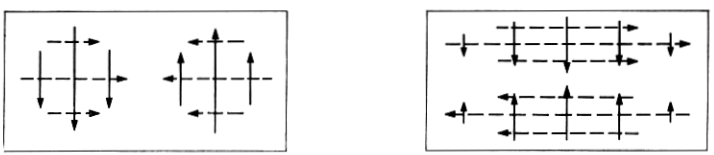

Fig. 8—Field configurations for the (a) E^y_{21} and (b) E^y_{12} modes far from cutoff.

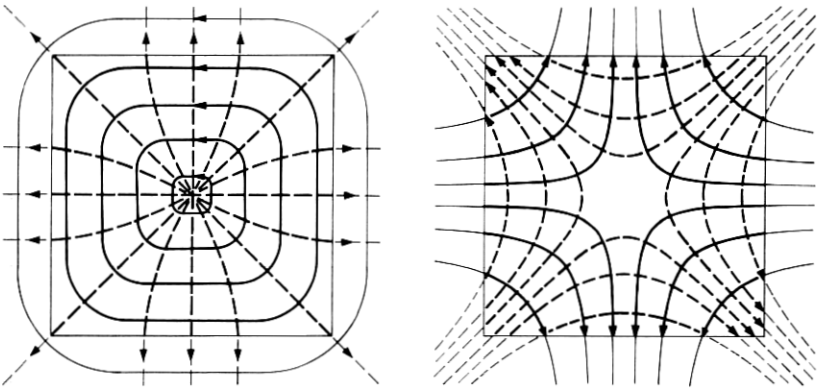

Fig. 9 — Field configurations for the square (a) E^y_{21} and (b) E^y_{12} modes.

13b its field pattern for unity aspect ratio. The field pattern inside the core is similar to a sum of the TE_{23} and TE_{32} of metallic waveguide, shown in Fig. 13c and d, respectively. Figure 13a demonstrates that the circular-harmonic analysis can generate complex field patterns with a relatively small number of harmonics.

Figures 14 and 15 show the variation of the intensity distribution with $\mathcal{P}^2$ for the E^y_{11} and E^y_{21} modes, respectively. As one would expect, for small values of $\mathcal{P}^2$ the radial extent of both modes increases very rapidly as $\mathcal{P}^2$ decreases. It is of significance, however, that most of the energy is contained within the waveguide core, even for relatively small values of $\mathcal{P}^2$ and Δn. Thus, Marcatili's assumption that very little energy propagates in the region of the corners is valid over a wide range.

3.3 *Propagation Curves*

In all cases of computed propagation curves, the normalized waveguide height $\mathcal{B}$, as given in equation (11), is plotted on the horizontal

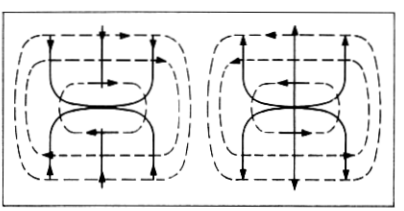

Fig. 10 — Field configuration of the E^y_{22} mode.

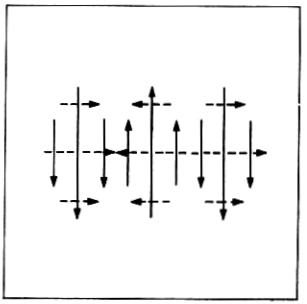

Fig. 11 — Field configuration of the E^y_{31} mode.

axis and the normalized propagation constant, $\mathcal{P}^2$, given in equation (16), along the vertical axis.

Figure 16 shows the case of vanishing index difference for an aspect ratio of one. The first 16 modes are shown. For this case the following six degenerate groups exist

$$E^y_{11}\ ,\ E^x_{11}$$

$$E^y_{12}\ ,\ E^x_{12}\ ,\ E^y_{21}\ ,\ E^x_{21}$$

$$E^y_{31}\ ,\ E^x_{13}$$

$$E^z_{31}\ ,\ E^y_{13}$$

$$E^z_{22}\ ,\ E^y_{22}$$

$$E^y_{32}\ ,\ E^x_{23}\ ,\ E^y_{23}\ ,\ E^x_{23}\ .$$

In addition, the E^y_{31} and the E^z_{31} modes are almost degenerate except

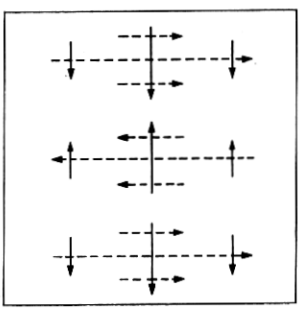

Fig. 12 — Field configuration of the E^y_{13} mode.

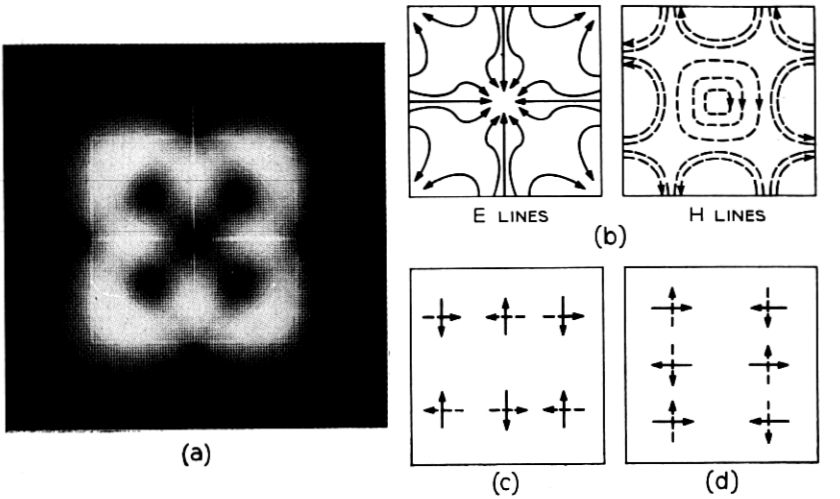

Fig. 13 — The E^{y}_{32} mode for unity aspect ratio: (a) intensity, (b) field configuration, (c) TE_{32}, and (d) TE_{23}.

near cutoff. The splitting of these modes can be accounted for by the differences of the field patterns shown in Fig. 11 and 12. Since the E^{x}_{31} mode reversals occur along the direction of the electric field lines, the electric field for this mode must have a larger longitudinal field component than for the E^{y}_{31} mode.

All degeneracies, except the $E^{y}_{mn} - E^{x}_{mn}$, are broken by a change in the aspect ratio as demonstrated in Fig. 17, which is drawn for the first 12 modes of a waveguide of aspect ratio 2. One interesting feature of this curve is the mode crossing of the E^{y}_{31} and E^{y}_{12} modes. Crossings of this type, which cannot occur in metallic waveguides, are possible because the field functions are frequency dependent. Qualitatively, it can be explained by noting that field reversals must take place in the core, therefore constraining the central lobe of the E^{y}_{31} more than any of the E^{y}_{12} mode lobes as cutoff is approached. Far from cutoff, however, all fields are well constrained and the E^{y}_{31} mode has a larger propagation constant than the E^{y}_{12} mode, as it does for the similar metallic waveguide mode with an aspect ratio of 2.

The effect of finite index difference on the modes can be observed by comparing Fig. 16, which is computed for unity aspect ratio and a vanishing index difference, with Fig. 18, which is computed for unity aspect ratio and a 0.5 index difference. The curves for modes whose

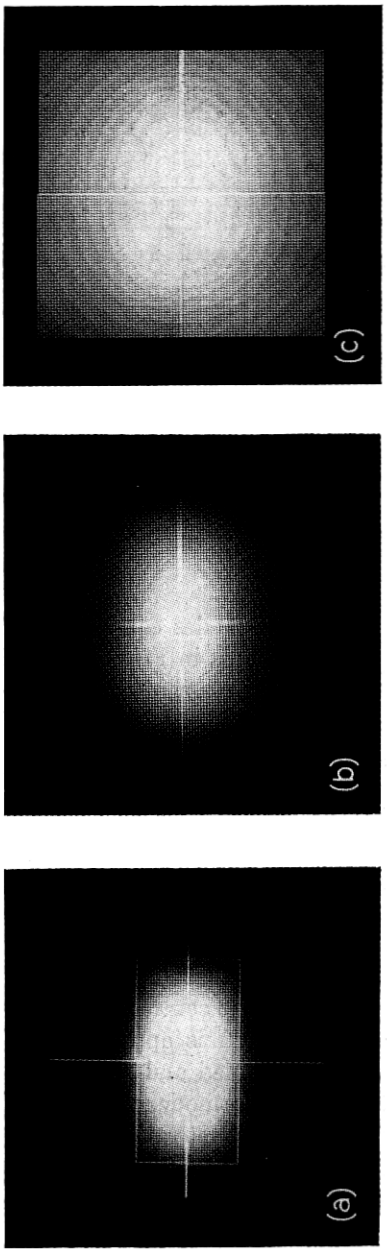

Fig. 14 — Intensity pictures of the E_{11}^{y} mode for (a) $\delta^2 = 0.81$, (b) $\delta^2 = 0.50$, and (c) $\delta^2 = 0.02$.

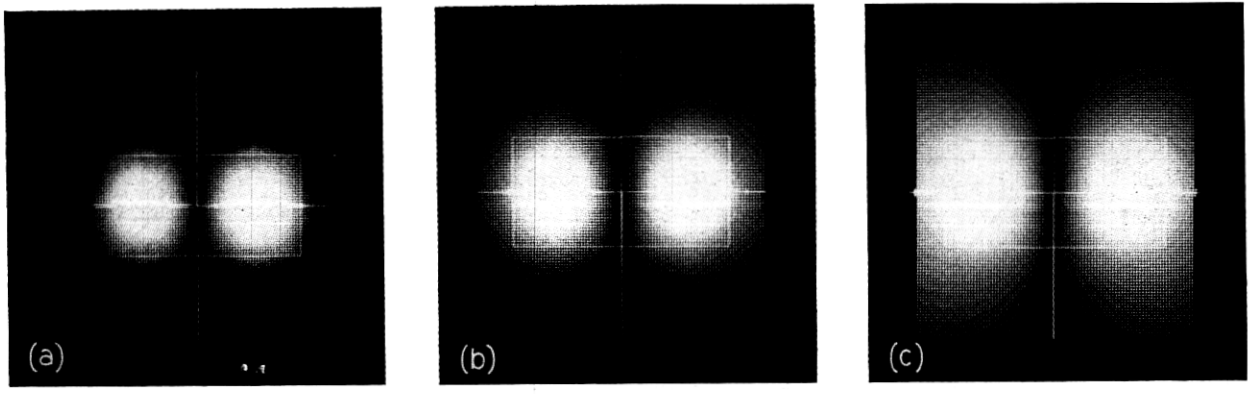

Fig. 15 — Intensity pictures of the E_{21}^{y} mode for (a) $\mathcal{P}^2 = 0.76$, (b) $\mathcal{P}^2 = 0.31$, and $\mathcal{P}^2 = 0.04$.

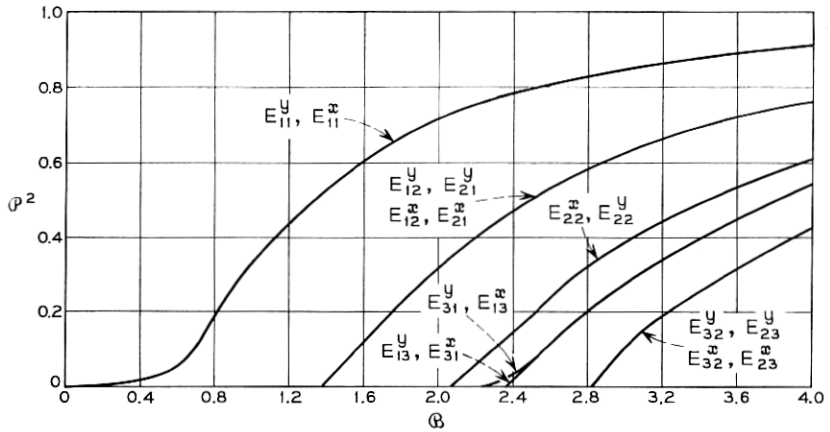

Fig. 16 — Propagation curves for the first 16 modes for unity aspect ratio and $\Delta n_r \to 0$.

field lines reverse direction across the origin are no longer degenerate, but those whose field lines do not reverse still are degenerate. For all degeneracies to be split, there must exist a finite index difference as well as an aspect ratio other than unity. Figure 19 illustrates one such case.

The effect of index difference on the degenerate principal modes for unity aspect ratio is examined in Fig. 20. The curve shows both a low and high index difference limit. In the range of interest for optical

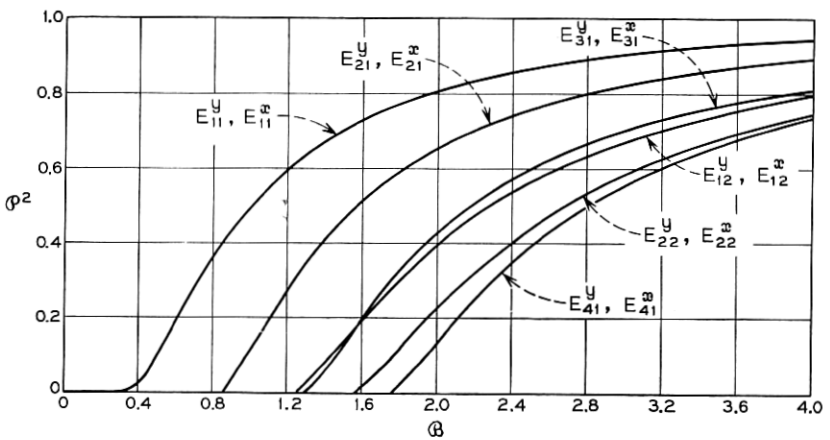

Fig. 17 — Propagation curves for the first 12 modes for $a/b = 2$ and $\Delta n_r \to 0$.

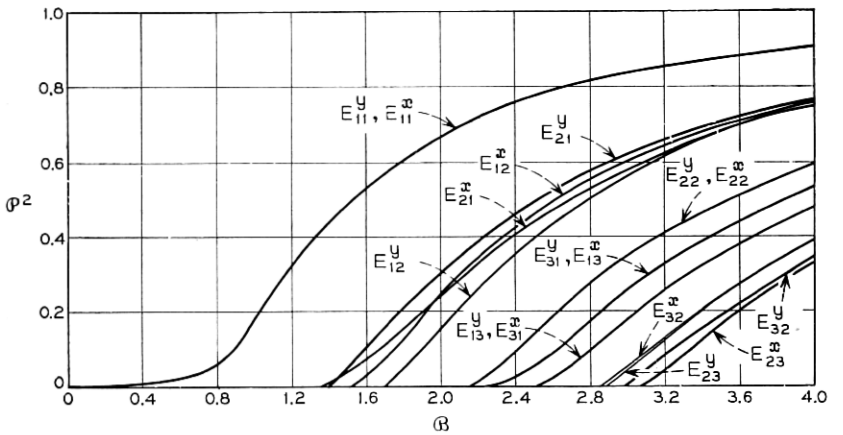

Fig. 18 — Propagation curves for the first 16 modes for unity aspect ratio and $\Delta n_r = 0.5$.

circuits (0 — 0.1) the vanishing difference curve is an excellent approximation. The greatest changes occur in the 0.1 — 10 range, which is the range of interest for some microwave problems.

Figure 21 presents the computed results for the effect of index changes on the principal modes for an aspect ratio of 2. The effect is much stronger on the E^y_{11} mode than the E^x_{11} mode. In fact, the effect on the E^x_{11} mode is comparatively small, except near cutoff.

The effect of aspect ratio on the principal modes is demonstrated for

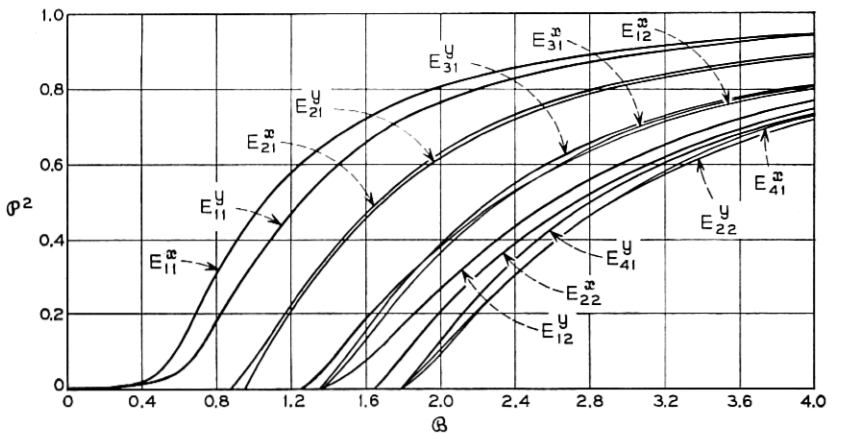

Fig. 19 — Propagation curves for the first 12 modes for $a/b = 2$ and $\Delta n_r = 0.5$.

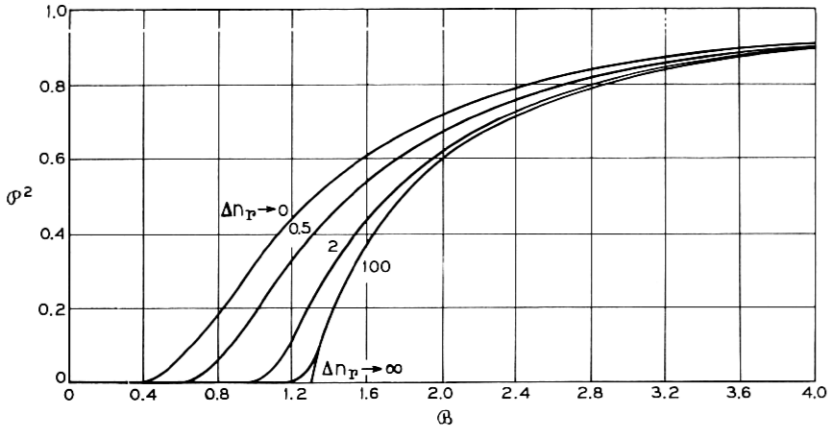

Fig. 20 — E_{11}^y and E_{11}^x mode propagation curves for several values of Δn_r with unity aspect ratio.

vanishing index difference in Fig. 22. The curve for infinite aspect ratio was obtained from the exact analysis of the slab case.[14]

IV. CONCLUSIONS

The results of the computations show that the circular harmonic method for analyzing rectangular dielectric waveguides gives excel-

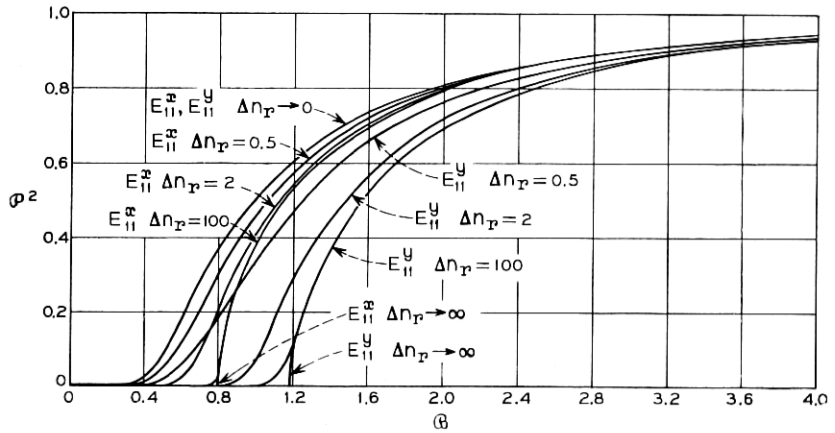

Fig. 21 — E_{11}^y and E_{11}^x mode propagation curves for several values of Δn_r with $a/b = 2$.

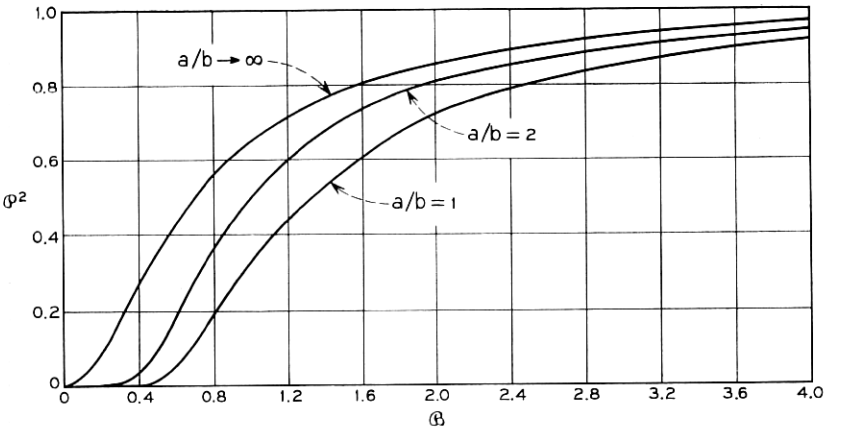

Fig. 22 — E_{11}^{y} and E_{11}^{x} mode propagation curves for several values of a/b with $\Delta n_r \to 0$.

lent results for waveguides of moderate aspect ratio. The convergence of the computed results was rapid and the results are in agreement with those of Marcatili's in the regions where his approximations apply. Furthermore, the results compare very well with Schlosser's curves for the principal mode.

Comparison of the results presented here with Marcatili's show that the two methods give values of the normalized propagation constant, $\mathcal{P}^2$, which are within a few percent for $\mathcal{P}^2 > 0.5$. Thus for $\mathcal{P}^2$ in this range his method is to be preferred since the calculations required are much simpler. However, for $\mathcal{P}^2 < 0.5$, and when it is desired to differentiate between modes for some of the near degenerate cases, another method must be used.

The circular harmonic analysis is attractive for small $\mathcal{P}^2$ because of the nature of the matching boundary. For large refractive index difference and moderate $\mathcal{P}^2$ both the method presented here and the one presented by Scholosser can be used.

V. ACKNOWLEDGMENTS

The author wishes to express his appreciation to T. Li and E. A. J. Marcatili for their valuable suggestions, to Mrs. C. L. Beattie for her aid in writing the plotting program, and to Mrs. E. Kershbaumer for her aid in writing the program for computing the propagation constants.

REFERENCES

1. Miller, S. E., "Integrated Optics: An introduction," this issue, pp. 2059–2069.
2. Kapany, N. S., *Fiber Optics,* New York: Academic Press, 1967, pp. 36–80.
3. Bracey, M. F., Cullen, A. L., Gillespie, E. F. F., and Staniforth, J. A., "Surface Wave Research in Sheffield," I.R.E. Trans. Antennas and Propagation, *AP-7,* No. 10 (December 1959), pp. 219–225.
4. Snitzer, E., "Cylindrical Dielectric Waveguide Modes," J. Opt. Soc. of Amer., *51,* No. 5 (May 1961), pp. 491–498.
5. Snitzer, E., and Osterberg, H., "Observed Dielectric Waveguide Modes in the Visible Spectrum," J. Opt. Soc. of Amer., *51,* No. 5 (May 1961), pp. 491–505.
6. Marcatili, E. A. J., "Dielectric Rectangular Waveguide and Directional Coupler for Integrated Optics," this issue, pp. 2071–2102.
7. Schlosser, W., and Unger, H. G., "Partially Filled Waveguides and Surface Waveguides of Rectangular Cross-Section," *Advances in Microwaves,* New York: Academic Press, 1966, pp. 319–387.
8. Shaw, C. B., French, B. T., and Warner, C. III, "Further Research on Optical Transmission Lines," Sci. Rep. No. 2, Contract AF449 (638)-1504 AD 625 501, Autonetics Report No. C7-929/501, pp. 13–44.
9. Stratton, J. A., "Electromagnetic Theory," New York: McGraw-Hill, 1941, p. 361.
10. Laxpati, S. R., and Mittra, R., "Energy Considerations in Open and Closed Waveguides," IEEE Trans. Antennas and Propagation, *AP-13,* No. 6 (November 1965), pp. 883–890.
11. Hamming, R. W., *Numerical Analysis for Scientists and Engineers,* New York: McGraw-Hill, 1962, pp. 81–82.
12. Freed, B. H., "Algorithm 41," Revision Evaluation of Determinant Comm. ACM, *6,* No. 9 (September 1963), p. 520.
13. "System/360 Scientific Subroutine Package," IBM, White Plains, N. Y., H20-0205-2, pp. 179–182.
14. Collin, R. E., "Field Theory of Guided Waves," New York: McGraw-Hill, 1960, pp. 480–495.