Propagation, Attenuation, and Dispersion Characteristics of Inhomogeneous Dielectric Slab Waveguides

EDWARD F. KUESTER, STUDENT MEMBER, IEEE, AND DAVID C. CHANG, MEMBER, IEEE

***Abstract*—A numerical method based upon invariant imbedding and the transverse impedance concept is applied to the problem of calculating various properties of inhomogeneous slab waveguides. The approach appears not only to be rapidly convergent but is also capable of giving arbitrary accuracy for any given mode. In particular, some interesting properties of a class of asymmetric profiles are pointed out, relating to discussions of lossy structures, temporal pulse distortion, and spatial broadening.**

I. INTRODUCTION

RECENTLY, some interest has been given to surface-wave modes guided by various inhomogeneous dielectric structures for use in optical communication and processing systems. Various transverse permittivity distributions have been used or proposed to improve delay distortion and radiation loss in dielectric waveguides over what is attainable in homogeneous structures [1], [2].

In the case of a single dielectric waveguide, analytical solutions are possible only for a few specific permittivity profiles in simple geometries [3], [4]. Asymptotic methods can also be applied in certain limiting cases [5], [6, pp. 193–215]. In general, however, the problem can only be solved numerically.

In this work, the properties of an inhomogeneous dielectric slab will be studied as a simple model of more realistic inhomogeneous structures. The slab considered is diagrammed in Fig. 1. The boundary regions 1 and 2 are homogenous and characterized by the constant permittivities ϵ_1 (for $x > 0$) and ϵ_2 ($x > d$), respectively. The slab, which lies between $0 \leq x \leq d$, has a dielectric constant which is an arbitrary function $\epsilon(x)$ of x on this interval, subject to the restrictions that $\epsilon(x) \geq \epsilon_0$ and only a finite number of discontinuities be allowed. All media have permeability μ_0 and are uniform and of infinite extent in the plane perpendicular to the x axis. We identify the z direction as the direction of propagation so that $\partial/\partial y \equiv 0$ for all quantities involved.

In a sourceless region, the fields are assumed to have a t and z dependence of $\exp[j(\omega t - \beta z)]$, where β is a yet-undetermined axial propagation constant. We have for a TE wave

Manuscript received April 1, 1974; revised July 29, 1974. This work was supported by the Air Force Office of Scientific Research (AFSC) under Grant AFOSR-72-2417.

The authors are with the Electromagnetics Laboratory, Department of Electrical Engineering, University of Colorado, Boulder, Colo. 80302.

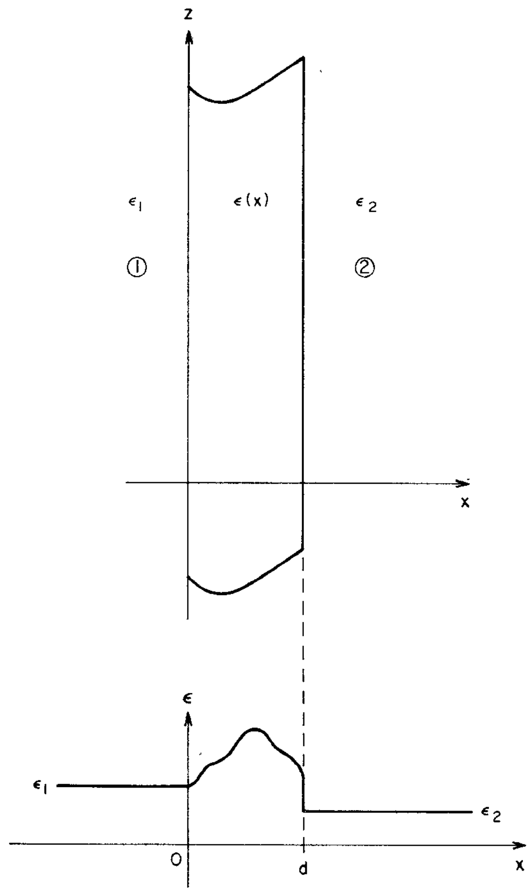

Fig. 1. Geometry of the problem.

$$\frac{d^2}{dx^2}E_y + (k^2 - \beta^2)E_y = 0 \tag{1}$$

$$H_x = -\frac{\beta}{\omega\mu_0} \qquad H_z = -\frac{1}{j\omega\mu_0}\frac{d}{dx}E_y \tag{2}$$

and for a TM wave

$$\frac{d}{dx}\left[\frac{1}{\epsilon(x)}\frac{d}{dx}H_y\right] + \frac{k^2 - \beta^2}{\epsilon(x)}H_y = 0 \tag{3}$$

$$E_x = \frac{\beta}{\omega\epsilon(x)}H_y \qquad E_z = \frac{1}{j\omega\epsilon(x)}\frac{d}{dx}H_y \tag{4}$$

where $k^2 = \omega^2\mu_0\epsilon$. At points where $\epsilon(x)$ is discontinuous, the tangential fields (E_y, H_z, H_y, E_z) are required to be continuous.

The numerical approach most often taken [7]–[9] is to approximate the inhomogeneous profile by a finite number of homogeneous slabs and to determine the modal characteristics from a system of simultaneous equations obtained from the wave solutions within the layers and the boundary conditions at the interfaces. Another method, using Hill's functions, has been implemented by Casey [10], but it ultimately is also an approximation since calculation of an infinite determinant is required, to be done numerically by suitable truncation. Marcuse [11] has used a kind of combination of these two approaches, approximating the profile by a piecewise linear one, and obtaining solutions within the linear sections in terms of Hankel functions of order one-third. In any of these methods, many layers (or terms of the determinant) may be used in a time-consuming solution without any satisfactory criterion for determining how many might be required to obtain a given accuracy.

In this paper, a different method, based upon the invariant imbedding principle and utilizing the transverse impedance concept, is developed to analyze the modal characteristics of dielectric slabs with arbitrary permittivity profiles. A first-order Ricatti differential equation for the impedance (or admittance) is formulated and numerically integrated across the slab using a fourth-order Runge–Kutta method with error estimation. The value of β of each propagating mode is then adjusted to satisfy the transverse resonance condition independently. It appears that this approach is not only rapidly convergent and highly accurate but can also provide useful insight into the design of dielectric waveguides.

The method will be derived for a lossless slab, then extended in a straightforward manner with the help of a perturbational formula to include guides with small loss. Finally, the propagation of pulses in inhomogeneous slabs will be investigated.

II. LOSSLESS SLABS

A. Formulation of the Modal Equation

We consider the surface-wave modes of the structure of Fig. 1 to be forward traveling and of arbitrary amplitude, which we will suitably normalize. We introduce the normalized quantities

$$\nu = \beta/k_0 \qquad W = k_0 x$$

where $k_0 = \omega(\mu_0\epsilon_0)^{1/2}$ and let $\epsilon_r = \epsilon(x)/\epsilon_0$, which allows us to write (1) and (3) with a single notation to cover both TE and TM modes.

$$\frac{d}{dW}\left(\frac{1}{K(W)}\frac{d}{dW}f(W)\right) + \frac{\gamma^2(W;\nu)}{K(W)}f(W) = 0 \quad (5)$$

where $\gamma^2(W;\nu) = \epsilon_r - \nu^2$, and $K(W)$ and $f(W)$ are given in Table I. Calling the boundary points $W_1 = 0$ and $W_2 = k_0 d$, we can express the solutions of (5) in regions (1) and (2) as

TABLE I
PARAMETER DEFINITIONS FOR TE AND TM WAVES

Symbol	TE	TM
$f(W)$	E_y	H_y
$f'(W)$	$-j\eta_0 H_z$	$j\epsilon_r E_z/\eta_0$
$K(W)$	1	ϵ_r
r_1	1	$\epsilon_r(W_1)/\epsilon_{r_1}$
r_2	-1	$-\epsilon_r(W_2)/\epsilon_{r_2}$
$R(W)$	$jE_y/(\eta_0 H_z)$	$-j\eta_0 H_y/E_z$
$S(W)$	$-j\eta_0 H_z/E_y$	$jE_z/(\eta_0 H_y)$

$$f_1 = A_1 \exp[\Gamma_1(W - W_1)], \qquad W < W_1$$
$$f_2 = A_2 \exp[-\Gamma_2(W - W_2)], \qquad W > W_2$$

where $\Gamma_i = +(\nu^2 - \epsilon_{r_i})^{1/2} > 0$, $i = 1,2$, assuring that the fields are exponentially decaying away from the slab, if $\nu^2 > \epsilon_{r_i}$, $i = 1,2$. At points of discontinuity of ϵ_r, we must have tangential field components continuous; this is equivalent to requiring that $f(W)$ and $f'(W)/K(W)$ be continuous. This gives rise to the boundary conditions

$$f'(W_i) = r_i\Gamma_i f(W_i), \qquad i = 1,2 \quad (6)$$

where the r_i are given in Table I. Equations (5) and (6) together constitute a two-point boundary/eigenvalue problem for the normalized propagation constant ν.

We now make us of a polar coordinate transformation due to Prüfer [12] by defining two new functions: $\Theta(W)$, the phase function and $\rho(W)$, the amplitude function:

$$f(W) = \rho(W)\sin\Theta(W)$$
$$f'(W)/K(W) = \rho(W)\cos\Theta(W) \quad (7)$$

which are uniquely defined if $\rho > 0$ and $0 \le \Theta(W_1) < \pi$ are specified. In addition, Θ and ρ are continuous regardless of discontinuities in ϵ_r.

Substituting (7) into (5) gives

$$\rho'(W) = \rho(W)\left\{\frac{\sin 2\Theta(W)}{2}\left[K(W) - \frac{\gamma^2(W;\nu)}{K(W)}\right]\right\} \quad (8)$$

$$\Theta'(W) = K(W)\cos^2\Theta(W) + \frac{\gamma^2(W;\nu)}{K(W)}\sin^2\Theta(W) \quad (9)$$

and the boundary conditions become

$$\Theta(W_1) = \operatorname{arccot}(q_1\Gamma_1), \qquad 0 < \Theta(W_1) \le \tfrac{1}{2}\pi \quad (10a)$$

$$\Theta(W_2) = \operatorname{arccot}(q_2\Gamma_2) + p\pi, \qquad \tfrac{1}{2}\pi \le \operatorname{arccot}(q_2\Gamma_2) < \pi \quad (10b)$$

where $q_i = r_i/K(W_i)$, $i = 1,2$, and p is (as yet) an arbitrary integer. The restrictions on the ranges of the inverse cotangents come from the assumptions made for the uniqueness of θ, as well as requiring that Γ_1 and Γ_2 be positive.

If we denote by $\Theta(W;\nu)$ the particular solution of (9) subject to the initial condition (10a), the boundary condition (10b) gives us the modal characteristic equation

$$P_p(\nu) = 0 \tag{11}$$

where the characteristic function $P_p(\nu)$ is given by

$$P_p(\nu) = \Theta(W_2;\nu) - \operatorname{arccot}(q_2\Gamma_2) - p\pi \tag{12}$$

such that the root ν of (11) is the propagation coefficient of the surface-wave mode corresponding to the integer p.

B. Maximum Number of Propagating Surface-Wave Modes

To examine the behavior of $P_p(\nu)$, it is convenient to define a function $\chi(W;\nu)$ to be $\partial\Theta(W;\nu)/\partial\nu$ so that

$$\chi'(W;\nu) = -2\left\{\frac{\rho'(W;\nu)}{\rho(W;\nu)}\chi(W;\nu) + \frac{\nu}{K(W)}\sin^2\Theta(W;\nu)\right\}. \tag{13}$$

In the last expression we have denoted by $\rho(W;\nu)$ the solution of (8) which satisfies the condition $\rho(W;\nu) = 1$. The initial condition for χ from (10a) is

$$\chi(W_1;\nu) = -\frac{\nu}{\Gamma_1}\frac{q_1}{1 + q_1^2\Gamma_1^2}. \tag{14}$$

Using the transformation $\chi = \psi/\rho^2$, we can show that

$$\chi(W;\nu) = -\frac{\nu}{\rho^2(W;\nu)}\left\{\frac{q_1}{\Gamma_1}\frac{1}{1 + q_1^2\Gamma_1^2} + 2\int_{W_1}^{W}\frac{\rho^2(\xi;\nu)}{K(\xi)}\sin^2\Theta(\xi;\nu)\,d\xi\right\}. \tag{15}$$

If $\nu > \nu_{\min} = \max\{(\epsilon_{r1})^{1/2}, (\epsilon_{r2})^{1/2}\}$, then Γ_1 and Γ_2 are both positive and real, and in particular $\chi(W;\nu)$ must then be negative for all W on the interval $[W_1,W_2]$ and all $\nu > \nu_{\min}$. This implies that

$$\Theta(W_2;\nu_{\min}) \geq \Theta(W_2;\nu)$$

for all $\nu > \nu_{\min}$, and furthermore, since for each p the boundary condition (10b) is an increasing function of ν, the characteristic function $P_p(\nu)$ is strictly decreasing for $\nu > \nu_{\min}$. Thus the integer

$$p_{\max} = \operatorname{int}[P_0(\nu_{\min})/\pi]$$

where "int" denotes the truncation or "greatest integer" function, gives the largest value of p for which (10b) can be satisfied (see Fig. 2).

If we further denote

$$\nu_{\max} = (\epsilon_{r\max})^{1/2} = \max_{0\leq x\leq d}[\epsilon_r(x)]^{1/2}$$

and suppose that $\nu > \nu_{\max}$, then $\gamma^2(W;\nu)$ is negative throughout the slab. Thus if Θ is to have a trajectory which crosses any of the lines $\Theta = (n + 1/2)\pi$, n an integer, then the slope at the point of crossing must be from (9):

$$\Theta' = \gamma^2(W;\nu)/K(W) < 0$$

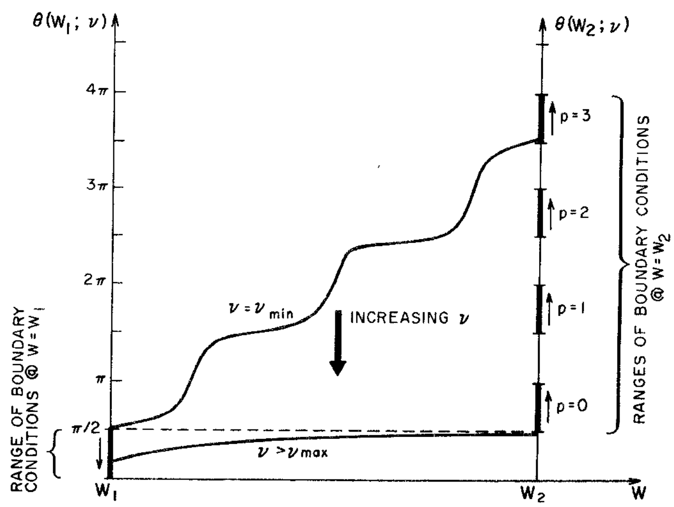

Fig. 2. Dependence of $\Theta(W)$ and the boundary conditions on the parameter ν. All arrows denote the direction of increasing ν.

but since $\Theta(W_1)$ is less than $\pi/2$ to begin with, all possible trajectories of Θ are bounded above by $\pi/2$. In a similar manner, if a Θ trajectory is to cross the line $\Theta = 0$, we must at the crossing point have

$$\Theta' = K(W) > 0$$

but since $\Theta(W_1) > 0$, all trajectories of Θ are bounded from below by 0. Thus for all $\nu > \nu_{\max}$, $0 < \Theta(W_2;\nu) < \pi/2$, and since no value assumed by the boundary condition (10b) falls within this range, no eigenvalues may exist for $\nu > \nu_{\max}$. Thus the only solutions of (11) will correspond to $p = 0,1,\cdots,p_{\max}$ and $p_{\max}$ is thus the order of the highest order propagating mode. For each p, since $f(W)$ has nodes when Θ is an integer multiple of π, an examination of Fig. 2 shows that exactly p such nodes will exist; consequently, the corresponding solution of (11) will be precisely the TE_p (or TM_p) node.

We further note that many other properties of the characteristic function are also obtainable from the behavior of χ and Θ. In particular, an important corollary to these properties is that an asymmetric waveguide with $\epsilon_1 \neq \epsilon_2$ must posses a nonzero cutoff frequency, even for the TE_0 and TM_0 nodes, as we show in [13]. This property turns out to have many interesting consequences as will be discussed later.

C. Transmission-Line Model

If we make the further transformations [14]

$$R(W) = \tan\Theta(W) = K(W)f(W)/f'(W)$$
$$S(W) = \cot\Theta(W) = f'(W)/[K(W)f(W)] \tag{16}$$

then, from Table I, it can be seen that R and S are a normalized wave impedance, and its reciprocal, a normalized wave admittance, the precise correspondence depending upon whether TE or TM nodes are being considered. It is easy to show from (9) that

$$S'(W) = -K(W)S^2(W) - \gamma^2(W;\nu)/K(W) \tag{17}$$
$$R'(W) = K(W) + \gamma^2(W;\nu)R^2(W)/K(W) \tag{18}$$

with initial condition from (10a):

$$S(W_1) = [R(W_1)]^{-1} = q_1\Gamma_1. \tag{19}$$

Equations (17) and (18) are Ricatti equations, which have appeared in many connections with wave propagation problems [6, pp. 215–233],[15]-[17]. It is, in fact, easy to show that within a layer of constant permittivity the solution to (17) is

$$S(W;\nu) = \frac{\gamma_i}{K_i}\frac{K_iS_i\cos\gamma_i(W-W_i) - \gamma_i\sin\gamma_i(W-W_i)}{K_iS_i\sin\gamma_i(W-W_i) + \gamma_i\cos\gamma_i(W-W_i)} \tag{20}$$

where

$$\gamma_i = [\gamma^2(W;\nu)]^{1/2} = (\epsilon_{r_i} - \nu^2)^{1/2}$$

$$K_i = K(W) \qquad S_i = S(W_i;\nu)$$

and ϵ_{r_i} is the relative permittivity of the layer, and W_i is any point in the layer, usually one of its boundary points. Equation (20) is the well-known impedance transformation for transmission lines, if we identify $-j\gamma_i/K_i$ as characteristic impedance, S_i as input impedance, and γ_i as phase constant. If we now regard the inhomogeneous slab as the limiting case of a sequence of piecewise constant layers whose thickness becomes infinitesimal, we may regard (17) and (18) as describing a nonuniform transmission line, and the boundary conditions $S(W_1) = q_1\Gamma_1$ and $S(W_2) = q_2\Gamma_2$ as terminating impedances, for which the eigenvalues ν represent a resonance condition between the line and its terminations.

D. Numerical Scheme for Determining ν

To solve the problem numerically, assume for the moment that $\Theta(W_2;\nu)$ and $\chi(W;\nu)$ are somehow known as functions of ν (we shall return to the question of how this is done later). Then we need to find the zero of $P_p(\nu)$ for the pth mode for each p between 0 and $p_{\max}$. Since

$$P_p'(\nu) = \chi(W_2;\nu) + \frac{\nu}{\Gamma_2}\frac{q_2}{1+q_2^2\Gamma_2^2} \tag{21}$$

we might attempt a solution by Newton's method, successively approximating the pth root $\nu = \nu^{(p)}$ by the formula

$$\nu_{n+1}^{(p)} = \nu_n^{(p)} - P_p(\nu_n)/P'(\nu_n) \tag{22}$$

where $\nu_n^{(p)}$ is the value of ν for the nth iteration. As is well known, unless ν_0 is initially close enough to $\nu^{(p)}$, this method may fail to converge. We thus modify Newton's method for this case as follows. Since $P_p'(\nu)$ is negative on the interval $[\nu_{\min},\nu_{\max}]$, the correction term in (22) is always of the correct sign. Thus at any stage of the procedure, an upper bound ν_u (at most, $\nu_{\max}$) and a lower bound ν_l (at least, $\nu_{\min}$) are available [12]. So if for some $\nu_n^{(p)}$, $\nu_{n+1}^{(p)}$ from (22) falls outside $[\nu_{\min},\nu_{\max}]$, instead of a step in Newton's method, a bisection step is performed:

$$\nu_{n+1}^{(p)} = (\nu_n^{(p)} + \nu_u)/2 \quad \text{or} \quad (\nu_l + \nu_n^{(p)})/2$$

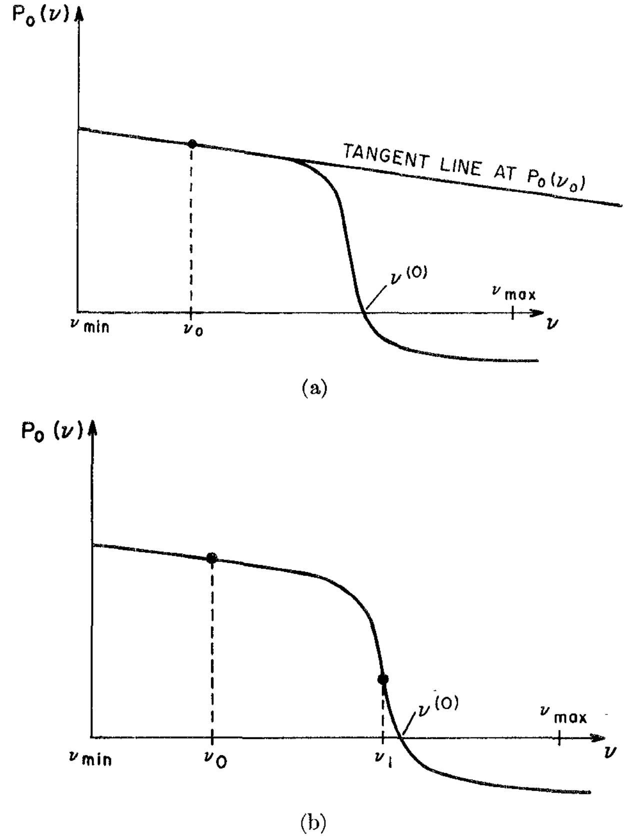

Fig. 3. Modification to Newton's method. (a) ν_1 from Newton's method overshoots $\nu_{\max}$. (b) ν_1 obtained from bisection.

according as the correction term from Newton's method is positive or negative in sign. At the same time ν_l or ν_u, respectively, is reset to $\nu_n^{(p)}$, so that the effect is to provide at least bisection-method convergence until the $\nu_n^{(p)}$ is close enough that Newton's method (with its quadratic convergence) takes over, finding $\nu^{(p)}$ to a specified accuracy in a minimum of computing time. This procedure is illustrated in Fig. 3 for a typical $P_0(\nu)$ curve.

To return to the task of evaluating $\Theta(W_2;\nu)$, we may calculate R and S from (17) and (18) to obtain Θ without the necessity of many time-consuming trigonometric function evaluations. Since R and S exhibit poles at $\Theta = (n + 1/2)\pi$ and $\Theta = n\pi$, respectively, by alternating between (17) and (18) so as to be dealing with R when S is near a pole and vice versa we may avoid poles and large values of R and S altogether [14]. To recover the multiple of π in Θ which is lost in (16), we note that S has only simple nonrepeated poles at $n\pi$, so that if m poles of S (that is, zeros of R) are encountered during a solution of (17) and (18), then

$$\Theta(W_2) = m\pi + \operatorname{arccot} S(W_2)$$

where $0 \le \operatorname{arccot} S < \pi$ is the correct reconstruction of Θ. Since also $\chi(W_2;\nu)$ is required in the Newton's method and since from (13) and (8) we have

$$\chi'(W;\nu) = \{[\gamma^2(W;\nu)/K(W) - K(W)]\cdot 2S(W;\nu)\chi(W;\nu) - 2\nu/K(W)\}/(1 + S^2(W;\nu)) \tag{23}$$

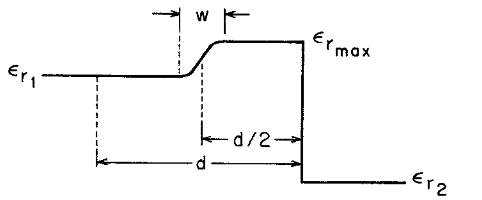

Fig. 4. Asymmetric permittivity profile of width d with cosinusoidal transition of width w.

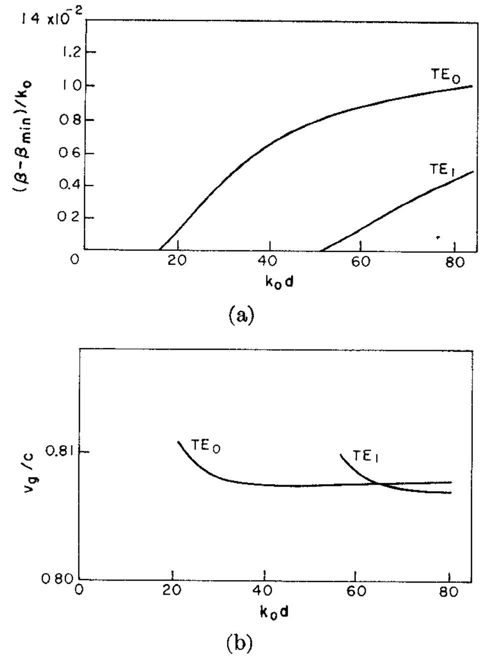

Fig. 5. TE modes for an asymmetric profile with cosinusoidal transition. $\epsilon_{r1} = 1.50$, $\epsilon_{r2} = 1.00$, $\epsilon_{r\max} = 1.53$, $d = 4 \times 10^{-6}$ $w/d = 0.2$, $\beta_{\min}/k_0 = 1.225$. (a) Propagation constant. (b) Group velocity.

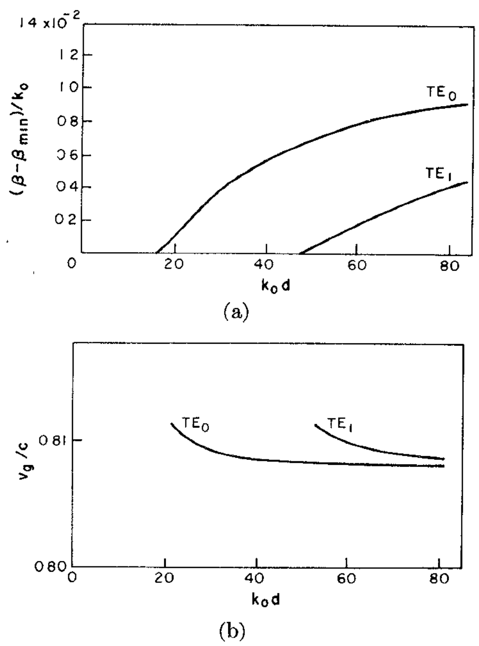

Fig. 6. TE modes for an asymmetric profile with cosinusoidal transition. $\epsilon_{r1} = 1.50$, $\epsilon_{r2} = 1.00$, $\epsilon_{r\max} = 1.53$, $d = 4 \times 10^{-6}$, $w/d = 0.4$, $\beta_{\min}/k_0 = 1.225$. (a) Propagation constant. (b) Group velocity.

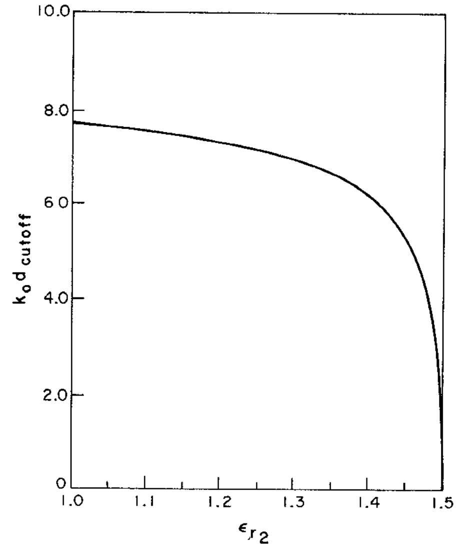

Fig. 7. TE_0 mode cutoff value of $k_0 d$ versus ϵ_{r2} for a uniform asymmetric profile ($w = 0$). $\epsilon_{r1} = 1.50$.

we can compute $\chi(W_2;\nu)$ in parallel with (17) and (18) which also requires no trigonometric function evaluations.

To obtain an error estimate for R and S (hence, eventually on $\nu^{(p)}$), a fourth-order Runge–Kutta method with a good error estimator [18] was chosen to numerically integrate the Ricatti equations (17) and (18). This estimate allows the step size to be adjusted automatically to maintain the truncation error at each step at an approximately constant preset value. In this way a definite overall error estimate is obtained, whereas conventional approximations by piecewise uniform slabs have no way of determining what errors are involved or how varying layer widths affects these errors.

E. Numerical Result for an Asymmetrical Slab

As mentioned previously, asymmetric structures must possess nonzero cutoff frequencies for TE_0 and TM_0 modes. In Fig. 4 is shown such a structure, and in Figs. 5 and 6 are shown dispersion curves and group velocity curves for the TE modes of this structure when $w = 0.2d$ and $0.4d$, respectively. (TM modes were found in all cases to differ only very slightly from the TE modes and are not presented here.) It is of particular interest to note that the group velocity curves for the two lowest order modes possess an intersection point in the case $w = 0.2d$ whereas this intersection has vanished for the case $w = 0.4d$. In fact, as w/d was varied from 0 to 1, the intersection point moved up in frequency until it disappeared at about $w/d = 0.3$, and the group velocity curves move farther apart as w/d approaches 1. This suggests that two-mode operation could be obtained at this intersection point with essentially only single-mode type distortion of signals, which is almost always negligible for practical purposes (see the following and [19]). It is conceivable that in the case of a dielectric waveguide with a diffused boundary, the actual diffusion length can be utilized for optimum multimode operation.

For $w = 0$, the dependence of the cutoff frequency of the TE_0 mode upon the degree of asymmetry was studied (Fig. 7). The extreme sensitivity of this dependence near symmetry suggests that in actual waveguides, the symmetry should be closely controlled; otherwise in single-mode operation, a slight asymmetry may produce (possibly intentionally) an effective "cutoff frequency" near the frequency of operation, so that the guided fields may radiate into the surrounding medium in the vicinity of the asymmetry. This asymmetry will also be seen to have a large effect upon lossy guides in the next section. We present results for other lossless profiles in [13].

III. LOSSY SLABS

When some or all of the media of the slab structure are allowed to be lossy, the propagation constants have imaginary parts which cause the modes to be attenuated as they propagate. Previous treatments have been limited to homogeneous structures [21], [22]. In [20], we extend the present method to lossy inhomogeneous structure of loss tangent up to about 10^{-3}, by means of an initial estimate of $\nu^{(p)}$ (now complex) provided by a perturbational formula making use of the fields of a corresponding lossless guide [23]. The following formula is obtained for a small perturbation $\Delta\epsilon_r(W)$ in the dielectric constant (in general, complex) for TE modes:

$$\nu \approx \nu_s + \frac{1}{2\nu_s}\left[\left(\int_{-\infty}^{+\infty} \Delta\epsilon_r E_{ys}^2\, dW\right) \Big/ \left(\int_{-\infty}^{+\infty} E_{ys}^2\, dW\right)\right] \tag{24}$$

with a similar, though more complicated, expression holding for TM modes. Here ν and ν_s are the perturbed and unperturbed propagation coefficients, respectively, and E_{ys} is the field in the unperturbed guide. If we stipulate that $\Delta\epsilon$ is constant in regions 1 and 2, the portions of the integrations in $W < W_1$ and $W > W_2$ may be done in closed form, and only a numerical integration across the slab need be done, with E_{ys} obtained from previous relations.

For the asymmetric permittivity profile of Fig. 4 with the loss profiles shown in Fig. 8, plots of the attenuation coefficient are given in Figs. 9 and 10 (the propagation constants remain essentially unaffected for loss tangents $<10^{-3}$). It is seen that the highest attenuation occurs when the fields are most highly concentrated in the region (s) of most loss. This is consistent with the results of [22] for homogeneous slabs. In Fig. 10, it is demonstrated that the attenuation characteristic may be altered by placing the maximum loss at other than the point of maximum dielectric constant. In principle one could produce almost any desired attenuation characteristic by adjusting the loss profile in this manner. It might prove useful, for instance, as a means of attenuating unwanted modes, but more practically, since loss in any material is inevitable, it merely serves to show that there are regions where it is least critical as to how much loss is present.

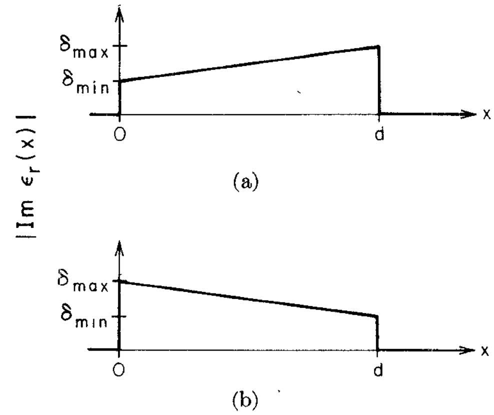

Fig. 8. Loss profiles used with permittivity profile of Fig. 4.

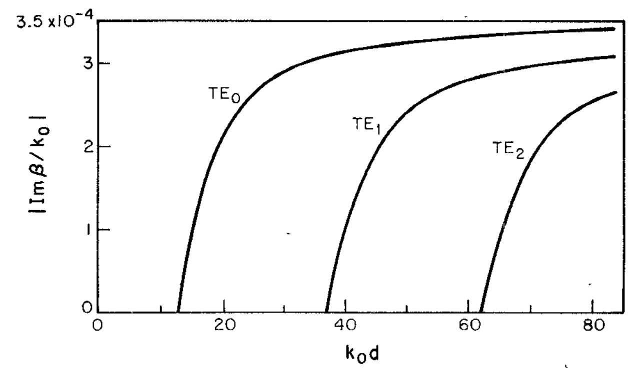

Fig. 9. TE mode attenuation coefficients for asymmetric profile with cosinusoidal transition, loss profile Fig. 8(a). $\epsilon_{r1} = 1.50$, $\epsilon_{r2} = 1.00$, $\epsilon_{r_{max}} = 1.53$, $\delta_{min} = 0.0005$, $\delta_{max} = 0.001$, $d = 4 \times 10^{-6}$ m, $w/d = 1.0$.

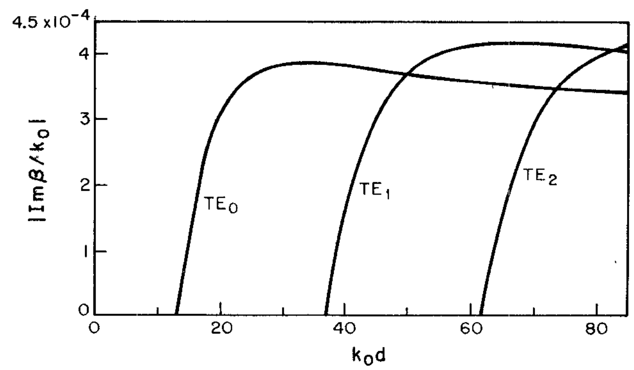

Fig. 10. TE mode attenuation coefficients for asymmetric profile with cosinusoidal transition, loss profile Fig. 8(b). $\epsilon_{r1} = 1.50$, $\epsilon_{r2} = 1.00$, $\epsilon_{r_{max}} = 1.53$, $\delta_{min} = 0.0005$, $\delta_{max} = 0.0015$, $d = 4 \times 10^{-6}$ m, $w/d = 1.0$.

To pursue this point further, we again let $w = 0$ and vary the extent of the asymmetry of the profile. In Figs. 11, 12, and 13, this asymmetry is plotted against the attenuation coefficient for the loss located in region 1, the slab, and region 2, respectively. As Re ϵ_{r2} ranges between 1.00 and 1.50, the attenuation decreases somewhat for the case where loss is assigned to region 1, increases by an order of magnitude for the loss in the slab, and by more than two orders of magnitude when the loss is in region 2. The field plots of an asymmetric slab given in

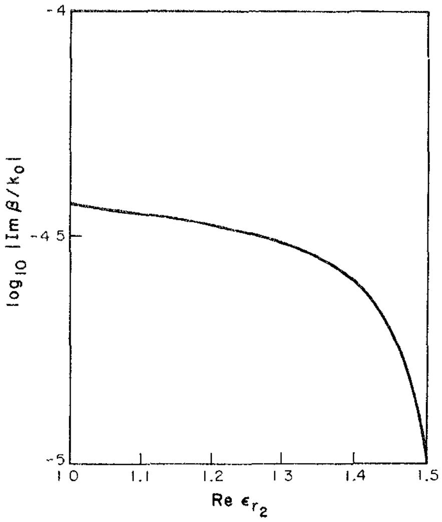

Fig. 11. TE_0 mode attenuation coefficient versus ϵ_{r2} for a uniform asymmetric profile ($w = 0$). $\epsilon_{r1} = 1.50$, $\epsilon_{r\max} = 1.53$, $k_0d = 8$, loss of $|\mathrm{Im}\,\epsilon_r| = 10^{-4}$ in region 1.

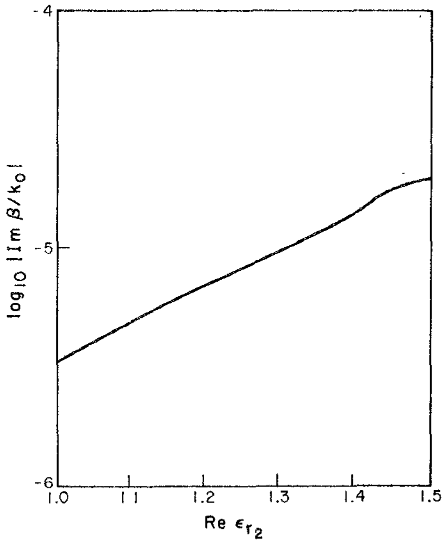

Fig. 12. TE_0 mode attenuation coefficient versus ϵ_{r2} for a uniform asymmetric profile ($w = 0$). $\epsilon_{r1} = 1.50$, $\epsilon_{r\max} = 1.53$, $k_0d = 8$, loss of $|\mathrm{Im}\,\epsilon_r| = 10^{-4}$ in slab.

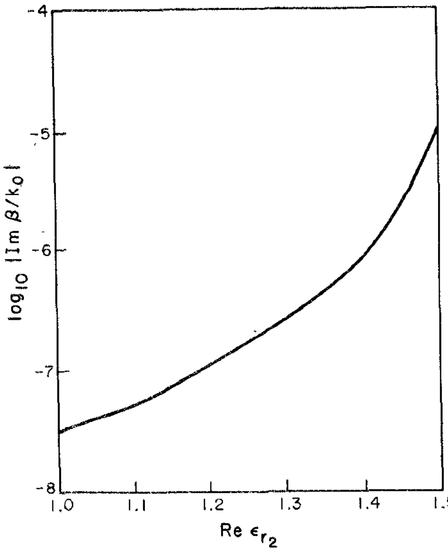

Fig. 13. TE_0 mode attenuation coefficient versus ϵ_{r2} for a uniform asymmetric profile ($w = 0$). $\epsilon_{r1} = 1.50$, $\epsilon_{r\max} = 1.53$, $k_0d = 8$, loss of $|\mathrm{Im}\,\epsilon_r| = 10^{-4}$ in region 2.

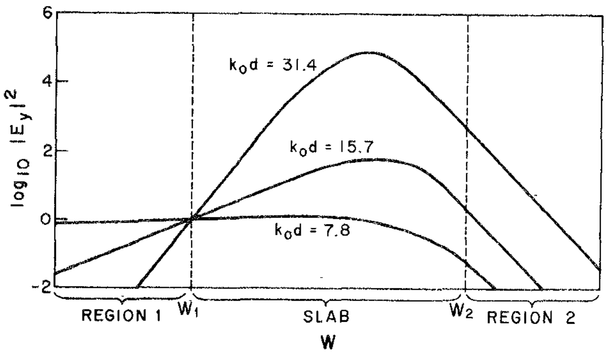

Fig. 14. Field plots of TE_0 mode normalized to $E_y = 1$ at $W = W_1$ for uniform asymmetric profile ($w = 0$). $\epsilon_{r1} = 1.50$, $\epsilon_{r2} = 1.00$, $\epsilon_{r\max} = 1.53$.

Fig. 14 show that at $W = W_2$ the field amplitude relative to that at $W = W_1$ can vary widely with the proximity to the cutoff frequency. It will be appreciated that similar disparity is not possible in a symmetric structure. This effect, which is most pronounced near cutoff, can be held accountable for the wide variations in attenuation coefficient caused by variation in ϵ_{r_2} (which is in a sense equivalent for fixed k_0d to moving k_0d relative to cutoff). It is therefore seen that in certain configurations the waveguide attenuation depends quite strongly on guide parameters not directly connected with the loss mechanism; rather than changing the amount of loss, this is accomplished by moving the field pattern relative to the locations of the losses.

IV. DISPERSION IN SLAB WAVEGUIDES

A. Temporal Dispersion

In principle, when one has found (either analytically or numerically) the dispersion characteristic (i.e., β as a function of ω) for a particular waveguide mode (we ignore the problems involved with multimode propagation, and refer the reader to [19]), one has solved the problem completely, since an arbitrary signal propagating down the guide in that mode may be analyzed using standard Fourier transform techniques in the frequency domain. In practice, however, one must nearly always resort to approximate methods. Kapron and Keck [24] have treated the propagation of a "quasi-monochromatic" pulse with Gaussian envelope (that is, of very narrow bandwidth) on an optical waveguide by a widely used method, but one whose limitations are not always pointed out. Subject to these limitations (see [29]) a Gaussian input pulse of width a is output (attenuated and with both group and phase delay) as a Gaussian pulse of width

$$b = [a_0^2 + (h_0''L/a_0)^2]^{1/2} \qquad a_0^2 = a^2 + \tfrac{1}{2}\alpha_0''L.$$

Here the propagation constant $\beta(\omega) = h(\omega) - (1/2)j\alpha(\omega)$, with $\alpha(\omega)$ and $h(\omega)$ both real and positive for $\omega > 0$, a subscript zero indicates evaluation at the carrier frequency ω_0, and the primes indicate differentiation with respect to ω. If $|(1/2)\alpha_0''| \ll |h_0''|$ (which was always true by at least a factor of 10 in the examples we studied [20]), then the length L_0 needed to achieve 10-percent pulse broadening is approximately

$$L_0 = a^2/2\,|\,h_0''\,|.$$

We present in [20] plots of $h''(\omega)$ and $\alpha''(\omega)$ for various lossless and lossy inhomogeneous slabs, and in all of these, $2\,|\,h''\,| = 10^{-24}$ s^2/m was not exceeded. For an acceptable transmission criterion of 10-percent broadening, or roughly $A_0 = 1/2\,|\,h_0''\,|$, a minimum pulse width of $a = 3 \times 10^{-11}$ s over a distance $L = 1$ km is possible—corresponding to a pulse rate of 30×10^9 pulses/s. At $L = 1$ m, the maximum pulse rate has already risen to 10^{12} pulses/s, so that the broadening attributable to the waveguiding structure for a single mode is unimportant for integrated optics, and only marginally important for long fiber guides.

Whenever a lossy structure was investigated, it was found that for a pulse to be broadened noticeably (say 10 percent) it was necessary for it to propagate such a long distance that it was essentially attenuated out of existence. Since in low permittivity contrast guides, the frequency dependence of the media themselves is much more important for single-mode pulse dispersion than the contribution from the waveguiding structure [28], one may use a simple atomic resonance model for this frequency dependence (as in [25] for example), again under narrow-bandwidth assumptions on the signal, to draw the same conclusion in a more general situation. We present a more detailed discussion of this point in [29].

Recently, a permittivity profile has been proposed [27] which seems to permit partial cancellation of the material dispersion by the structural dispersion; however, in view of the more serious attenuation and multimode distortion problems, this seems to be of largely academic interest.

B. Spatial Dispersion

Throughout this analysis, it has everywhere been assumed that guide properties and fields were invariant in the y direction. Since no real structure satisfies these conditions, we have analyzed the slab primarily as a model for more realistic three-dimensional structures. But recently, structures closely approximating slabs (i.e., with very large y dimensions) have been increasingly employed as waveguides in their own right. Thus it is no longer realistic to assume fields of infinite extent in the y direction, and it becomes more important to determine the behavior of fields localized in this dimension on a slab.

It is well known [30] that a field which has a Gaussian spatial distribution (in this case, in the y direction) of characteristic width $a = w/\sqrt{2}$, where w is the commonly defined half-width of the Gaussian beam, is broadened according to the formula

$$a(z) = a[1 + \delta^4]^{1/2} \tag{25}$$

where $\delta = (\beta r)^{1/2}/\beta a \simeq (\beta z)^{1/2}/\beta a$ is a measure of the distance the signal has traveled. Thus for $\delta \ll 1$, the signal is essentially unchanged, whereas for $\delta \gg 1$, the broadening is approximately linear:

$$a(z) \simeq z/\beta a.$$

These two limiting statements are true as well when the initial signal is exponentially distributed [20] (and it is interesting to note that the diffracted field for $\delta \gg 1$ is no longer exponential but Lorentzian). In fact, this type of dependence on the parameter δ is a quite general property of signals with a recognizable characteristic width a [26].

The function (25) is minimized when $\delta = 1$ [that is $a = (z/\beta)^{1/2}$] and takes the value $a\sqrt{2}$ at this point. Any wider or narrower initial signal gives a wider output signal at z. Thus if a detector with an effective acceptance width of $a_1\sqrt{2}$ is given, a "dispersion length" (or "diffraction length") L can be given as $L = \beta a_1^2$, beyond which a significant portion of the signal will be lost. To give some typical numbers, suppose $\beta = 10^7$ m^{-1} and $a_1 = 10^{-3}$ m. Then $L = 10$ m, which is relatively large for integrated optics considerations. However, reducing a_1 by a factor of 10 to 10^{-4} m reduces L to 10 cm, at which point spatial broadening begins to be of more serious concern.

V. CONCLUSION

The propagation characteristics of inhomogeneous dielectric slabs have been discussed. It was found that an asymmetric profile with a transition region exhibited extreme sensitivity to the degree of asymmetry present in the structure, and that for certain values of transition region width, two-mode operation at a single group velocity was possible. Loss profiles have also been studied, and it appears possible in nonsymmetrical structures to adjust attenuation by varying essentially nonloss-related parameters, in some cases over a very broad range. With relation to temporal pulse distortion, it is found that in nearly all practical cases, attenuation will be a more serious consideration than will single-mode broadening of the pulse. Finally, a criterion is given for the use of slablike structures for propagation of narrow-beam signals which suffer spatial dispersion in such guides.

ACKNOWLEDGMENT

The authors wish to thank Prof. F. S. Barnes, Prof. S. W. Maley, and Prof. L. Lewin for many helpful discussions. They would also like to thank M. Kindgren, M. Pan, and D. Hedden who at various times prepared the manuscript.

REFERENCES

[1] D. Gloge, "Optical waveguide transmission," *Proc. IEEE (Special Issue on Optical Communication)*, vol. 58, pp. 1513–1522, Oct. 1970.

[2] D. C. Chang and F. S. Barnes, "Radiation loss of curved axis dielectric slab waveguides," Univ. Colorado, Boulder, Tech. Rep. 2, Air Force Contract AF OSR-72-2417, 1973.

[3] H. Kirchhoff, "The solution of Maxwell's equation for inhomogeneous dielectric slabs," *Arch. Elek Übertragungstech.*, vol. 26, pp. 537–541, 1972.

[4] C. N. Kurtz and W. Streifer, "Guided waves in inhomogeneous focusing media, Part I: Formulation, solution for quadratic inhomogeneity," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 11–15, Jan. 1969.

[5] ——, "Guided waves in inhomogeneous focusing media—Part II: Asymptotic solution for general weak inhomogeneity," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 250–253, May 1969.

[6] L. M. Brekhovskikh, *Waves in Layered Media.* New York: Academic, 1960.

[7] W. R. Graham and W. J. Karzas, "Electromagnetic wave propagation along stratified media," Air Force Naval Lab., EMP Theoretical Note 67, Oct. 1971.

[8] P. J. B. Clarricoats and K. B. Chan, "Electromagnetic wave propagation along radially inhomogeneous dielectric cylinders," *Electron. Lett.*, vol. 6, pp. 694–695, 1970.
[9] Y. Suematsu and K. Furuya, "Propagation mode and scattering loss of a two-dimensional dielectric waveguide with gradual distribution of refractive index," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-20, pp. 524–531, Aug. 1972.
[10] K. F. Casey, "Radiation from a line source near an inhomogeneous layer," *IEEE Trans. Antennas Propagat.*, vol. AP-21, pp. 640–648, Sept. 1973.
[11] D. Marcuse, "TE modes of graded-index slab waveguides," *IEEE J. Quantum Electron.*, vol. QE-9, pp. 1000–1006, Oct. 1973.
[12] M. Godart, "An iterative method for the solution of eigenvalue problems," *Math. Comput.*, vol. 20, pp. 399–406, 1966.
[13] E. F. Kuester and D. C. Chang, "Modal and coupling characteristics of inhomogeneous dielectric slab waveguides—Part I: Lossless slabs," Univ. Colorado, Boulder, Sci. Rep. 1, Air Force Contract AFOSR-72-2417, 1973.
[14] M. R. Scott *et al.*, "Invariant imbedding and the calculation of eigenvalues for Sturm-Liouville systems," *Computing*, vol. 4, pp. 10–23, 1969.
[15] R. J. Garbacz, "Electromagnetic scattering from radially inhomogeneous spheres," *Proc. IRE* (Corresp.), vol. 50, p. 1837, Aug. 1962.
[16] G. Millington, "Ricatti form of the wave equation," *Electron. Lett.*, vol. 1, pp. 184–185, 1965.
[17] R. Bellman and R. Kalaba, "Functional equations, wave propagation, and invariant imbedding," *J. Math. Mech.*, vol. 8, pp. 683–704, 1959.
[18] L. F. Shampine and H. A. Watts, "Computing error estimates for Runge-Kutta methods," *Math. Comput.*, vol. 25, pp. 445–455, 1971.
[19] D. Gloge, A. R. Tynes, M. A. Duguay, and J. W. Hansen, "Picosecond pulse distortion in optical fibers," *IEEE J. Quantum Electron.* (*Special Issue on 1971 IEEE/OSA Conference on Laser Engineering and Applications: Part II of Two Parts*), vol. QE-8, pp. 217–221, Feb. 1972.
[20] E. F. Kuester and D. C. Chang, "Modal and coupling characteristics of inhomogeneous dielectric slab waveguides—Part II: Lossy slabs and pulse dispersion," Univ. Colorado, Boulder, Sci. Rep. 7, Air Force Contract AFOSR-72-2417, 1974.
[21] J. J. Burke, "Propagation constants of resonant waves on homogeneous isotropic slab waveguides," *Appl. Opt.*, vol. 9, pp. 2444–2452, 1970.
[22] J. G. Doidge, "Investigation of the theory of a thin film optical transmission line," Autonetics Div., NAA, Anaheim, Calif., Sci. Rep. 1, Air Force Contract AF 49(638)-1504, 1966.
[23] G. J. Gabriel and M. E. Brodwin, "The solution of guided waves in inhomogeneous anisotropic media by perturbation and variational methods," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-13, pp. 364–370, May 1965.
[24] F. P. Kapron and D. B. Keck, "Pulse transmission through a dielectric optical waveguide," *Appl. Opt.*, vol. 10, pp. 1519–1523, 1971.
[25] C. G. B. Garrett and A. E. McCumber, "Propagation of a Gaussian light pulse through an anomalous dispersion medium," *Phys. Rev. A*, vol. 1, pp. 305–313, 1970.
[26] P. C. Clemmow, *The Plane Wave Spectrum Representation of Electromagnetic Fields*. London, England: Pergamon, 1966, p. 76.
[27] S. Kawakami and S. Nishida, "Anomalous dispersion of new doubly clad optical fibre," *Electron. Lett.*, vol. 10, pp. 38–40, 1974.
[28] D. Gloge, "Dispersion in weakly guiding fibers," *Appl. Opt.*, vol. 10, pp. 2442–2445, 1971.
[29] E. F. Kuester and D. C. Chang, "Single-mode pulse dispersion in optical waveguides," to be published.
[30] D. Marcuse, *Light Transmission Optics*. New York: Van Nostrand Reinhold, 1972, ch. 6.

Propagation Effects in Optical Fibers

DETLEF GLOGE, MEMBER, IEEE

(Invited Paper)

***Abstract*—The round dielectric waveguide exhibits a surprising variety of characteristics that are not accurately inferable from the slab model. The forceful effort of recent years has grately extended the knowledge of these structures and added new and exciting modifications. An attempt to unify these results in a simplified picture is made. Specific phenomena relevant to optical fiber design and fabrication are then brought into focus. Some of the problems discussed are cross sectional loss variations, various core index profiles and the tolerances required in their preparation, the necessary cladding thickness, directional changes, and sources of mode coupling affecting signal distortion and loss.**

I. INTRODUCTION

THE LAST few years have seen a rapid increase both in technological know-how and theoretical understanding of optical fibers and, along with it, a new variety of fiber structures. At the same time, the issue of material loss, which had barred fibers from the communications field longer than necessary, was so convincingly solved that other aspects are now becoming a prime concern of optical communications research. This then seems to be a good time to take stock, to organize the knowledge gained, and to assess the available options. Accordingly, a great number of review articles have appeared in a rapid sequence all over the world. To name only the most recent in the order of their appearance, there is an article by Maurer [1] addressed mainly to the technology of fiber preparation. *Opto-Electronics* devoted its July and September issues of 1973 to the subject of fiber optics featuring reviews of the state of the art in Britain [2] and Germany [3]. An article by Ohnesorge [4] advanced some of the less conventional ideas of communication systems application for optical fibers. Miller *et al.* [5] have prepared a very comprehensive review of the current knowledge relating it to potential applications in the conventional communications network. The quite different though equally immediate potential of fibers for military applications becomes apparent in an article [6] which appeared in this TRANSACTIONS in December, 1973. The conventional technology of fiber bundles [7] seems to be more

Manuscript received March 27, 1974; revised August 15, 1974.
The author is with the Crawford Hill Laboratory, Bell Laboratories, Holmdel, N. J.