A General Transform for Regularizing Planar Open Waveguide Dispersion Relation

Aref Bakhtazad, Habibollah Abiri, and R. Ghayour, *Member, IEEE*

***Abstract*— Recently, conformal mapping technique has been introduced to unfold the characteristic equation of multilayered dielectric waveguides. The previous transformations have various deficiencies, i.e., the corresponding transformed functions are either multivalued, or have a singular point. Also one of them is not generally applicable. We introduce a new mapping that has none of the above deficiencies. Furthermore, in the new plane, the modal zones are rectangular strips which facilitate the implementation of the corresponding numerical routines. By this transformation, various modes of general planar dielectric waveguides can be traced easily.**

***Index Terms*— Dielectric slab waveguides, unfolding the dispersion equation.**

I. Introduction

WAVE propagation in planar multilayered media, can be formulated conveniently by multiplication of 2×2 transfer matrices of the layers [1]. However, due to different modal behavior of lateral layers (substrate and cadding), the resultant characteristic equation is multivalued [2]. An efficient procedure for finding the zeros of the characteristic equation (propagation constants) is known as Argument Principal Method (APM) [2], [3]. This method uses integration over a closed loop on the complex plane, encircling the zeros of the characteristic equation. Note that APM method can be implemented only on contours over which the characteristic function is regular (i.e., analytic and single valued).

It is conventional to make the dispersion function single-valued by imposing constraints on the sign of either the real or the imaginary parts of the outer layers transverse wave numbers $\alpha_{\rm cl}$ and $\alpha_{\rm sub}$, referring to cladding and substrate, respectively. These constraints correspond to a particular physical behavior (growing or decaying field amplitude, or the direction of transverse power flow) in the outer layers. This is equivalent to selecting one of the four sheets of the Reimann surface and discarding the other three. Although such a selection yields a single-valued function, it has undesired consequences.

The first consequence is the discontinuity of the selected single-valued function along the lines (branch points) where it was separated from the four-valued surface. Therefore, the application of numerical techniques which rely upon the function analyticity will be difficult or at least inconvenient due to the presence of this discontinuous behavior (e.g., any contour of APM method must avoid such lines [3]). These discontinuities are specially troublesome since "cuton" will occur across these lines [5]. New solutions (modes) will usually appear from these discontinuities when the physical configuration of the guide or the wavelength is varied.

Another problem is that no single branch cut can simultaneously allow for the determination of all solutions (e.g., it is not possible to determine both the leaky and absorptive modes in a guide containing an absorbing material by selecting a single branch cut). Therefore, we must use one cut to look for one type of mode and a second cut to look for another. However, it usually causes confusion which may lead to missed or misinterpreted solutions.

These problems can be avoided, if a conformal mapping which unfolds the four-sheeted Reimann surface, is used. The unfolding retains all the regions found on the four sheeted form and results in a new single valued form for the dispersion function [2], [4], [5]. Therefore, in the view of APM method, the integration over any contours in the Reimann surface is mapped to the integration over one in the conventional unfolded surface.

So far, the authors who have proposed the conformal mapping technique, have used transformations leading to functions, which are not single valued in the entire complex plane [2], or have a singular point [4], and or are not generally applicable [5]. In Section II, we have introduced a transformation without the above deficiencies. The level curves are also plotted to illustrate the relation between the two spaces. In this way, any contour of $X =$ const. and $Y =$ const. in the new plane, can be easily traced to a contour in the four folded Reimann surface. In Section III, the modal behaviors in various regions of the new plane are discussed. It is shown that the new plane can provide a whole picture of the modal behavior. This view hides itself in various folds of the ordinary β^2 plane. In Section IV, the mode migration plots [5] in the new plane for simple symmetrical and asymmetrical three layers dielectric waveguides with or without core attenuation, are plotted. We have shown that the unique property of the new transform is that it can be used for any type of planar dielectric waveguides. Finally, conclusion is presented in Section V.

II. The General Regularizer Conformal Mapping

An $N + 2$ layers dielectric waveguide shown in Fig. 1, consists of a stack of N layers (called core region) sandwiched between two semi-infinite substrate and cladding regions. Assume that the propagation is along the z axis. The waveguide

Manuscript received February 16, 1996; revised August 23, 1996. This work was supported by the Board of Research of Shiraz University, Shiraz, Iran, under Grant 67-EN-519-276.

The authors are with the Electrical Engineering Department, School of Engineering, Shiraz University, Shiraz, Iran.

Publisher Item Identifier S 0733-8724(97)01375-3.

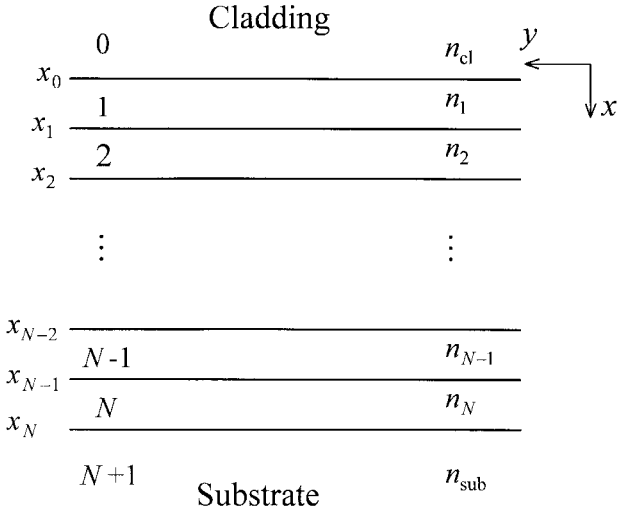


Fig. 1. $N+2$ layers dielectric stack, laying between the semi-infinite cladding and substrate media.

is uniform along the y axis (the fields are independent of y). Furthermore, we assume that the refractive index of each layer n_j, is a constant independent of x (it may be complex), thus the wave amplitude in the jth layer can be written as

$$\Psi_j = A_j \exp[\alpha_j(x-x_j)] + B_j \exp[-\alpha_j(x-x_j)] \qquad (1)$$

where the factor $\exp[i(\beta z - \omega t)]$ has been omitted throughout and $\alpha_j = \sqrt{\beta^2 - k_0^2 n_j^2}$, with k_0 as the wave number in vacuum. The wave behaviors in the zero (cladding) and the $(N+1)$st (substrate) layers are assumed to be in the form of $\exp[\alpha_{\rm cl}(x-x_0)]$, and $\exp[-\alpha_{\rm sub}(x-x_N)]$, respectively. The transfer matrix $\mathbf{M}_j$ relates the field amplitudes $[U\ V]^T$ at the beginning of the succeeding layers, according to the relation $[U_{j-1}\ V_{j-1}]^T = \mathbf{M}_j[U_j\ V_j]^T$. We take $U = E_y$ and $V = H_z$ for TE modes while $U = H_y$ and $V = -E_z$ for TM modes (These selections have been confused in [2]). Now it can be shown that $\mathbf{M}_j$, for a uniform layer of width $\Delta_j = x_j - x_{j-1}$, is written as

$$\mathbf{M}_j = \begin{bmatrix} \cosh(\alpha_j\Delta_j) & -\gamma_j^{-1}\sinh(\alpha_j\Delta_j) \\ -\gamma_j \sinh(\alpha_j\Delta_j) & \cosh(\alpha_j\Delta_j) \end{bmatrix} \qquad (2)$$

where $\gamma_j = -i\alpha_j\sqrt{\varepsilon_0/\mu_0}/k_0$ for TE modes and $\gamma_j = -i\alpha_j\sqrt{\mu_0/\varepsilon_0}/n_j^2 k_0$ for TM modes. Using the elements of the total transfer matrix $\mathbf{m}, (\mathbf{m} = \prod_{j=1}^{N}\mathbf{M}_j)$ the characteristic equation can be written as [1]

$$F(\beta^2) = \gamma_{\rm cl} m_{11} - \gamma_{\rm cl}\gamma_{\rm sub} m_{12} - m_{21} + \gamma_{\rm sub} m_{22}. \qquad (3)$$

In the above equation β appears only in β^2 form, so we choose β^2 as an independent complex variable [2].

The four-valuedness of $F(\beta^2)$ in (3) arises from the square-root function defining $\gamma_{\rm cl}$ and $\gamma_{\rm sub}$. For a given value of β^2, the four combinations of signs of $\pm\alpha_{\rm cl}$ and $\pm\alpha_{\rm sub}$ correspond to the four values of the function. The elements of $\mathbf{M}_j$ [in (2)] are all even functions of α_j, thus the sign of the square-root function giving α_j has no effect on the dispersion function. Note also that the elements of the total transfer matrix $\mathbf{m} = \prod_{j=1}^{N}\mathbf{M}_j$ are well defined when α_j goes to zero, and it shows that the points $\beta^2 = \alpha_j^2$ are removable singularities of $F(\beta^2)$. However, the signs of $\gamma_{\rm cl}$ and $\gamma_{\rm sub}$ in the dispersion function (3) remain ambiguous, and therefore $\beta^2 = \alpha_{\rm cl}^2$ and $\beta^2 = \alpha_{\rm sub}^2$ remain branch points of $F(\beta^2)$. More rigorously, the β^2 complex plane can be considered as a four folded Reimann surface with $\beta^2 = n_{\rm cl}^2 k_0^2$ and $\beta^2 = n_{\rm sub}^2 k_0^2$ as its branch points.

Note that in this folded plane, the characteristic function is single valued, and APM can be implemented over any contour in this plane. However, the complex integration in this folded plane is not an easy task. Now we look for a conformal transformation that maps this folded plane (β^2 plane) to a new unfolded one (Z plane) in such a manner that the characteristic equation be single valued, while it remains analytic over the entire Z plane. Indeed, in this way the complex integration over any contours on a the folded plane is replaced by an integration over a conventional plane.

For convenience, we normalize $\beta, \alpha_{\rm cl}$, and $\alpha_{\rm sub}$ versus k_0, i.e.,

$$\bar{\beta} = \beta/k_0 \equiv n_{\rm eff}, \quad \bar{\alpha}_{\rm cl} = \alpha_{\rm cl}/k_0 \ \text{ and } \bar{\alpha}_{\rm sub} = \alpha_{\rm sub}/k_0. \qquad (4)$$

The mapping $\bar{\beta}^2 = w(Z)$ must be chosen so that $g(Z)$ and $h(Z)$ defined by the following two relations

$$g^2(Z) = \bar{\alpha}_{\rm cl}^2 = w(Z) - n_{\rm cl}^2 \qquad (5a)$$

$$h^2(Z) = \bar{\alpha}_{\rm sub}^2 = w(Z) - n_{\rm sub}^2 \qquad (5b)$$

are regular over the entire Z plane. The following functions satisfy the above conditions:

$$g(Z) = \exp(\eta Z) + R\exp(-\eta Z) \qquad (6a)$$

$$h(Z) = \exp(\eta Z) - R\exp(-\eta Z) \qquad (6b)$$

where η is a positive arbitrary constant and

$$R = \frac{n_{\rm sub}^2 - n_{\rm cl}^2}{4}. \qquad (6c)$$

Thus, the function $w(Z)$ will be

$$w(Z) = \exp(2\eta Z) + R^2\exp(-2\eta Z) + S \qquad (7a)$$

where

$$S = \frac{n_{\rm sub}^2 + n_{\rm cl}^2}{2}. \qquad (7b)$$

Now, the inverse transformation can be written as

$$Z = w^{-1}(\bar{\beta}^2) = \frac{1}{\eta}\ln\left(\frac{\bar{\alpha}_{\rm sub} + \bar{\alpha}_{\rm cl}}{2}\right) + i\frac{2\pi k}{\eta}, \quad (k = 0, \pm 1, \pm 2, \cdots). \qquad (8)$$

The above mapping is periodic with respect to the imaginary part of Z with period $2\pi/\eta$. Taking the principal value of the complex logarithm ($k = 0$), we can confine the case to the first period of the transformation. The domain of $\Phi(Z) = F(w(Z))$ can be considered as a horizontal strip with the width of $2\pi/\eta$. It is more convenient to connect the top of the strip to its bottom in order to form a cylinder, on which the function $\Phi(Z)$ is regular. Each point in this domain unambiguously corresponds to a point on the Reimann surface of the multivalued dispersion function $F(\bar{\beta}^2)$ and vice versa.

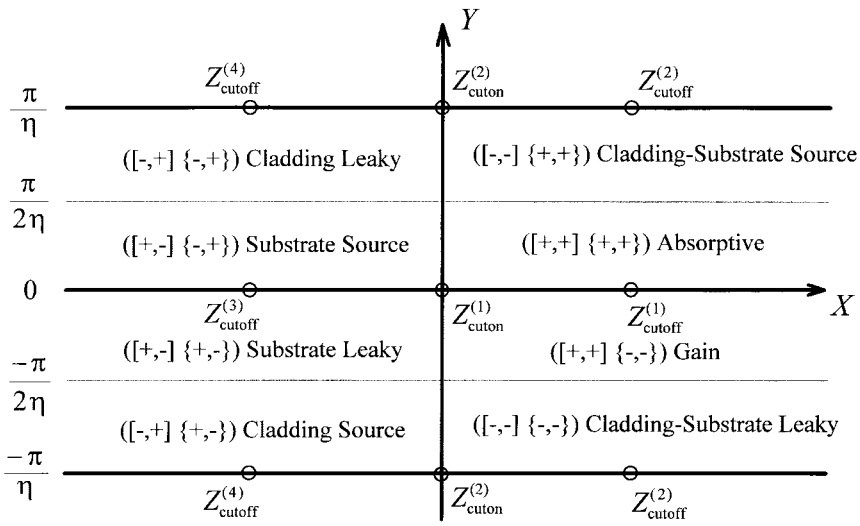


Fig. 2. Complex Z plane labeled in terms of the types of modes in various regions of β^2 plane.

Now, it is sufficient to find the zeros of the function $\Phi(Z)$, which is regular on the entire complex Z plane. The inverse of the transformation (7a) can also be written as

$$Z = \frac{1}{2\eta}\ln\left(\frac{\bar{\beta}^2 - S \pm \sqrt{(\bar{\beta}^2 - S)^2 - 4R^2}}{2}\right) \pm i\frac{\pi k}{\eta}, \quad (k = 0, 1, 2, 3, \cdots) \tag{9}$$

As it is easily verified, the above transformation leads to four distinct Z values in each period $(2\pi/\eta)$ for each value of $\bar{\beta}^2$. This is originated from the four-valuedness of the original plane. Generally, each of these distinct Z values corresponds to a mode located on a fold of the Reimann surface. In Fig. 2 the domain is labeled in terms of the signs of real and imaginary parts of α_{sub} and α_{cl}, where the notations in the parentheses are the signs of $([\text{Re}(\alpha_{\text{cl}}), \text{Re}(\alpha_{\text{sub}})]\{\text{Im}(\alpha_{\text{cl}}), \text{Im}(\alpha_{\text{sub}})\})$. In this figure we have assumed that $\text{Im}(n^2_{\text{sub}}) = \text{Im}(n^2_{\text{cl}}) = 0$, and $\text{Re}(n^2_{\text{sub}}) > \text{Re}(n^2_{\text{cl}})$. Note that due to the periodic nature of the transformation, the points located on the lower and upper boundaries are not distinct, but correspond to the same mode.

The cuton point (or the lower cutoff point) of the bound modes for a lossless waveguide is located at

$$Z^{(1)}_{\text{cuton}} = \frac{\ln R}{2\eta} \tag{10a}$$

which corresponds to $n_{\text{eff}} = n_{\text{sub}}$. Another cuton point can be distinguished as

$$Z^{(2)}_{\text{cuton}} = Z^{(1)}_{\text{cuton}} \pm i\pi/\eta \tag{10b}$$

where the principal period is assumed. The cutoff point (or the upper cutoff point) of the bound modes is located at

$$Z^{(1)}_{\text{cutoff}} = \frac{1}{2\eta}\ln\left(\frac{n^2_{\max} - S + \sqrt{(n^2_{\max} - S)^2 - 4R^2}}{2}\right) \tag{11a}$$

where $n_{\max}$ denotes the maximum of the refractive index profile (we assume further that $\text{Re}(n^2_{\max}) > n^2_{\text{sub}}$). The cutoff points of the other types of modes ($\text{Re}(\alpha_{\text{cl}})$ and $\text{Re}(\alpha_{\text{sub}}) < 0$) can be distinguished as

$$Z^{(2)}_{\text{cutoff}} = Z^{(1)}_{\text{cutoff}} \pm i\pi/\eta. \tag{11b}$$

Cutoff point of the leaky modes with $\text{Re}(\alpha_{\text{cl}}) > 0$ and $\text{Re}(\alpha_{\text{sub}}) < 0$ is located at

$$Z^{(3)}_{\text{cutoff}} = \frac{1}{2\eta}\ln\left(\frac{n^2_{\max} - S - \sqrt{(n^2_{\max} - S)^2 - 4R^2}}{2}\right) \tag{12a}$$

and similarly for modes with $\text{Re}(\alpha_{\text{cl}}) < 0$ and $\text{Re}(\alpha_{\text{sub}}) > 0$

$$Z^{(4)}_{\text{cutoff}} = Z^{(3)}_{\text{cutoff}} \pm i\pi/\eta. \tag{12b}$$

These points are marked in Fig. 2 for a transparent waveguide. The following relations also hold

$$Z^{(1)}_{\text{cutoff}} + Z^{(3)}_{\text{cutoff}} = 2Z^{(1)}_{\text{cuton}} \tag{13a}$$

$$Z^{(2)}_{\text{cutoff}} + Z^{(4)}_{\text{cutoff}} = 2Z^{(2)}_{\text{cuton}}. \tag{13b}$$

The maps of the branch cuts in the new plane are the following semi-lines

$$\begin{cases} \text{Im}(Z) = 0, & \text{Re}(Z) \geq Z^{(1)}_{\text{cuton}} \\ \text{Im}(Z) = 0, & \text{Re}(Z) < Z^{(1)}_{\text{cuton}} \\ \text{Im}(Z) = \pm\pi/\eta, & \text{Re}(Z) \geq Z^{(1)}_{\text{cuton}} \\ \text{Im}(Z) = \pm\pi/\eta, & \text{Re}(Z) < Z^{(1)}_{\text{cuton}}. \end{cases} \tag{14}$$

The four regions (quadrants) between these semi-lines and the vertical line of $\text{Re}(Z) = Z^{(1)}_{\text{cuton}} = \ln R/2\eta$ (see Fig. 2), correspond to the folds of the four folded $\bar{\beta}^2$ Reimann surface. The path of integration between any of the above quadrants, is equivalent to the path between folds in the previous $\bar{\beta}^2$ plane. This justifies the simplifying nature of mapping for contour integration in APM method. Furthermore, the strip modal zones facilitate the APM implementation [3].

If $u = \text{Re}(\bar{\beta}^2)$ and $v = \text{Im}(\bar{\beta}^2)$ are variables in Cartesian coordinates, then $X = \text{Re}(Z)$ and $Y = \text{Im}(Z)$ in transformation (7) would be the variables in the elliptic coordinates

as

$$u = 2R\cosh\left[2\eta\left(X - Z^{(1)}_{\text{cuton}}\right)\right]\cos 2\eta Y + S \quad (15a)$$

$$v = 2R\sinh\left[2\eta\left(X - Z^{(1)}_{\text{cuton}}\right)\right]\sin 2\eta Y. \quad (15b)$$

The level curves corresponding to $X = \text{const.}$ are confocal ellipses with the two principal axes parallel to u and v axes. The focuses are located at $(n^2_{\text{sub}}, 0)$ and $(n^2_{\text{cl}}, 0)$, where the lengths of the major and minor axes are $2[\exp(2\eta X) + R^2\exp(-2\eta X)]$ and $2|\exp(2\eta X) - R^2\exp(-2\eta X)|$, respectively. Note that at $X = Z^{(1)}_{\text{cuton}} = \ln \cdot R/2\eta$, the minor axis of the level ellipse is zero. The level ellipses corresponding to $X > Z^{(1)}_{\text{cuton}}$, belong to the points on the first and fourth quadrants of complex Z plane (see Fig. 2), whereas the level ellipses corresponding to $X < Z^{(1)}_{\text{cuton}}$ map the second and third quadrants. The level curves for $Y = \text{const.}$ are hyperbolas with the same focuses as the level ellipses, and the length of transverse and conjugate axes are $4R|\cos(2\eta Y)|$ and $4R|\sin(2\eta Y)|$, respectively.

In Fig. 3, with $\eta = 1, n_{\text{cl}} = 1$, and $n_{\text{sub}} = 1.5$, the level curves of transformation (7) for $Y \le \pi/4$ are plotted in the $\bar{\beta}^2$ plane (the conventional $\bar{\beta}^2$ plane). Note that the level curves are symmetrical around $u = S$ (which corresponds to the level curve of $Y = \pi/4$). Generally, the level ellipses for $X - Z^{(1)}_{\text{cuton}} > 0$ and $X - Z^{(1)}_{\text{cuton}} < 0$ are not identical. However, in Fig. 3 $|X - Z^{(1)}_{\text{cuton}}| \ll 1$, so they are not distinguishable. It is interesting to note that such $\bar{\beta}^2$ plane is an image of all folds of the $\bar{\beta}^2$ Reimann surface into an arbitrary branch cut.

For the special case of $n^2_{\text{sub}} = n^2_{\text{cl}}$, even though $R = 0$, the transformation (7) remains applicable, i.e.,

$$w(Z) = \exp(2\eta Z) + S \quad (16)$$

whereas, the transformation used in [5], can not be applied. For this case, the cuton points together with the third and the fourth cutoff points move to $-\infty$ as R approaches zero [see (10) and (12)]. Indeed in this case, the $\bar{\beta}^2$ Reimann plane has only two folds and the modal zones are limited to the right handed four regions of Fig. 2. The inverse transformation is then

$$Z = \frac{\ln \alpha_{\text{sub}}}{\eta} = \frac{\ln \alpha_{\text{cl}}}{\eta} \quad (17)$$

where the principal value of the complex logarithm is used. The mapping has the same periodicity as that of the previous case ($R \neq 0$). The back transformation of

$$Z = \frac{1}{2\eta}\ln(\bar{\beta}^2 - S) \pm i\frac{\pi k}{\eta}, \quad (k = 0, 1, 2, 3, \cdots) \quad (18)$$

leads to only two distinct Z values in each period $(2\pi/\eta)$ for each β^2 (compared to four in the previous case). The bound modes, are spread to the left of the point

$$Z^{(1)}_{\text{cutoff}} = \frac{\ln\left(n^2_{\max} - S\right)}{2\eta}. \quad (19)$$

Cutoff points for the other types of modes can be obtained similarly from (11b).

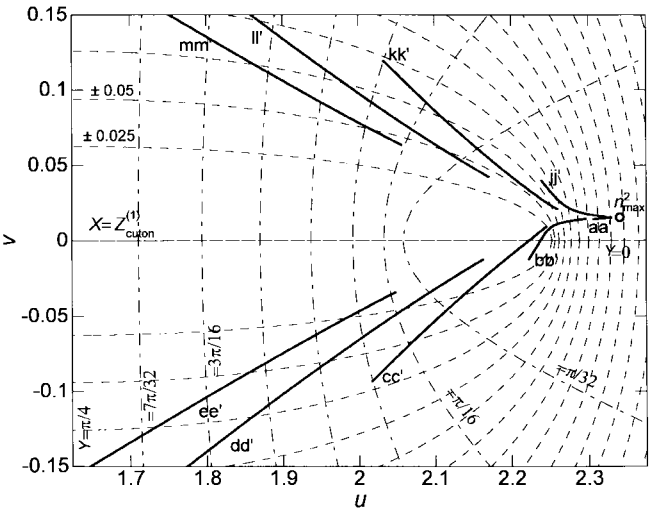


Fig. 3. The level curves of transformation (7a), with $\eta = 1, \eta_{\text{cl}} = 1$ and $\eta_{\text{sub}} = 1.5$.

The branch cut of $\bar{\beta}^2 = n^2_{\text{cl}} = n^2_{\text{sub}}$ maps to the following lines

$$\begin{cases} \text{Im}(Z) = 0 \\ \text{Im}(Z) = \pm\pi/\eta \end{cases}. \quad (20)$$

The regions between these lines (upper and lower strips) correspond to the two folds of the Reimann $\bar{\beta}^2$ surface, respectively.

In this case, at the first glance, there exists a difficulty in finding the modes near the cuton frequency modes, since near the cuton as $\bar{\beta}^2 \to S = n^2_{\text{sub}} = n^2_{\text{cl}}$, the zeros of the dispersion function approach Z_{cuton}, which is at $-\infty$ [see (10)]. Indeed, we can not determine the cuton frequencies exactly. However, we can approach them sufficiently, without any significant problem. Assume that by approaching a cuton frequency, $\bar{\beta}^2$ approaches S by an infinitesimal difference value of δ. Then the zeros in the new plane will approach $\ln \delta/2\eta$, which is not a very large number. For example, if we take $|\delta| = 3.8\times10^{-11}$, then with $\eta = 1$, the amplitude of the zeros in the new plane would be -12.

In this case, the transformation (16) can be viewed as a mapping of the polar coordinates X and Y to the Cartesian coordinates u and v by the following relations:

$$u = \exp(2\eta X)\cos 2\eta Y + S \quad (21a)$$

$$v = \exp(2\eta X)\sin 2\eta Y. \quad (21b)$$

The level curves corresponding to $X = \text{const.}$ are concentric circles with the radii of $\exp(2\eta X)$ centered at $(S, 0)$. On the other hand, the level curves $Y = \text{const.}$ are semi-lines originated from the point $(S, 0)$ with the slope of $\tan(2\eta Y)$.

For $\eta = 1$ and $n_{\text{sub}} = n_{\text{cl}} = 1.5$ the level curves of transformation (16) are plotted in Fig. 4.

III. Characterization and Classification of Modes in the New Plane

Since the transform is conformal, any incremental configuration in one plane conforms to its image in the other plane (in the sense that they are approximately similar). However,

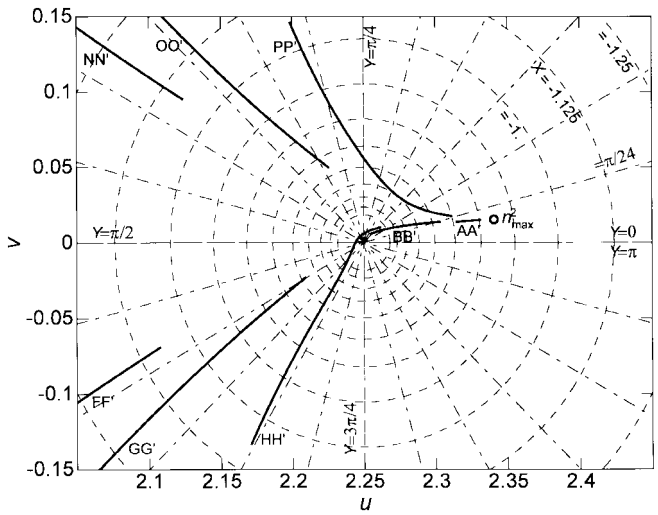


Fig. 4. The level curves of transformation (16), with $\eta = 1$ and $\eta_{\rm cl} = \eta_{\rm sub} = 1.5$.

this is not true for large configurations. Therefore, the real and imaginary directions on the complex Z plane will often have local physical meaning, and the interpretation of the modes in the new Z plane is simple, but it deserves to be discussed in more details. Furthermore, the new plane can present the whole modal characteristic in a single view.

To obtain a better insight into the new plane, first we consider an asymmetric planar dielectric waveguide (i.e., $n_{\rm cl} \neq n_{\rm sub}$) with the previous assumptions i.e., $\mathrm{Im}(n^2_{\rm sub}) = \mathrm{Im}(n^2_{\rm cl}) = 0$, and $\mathrm{Re}(n^2_{\rm sub}) > \mathrm{Re}(n^2_{\rm cl})$. As it was mentioned in the previous section, four quadrants of the new plane (see Fig. 2) correspond to the four folds of β^2 Reimann surface. $X = Z_{\rm cuton}$, and $Y = 0$ separate the quadrants, and each quadrant is also divided into two halves.

The power flow in any direction is given by the time average of Poyinting vector P in the given direction. It can be shown that in the cladding region, the signs of P_x (P in the x direction) and $\mathrm{Im}(\alpha_{\rm cl})$ are the same, whereas in the substrate region the signs of P_x and $\mathrm{Im}(\alpha_{\rm sub})$ are different [1]. In the other hand the sign of P_z (P in the z direction) is the same as $\mathrm{Re}(\beta)$ that we have taken arbitrarily positive.

The upper left quadrant (the second quadrant) is for modes with $\mathrm{Im}(\alpha_{\rm cl}) < 0$, and $\mathrm{Im}(\alpha_{\rm sub}) > 0$ (see Fig. 2). In this quadrant the direction of power flow is from substrate toward cladding. Modes in the upper part of this quadrant (the "cladding leaky" mode) direct the wave excited at $z = -\infty$ through the core region toward the cladding, and the power stored in the substrate toward the core region. This process causes the accumulation of the radiated power at far cladding, while diminishing of power at far substrate. For the "cladding leaky" modes the net power increases with z in the guide ($\mathrm{Im}(\beta) > 0$). The lower part of this quadrant (the "substrate source" modes) might be excited from sources at far substrate. These modes direct the radiated power toward the cladding, whereas they may be partly dissipated or enhanced in the core region. In this region, $\mathrm{Im}(\beta) < 0$, i.e., the net power accumulates into the guide as the mode proceeds toward the positive z direction.

The lower left quadrant (the third quadrant) corresponds to the modes with $\mathrm{Im}(\alpha_{\rm cl}) > 0$ and $\mathrm{Im}(\alpha_{\rm sub}) < 0$. The direction of power flow is from substrate toward cladding. The modes in the upper half of this quadrant (the "substrate leaky" modes) might be excited in the core region at $z = -\infty$. The modes direct the power in the core together with the power stored in the cladding region, toward the substrate. The power at the far substrate accumulates to infinity, and $\mathrm{Im}(\beta) > 0$, which means that the total power escapes from the core region while the core itself may be dissipative, transparent or active. In the lower part of this quadrant ("cladding source" mode), the modes direct the power excited at far cladding toward the substrate. These modes enhance the power carried by the waveguide ($\mathrm{Im}(\beta) < 0$).

The cuton and cutoff points indicated in Fig. 2 are for transparent waveguides with $n^2_{\rm max} > n^2_{\rm sub}$. The interval between $Z^{(2)}_{\rm cuton}$ and $Z^{(4)}_{\rm cutoff}$ for lossless waveguides corresponds to a kind of propagating mode. This mode has interesting properties: there is no power flow in the x direction, (i.e., $\mathrm{Im}(\alpha_{\rm cl}) = \mathrm{Im}(\alpha_{\rm sub}) = 0$); the power is accumulated at far cladding and diminishes at far substrate (i.e., $\mathrm{Re}(\alpha_{\rm cl}) < 0$, and $\mathrm{Re}(\alpha_{\rm sub}) > 0$). However, there is not any power escaping from the core region as the modes travel toward positive z direction (i.e., $\mathrm{Im}(\beta) = 0$). Similarly, the modes in the interval between $Z^{(1)}_{\rm cuton}$ and $Z^{(3)}_{\rm cutoff}$ for lossless waveguides are propagating modes with similar property, except that the power is accumulated in the substrate instead of the cladding.

The upper right quadrant (the first quadrant) is for the modes with $\mathrm{Im}(\alpha_{\rm cl}) > 0$, and $\mathrm{Im}(\alpha_{\rm sub}) > 0$. The direction of power flow is from the substrate and cladding toward the core. Modes in the upper region of this quadrant ("cladding-substrate source") might be excited by the sources located far at both the cladding and the substrate layers. In this region $\mathrm{Im}(\beta) < 0$, that is the power is totally accumulated in the guide. Note that this behavior is independent of the characteristic of the core region (i.e., the core region may be lossless, lossy or even active). In the lower part of this quadrant ("absorptive") the power stored in the substrate and in the cladding regions due to an exciting source at $z = -\infty$, is directed toward the core region. The direction of power flow is toward the guide, with $\mathrm{Im}(\beta) > 0$, so the absorptive modes are modes of a lossy waveguide.

Finally, the lower right quadrant (the forth quadrant) is for modes with $\mathrm{Im}(\alpha_{\rm cl})$ and $\mathrm{Im}(\alpha_{\rm sub})$ negative. The direction of power flow is outward the core region and toward the cladding and the substrate regions. The upper half of this quadrant is for "gain" modes, with $\mathrm{Im}(\beta) < 0$. These modes accumulate power in the guide as the they propagate along the positive z direction. This is the characteristic of an active waveguide. The lower half of this quadrant ("cladding-substrate leaky") corresponds to the modes leaking of the guides from the core region toward both the cladding and the substrate. The source of excitation would be located at $z = -\infty$, and the guide radiates this power through the cladding and substrate.

Modes between $Z^{(1)}_{\rm cuton}$ and $Z^{(1)}_{\rm cutoff}$ for lossless waveguides are propagating "bound" modes ($\mathrm{Im}(\beta) = 0$), with $\mathrm{Im}(\alpha_{\rm sub}) = \mathrm{Im}(\alpha_{\rm cl}) = 0$. Modes between $Z^{(2)}_{\rm cuton}$ and $Z^{(2)}_{\rm cutoff}$

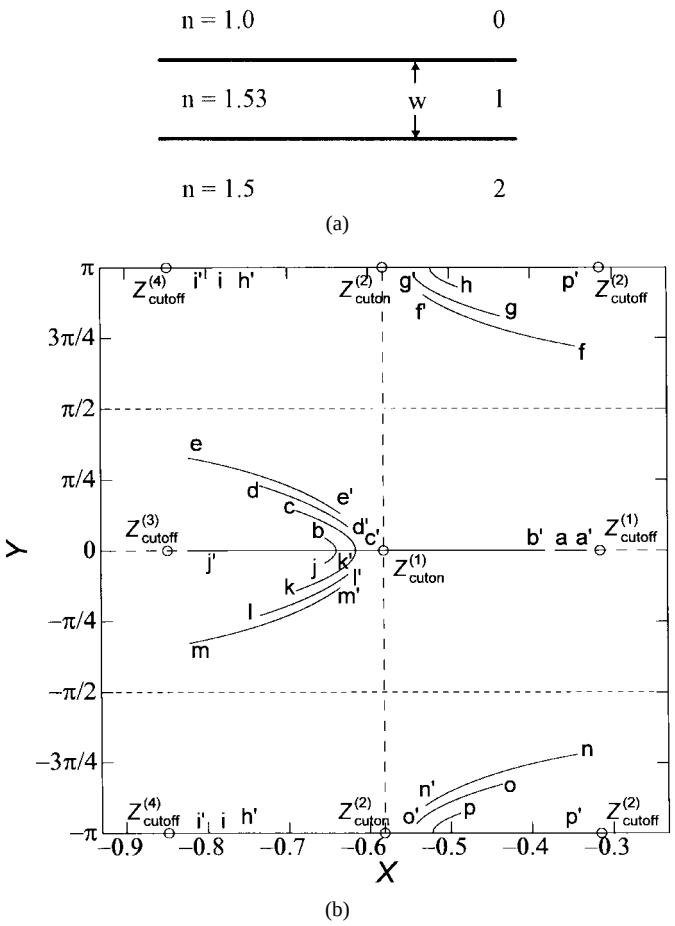


Fig. 5. (a) A transparent asymmetrical three-layered waveguide. (b) The modal migration curves of structure (5a) from x to x' for the variation of normalized core width ($\bar{w} = w/\lambda_0$) from 2 to 4.

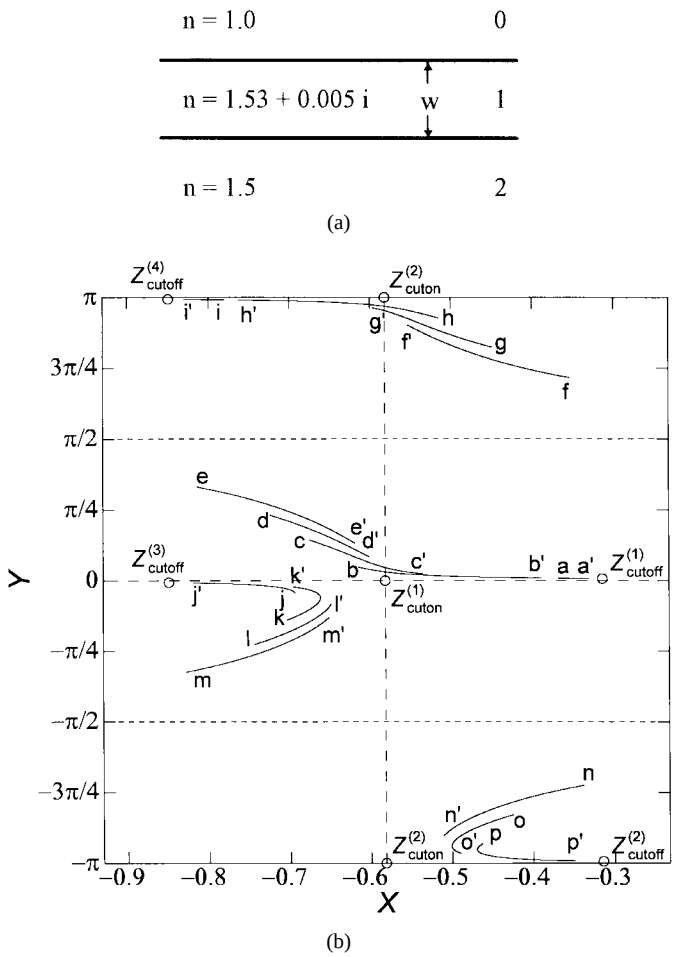


Fig. 6. (a) An asymmetrical three-layered waveguide with absorptive core. (b) The modal migration curves of structure (6a) from x to x' for the variation of normalized core width ($\bar{w} = w/\lambda_0$) from 2 to 4.

for lossless waveguides are the other kinds of propagating mode ($\mathrm{Im}(\beta) = 0$). In latter case, there is no power flow in the x direction, (i.e., $\mathrm{Im}(\alpha_{\mathrm{cl}}) = \mathrm{Im}(\alpha_{\mathrm{sub}}) = 0,$), whereas the power is accumulated at far cladding and substrate (i.e., $\mathrm{Re}(\alpha_{\mathrm{cl}})$ and $\mathrm{Re}(\alpha_{\mathrm{sub}})$ are negative).

For a symmetric waveguide, the new plane is divided into two halves, separated at $Y = 0$. Indeed the left quadrants for asymmetrical waveguides can not exist, since the modal behavior in the cladding and the source regions must be the same. The modal characterization of this case is similar to that of the modes in the right hand side quadrants of the previous case [6].

IV. Examples and Discussion

The lossless waveguides have real refractive indexes, so if β is a propagation constant of a mode in such waveguides, then β^* is also a one [see (3)]. However, for an absorptive waveguide with complex refractive indexes, if β is a propagation constant, then β^* is the propagation constant of modes of the counterpart waveguide with complex conjugate refractive indexes [see (3)]. These symmetry properties also hold in the new plane [see (9)]. In this way, any active waveguide can be analyzed by analyzing its absorptive counterpart.

For the simple three layer asymmetric waveguides shown in Figs. 5(a) and 6(a) (examples of [5]), the migration curves of TE modes are plotted in Figs. 5(b) and 6(b), respectively. We have taken $\eta = 1$ for convenience. The trace of each mode is denoted by xx', where x is a small letter. The direction of migrations is from x to x' (unprimed to primed). The modes move (migrate) due to variation of the normalized core width from 2 to 4 (the core width is normalized versus the wavelength, i.e., $\bar{w} = w/\lambda_0$). Note that in the lossless case [Fig. 5(b)], the directions of migration of the "substrate source" modes (bb′, cc′, dd′ and ee′) and "substrate leaky" modes (jj′, kk′, ll′ and mm′) appearing in complex conjugate pairs, are toward the propagating modes between $Z^{(1)}_{\mathrm{cuton}}$ and $Z^{(3)}_{\mathrm{cutoff}}$. Each of these pairs, after reaching this interval (e.g., bb′ and jj′), becomes indistinguishable from each other. Then one of them (say jj′) approaches toward the third cutoff point, whereas the other (bb′) goes toward the first cuton point, and after passing through this point, it becomes a bound mode and approaches toward the first cutoff point. The "cladding-substrate source" modes (ff′, gg′ and hh′) complex conjugately paired with the "cladding-substrate leaky" modes (nn′, oo′ and pp′) migrate toward the propagating modes between $Z^{(2)}_{\mathrm{cuton}}$ and $Z^{(2)}_{\mathrm{cutoff}}$. After reaching this interval, one of them (say pp′) approaches toward the second cutoff point, whereas the other (hh′) goes toward the second cuton point, and after passing, it approaches to the forth cutoff point. There are also the propagating modes, aa′ in the "bound" mode interval (between

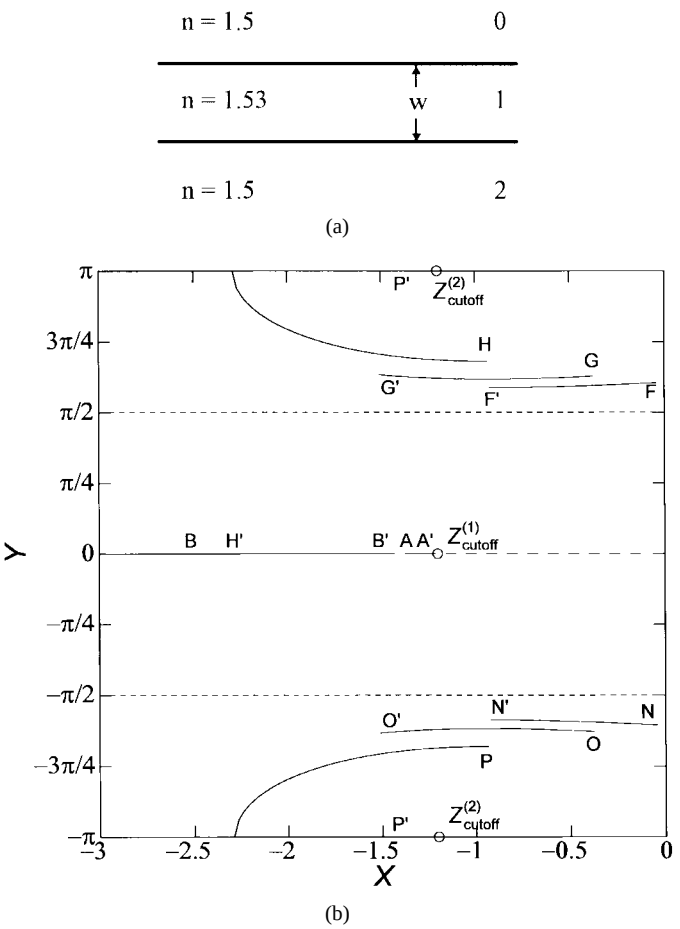


Fig. 7. (a) A transparent symmetrical three-layered waveguide,(b) The modal migration curves of structure (7a) from X to X' for the variation of normalized core width ($\bar{w} = w/\lambda_0$) from 2 to 4.

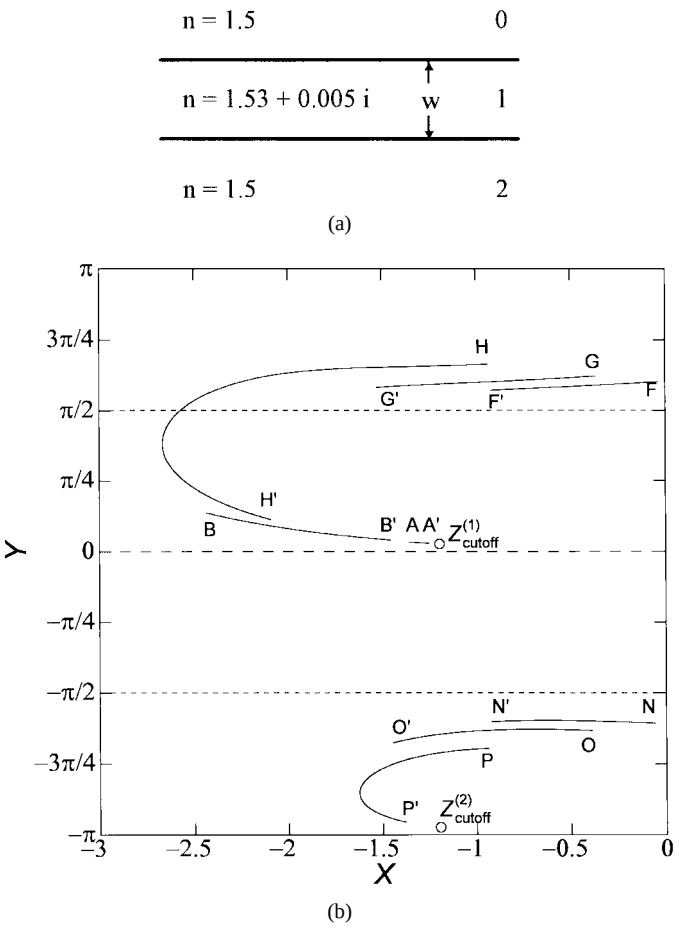


Fig. 8. (a) A symmetrical three-layered waveguide with absorptive core, (b) The modal migration curves of structure (8a) from X to X' for the variation of normalized core width ($\bar{w} = w/\lambda_0$) from 2 to 4.

$Z^{(1)}_{\text{cuton}}$ and $Z^{(2)}_{\text{cutoff}}$) and ii$'$ (between $Z^{(4)}_{\text{cutoff}}$ and $Z^{(2)}_{\text{cuton}}$) which usually approaches to the first and the forth cutoff point, respectively.

In the lossy case, [Fig. 6(b)] the directions of migration of the "substrate source" modes (bb$'$, cc$'$, dd$'$, and ee$'$) are toward the first cutoff point through the "absorptive" mode region, while the "substrate leaky" modes (jj$'$, kk$'$, ll$'$, and mm$'$) move toward the third cutoff point. The "cladding-substrate source" modes (ff$'$, gg$'$, and hh$'$) move always toward the forth cutoff point through the "cladding leaky" mode region, whereas the "cladding-substrate leaky" modes migrates always toward the second cutoff point. The "absorptive" mode (aa$'$) and "cladding leaky" mode (ii$'$) travel always toward the first and the forth cutoff point, respectively.

For the three layer symmetric waveguide with lossless and lossy cores shown in Figs. 7(a) and 8(a), the migration curves are plotted in Figs. 7(b) and 8(b), respectively. The mode traces are denoted by XX' where X is a capital letter. The normalized core width variation is the same as that of the above examples, and the mode migrations are from X to X'. For the lossless case [Fig. 7(b)] the migration of "cladding-substrate source" modes (FF$'$, GG$'$ and HH$'$) with their complex conjugate "cladding-substrate leaky" modes (NN$'$, OO$'$ and PP$'$) are the same as that of the corresponding modes of asymmetric waveguide [see Fig. 5(b)]. Note that in this case the second cuton point is at infinity and there is a difference: the modes after reaching the second cuton point at infinity ($Z^{(2)}_{\text{cuton}} = -\infty \pm i\pi$), jump to the first cuton point at infinity ($Z^{(1)}_{\text{cuton}} = -\infty + i0$). Then they appear as bound modes and approach to the first cutoff point (e.g., mode HH$'$). In the lossy case (see Fig. 8(b)), the behavior of the "cladding-substrate leaky" modes (NN$'$, OO$'$ and PP$'$) is the same as that of the previous case (i.e., they move toward the second cutoff point). However, the "cladding-substrate source" modes (FF$'$, GG$'$ and HH$'$) travel toward the first cutoff point through the "absorptive" mode region, without going to $-\infty$. The absorptive modes AA$'$ and BB$'$ move always toward the first cutoff point.

Fig. 9 shows the migration plot of the modes of the three layer lossless waveguide shown in Fig. 5(a) with $\bar{w} = 2$. The source of the mode migration is the variation of the cladding refractive index from 1. to 1.5 (or R from 0.3125 to zero). The trace of the modes are denoted by xX, and the migrations are from x (corresponding to the asymmetric waveguide with $R = 0.3125$) to X (corresponding to the symmetric waveguide with $R = 0$). The possibility of such investigations in a single transformed plane is the unique property of the proposed transform (a general transform for symmetric and asymmetric waveguides). Referring to Fig. 9 the characteristic of "cladding-substrate source" and "cladding-substrate leaky" modes do not change (modes hH, gG, fF and nN, oO, pP),

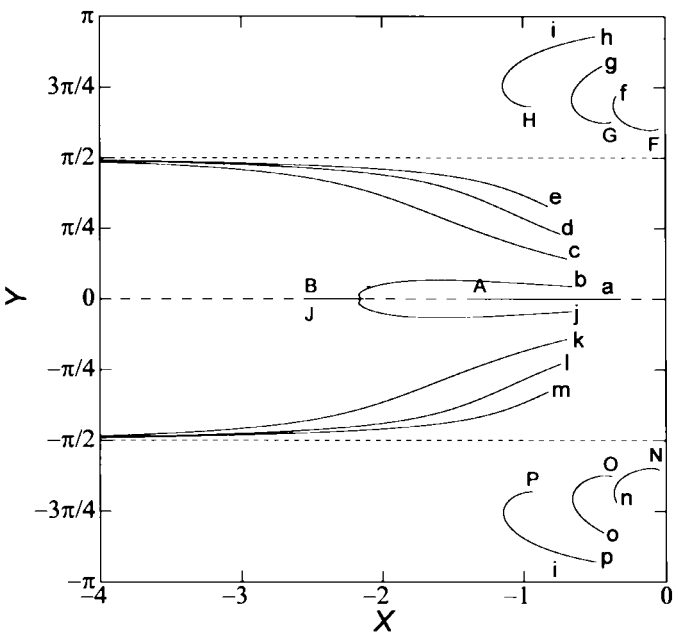


Fig. 9. The migration curves of three layer dielectric waveguide of Fig. 5(a) from x to X for the variation of R from 0 to 0.3125, with $\bar{w} = 2$.

while some of the "substrate source" and "substrate leaky" modes become bound modes (modes bB and jJ) and the others approach to $(-\infty, \pm\pi/2)$ as R approaches zero. The bound mode (mode aA) remains bound, while the other propagating mode (mode iI, with I at $-\infty$) approaches toward $(-\infty, \pm\pi)$.

We also have transferred some modal migration traces of Figs. 6(b) and 8(b) into the $\bar{\beta}^2$ planes of Figs. 3 and 4, respectively. The migration of all modes regardless of their types are usually toward $n^2_{\max}$. Note that as stated previously, a single branch cut is not sufficient to determine all modes, and previous discussions about modal behavior in each branch cut is needed. Furthermore, a single branch cut of the $\bar{\beta}^2$ folded surface may not be sufficient to trace the whole migration of modes, since there may be migration of modes from one branch cut to another (e.g., modes bb′ and dd′ in Fig. 6(b) pass from the second quadrant to the first). The crossing of traces with the branch points corresponds to the crossing of the relevant traces with $X = Z^{(1)}_{\text{cuton}}$ ellipse (line) in Fig. 3.

V. Conclusion

A new conformal mapping is proposed to solve the problem of multi-valuedness of the dispersion function in multilayered dielectric waveguides. We have pointed out the problems and restrictions of the previous transformations. The main feature of this transformation is that, the transformed dispersion function is regular, i.e., analytic without any singularities in the entire complex plane. This feature makes this plane more suitable for modal analysis of planar dielectric waveguides. Furthermore, this leads us to a single and clear view of the folded complex β^2 plane. Based on this view, we have clarified the modal behaviors of the corresponding structures. The transformation is general and can be applied to any types of layered isotropic dielectric waveguides, i.e., active, passive and lossless cases with symmetric or asymmetric configurations. In the new plane the modal zones which are strips, facilitate APM implementation. Any rectangular integration contour in the new plane can be easily traced to a contour in the original folded surface by the aids of the corresponding level curves.

References

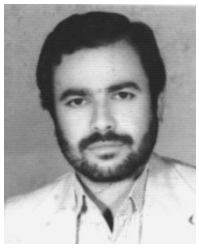
[1] J. Chilwell and I. Hodgkinson, "Thin-films field-transfer matrix theory of planar multilayer waveguides and reflection from prism-loaded waveguides," *J. Opt. Soc. Amer. A*, Vol. 1, no. 7, pp. 742–753, 1984.
[2] R. E. Smith, S. N. Houde-Walter, and G. W. Forbes, "Mode determination for planar waveguides using the four-sheeted dispersion relation," *IEEE J. Quantum Electron.*, vol. 28, pp. 1520–1526, June 1992.
[3] E. Anemogiannis and E. N. Glytsis, "Multilayer waveguides: Efficient numerical analysis of general structure," *J. Lightwave Technol.*, vol. 10, pp. 1344–1351, Oct. 1992.
[4] R. E. Smith, G. W. Forbes, and S. N. Houde-Walter, "Unfolding the multivalued planar waveguide dispersion relation," *IEEE J. Quantum Electron.*, vol. 29, pp. 1031–1034, Apr. 1993.
[5] R. E. Smith and S. N. Houde-Walter, "The migration of bound and leaky solutions of the waveguide dispersion relation," *J. Lightwave Technol.*, vol. 11, pp. 1760–1768, Nov. 1993.
[6] D. Marcuse, *Theory of Dielectric Optical Waveguides*. New York: Academic, 1991, ch. 1, 2nd ed.

Aref Bakhtazad was born in Jahrom, Iran, in 1964. He received the B.Sc. degree in electrical engineering from Shiraz University, Shiraz, Iran, and the M.Sc. degree in electrical engineering in 1995 from Tousi University of Technology, Tehran, Iran. He was selected as a Distinguished Graduate Student in 1993 and 1994 from Shiraz University, where he was preparing his M.Sc. thesis.

His research interests are in the area of integrated optics and numerical modeling. He is now working on the various aspects of coupled mode theory for dielectric waveguides.

Habibollah Abiri was born in Jahrom, Iran. He received the B.S. degree in electrical engineering from Shiraz University, Shiraz, Iran, in 1978. He received the D.E.A. and Docteur d'Ingénieur degree (Ph.D.) from the National Polytechnique Institute of Grenoble (INPG), France, in 1981 and 1984, respectively.

During 1985, he was an Assistant Professor at the University of Savoie, Chambéry, France. Also in 1985, he was with the Electrical Engineering Department of Shiraz University, where he is now an Associate Professor. In 1995, he was on sabbatical leave at the Electrical Engineering Department, Colorado State University, Fort Collins, CO. His research interests include numerical methods in electromagnetic theory, microwave circuits and integrated optics.



R. Ghayour (M'95) was born in Iran in 1953. He received the M.S. degree from Shiraz University, Shiraz, Iran, in 1985, and the Ph.D. degree from the University of Birmingham, England, in 1989, both in electronics.

He has been with the Department of Electrical Engineering, Shiraz University, as Assistant Professor since 1989. His research interests are in optoelectronics, modeling and design of devices for microelectronics and power electronics, and analysis and modeling of solid-state lasers.