

Introduction to Slab Dielectric Waveguides

Prof. Elias N. Glytsis
November 19, 2020



School of Electrical & Computer Engineering
National Technical University of Athens

This page was intentionally left blank.....

Contents

1	Introduction	1
2	Ray Approach for Guided Modes	5
2.1	Alternate Approach	8
3	Electromagnetic Approach for Guided Modes	9
3.1	TE Guided Modes	12
3.2	TM Guided Modes	14
4	Substrate Modes	18
5	Radiation Modes	20
6	Unphysical Modes (or Nonphysical Modes)	22
7	Evanescent Modes	23
8	Normalized Slab-Waveguide Variables	26
9	Cutoff Conditions	28
10	Power Considerations	32
10.1	TE Modes	32
10.2	TM Modes	34
11	Multi-Layered Slab Waveguides	36
11.1	TE Guided Modes	37
11.2	TM Guided Modes	40
11.3	TE and TM Guided Modes using Chen's Approach	42
12	Finite-Difference Frequency-Domain Method	44
12.1	FDFD Method Based on Yee's Cell	47
12.1.1	Average-Scheme Modification for Arbitrary Boundary Location	55
13	Graded-Index Slab Waveguides	61
14	Lossy, Leaky, and Active Planar Multilayer Waveguides	72
14.1	Basic Theory of the Principle Argument Method (APM)	72
14.2	Basic Theory of the Abd-ellal, Delves, and Reid (ADR) Method	74
14.3	Basic Theory of the Derivative-Free Zero-Extraction by Phase-based Enclosure Method (DFZEPE)	76
14.4	Application of APM, ADR, and DFZEPE methods to Planar Multilayer Waveguides	82
14.5	Results	85
14.5.1	Lossless Dielectric Waveguide	86

14.5.2	Low-Loss Dielectric Waveguide	88
14.5.3	Active Waveguide	90
14.5.4	Leaky Modes of an ARROW Waveguide	91
References		97

This page was intentionally left blank.....

INTRODUCTION TO SLAB DIELECTRIC WAVEGUIDES

1. Introduction

Planar waveguides are critical elements in present day optical devices such as modulators, semiconductor lasers, couplers, wavelength filters, polarizers, and optical interconnects. Passive waveguides, electro-optic components, transmitters, receivers, active waveguides, and driving electronics can be integrated into an optical/electronics chip using planar technology, similar to microelectronics. Optical waveguides are the fundamental elements that interconnect the various devices of an photonic integrated circuit, just as a metallic strips make interconnections in an electronic integrated circuit. Although the operation of waveguide devices is well researched and understood, their particular performance relies on many parameters such as geometry, freespace wavelength, initial field distribution, material characteristics, as well as electro-optic, acousto-optic, or other driving conditions. The device parameters must be optimized before fabricating a well operating device. With large-scale optoelectronic circuits, accurate modeling is predominant because of the numerous resources required to fabricate a chip. Optical waveguide design relies on simulating the propagation of light signals, determining the waveguide modes, evaluating the mode coupling, and assessing the loss and gain of the structure. Unlike electrical current that flows through a metal strip (wire) according to Ohm's law, optical waves travel in the waveguide in characteristic optical modes. A mode can be defined as a spatial distribution of optical energy in one or more dimensions that remains constant in space perpendicularly to the direction of propagation. A mode is an eigenfunction of the Maxwell's equations for the waveguiding structure.

For the successful design and optimization of waveguide-based devices, robust and reliable numerical methods are needed in order to determine accurately the propagation characteristics of their modes. For example, the determination of guided-mode characteristics in optical waveguides comprised of lossless and/or lossy materials, results in the solution of a transcendental complex equation [1]. In addition, the design of semiconductor lasers/amplifiers require the utilization of both lossy and active (for gain) materials [2]. Recently, there has been a lot of interest in non-Hermitian optical structures where both gain and loss are present [3,4].

Furthermore, in waveguiding structures, the determination of leaky-mode characteristics that can compactly model the radiation field could be important. A practical example of this is found in the antiresonant reflecting optical waveguides (ARROW) [5, 6]. ARROW structures have received significant attention due to their low transmission loss, high polarization sensitivity, as well as compatibility with optical fibers for telecommunications applications. The determination of complex, in general, eigenvalues facilitates the evaluation of the corresponding field distributions, power density, radiation losses, modal gain (for active waveguides) and possibly other parameters of interest. For this reason it has been a great interest over the last few decades to find reliable and robust techniques for calculating the zeros of an analytic function in the complex plane. Common general numerical analysis methods that are used are the Newton's method (where the function's derivative is used) and Müller's method (where the function's derivative is not needed) [7–9]. However, these methods require a good estimate of the root of the complex function which in most practical problems is not available. Furthermore, usually not even the number of zeros is known. A successful numerical method should be able to extract all the zeros of the complex function in a domain of interest without any prior knowledge of either their number or their approximate location.

One of the first numerical methods for the solution of complex transcendental equations is based on the Cauchy integration method introduced by Delves and Lyness [10]. This technique finds the number of zeros and approximates the roots of the function in a specified domain by evaluating a polynomial with the same roots. The polynomial roots could be used as estimates for further root improvements via iterative methods. The first applications of the approach to multilayer planar waveguides using the transfer matrix approach [11–13] was presented by Smith *et al.* [14, 15] and Anemogiannis and Glytsis (known as the argument principal method - APM) [16]. Since then similar methods appear in the literature. Chen *et al.* [17] used APM for multilayer lossy anisotropic waveguides, Kwon and Shin applied APM to isolated roots [18, 19], and Michalski and Mustafa also utilized APM in conjunction to either transfer- or scatter-matrix formulation and automatic differentiation [20]. These methods are based on the numerical evaluation of the Cauchy integrals of the waveguide dispersion equation and require the derivative of the dispersion equation that can be calculated either numerically [21] or analytically. Some alternative methods that have been proposed in the literature are the reflection-pole method [22, 23], the wavevector density method [23], a variational method [24],

a method based on smooth transition between bound modes of a closed waveguide and leaky modes of an open waveguide [25], an iterative "downhill" method [26], and a two-dimensional minimization [27], to mention a few. Of course the fully numerical methods based on finite elements [28, 29] or finite differences [30, 31] should also be mentioned. These latter methods can also be applied to three-dimensional waveguides (channel waveguides) since they represent completely numerical solution of the Maxwell's equations. However, they require specialized boundary layers (such as Perfectly Matched Layers [32]) around the structure in order to treat open dielectric waveguides.

One of the drawbacks of the APM method is the use of the derivative of the function which makes it rather cumbersome especially when the derivative cannot be calculated analytically. Therefore, there has been significant effort to focus on globally reliable methods that find the zeros of a complex function without the use of its derivative. Anemogiannis *et al.* [33] first proposed the ADR method which is based on an algorithm provided by Abd-ellal, Delves, and Reid [34] (the ADR acronym was derived from their last names initials). The ADR method uses integrals (representing function moments) on a contour enclosing a specified domain that do not involved the derivative of the function. The ADR method requires though a double number of integrals as compared to APM method in order to determine the coefficients of the approximating polynomial. In addition, it has some difficulty in estimating the actual number of zeros inside the complex-plane domain under investigation. Ying and Katz [35] proposed a simple method to determine the exact number of zeros of an analytic function based on the its winding number in a bounded domain. Based on Ying and Katz method Kravanja and Van Barel [36] and Gillan *et al.* [37] introduced a reliable derivative-free method for determining the roots of an analytic function in a bounded complex domain. Their method is based on calculating similar integrals as in ADR and using orthogonal polynomial theory to estimate the zeros of the function by solving a general eigenvalue problem. Refinement of the zeros is achieved using the Halley/Aitken method [37]. Recently Semwal and Rastogi applied the Gillan's method to multilayer planar waveguides [38].

Another simple method proposed recently [39], is based on Ying and Katz [35] winding number evaluation and uses zero enclosing by evaluating only the winding number of the function in successively shrinking rectangles until every root (function's zero) has been surrounded by a pre-specified dimensions corresponding rectangle in the complex plane. Then the center of each

rectangle can serve as an estimate of the zero of the function in the complex plane. The winding number around each shrinking rectangle is calculated by using only the phase of the function and without any integral evaluation or derivative knowledge. Then, the Müller's algorithm [9] with deflation [8] is used for the iterative refinement of the root. Root bracketing by using the winding number was also proposed by Ying and Katz [40], Dellnitz *et al.* [41] and Kwon and Shin [18] but in conjunction to the Cauchy's integrals. Furthermore, very recently Kowalczyk and Marynowski [42] have used the winding number approach on a self-adaptive triangular-based mesh to find the roots of an analytic function for radiation and propagation electromagnetic problems. This approach is purely based on the successive enclosing by shrinking rectangles of each zero (complex propagation constant) of the dispersion function of multilayer planar waveguides and its based exclusively on the phase only of the function without using any integrals or approximating polynomials. For brevity this method will be referred as Derivative-Free Zero-Extraction by Phase-based Enclosure (DFZEPE). The DFZEPE approach is much simpler than the adaptive Delaunay triangulation method proposed by Kowalczyk [42, 43].

In this chapter, the fundamental theory of optical modes in a waveguiding structures is developed in order to acquire a general understanding of the nature of light propagation in an optical waveguides. In the first part of the chapter the simple single-film planar waveguide will be analyzed using both a ray approach as well as a fully electromagnetic approach. Then the classification of the modes of a planar optical waveguide into TE and TM will be introduced. Guided modes, substrate modes, radiation modes, evanescent modes, and unphysical modes will be discussed. Then, the normalized variable will be introduced along with the cut-off conditions. Power considerations will also be discussed. Later in the chapter, the multilayer planar waveguides are presented in conjunction to the transfer-matrix technique. The usage of the Finite-Differences Frequency-Domain (FDFD) method for the study of planar multilayer waveguides is also explained with example cases. Then the graded-index waveguide analysis is shown. Finally, a short review of the most useful methods for extracting complex in general propagation constants for lossy and active waveguides will be presented along with the definition of the leaky modes which are not formal eigen-solutions of the Maxwell's equations. Some excellent books exist on the topic on optical waveguides including [1, 13, 44–50] that may include parts of the topics discussed in this chapter.

2. Ray Approach for Guided Modes

The basic geometry of a dielectric slab waveguide is shown in Fig. 1. The light is confined along the x -axis and propagates along the z -axis in the coordinate system shown. The waveguide is assumed to be uniform along the y -axis and therefore this comprises a two-dimensional electromagnetic problem. The basic principle behind the concept of light guiding in dielectric waveguide is the phenomenon of total internal reflection. If light can be launched into the film layer of the waveguide it can remain in the film, provided that the angle of incidence θ between the film-cover and the film-substrate regions is greater than both of the corresponding critical angles. The refractive indices of the three regions are n_c , n_f , and n_s for cover, film, and substrate regions, respectively while the film thickness is h and the free-space wavelength of the light that can travel guided within this structure is λ_0 . Of course it is necessary that $n_f > \max\{n_c, n_s\}$ in order to guarantee that total internal reflection can occur in both film-substrate and film-cover boundaries. The angle θ , known also as the zig-zag angle, for a specific mode should satisfy the condition $\max\{\theta_{cr,fs}, \theta_{cr,fc}\} < \theta < \pi/2$, where $\theta_{cr,fs} = \sin^{-1}(n_s/n_f)$ and $\theta_{cr,fc} = \sin^{-1}(n_c/n_f)$, are the critical angles for the film-substrate and film-cover boundaries, respectively. Without loss of generality, in the following discussion, it is assumed that $n_s \geq n_c$ (which is usually the most practical case). In the special case that $n_s = n_c$ the waveguide is characterized as a symmetric slab waveguide.

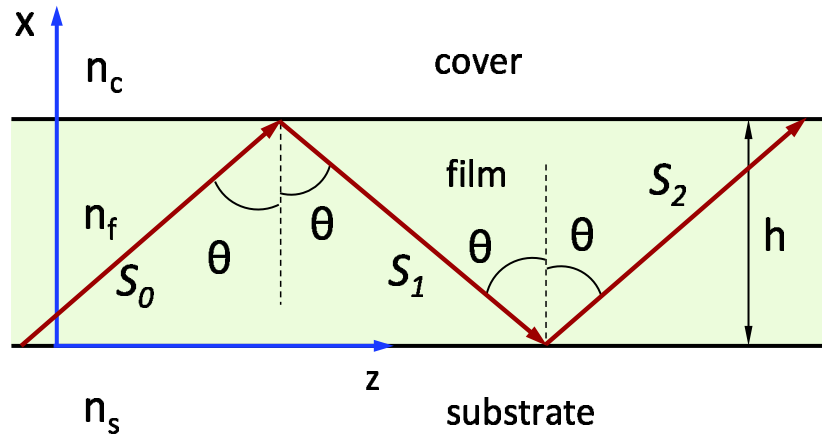


Figure 1: The geometric configuration of a three-region slab dielectric waveguide.

It is well known that there is an induced phase shift upon total internal reflection at both

boundaries. Therefore there should be a self-consistency condition that must be satisfied in order for the plane waves shown in the ray picture of Fig. 1 to comprise a valid solution for a guided mode.

In order to quantify the self-consistency condition Fig. 2 will be used [47]. In this figure two possible rays are shown along with two wavefronts the one specified by AB and the other by DC . The wavefronts AB and DC are selected to be at an infinitesimal distance away from the film-cover and film-substrate boundaries, respectively. This implies that the ray from point B to point D has not suffered any total internal reflections. In contrast, the ray from point A to point C has suffered two total internal reflections at the two boundaries. Following the phases accumulated by points A and B of the initial wavefront as the rays move towards to their positions C and D of the second wavefront shown, it is necessary to require that the accumulated phase from $A \rightarrow C$ and from $B \rightarrow D$ should differ at most by $2\pi\nu$ where ν is an integer. This condition will guarantee that the second wavefront will remain a valid wavefront of the guided mode. Mathematically, the previous conditions can be written as follows

$$\underbrace{\{-k_0 n_f(AC) + 2\phi_{fs}^p + 2\phi_{fc}^p\}}_{\text{phase from } A \rightarrow C} - \underbrace{\{-k_0 n_f(BD)\}}_{\text{phase from } B \rightarrow D} = 2\pi\nu, \quad \nu = 0, \pm 1, \pm 2, \dots \quad (1)$$

where $k_0 = 2\pi/\lambda_0$, and ϕ_{fs}^p and ϕ_{fc}^p ($p = \text{TE or TM}$) are the phase shifts that occur upon total internal reflection at the film-substrate and film-cover boundaries respectively. These phase shifts are functions of the angle θ and are given by the equations (as it was presented in the *Review of Electromagnetic Principles* notes)

$$\phi_{fw}^{TE}(\theta) = \tan^{-1} \left\{ \frac{\sqrt{n_f^2 \sin^2 \theta - n_w^2}}{n_f \cos \theta} \right\}, \quad \text{for TE Polarization,} \quad (2)$$

$$\phi_{fw}^{TM}(\theta) = \tan^{-1} \left\{ \frac{n_f^2 \sqrt{n_f^2 \sin^2 \theta - n_w^2}}{n_w^2 n_f \cos \theta} \right\}, \quad \text{for TM Polarization,} \quad (3)$$

where $w = c$ or s .

The distances (AC) and (BD) can be easily determined from the geometry shown in Fig. 2. Specifically, $(AC) = h/\cos \theta$ and $(BD) = (AD) \sin \theta = [(OA) - (OD)] \sin \theta = [h \tan \theta - h/\tan \theta] \sin \theta = h[\sin^2 \theta - \cos^2 \theta]/\cos \theta$. Replacing the previous expressions in Eq. (1) and

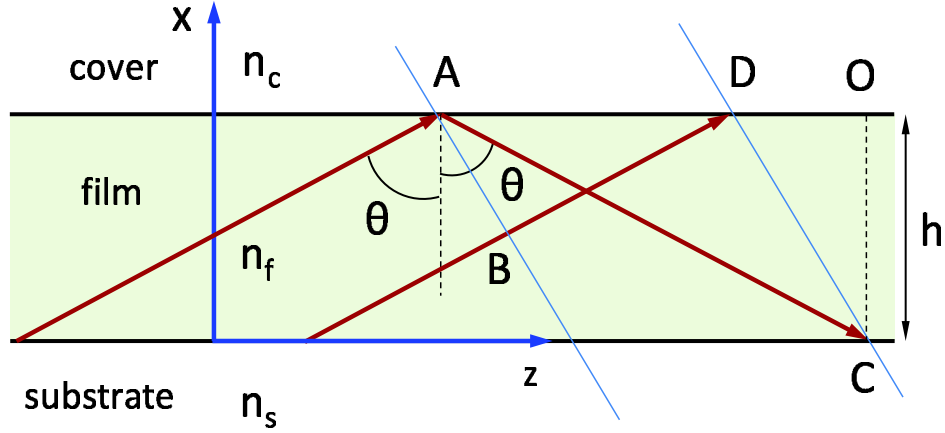


Figure 2: The slab waveguide geometric configuration showing two rays and two selective wavefronts (AB and DC). For self-consistency it is necessary that the accumulated phase from $A \rightarrow C$ and from $B \rightarrow D$ should differ by $2\pi\nu$ where ν is an integer. The points A and C are assumed to be infinitesimally before/after the film-cover/film-substrate boundaries respectively.

changing the signs assigning only positive (or zero) values for ν the following equation is derived

$$2k_0 n_f h \cos \theta - 2\phi_{fs}^p(\theta) - 2\phi_{fc}^p(\theta) = 2\pi\nu, \quad \nu = 0, 1, 2, \dots \quad (4)$$

where $p = \text{TE or TM}$. Of course solution for θ of the previous equation has meaning for a guided mode only when $\max\{\theta_{cr,fs}, \theta_{cr,fc}\} < \theta < \pi/2$. This equation can be solved only numerically (using for example the bisection method). For every value of ν solutions for θ_ν^p ($p = \text{TE or TM}$) can be determined. There can be none, one, or multiple solutions depending on the parameters of the waveguide and the free-space wavelength. A graphical representation of the solution is shown in Fig. 3 for the example case where $\lambda_0 = 1.0\mu\text{m}$, $h = 1.2\mu\text{m}$, $n_c = 1.0$, $n_f = 2.2$, and $n_s = 1.5$. The solutions for the zig-zag angles θ_ν^p can be visualized as the intersection of two curves representing the functions $f_1(\theta) = 2k_0 n_f h \cos \theta$ and $f_2(\theta) = 2\phi_{fs}^p(\theta) + 2\phi_{fc}^p(\theta) + 2\pi\nu$ ($p = \text{TE or TM}$). In this particular example there are 4 TE modes ($\text{TE}_0, \text{TE}_1, \text{TE}_2$, and TE_3) and 4 TM modes ($\text{TM}_0, \text{TM}_1, \text{TM}_2$, and TM_3) that can be supported. It can be observed that the solutions θ_ν^{TE} and θ_ν^{TM} satisfy the inequality $\theta_\nu^{\text{TE}} > \theta_\nu^{\text{TM}}$ (for any ν) due to the larger phase shift $[\phi_{fw}^{\text{TE}}(\theta) < \phi_{fw}^{\text{TM}}(\theta), w = c \text{ or } s]$ upon total internal reflection for the TM polarization.

In general, the solution for the zig-zag angle $\theta = \theta_\nu^p$ (where $p = \text{TE or TM}$) is characteristic of the guided mode. The associated to θ_ν^p electromagnetic field represents the guided mode TE_ν or TM_ν and has a characteristic profile. The complete electromagnetic field will be determined

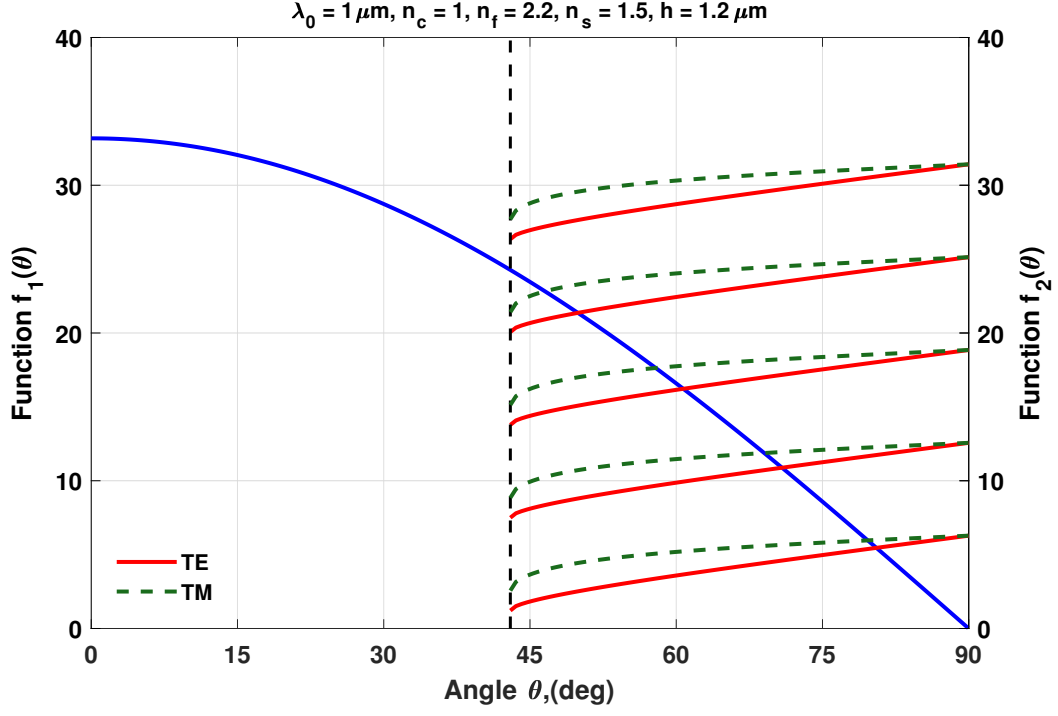


Figure 3: The graphical representation of the solution of the dispersion equation $2k_0hn_f \cos \theta - 2\phi_{fc}^p - 2\phi_{fs}^p = 2\nu\pi$ ($p = \text{TE or TM}$, and $\nu = 0, 1, \dots$). The solution for the zig-zag angle θ is shown as the intersection of two simple curves, the $f_1(\theta) = 2k_0hn_f \cos \theta$ and $f_2(\theta) = 2\phi_{fc}^p(\theta) + 2\phi_{fs}^p(\theta) + 2\nu\pi$. For the example case shown $\lambda_0 = 1.0\mu\text{m}$, $h = 1.2\mu\text{m}$, $n_c = 1.0$, $n_f = 2.2$, and $n_s = 1.5$. The solutions for the zig-zag angles are $\theta_0^{TE} = 80.5307^\circ$, $\theta_1^{TE} = 70.8514^\circ$, $\theta_2^{TE} = 60.7247^\circ$, and $\theta_3^{TE} = 49.9121^\circ$ for the TE modes, and $\theta_0^{TM} = 79.6444^\circ$, $\theta_1^{TM} = 69.0254^\circ$, and $\theta_2^{TM} = 57.8963^\circ$, and $\theta_3^{TM} = 46.5355^\circ$ for the TM modes. The corresponding critical angles are $\theta_{cr,fc} = 27.0357^\circ$ and $\theta_{cr,fs} = 42.9859^\circ$.

in the next section when the full electromagnetic approach will be presented. From the angle θ_ν^p , the effective refractive index of the mode can be defined as $N_\nu^p = n_f \sin \theta_\nu^p$. Furthermore, the effective propagation constant is related to the effective index as $\beta_\nu^p = k_0 N_\nu^p = k_0 n_f \sin \theta_\nu^p$ ($p = \text{TE or TM}$).

2.1 Alternate Approach

A similar approach based on a self-consistency condition is presented in this section and is based on the geometry shown in Fig. 1. In this ray diagram of a guided mode in the waveguide a plane wave is associated with each ray. Specifically (neglecting polarization) the three plane

waves, S_0 , S_1 , and S_2 , can be described by the following expressions

$$S_0 = S_0(x, z) = E_0 e^{-jk_0 n_f (x \cos \theta + z \sin \theta)}, \quad (5)$$

$$S_1 = S_1(x, z) = E_1 e^{-jk_0 n_f (-x \cos \theta + z \sin \theta)}, \quad (6)$$

$$S_2 = S_2(x, z) = E_2 e^{-jk_0 n_f (x \cos \theta + z \sin \theta)}. \quad (7)$$

The plane wave S_1 is produced by a reflection of S_0 at the film-cover boundary. Similarly, plane wave S_2 is produced by reflection of S_1 at the film-substrate boundary. These arguments can be quantified as follows

$$S_1(x = h, z) = r_{fc} S_0(x = h, z) \Rightarrow E_1 e^{+jk_0 n_f h \cos \theta} = e^{j2\phi_{fc}^p} E_0 e^{-jk_0 n_f h \cos \theta}, \quad (8)$$

$$S_2(x = 0, z) = r_{fs} S_1(x = h, z) \Rightarrow E_2 = e^{j2\phi_{fs}^p} E_1, \quad (9)$$

where r_{fc} and r_{fs} are the reflection coefficients upon total internal reflection at the film-cover and film-substrate interfaces, respectively, and $p = \text{TE or TM}$. From the last two equations it can be deduced that

$$E_2 = E_0 e^{j(-2k_0 n_f h \cos \theta + 2\phi_{fs}^p + 2\phi_{fc}^p)}. \quad (10)$$

The last equation requires that in order for S_2 to be equivalent to S_0 their amplitudes must differ by at most by $2\pi\nu$, i.e.

$$E_2 = E_0 e^{j2\pi\nu} = E_0 e^{j(-2k_0 n_f h \cos \theta + 2\phi_{fs}^p + 2\phi_{fc}^p)}, \quad (11)$$

and from the last expression Eq. (4) is derived again.

3. Electromagnetic Approach for Guided Modes

The electromagnetic fields representing a mode for the slab waveguide problem (using the coordinate axes system shown in Fig. 1) can be written in the form

$$\vec{E} = [E_x(x)\hat{x} + E_y(x)\hat{y} + E_z(x)\hat{z}] \exp(-j\beta z), \quad (12)$$

$$\vec{H} = [H_x(x)\hat{x} + H_y(x)\hat{y} + H_z(x)\hat{z}] \exp(-j\beta z), \quad (13)$$

where E_w and H_w ($w = x, y, z$) represent the electric and magnetic field components and β is a propagation constant. Using Maxwell equations in their differential time-harmonic form, for a homogeneous, linear, lossless, and isotropic material, of permittivity ϵ and permeability μ_0

(non-magnetic) with the above field expressions results in the following sets of equations (ω is the radial frequency)

$$\frac{d}{dx} \begin{bmatrix} E_y \\ H_z \\ H_y \\ E_z \end{bmatrix} = \begin{bmatrix} 0 & -j\omega\mu_0 & 0 & 0 \\ -j\omega\epsilon + j\frac{\beta^2}{\omega\mu_0} & 0 & 0 & 0 \\ 0 & 0 & 0 & j\omega\epsilon \\ 0 & 0 & j\omega\mu_0 - j\frac{\beta^2}{\omega\epsilon} & 0 \end{bmatrix} \begin{bmatrix} E_y \\ H_z \\ H_y \\ E_z \end{bmatrix}, \quad (14)$$

$$\begin{bmatrix} H_x \\ E_x \end{bmatrix} = \begin{bmatrix} -\frac{\beta}{\omega\mu_0} & 0 \\ 0 & \frac{\beta}{\omega\epsilon} \end{bmatrix} \begin{bmatrix} E_y \\ H_y \end{bmatrix}. \quad (15)$$

From the above equations it is straightforward to distinguish two set of field components, the $\{E_y, H_x, H_z\}$ and $\{H_y, E_x, E_z\}$, which are independent from each other due to the zeros appearing in the matrices of Eqs. (14) and (15). These two sets can be used to define the TE-modes solutions for $\{E_y, H_x, H_z\}$, and the TM-modes solutions for $\{H_y, E_x, E_z\}$. It is mentioned that these two triplets of field components are independent only in the isotropic case. If the material becomes anisotropic the zero elements are replaced by non-zero ones and all six field components become coupled. Then a waveguide mode, usually called hybrid mode, can be characterized by all six field components as $\{E_x, E_y, E_z, H_x, H_y, H_z\}$.

In a homogeneous, isotropic, lossless, non-magnetic, and linear material the solutions to the Helmholtz's equation are in the form of plane waves. For example, if solutions of the form of Eqs. (12), (13) are sought the Helmholtz's equation becomes

$$\frac{d^2 \vec{U}}{dx^2} + (k_0^2 n^2 - \beta^2) \vec{U} = 0, \quad (16)$$

where $\vec{U} = \vec{E}$ or \vec{H} for electric or magnetic field, respectively, and $n = \sqrt{\epsilon} = \sqrt{\epsilon/\epsilon_0}$ is the refractive index of the material (ϵ is the relative permittivity of the material). The solutions of the above equation have the form

$$\vec{U} = \vec{U}_+ \exp(-j\vec{k}_+ \cdot \vec{r}) + \vec{U}_- \exp(-j\vec{k}_- \cdot \vec{r}), \quad (17)$$

where \vec{U}_+ and \vec{U}_- are the vector amplitudes of the two plane wave solutions and \vec{k}_+ and \vec{k}_- are their corresponding wave-vectors. The “ k_+/k_- ” terms represent the waves propagating towards

the positive/negative x direction. By applying the above solutions to each of the three separate regions of the slab waveguide of Fig. 1, one can write the following equations

$$\vec{U} = \begin{cases} \vec{U}_{c1}e^{-j(k_{cx}x+k_{cz}z)} + \vec{U}_{c2}e^{-j(-k_{cx}x+k_{cz}z)}, & x > h, \\ \vec{U}_{f1}e^{-j(k_{fx}x+k_{fz}z)} + \vec{U}_{f2}e^{-j(-k_{fx}x+k_{fz}z)}, & 0 < x < h, \\ \vec{U}_{s1}e^{-j(k_{sx}x+k_{sz}z)} + \vec{U}_{s2}e^{-j(-k_{sx}x+k_{sz}z)}, & x < 0, \end{cases} \quad (18)$$

where the subscripts c , f , and s denote fields as well as wavevector components in cover, film, and substrate, respectively. From the phase matching condition along the boundaries ($x = 0$ for film-substrate and $x = h$ for film-cover) the following relation is necessary

$$k_{cz} = k_{fz} = k_{sz} = \beta. \quad (19)$$

At the same time the wave-vector components should satisfy the plane wave dispersion equation of the form $k_{wx}^2 + \beta^2 = k_0^2 n_w^2$ (for $w = c, f$, and s). In order to warrant total internal reflection at the film-cover and film-substrate boundaries β must satisfy the inequality $k_0 n_c \leq k_0 n_s < \beta < k_0 n_f$. Then the x components of the wavevectors become

$$k_{cx}^2 = k_0^2 n_c^2 - \beta^2 < 0 \Rightarrow k_{cx} = \pm j \sqrt{\beta^2 - k_0^2 n_c^2} = \pm j \gamma_c, \quad (20)$$

$$k_{fx}^2 = k_0^2 n_f^2 - \beta^2 > 0 \Rightarrow k_{fx} = \pm \sqrt{k_0^2 n_f^2 - \beta^2}, \quad (21)$$

$$k_{sx}^2 = k_0^2 n_s^2 - \beta^2 < 0 \Rightarrow k_{sx} = \pm j \sqrt{\beta^2 - k_0^2 n_s^2} = \pm j \gamma_s, \quad (22)$$

where the signs of the imaginary wave-vector components must be selected in such a way in order to warrant exponentially decaying solutions in the cover and the substrate regions. Thus, from the two possible solutions in the cover and the substrate regions only the exponentially decaying one are retained. Therefore the field solutions can be summarized as

$$\vec{U} = \begin{cases} \vec{U}_c e^{-\gamma_c(x-h)} e^{-j\beta z}, & x > h, \\ [\vec{U}_{f1} e^{-jk_{fx}x} + \vec{U}_{f2} e^{+jk_{fx}x}] e^{-j\beta z}, & 0 < x < h, \\ \vec{U}_s e^{\gamma_s x} e^{-j\beta z}, & x < 0, \end{cases} \quad (23)$$

The unknown constants of the above equations (field amplitudes and β) should be determined by suitable application of the boundary conditions. The boundary conditions are the continuity of

the tangential electric field and tangential magnetic field components at the film-cover and film-substrate boundaries. In order to proceed with the boundary conditions and the determination of the unknowns the two distinguished families of modes (TE and TM) will be treated separately since they are decoupled (in isotropic regions).

3.1 TE Guided Modes

As it was explained in the previous section the TE mode solutions include the field triplet $\{E_y, H_x, H_z\}$. Then the electric field described in Eq. (23) can be written in the form

$$\vec{E} = \hat{y} \begin{cases} E_c e^{-\gamma_c(x-h)} e^{-j\beta z}, & x \geq h, \\ [E_{f1} e^{-jk_{fx}x} + E_{f2} e^{+jk_{fx}x}] e^{-j\beta z}, & 0 \leq x \leq h, \\ E_s e^{\gamma_s x} e^{-j\beta z}, & x \leq 0, \end{cases} \quad (24)$$

The magnetic field components can be determined from Maxwell's equations, $H_x = (1/j\omega\mu_0)(dE_y/dz)$ and $H_z = -(1/j\omega\mu_0)(dE_y/dx)$, and are given by

$$H_x = -\frac{\beta}{\omega\mu_0} \begin{cases} E_c e^{-\gamma_c(x-h)} e^{-j\beta z}, & x \geq h, \\ [E_{f1} e^{-jk_{fx}x} + E_{f2} e^{+jk_{fx}x}] e^{-j\beta z}, & 0 \leq x \leq h, \\ E_s e^{\gamma_s x} e^{-j\beta z}, & x \leq 0, \end{cases} \quad (25)$$

and by

$$H_z = \frac{1}{\omega\mu_0} \begin{cases} -j\gamma_c E_c e^{-\gamma_c(x-h)} e^{-j\beta z}, & x \geq h, \\ [k_{fx} E_{f1} e^{-jk_{fx}x} - k_{fx} E_{f2} e^{+jk_{fx}x}] e^{-j\beta z}, & 0 \leq x \leq h, \\ j\gamma_s E_s e^{\gamma_s x} e^{-j\beta z}, & x \leq 0. \end{cases} \quad (26)$$

Using the continuity of the tangential electric and magnetic field components across the film-cover $[E_y(x = h^+) = E_y(x = h^-)$ and $H_z(x = h^+) = H_z(x = h^-)]$ and film-substrate boundaries $[E_y(x = 0^+) = E_y(x = 0^-)$ and $H_z(x = 0^+) = H_z(x = 0^-)]$ the following system of equations is formed

$$\underbrace{\begin{bmatrix} -1 & e^{-jk_{fx}h} & e^{+jk_{fx}h} & 0 \\ j\gamma_c & k_{fx}e^{-jk_{fx}h} & -k_{fx}e^{+jk_{fx}h} & 0 \\ 0 & 1 & 1 & -1 \\ 0 & k_{fx} & -k_{fx} & -j\gamma_s \end{bmatrix}}_{\hat{\mathcal{A}}_{TE}(\beta^2)} \begin{bmatrix} E_c \\ E_{f1} \\ E_{f2} \\ E_s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (27)$$

In order to have nontrivial solutions of the previous equation it is necessary that the determinant of $\tilde{\mathcal{A}}_{TE}$ be set to zero. After some manipulations the following dispersion equation is derived

$$\det\{\tilde{\mathcal{A}}_{TE}(\beta^2)\} = 0 \implies \tan(k_{fx}h) = \frac{\frac{\gamma_s}{k_{fx}} + \frac{\gamma_c}{k_{fx}}}{1 - \frac{\gamma_s}{k_{fx}} \frac{\gamma_c}{k_{fx}}}. \quad (28)$$

A graphical representation of Eq. (28) is shown in Fig. 4. There can be none, one, or multiple solutions depending on the waveguide parameters and the free-space wavelength. For this particular example (the waveguide parameters are given in the figure caption) there 4 TE and 4 TM solutions. This example case is the same with the one used for the graphical representation of Eq. (4).

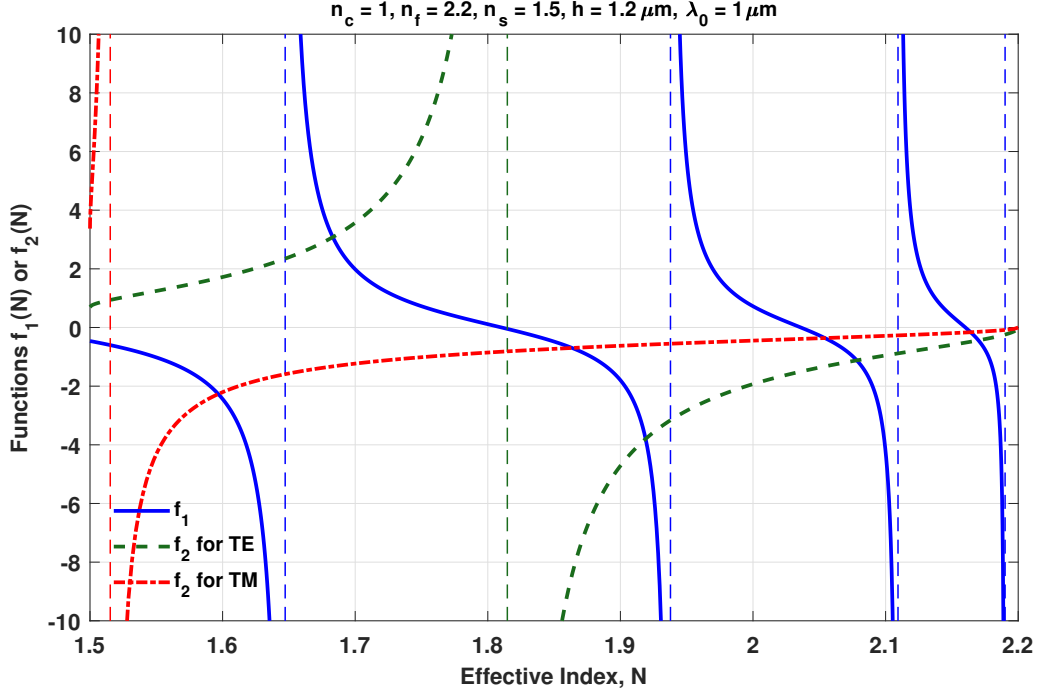


Figure 4: A graphical representation of the solution of Eq. (28) for TE modes. The functions f_1 and f_2 are given by $f_1(N) = \tan(k_{fx}h)$ and $f_2(N) = [(\gamma_s/k_{fx}) + (\gamma_c/k_{fx})]/[1 - (\gamma_c\gamma_s/k_{fx}^2)]$, where $k_{fx} = k_0(n_f^2 - N^2)^{1/2}$, $\gamma_c = k_0(N^2 - n_c^2)^{1/2}$, $\gamma_s = k_0(N^2 - n_s^2)^{1/2}$, and N is the effective index. The case parameters are $n_c = 1$, $n_f = 2.2$, $n_s = 1.5$, $h = 1.2\mu\text{m}$, and $\lambda_0 = 1.0\mu\text{m}$. The points along the effective index axis where f_1 is discontinuous and infinite are given by $N(m) = [n_f^2 - ((2m+1)/4)^2(\lambda_0/h)^2]^{1/2}$ ($m = 0, 1, \dots$). Similarly f_2 becomes infinite and discontinuous at $N = [(n_f^4 - n_s^2 n_c^2)/(2n_f^2 - n_c^2 - n_s^2)]^{1/2}$. In the case of TM modes the solution of Eq. (34) is also shown. In the latter case f_1 remains the same but f_2 becomes $f_2(N) = [(n_f^2/n_s^2)(\gamma_s/k_{fx}) + (n_f^2/n_c^2)(\gamma_c/k_{fx})]/[1 - (n_f^4/n_c^2 n_s^2)(\gamma_c\gamma_s/k_{fx}^2)]$. In the TM case f_2 becomes discontinuous at N being the solution of $(1 - a^2)N^4 + (2n_f^2 a^2 - n_s^2 - n_c^2)N^2 + (n_c^2 n_s^2 - n_f^4 a^2) = 0$, that lies in the interval $n_s < N < n_f$ (where $a = n_s^2 n_c^2 / n_f^4$).

It is straightforward to show that the last equation is equivalent to Eq. (4) (for $p = \text{TE}$). It is worth mentioning that the dispersion equation is actually a function of β^2 (or N^2) and not of just β (or N). This implies that if β_ν is a solution of Eq. (28) then $-\beta_\nu$ is also a solution that corresponds to the same mode propagating backwards (along the $-z$ axis). If β_ν satisfies Eq. (28) then the matrix $\tilde{\mathcal{A}}_{TE}$ becomes singular and correspondingly the boundary conditions contained in Eq. (27) become dependent. This fact provides the flexibility of selecting a free parameter and then solve for the electric field amplitudes E_c , E_{f1} , E_{f2} , and E_s as functions of this parameter. It is a simple task to perform this procedure and find that $E_c = E_0 \cos(k_{fx}h - \phi_{fs}^{TE})$, $E_{f1} = (E_0/2) \exp(+j\phi_{fs}^{TE})$, $E_{f2} = (E_0/2) \exp(-j\phi_{fs}^{TE})$, and $E_s = E_0 \cos \phi_{fs}^{TE}$ where E_0 is the free parameter and $\phi_{fs}^{TE} = \tan^{-1}(\gamma_s/k_{fx})$. The amplitude E_0 is the free parameter and can also be determined if information is known about the power that the mode carries. It is mentioned that all amplitudes are calculated for $\beta = \beta_\nu$ and therefore are characteristic of the TE_ν mode. Then the final form of the electric field of the TE_ν mode is given by

$$\vec{E}_\nu = \hat{y} \mathcal{E}_{y\nu}(x) e^{-j\beta_\nu z} = \hat{y} E_0 \begin{cases} \cos(k_{fx}h - \phi_{fs}^{TE}) e^{-\gamma_c(x-h)} & (x \geq h), \\ \cos(k_{fx}x - \phi_{fs}^{TE}) & (0 \leq x \leq h), \\ \cos \phi_{fs}^{TE} e^{\gamma_s x} & (x \leq 0), \end{cases} e^{-j\beta_\nu z} \quad (29)$$

where for simplicity the subscript ν , which is characteristic of the TE_ν mode, has been omitted from E_0 , k_{fx} , γ_c , γ_s and ϕ_{fs}^{TE} . The term $\mathcal{E}_{y\nu}(x)$ denotes the characteristic profile of the TE_ν mode. Example TE_ν mode electric field patterns are shown in Fig. 5 where the effective index solutions are also included. It can be observed that the number of zero-crossings of the electric field profile for TE_ν mode is ν .

3.2 TM Guided Modes

The TM mode solutions include the field triplet $\{H_y, E_x, E_z\}$. Then the magnetic field represented in Eqs. (23) can be written in the form

$$\vec{H} = \hat{y} \begin{cases} H_c e^{-\gamma_c(x-h)} e^{-j\beta z}, & x \geq h, \\ [H_{f1} e^{-jk_{fx}x} + H_{f2} e^{+jk_{fx}x}] e^{-j\beta z}, & 0 \leq x \leq h, \\ H_s e^{\gamma_s x} e^{-j\beta z}, & x \leq 0, \end{cases} \quad (30)$$

The electric field components of the TM guided mode can be easily determined from Maxwell's equations, i.e. $E_x = -(1/j\omega\epsilon)(dH_y/dz)$ and $E_z = (1/j\omega\epsilon)(dH_y/dx)$, and are given by

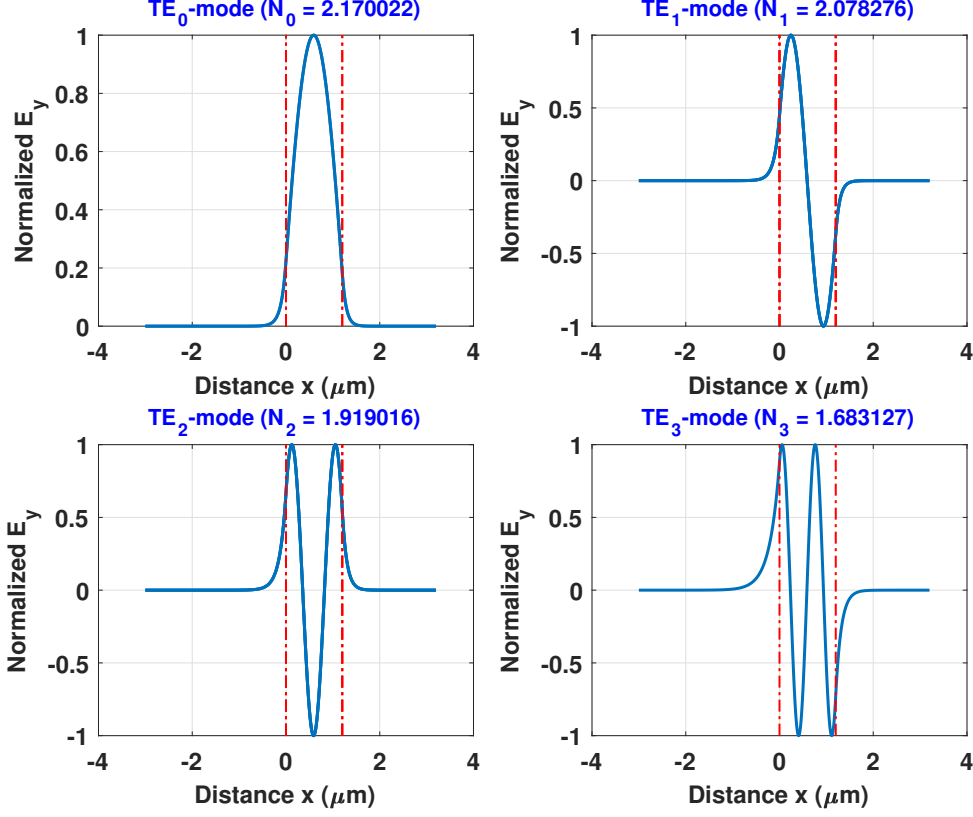


Figure 5: The normalized electric field profiles, $\mathcal{E}_{y\nu}(x)$, for TE_ν modes ($\nu = 0, 1, 2$, and 3). The slab waveguide parameters are $n_c = 1$, $n_f = 2.2$, $n_s = 1.5$, $h = 1.2\mu\text{m}$, and $\lambda_0 = 1.0\mu\text{m}$. The vertical lines indicate the film-substrate and film-cover boundaries. The effective indices are also shown on the top of each plot.

$$E_x = \frac{\beta}{\omega\epsilon_0} \begin{cases} \frac{1}{n_c^2} H_c e^{-\gamma_c(x-h)} e^{-j\beta z}, & x > h, \\ \frac{1}{n_f^2} [H_{f1} e^{-jk_{fx}x} + H_{f2} e^{+jk_{fx}x}] e^{-j\beta z}, & 0 < x < h, \\ \frac{1}{n_s^2} H_s e^{\gamma_s x} e^{-j\beta z}, & x < 0, \end{cases} \quad (31)$$

and by

$$E_z = \frac{1}{\omega\epsilon_0} \begin{cases} +j\frac{\gamma_c}{n_c^2} H_c e^{-\gamma_c(x-h)} e^{-j\beta z}, & x \geq h, \\ \left[\frac{k_{fx}}{n_f^2} H_{f1} e^{-jk_{fx}x} - \frac{k_{fx}}{n_f^2} H_{f2} e^{+jk_{fx}x} \right] e^{-j\beta z}, & 0 \leq x \leq h, \\ -j\frac{\gamma_s}{n_s^2} H_s e^{\gamma_s x} e^{-j\beta z}, & x \leq 0. \end{cases} \quad (32)$$

Using the continuity of the tangential electric and magnetic field components across the

film-cover [$H_y(x = h^+) = H_y(x = h^-)$ and $E_z(x = h^+) = E_z(x = h^-)$] and film-substrate boundaries [$H_y(x = 0^+) = H_y(x = 0^-)$ and $E_z(x = 0^+) = E_z(x = 0^-)$] the following system of equations is formed

$$\underbrace{\begin{bmatrix} -1 & e^{-jk_{fx}h} & e^{+jk_{fx}h} & 0 \\ -j\frac{\gamma_c}{n_c^2} & -\frac{k_{fx}}{n_f^2}e^{-jk_{fx}h} & \frac{k_{fx}}{n_f^2}e^{+jk_{fx}h} & 0 \\ 0 & 1 & 1 & -1 \\ 0 & -\frac{k_{fx}}{n_f^2} & \frac{k_{fx}}{n_f^2} & j\frac{\gamma_s}{n_s^2} \end{bmatrix}}_{\tilde{\mathcal{A}}_{TM}(\beta^2)} \begin{bmatrix} H_c \\ H_{f1} \\ H_{f2} \\ H_s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (33)$$

In order to have nontrivial solutions of the previous equation it is necessary that the determinant of $\tilde{\mathcal{A}}_{TM}$ be set to zero. After some manipulations, similar to the TE polarization case, the following dispersion equation is derived

$$\det\{\tilde{\mathcal{A}}_{TM}(\beta^2)\} = 0 \implies \tan(k_{fx}h) = \frac{\frac{n_f^2}{n_s^2} \frac{\gamma_s}{k_{fx}} + \frac{n_f^2}{n_c^2} \frac{\gamma_c}{k_{fx}}}{1 - \frac{n_f^4}{n_s^2 n_c^2} \frac{\gamma_c \gamma_s}{k_{fx}^2}}. \quad (34)$$

It is straightforward to show again that the last equation is equivalent to Eq. (4) (for $p = TM$). A graphical representation of Eq. (34) is also shown in Fig. 4. Similarly to the TE polarization case a solution $\pm\beta_\nu$ of Eq. (34) corresponds to the $\pm z$ -axis propagating mode (assuming that $k_0 n_s < \beta_\nu < k_0 n_f$). If β_ν satisfies Eq. (34) then the matrix $\tilde{\mathcal{A}}_{TM}$ becomes singular and correspondingly the boundary conditions contained in Eq. (33) become dependent. This provide the flexibility of selecting a free parameter and then solve for the magnetic field amplitudes H_c , H_{f1} , H_{f2} , and H_s as functions of this parameter. It is a simple task to perform this procedure and find that $H_c = H_0 \cos(k_{fx}h - \phi_{fs}^{TM})$, $H_{f1} = (H_0/2) \exp(+j\phi_{fs}^{TM})$, $H_{f2} = (H_0/2) \exp(-j\phi_{fs}^{TM})$, and $H_s = H_0 \cos \phi_{fs}^{TM}$ where H_0 is the free parameter and $\phi_{fs}^{TM} = \tan^{-1}[n_f^2 \gamma_s / (n_s^2 k_{fx})]$. The amplitude H_0 is the free parameter and can also be determined if information about the power that the mode carries is known. It is mentioned that all amplitudes are calculated for $\beta = \beta_\nu$ and therefore are characteristic of the TM_ν mode. Then the final form of the magnetic field of

the TM_ν mode is given by

$$\vec{H}_\nu = \hat{y}\mathcal{H}_{y\nu}(x)e^{-j\beta_\nu z} = \hat{y}H_0 \begin{cases} \cos(k_{fx}h - \phi_{fs}^{TM}) e^{-\gamma_c(x-h)} & (x \geq h), \\ \cos(k_{fx}x - \phi_{fs}^{TM}) & (0 \leq x \leq h), \\ \cos\phi_{fs}^{TM} e^{\gamma_s x} & (x \leq 0), \end{cases} \quad (35)$$

where for simplicity the subscript ν , which is characteristic of the TM_ν mode, has been omitted from H_0 , k_{fx} , γ_c , γ_s and ϕ_{fs}^{TM} . The term $\mathcal{H}_{y\nu}(x)$ denotes the profile of the TM_ν mode. Example TM_ν mode electric field patterns are shown in Fig. 6 where the effective index solutions are also shown. Again it is observed that the number of zero-crossings of the magnetic field profile for TM_ν mode is ν .

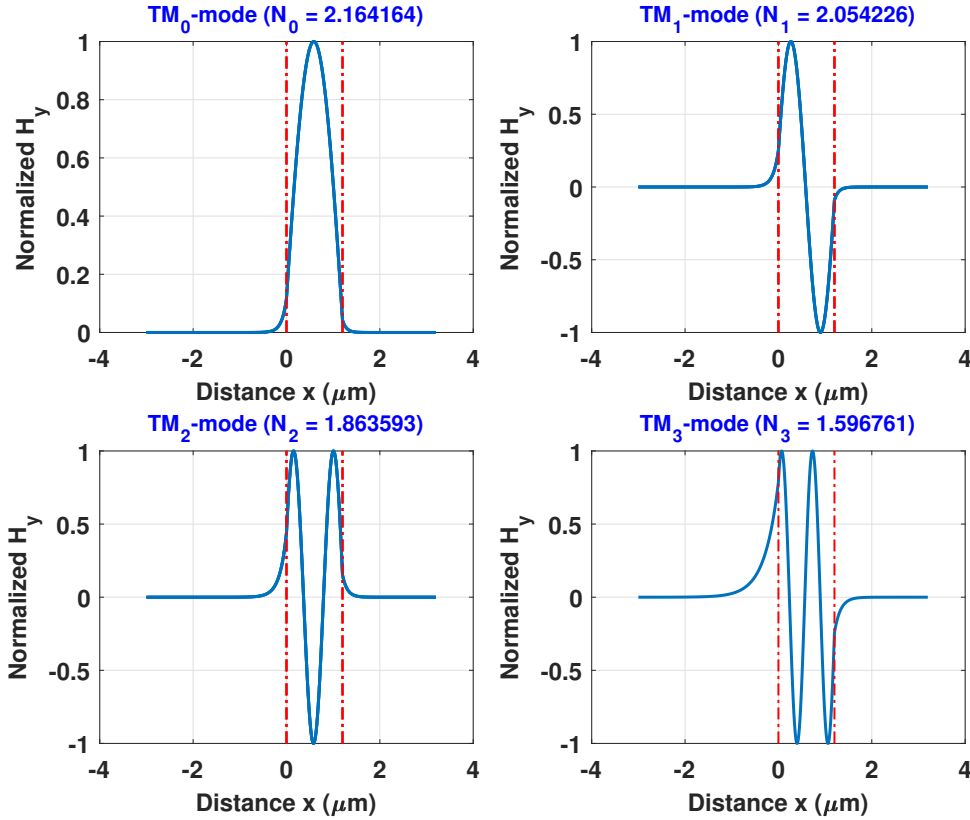


Figure 6: The normalized magnetic field profiles, $\mathcal{H}_{y\nu}(x)$, for TM_ν modes ($\nu = 0, 1, 2$, and 3). The slab waveguide parameters are $n_c = 1$, $n_f = 2.2$, $n_s = 1.5$, $h = 1.2\mu\text{m}$, and $\lambda_0 = 1.0\mu\text{m}$. The vertical lines indicate the film-substrate and film-cover boundaries. The effective indices are also shown on the top of each plot.

4. Substrate Modes

In the previous sections solutions for guided modes were determined. This was based on the assumption that $k_0 n_s < \beta < k_0 n_f$ (or equivalently $n_s < N < n_f$) in order to guarantee that total internal reflection occurs across the film-cover and film-substrate boundaries. When $k_0 n_c < \beta < k_0 n_s$ (or $n_c < N < n_s$) then the electromagnetic field in the substrate region is propagating instead of evanescent. The modes that have this property are called substrate modes and radiate power into the substrate as they propagate. For guided modes the solutions for β (or N) were discretized. In the case of the substrate modes the solutions for β form a continuum. Thus, any β in the interval $k_0 n_c < \beta < k_0 n_s$ can be a solution for a substrate mode. It is straightforward to determine the electromagnetic field that correspond to a substrate mode. The electric or magnetic field of a TE or TM substrate mode can be written as

$$\vec{U} = U(x, z)\hat{y} = \hat{y} \begin{cases} U_c e^{-\gamma_c(x-h)} e^{-j\beta z}, & x \geq h, \\ [U_{f1} e^{-jk_{fx}x} + U_{f2} e^{+jk_{fx}x}] e^{-j\beta z}, & 0 \leq x \leq h, \\ [U_{s1} e^{-jk_{sx}x} + U_{s2} e^{+jk_{sx}x}] e^{-j\beta z}, & x \leq 0, \end{cases} \quad (36)$$

where $U = E_y$ for TE substrate modes and $U = H_y$ for TM substrate modes, and $k_{sx} = (k_0^2 n_s^2 - \beta^2)^{1/2} > 0$ for the range of β that is valid for substrate modes. Using the continuity of the tangential electric and magnetic field components 4 equations can be specified. However, the number of unknowns is six, i.e. U_c , U_{f1} , U_{f2} , U_{s1} , U_{s2} , and β . Therefore, there is some flexibility in satisfying the boundary conditions. For example, any β in the interval $k_0 n_c < \beta < k_0 n_s$ can satisfy the boundary conditions. After some manipulations of the resulting boundary conditions it can be shown that the electric field of a TE substrate mode is given by

$$\vec{E}_\beta = \hat{y} \mathcal{E}_{y\beta}(x) e^{-j\beta z} = \hat{y} E_0 \begin{cases} \cos \phi_{fc}^{TE} e^{-\gamma_c(x-h)} & (x \geq h), \\ \cos[k_{fx}(x-h) + \phi_{fc}^{TE}] & e^{-j\beta_\nu z} \quad (0 \leq x \leq h), \\ \cos(k_{fx}h - \phi_{fc}^{TE}) \cos(k_{sx}x) + \frac{k_{fx}}{k_{sx}} \sin(k_{fx}h - \phi_{fc}^{TE}) \sin(k_{sx}x) & (x \leq 0), \end{cases} \quad (37)$$

where E_0 is a free parameter, $\mathcal{E}_{y\beta}(x)$ is the substrate TE mode profile, and $\phi_{fc}^{TE} = \tan^{-1}(\gamma_c/k_{fx})$. Similarly the magnetic field for a TM substrate mode is given by

$$\vec{H}_\beta = \hat{y}\mathcal{H}_{y\beta}(x)e^{-j\beta z} = \hat{y}H_0 \begin{cases} \cos \phi_{fc}^{TM} e^{-\gamma_c(x-h)} & (x \geq h), \\ \cos[k_{fx}(x-h) + \phi_{fc}^{TM}] & (0 \leq x \leq h), \\ \cos(k_{fx}h - \phi_{fc}^{TM}) \cos(k_{sx}x) + \frac{k_{fx}/n_f^2}{k_{sx}/n_s^2} \sin(k_{fx}h - \phi_{fc}^{TM}) \sin(k_{sx}x) & (x \leq 0), \end{cases} e^{-j\beta_\nu z} \quad (38)$$

where H_0 is a free parameter, $\mathcal{H}_{y\beta}(x)$ is the substrate TM mode profile, and $\phi_{fc}^{TM} = \tan^{-1}(\gamma_c n_f^2 / k_{fx} n_c^2)$. Sample normalized electric/magnetic field profiles for TE/TM substrate modes are shown in Fig. 7.

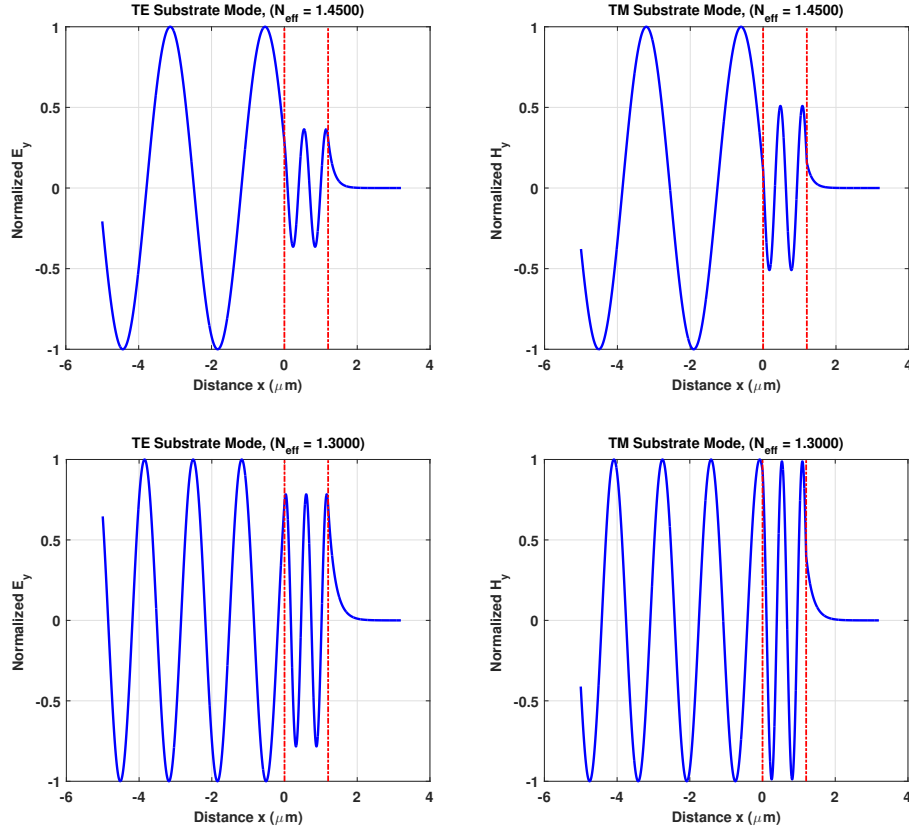


Figure 7: The normalized electric/magnetic $\mathcal{E}_{y\beta}(x)/\mathcal{H}_{y\beta}(x)$ field profiles for TE/TM substrate modes. The slab waveguide parameters are $n_c = 1$, $n_f = 2.2$, $n_s = 1.5$, $h = 1.2\mu\text{m}$, and $\lambda_0 = 1.0\mu\text{m}$. The vertical lines indicate the film-substrate and film-cover boundaries. The effective indices (selected arbitrarily in the interval $n_c = 1 < N < 1.5 = n_s$) are also shown on the top of each plot.

5. Radiation Modes

When $0 < \beta < k_0 n_c$ (or equivalently $0 < N < n_c$) then the field in all three regions (cover, film, and substrate) is propagating. The modes that have this property are called *radiation modes* and radiate power into both the cover and the substrate regions as they propagate. As in the case of the substrate modes, the solutions for β form a continuum and for the radiation modes. Thus, any β in the interval $0 < \beta < k_0 n_c$ can be a solution for a radiation mode. It is straightforward to determine the electromagnetic field of a radiation mode. The electric or magnetic field of a TE or TM radiation mode can be written as

$$\vec{U} = U(x, z)\hat{y} = \hat{y} \begin{cases} [U_{c1}e^{-jk_{cx}(x-h)} + U_{c2}e^{+jk_{cx}(x-h)}] e^{-j\beta z}, & x \geq h, \\ [U_{f1}e^{-jk_{fx}x} + U_{f2}e^{+jk_{fx}x}] e^{-j\beta z}, & 0 \leq x \leq h, \\ [U_{s1}e^{-jk_{sx}x} + U_{s2}e^{+jk_{sx}x}] e^{-j\beta z}, & x \leq 0, \end{cases} \quad (39)$$

where $U = E_y$ for TE radiation modes and $U = H_y$ for TM radiation modes, and $k_{cx} = (k_0^2 n_c^2 - \beta^2)^{1/2} > 0$ for the range of β that is valid for radiation modes. Using the continuity of the tangential electric and magnetic field components 4 equations can be specified. However, the number of unknowns is seven, i.e. U_{c1} , U_{c2} , U_{f1} , U_{f2} , U_{s1} , U_{s2} , and β . Therefore, there is even more flexibility in satisfying the boundary conditions. For example, any β in the interval $0 < \beta < k_0 n_c$ can satisfy the boundary conditions. After some manipulations of the resulting boundary conditions it can be shown that the electric field of a TE radiation mode is given by

$$\vec{E}_\beta = \hat{y}\mathcal{E}_{y\beta}(x)e^{-j\beta z} = \quad (40)$$

$$\hat{y}E_0 \begin{cases} \frac{1}{2} \left[\left(1 + \frac{k_{fx}}{k_{cx}} \right) \cos[k_{cx}(x-h) + k_{fx}h - \phi] + \left(1 - \frac{k_{fx}}{k_{cx}} \right) \cos[k_{cx}(x-h) - k_{fx}h + \phi] \right] & (x \geq h), \\ \cos(k_{fx}x - \phi) & (0 \leq x \leq h), \\ \frac{1}{2} \left[\left(1 + \frac{k_{fx}}{k_{sx}} \right) \cos[k_{sx}x - \phi] + \left(1 - \frac{k_{fx}}{k_{sx}} \right) \cos[k_{sx}x + \phi] \right] & (x \leq 0), \end{cases} e^{-j\beta z}$$

where E_0 and ϕ are two free parameters. Similarly the magnetic field for a TM radiation mode

is given by

$$\vec{H}_\beta = \hat{y}\mathcal{H}_{y\beta}(x)e^{-j\beta z} = \quad (41)$$

$$\hat{y}H_0 \left\{ \begin{array}{ll} \frac{1}{2} \left[\left(1 + \frac{k_{fx}/n_f^2}{k_{cx}/n_c^2} \right) \cos[k_{cx}(x-h) + k_{fx}h - \phi] + \right. & (x \geq h), \\ \left. \left(1 - \frac{k_{fx}/n_f^2}{k_{cx}/n_c^2} \right) \cos[k_{cx}(x-h) - k_{fx}h + \phi] \right] & \\ \cos(k_{fx}x - \phi) & (0 \leq x \leq h), \\ \frac{1}{2} \left[\left(1 + \frac{k_{fx}/n_f^2}{k_{sx}/n_s^2} \right) \cos[k_{sx}x - \phi] + \left(1 - \frac{k_{fx}/n_f^2}{k_{sx}/n_s^2} \right) \cos[k_{sx}x + \phi] \right] & (x \leq 0), \end{array} \right\} e^{-j\beta z}$$

where H_0 and ϕ are two free parameters. Sample normalized electric/magnetic field profiles for TE/TM radiation modes are shown in Fig. 8.

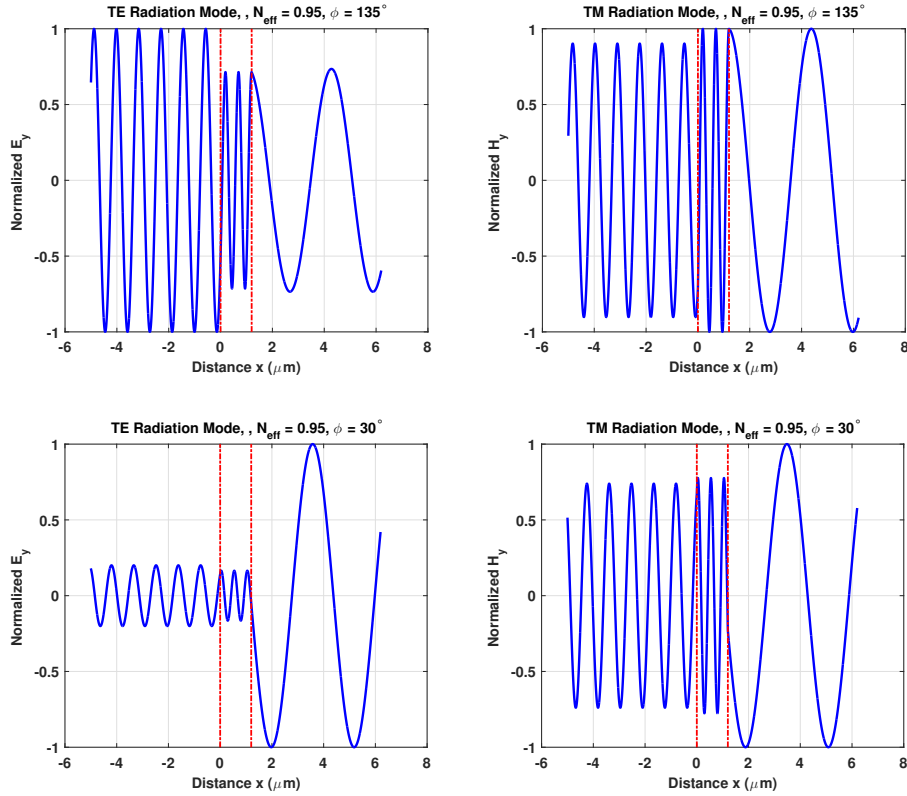


Figure 8: The normalized electric/magnetic field profiles, $\mathcal{E}_{y\beta}(x)/\mathcal{H}_{y\beta}(x)$, for TE/TM radiation modes. The slab waveguide parameters are $n_c = 1$, $n_f = 2.2$, $n_s = 1.5$, $h = 1.2\mu\text{m}$, and $\lambda_0 = 1.0\mu\text{m}$. The vertical lines indicate the film-substrate and film-cover boundaries. The effective indices (arbitrarily chosen in the interval $0 < N < 1 = n_c$) are also shown on the top of each plot along with the choice of the free parameter ϕ .

6. Unphysical Modes (or Nonphysical Modes)

When $k_0 n_f < \beta < \infty$ (or equivalently $n_f < N < \infty$) then the field in all three regions (cover, film, and substrate) is comprised of evanescent terms, usually including both exponentially increasing and exponentially decreasing terms. The modes that have this property are called *unphysical* (or nonphysical) modes since there are valid solutions of the Maxwell's equations but they cannot be excited since they have infinite power. For this reason they are called unphysical modes and are not usually referred to most of the textbooks. The analysis of these modes is similar with the case of radiation modes with the exception that all field components are comprised of both increasing and decreasing exponential terms. As in the case of the substrate/radiation modes, their solutions for β form a continuum. Thus, any β in the interval $k_0 n_f < \beta < \infty$ can be a solution for an unphysical mode. It is straightforward to determine the electromagnetic field of an unphysical mode. The electric or magnetic field of a TE or TM unphysical mode can be written as

$$\vec{U} = U(x, z)\hat{y} = \hat{y} \begin{cases} [U_{c1}e^{-\gamma_c(x-h)} + U_{c2}e^{+\gamma_c(x-h)}] e^{-j\beta z}, & x \geq h, \\ [U_{f1}e^{-\gamma_f x} + U_{f2}e^{+\gamma_f x}] e^{-j\beta z}, & 0 \leq x \leq h, \\ [U_{s1}e^{-\gamma_s x} + U_{s2}e^{+\gamma_s x}] e^{-j\beta z}, & x \leq 0, \end{cases} \quad (42)$$

where $U = E_y$ for TE unphysical modes and $U = H_y$ for TM unphysical modes, and $\gamma_f = (\beta^2 - k_0^2 n_f^2)^{1/2} > 0$ for the range of β that is valid for unphysical modes. If one tries to find a solution without the increasing exponentials in the cover and the substrate regions (i.e., for $U_{c2} = U_{s1} = 0$), one ends with a dispersion equation of the form (for TE polarization) $\tanh(\gamma_f h) = -[(\gamma_s/\gamma_f) + (\gamma_c/\gamma_f)]/[1 + (\gamma_c\gamma_s/\gamma_f^2)]$ that does not have a real solution for β in the range of the unphysical modes (a similar equation, that does not have a real solution, exists for the TM polarization too). Therefore, it is required to retain all the terms appearing in the above equation similarly to the radiation modes case. Using the continuity of the tangential electric and magnetic field components 4 equations can be specified. However, the number of unknowns is seven, i.e. U_{c1} , U_{c2} , U_{f1} , U_{f2} , U_{s1} , U_{s2} , and β (similar to the radiation modes case). Therefore, there is a lot of flexibility in satisfying the boundary conditions. For example, any β in the interval $k_0 n_f < \beta < \infty$ can satisfy the boundary conditions. After some manipulations of the resulting boundary conditions it can be shown that the electric field of a TE unphysical

mode is given by

$$\vec{E}_\beta = \hat{y}\mathcal{E}_{y\beta}(x)e^{-j\beta z} = \quad (43)$$

$$\hat{y}E_0 \left\{ \begin{array}{ll} \frac{1}{2} \left[\left(1 + \frac{\gamma_f}{\gamma_c} \right) \cosh[\gamma_c(x-h) + \gamma_f h - \phi] + \right. & (x \geq h), \\ \left. \left(1 - \frac{\gamma_f}{\gamma_c} \right) \cosh[\gamma_c(x-h) - \gamma_f h + \phi] \right] & \\ \cosh(\gamma_f x - \phi) & (0 \leq x \leq h), \\ \frac{1}{2} \left[\left(1 + \frac{\gamma_f}{\gamma_s} \right) \cosh[\gamma_s x - \phi] + \left(1 - \frac{\gamma_f}{\gamma_s} \right) \cosh[\gamma_s x + \phi] \right] & (x \leq 0), \end{array} \right\} e^{-j\beta z}$$

where E_0 and ϕ are two free parameters. Similarly the magnetic field for a TM unphysical mode is given by

$$\vec{H}_\beta = \hat{y}\mathcal{H}_{y\beta}(x)e^{-j\beta z} = \quad (44)$$

$$\hat{y}H_0 \left\{ \begin{array}{ll} \frac{1}{2} \left[\left(1 + \frac{\gamma_f/n_f^2}{\gamma_c/n_c^2} \right) \cosh[\gamma_c(x-h) + \gamma_f h - \phi] + \right. & (x \geq h), \\ \left. \left(1 - \frac{\gamma_f/n_f^2}{\gamma_c/n_c^2} \right) \cosh[\gamma_c(x-h) - \gamma_f h + \phi] \right] & \\ \cosh(\gamma_f x - \phi) & (0 \leq x \leq h), \\ \frac{1}{2} \left[\left(1 + \frac{\gamma_f/n_f^2}{\gamma_s/n_s^2} \right) \cosh[\gamma_s x - \phi] + \left(1 - \frac{\gamma_f/n_f^2}{\gamma_s/n_s^2} \right) \cosh[\gamma_s x + \phi] \right] & (x \leq 0), \end{array} \right\} e^{-j\beta z}$$

where H_0 and ϕ are two free parameters. Sample normalized electric/magnetic field profiles for TE/TM radiation modes are shown in Fig. 9.

7. Evanescent Modes

In some cases there might be a need to use waveguide modes that have a purely imaginary propagation constant $\tilde{\beta}$, such that $\tilde{\beta} = \pm j\beta$ (where $\beta > 0$). In this case the waveguide field could be written as $\vec{U} = \vec{U}(x) \exp(-j\tilde{\beta}z) = \vec{U}(x) \exp(-\beta z)$, where for forward evanescent modes the $\tilde{\beta} = -j\beta$ was selected in order for the evanescent mode to decay along the propagation direction z . Strictly speaking, the evanescent modes are not propagating since they decay along the propagation direction. However, they are proper modal solutions of the Helmholtz wave equation as it is applied to the slab waveguide problem. The acceptable imaginary propagation

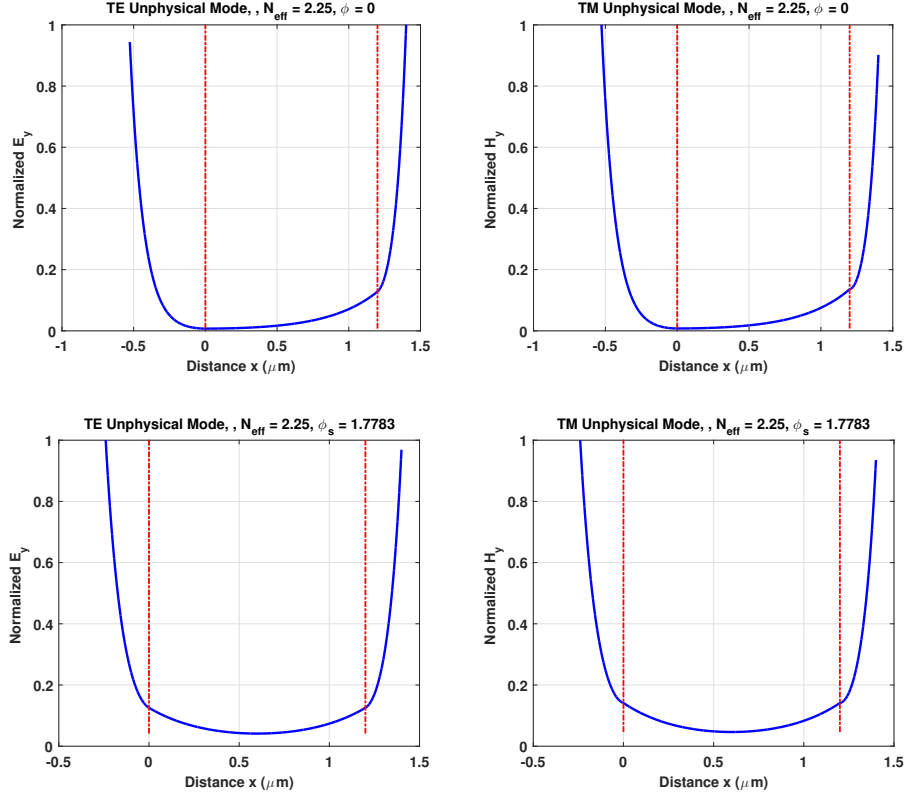


Figure 9: The normalized electric/magnetic field profiles, $\mathcal{E}_{y\beta}(x)/\mathcal{H}_{y\beta}(x)$, for TE/TM unphysical modes. The slab waveguide parameters are $n_c = 1$, $n_f = 2.2$, $n_s = 1.5$, $h = 1.2\mu\text{m}$, and $\lambda_0 = 1.0\mu\text{m}$. The vertical lines indicate the film-substrate and film-cover boundaries. The effective indices are also shown on the top of each plot along with the choice of the free parameter ϕ . For symmetry of the fields in the film layer the ϕ parameter was chosen as $\phi = \phi_s = \gamma_f h/2$ in the bottom two plots.

constants are of the form $\tilde{\beta} = -j\beta$ (for the forward evanescent modes) and $\tilde{\beta} = +j\beta$ (for the backward evanescent modes), i.e. they all lie along the imaginary $\tilde{\beta}$ axis. These evanescent modes form a continuum. For example any value of $\beta > 0$ gives a propagation constant $\tilde{\beta} = -j\beta$ which is an acceptable solution to Maxwell equations in the slab waveguide geometry.

It is straightforward to determine the electromagnetic field of an evanescent mode. The electric or magnetic field of a TE or TM forward propagating evanescent mode can be written as

$$\vec{U} = U(x, z)\hat{y} = \hat{y} \begin{cases} [U_{c1}e^{-jk'_{cx}(x-h)} + U_{c2}e^{+jk'_{cx}(x-h)}] e^{-\beta z}, & x \geq h, \\ [U_{f1}e^{-jk'_{fx}x} + U_{f2}e^{+jk'_{fx}x}] e^{-\beta z}, & 0 \leq x \leq h, \\ [U_{s1}e^{-jk'_{sx}x} + U_{s2}e^{+jk'_{sx}x}] e^{-\beta z}, & x \leq 0, \end{cases} \quad (45)$$

where $U = E_y$ for TE forward evanescent modes and $U = H_y$ for TM forward evanescent

modes, and $k'_{cx} = (k_0^2 n_c^2 + \beta^2)^{1/2} > 0$, $k'_{fx} = (k_0^2 n_f^2 + \beta^2)^{1/2} > 0$, and $k'_{sx} = (k_0^2 n_s^2 + \beta^2)^{1/2} > 0$, for the range of β that is valid for forward evanescent modes (i.e. the negative imaginary axis of the propagation constant region). Using the continuity of the tangential electric and magnetic field components 4 equations can be specified. However, the number of unknowns is seven, i.e. U_{c1} , U_{c2} , U_{f1} , U_{f2} , U_{s1} , U_{s2} , and β . Therefore, there is even more flexibility in satisfying the boundary conditions. For example, any β in the interval $0 < \beta < k_0 n_c$ can satisfy the boundary conditions. After some manipulations of the resulting boundary conditions it can be shown that the electric field of a TE evanescent mode is given by

$$\vec{E}_\beta = \hat{y} \mathcal{E}_{y\beta}(x) e^{-\beta z} = \quad (46)$$

$$\hat{y} E_0 \left\{ \begin{array}{ll} \frac{1}{2} \left[\left(1 + \frac{k'_{fx}}{k'_{cx}} \right) \cos[k'_{cx}(x-h) + k'_{fx}h - \phi] + \right. & (x \geq h), \\ \left. \left(1 - \frac{k'_{fx}}{k'_{cx}} \right) \cos[k'_{cx}(x-h) - k'_{fx}h + \phi] \right] & \\ \cos(k'_{fx}x - \phi) & (0 \leq x \leq h), \\ \frac{1}{2} \left[\left(1 + \frac{k'_{fx}}{k'_{sx}} \right) \cos[k'_{sx}x - \phi] + \left(1 - \frac{k'_{fx}}{k'_{sx}} \right) \cos[k'_{sx}x + \phi] \right] & (x \leq 0), \end{array} \right\} e^{-\beta z}$$

where E_0 and ϕ are two free parameters (similar to the radiation modes). In an analogous manner, the magnetic field for a TM forward evanescent mode is given by

$$\vec{H}_\beta = \hat{y} \mathcal{H}_{y\beta}(x) e^{-\beta z} = \quad (47)$$

$$\hat{y} H_0 \left\{ \begin{array}{ll} \frac{1}{2} \left[\left(1 + \frac{k'_{fx}/n_f^2}{k'_{cx}/n_c^2} \right) \cos[k'_{cx}(x-h) + k'_{fx}h - \phi] + \right. & (x \geq h), \\ \left. \left(1 - \frac{k'_{fx}/n_f^2}{k'_{cx}/n_c^2} \right) \cos[k'_{cx}(x-h) - k'_{fx}h + \phi] \right] & \\ \cos(k'_{fx}x - \phi) & (0 \leq x \leq h), \\ \frac{1}{2} \left[\left(1 + \frac{k'_{fx}/n_f^2}{k'_{sx}/n_s^2} \right) \cos[k'_{sx}x - \phi] + \left(1 - \frac{k'_{fx}/n_f^2}{k'_{sx}/n_s^2} \right) \cos[k'_{sx}x + \phi] \right] & (x \leq 0), \end{array} \right\} e^{-\beta z}$$

where H_0 and ϕ are again the two free parameters. Sample normalized electric/magnetic field profiles for TE/TM forward evanescent modes are shown in Fig. 10.

The various modes of the lossless slab waveguide problem (the example is for the single-film layer slab waveguide but it can be easily generalized to multilayer-film slab waveguides) are shown in the complex propagation constant ($\tilde{\beta} = \beta_r + j\beta_i$) diagram in Fig. 11. The only discrete

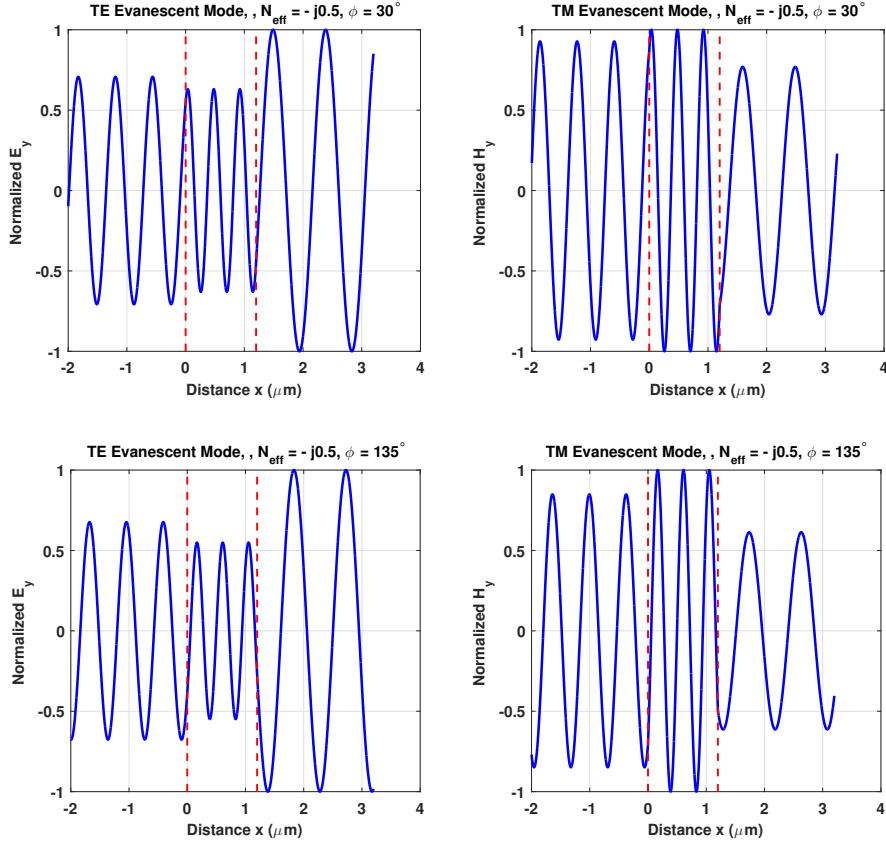


Figure 10: The normalized electric/magnetic field profiles, $\mathcal{E}_{y\beta}(x)/\mathcal{H}_{y\beta}(x)$, for TE/TM forward evanescent modes. The slab waveguide parameters are $n_c = 1$, $n_f = 2.2$, $n_s = 1.5$, $h = 1.2\mu\text{m}$, and $\lambda_0 = 1.0\mu\text{m}$. The vertical lines indicate the film-substrate and film-cover boundaries. The effective indices are also shown on the top of each plot along with the choice of the free parameter ϕ .

spectrum is of the guided modes as it can be seen. The substrate, radiation, unphysical, and evanescent modes from a continuum. In the same diagram the leaky modes are also shown. The leaky modes is an approximation of the radiation field by a series of discrete modes of complex propagation constant. These leaky modes will be discussed at a later section of this chapter. In the case of loss or gain in the film region, the guided modes will have nonzero component of their complex propagation constant, $\beta_i < 0$ for loss and $\beta_i > 0$ for gain (for the forward propagating modes and the reverse of this for the backward propagation modes).

8. Normalized Slab-Waveguide Variables

The dispersion Eq. (4) can also be written in terms of normalized variables. Such a representation is useful for any three-region slab waveguide. The normalized parameters are the

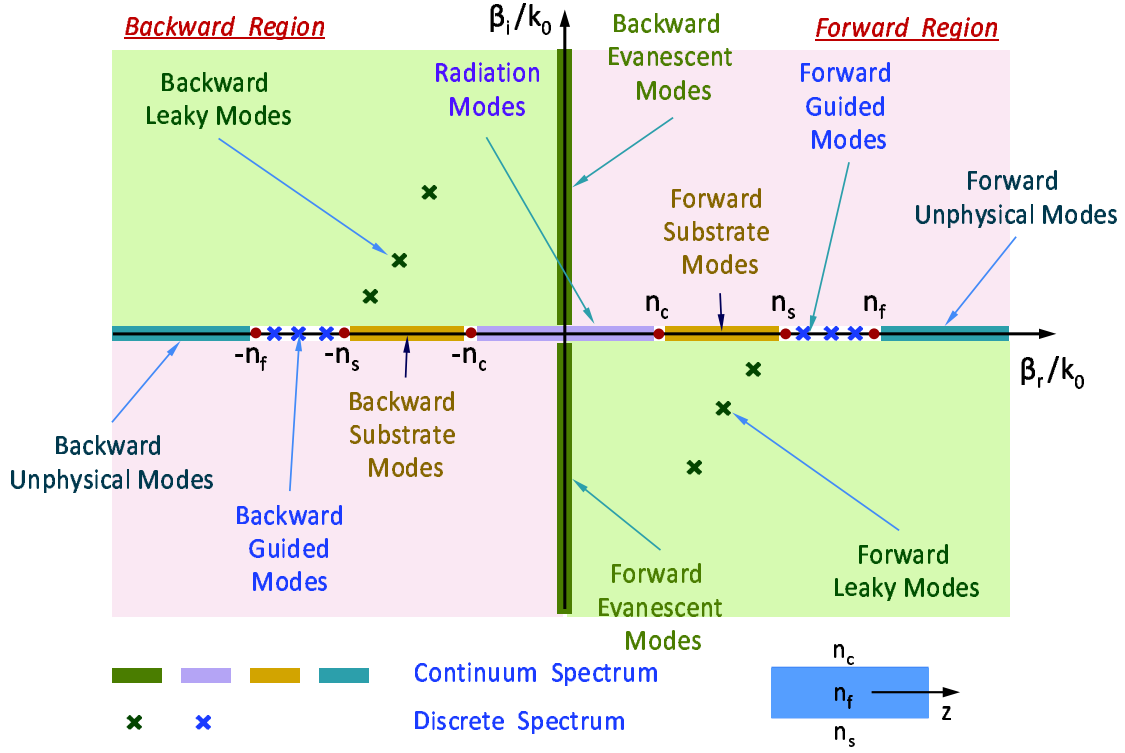


Figure 11: A diagram of the various modes of a lossless slab waveguide. The guided modes (forward or backward) form a discrete spectrum while the radiation, substrate, and evanescent modes (forward or backward) form a continuum spectrum. The unphysical mode solutions also have a continuum spectrum (forward or backward). The leaky modes are not exactly formal modes but they represent some form of normalization of the radiation spectrum (forward or backward).

normalized frequency, V , the normalized effective propagation constant (or normalized effective index), b , and are defined as follows

$$V = \frac{2\pi}{\lambda_0} h \sqrt{n_f^2 - n_s^2}, \quad (48)$$

$$b = \frac{N^2 - n_s^2}{n_f^2 - n_s^2} = \frac{n_f^2 \sin^2 \theta - n_s^2}{n_f^2 - n_s^2}, \quad (49)$$

$$a_{TE} = \frac{n_s^2 - n_c^2}{n_f^2 - n_s^2}, \quad (50)$$

$$a_{TM} = \frac{n_f^4}{n_c^4} \frac{n_s^2 - n_c^2}{n_f^2 - n_s^2} = \frac{n_f^4}{n_c^4} a_{TE}, \quad (51)$$

where the a_{TE} and a_{TM} are defined as the asymmetry factors for TE and TM modes respectively. Using these normalized variables in Eq. (4) it is straightforward to derive the following dispersion equation for the case of TE polarization

$$V\sqrt{1-b} - \tan^{-1} \left\{ \sqrt{\frac{b}{1-b}} \right\} - \tan^{-1} \left\{ \sqrt{\frac{b+a_{TE}}{1-b}} \right\} = \nu\pi, \quad \nu = 0, 1, \dots \quad (52)$$

A similar normalized dispersion equation can not be done for the TM modes. It can be shown that by defining the normalized variables, $b_{TM} = (n_f^2/qn_s^2)[(N^2 - n_s^2)/(n_f^2 - n_s^2)]$ and $q = (N^2/n_f^2) + (N^2/n_s^2) - 1$, the resulting normalized dispersion equation for the TM modes is given by

$$V\sqrt{q}\frac{n_f}{n_s}\sqrt{1-b_{TM}} - \tan^{-1} \left\{ \sqrt{\frac{b_{TM}}{1-b_{TM}}} \right\} - \tan^{-1} \left\{ \sqrt{\frac{b_{TM}+a_{TM}(1-b_{TM}d)}{1-b_{TM}}} \right\} = \nu\pi, \quad (53)$$

where $\nu = 0, 1, \dots$ and $d = [1 - (n_s/n_f)^2][1 - (n_c/n_f)^2]$. It is obvious that the above normalized equation is not completely normalized since the ratios n_s/n_f , n_c/n_f , and n_s , n_f , are necessary for the evaluation of b_{TM} . However, if $n_f - n_s \simeq 0$ then the normalized dispersion equation for the TM modes becomes similar to the one for the TE modes if the corresponding to the TM modes asymmetry factor, a_{TM} , is used.

Solving Eq. (52) for various values of the normalized frequency, V , of the asymmetry factor, a (a_{TE} or a_{TM}), and of ν , normalized diagrams such the one shown in Fig. 12 can be generated. These are exact for TE modes and only approximately correct for TM modes as it was explained above. Measuring from the diagram the value(s) of b_ν , for given V and a , the effective index of the TE_ν mode(s) (and approximately for TM_ν mode(s) with $a = a_{TM}$) can be determined as $N_\nu = [n_s^2 + b_\nu(n_f^2 - n_s^2)]^{1/2}$.

9. Cutoff Conditions

When the effective index of a guided mode becomes smaller than the substrate refractive index ($N_\nu < n_s$) then the mode starts radiating into the substrate and becomes a substrate mode as it was discussed in the previous sections. The condition $N_\nu = n_s$ (or equivalently $\beta_\nu = k_0 n_s$) represents a critical condition below which the guided field is cut-off and from guided mode becomes a substrate (or in general radiation) mode. Equivalently, this occurs when the zig-zag angle θ becomes equal to the critical angle at the film-substrate interface, $\theta_{cr,fs}$ (it is reminded

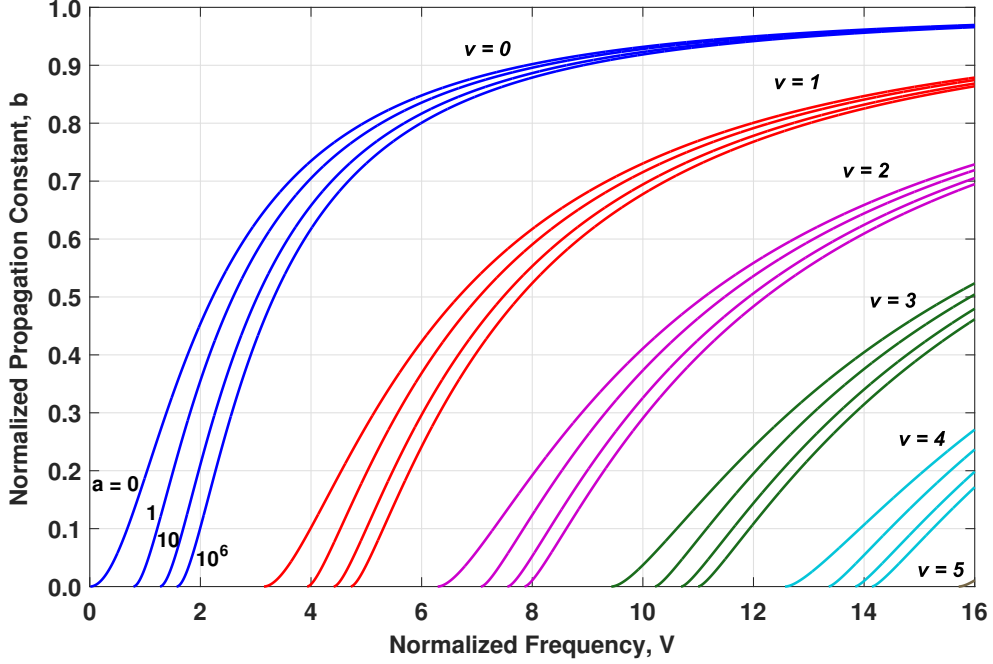


Figure 12: The normalized propagation constant b versus normalized frequency V for four selected asymmetry factors ($a = 0, 1, 10$, and 10^6) for the first 5 modes ($\nu = 0, 1, 2, 3$, and 4). A small segment for $\nu = 5$ and $a = 0$ is shown on the right bottom corner of the diagram.

that $n_c < n_s < n_f$). Then, Eq. (4) becomes

$$k_0 h \sqrt{n_f^2 - n_s^2} - \tan^{-1}(\sqrt{a_w}) = \nu\pi, \quad w = TE, TM. \quad (54)$$

The last equation is a cutoff condition and can be solved for one parameter retaining the others constant. For example, if it solved with respect to the waveguide thickness, h , the cutoff thicknesses can be determined as follows

$$h_{cut,\nu}^{TE} = \frac{\nu\pi + \tan^{-1}(\sqrt{a_{TE}})}{k_0 \sqrt{n_f^2 - n_s^2}}, \quad (55)$$

$$h_{cut,\nu}^{TM} = \frac{\nu\pi + \tan^{-1}(\sqrt{a_{TM}})}{k_0 \sqrt{n_f^2 - n_s^2}}, \quad (56)$$

where a_{TE} and a_{TM} are the asymmetry factors for TE and TM polarization respectively and are defined in Eqs. (50) and (51). It is interesting to notice that for symmetric waveguides, $a_{TE} = a_{TM} = 0$ the cutoff thicknesses for the lowest order mode ($\nu = 0$) are zero. I.e., for symmetric waveguides, any film thickness, independently of how small it is, can support the

lowest order modes TE_0 and TM_0 . In Fig. 13 the effective index of various modes is shown as a function of the film layer thickness. From these plots the cutoff thicknesses are the points from which each of the β_ν curves starts.

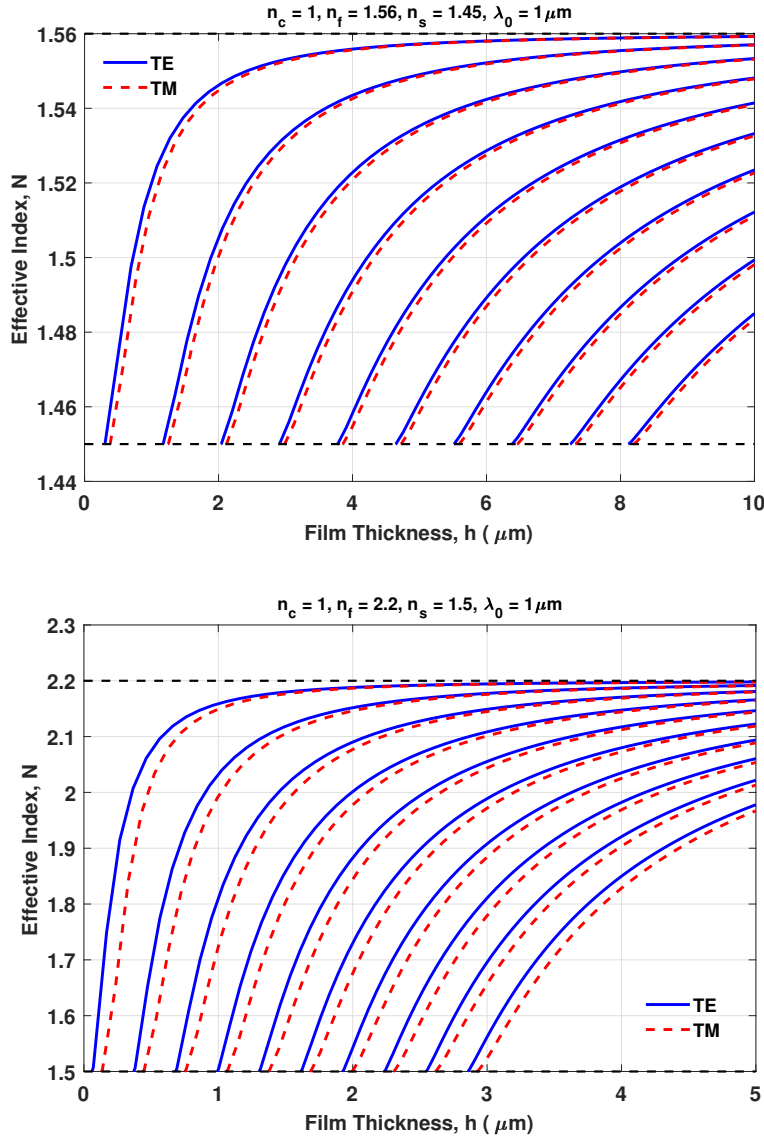


Figure 13: The effective index variation as a function of the waveguide film thickness. The slab waveguide parameters are shown on the top of each plot. It is mentioned that the effective indices of up to the tenth TE or TM are included.

If Eq. (54) is solved with respect to the free-space wavelength or frequency the following

cutoff wavelengths and frequencies can be defined for each mode and polarization

$$\lambda_{0,cut,\nu}^{TE} = \frac{2\pi h \sqrt{n_f^2 - n_s^2}}{\nu\pi + \tan^{-1}(\sqrt{a_{TE}})}, \quad (57)$$

$$\lambda_{0,cut,\nu}^{TM} = \frac{2\pi h \sqrt{n_f^2 - n_s^2}}{\nu\pi + \tan^{-1}(\sqrt{a_{TM}})}, \quad (58)$$

$$(59)$$

$$\omega_{cut,\nu}^{TE} = c \frac{\nu\pi + \tan^{-1}(\sqrt{a_{TE}})}{h \sqrt{n_f^2 - n_s^2}}, \quad (60)$$

$$\omega_{cut,\nu}^{TM} = c \frac{\nu\pi + \tan^{-1}(\sqrt{a_{TM}})}{h \sqrt{n_f^2 - n_s^2}}. \quad (61)$$

It is evident again that for symmetric waveguides the cutoff wavelength is infinite and the corresponding cutoff frequencies are zero for the zero order modes. An example $\omega - \beta$ diagram is shown in Fig. 14 where the cutoff frequencies can be shown for the first five TE and TM modes. The waveguide parameters are included in the figure caption.

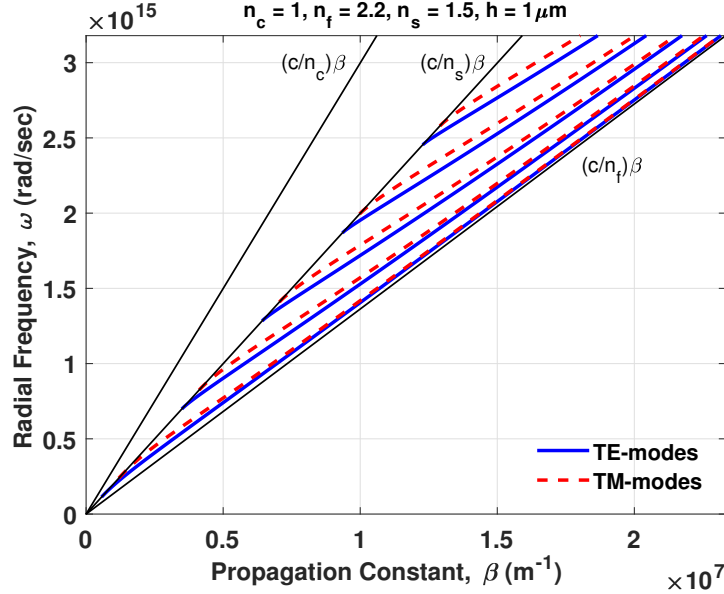


Figure 14: A sample $\omega - \beta$ diagram for a three layer slab waveguide with $n_c = 1.0$, $n_f = 2.2$, $n_s = 1.5$, and $h = 1\mu\text{m}$ for both TE and TM polarizations.

10. Power Considerations

In order to determine the time-average power that each mode can carry along the propagation direction z , it is necessary to evaluate the z -component of the time-average Poynting vector and integrate it over the xy -plane. Then, the power carried by the ν -th mode can be expressed as

$$P = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{2} \text{Re}\{\vec{E} \times \vec{H}^*\} \cdot \hat{z} dx dy, \quad (62)$$

where the integration is over the entire xy plane. However, in slab waveguides, there is no y -dependence, and therefore the integration over the y -axis should give infinite power! As a result in slab waveguides the power per unit y length is of interest and the previous equation is written as

$$\begin{aligned} P &= \int_{-\infty}^{+\infty} \frac{1}{2} \text{Re}\{\vec{E} \times \vec{H}^*\} \cdot \hat{z} dx = \\ &= \underbrace{\int_{-\infty}^0 \frac{1}{2} \text{Re}\{\vec{E} \times \vec{H}^*\} \cdot \hat{z} dx}_{P_s} + \underbrace{\int_0^h \frac{1}{2} \text{Re}\{\vec{E} \times \vec{H}^*\} \cdot \hat{z} dx}_{P_f} + \underbrace{\int_h^{+\infty} \frac{1}{2} \text{Re}\{\vec{E} \times \vec{H}^*\} \cdot \hat{z} dx}_{P_c}, \end{aligned} \quad (63)$$

where it is implied that P is expressed in Watts/per unit y length and the integral has been separated in three parts P_s , P_f , and P_c each representing the power per unit y length in each of the three waveguide regions, substrate, film, and cover, respectively. The x -component of the average Poynting vector should be zero since no power is propagating in the x -direction for guided modes.

10.1 TE Modes

Using the field Eqs. (29) and (25) in the case of TE modes the Poynting vector can be written as $(1/2)\text{Re}\{-E_y H_x^*\}$ and the resulting power components become

$$P_s = \frac{N_\nu^{TE}}{4} \sqrt{\frac{\epsilon_0}{\mu_0}} |E_0|^2 \frac{\cos^2(\phi_{fs}^{TE})}{\gamma_s}, \quad (64)$$

$$P_f = \frac{N_\nu^{TE}}{4} \sqrt{\frac{\epsilon_0}{\mu_0}} |E_0|^2 \left[h + \frac{\sin(2k_{fx}h - 2\phi_{fs}^{TE}) + \sin(2\phi_{fs}^{TE})}{2k_{fx}} \right], \quad (65)$$

$$P_c = \frac{N_\nu^{TE}}{4} \sqrt{\frac{\epsilon_0}{\mu_0}} |E_0|^2 \frac{\cos^2(k_{fx}h - \phi_{fs}^{TE})}{\gamma_c}, \quad (66)$$

$$P = \frac{N_\nu^{TE}}{4} \sqrt{\frac{\epsilon_0}{\mu_0}} |E_0|^2 \left[h + \frac{1}{\gamma_s} + \frac{1}{\gamma_c} \right] = \frac{N_\nu^{TE}}{4} \sqrt{\frac{\epsilon_0}{\mu_0}} |E_0|^2 h_{eff,\nu}^{TE}, \quad (67)$$

where $h_{eff,\nu}^{TE}$ is the effective thickness of the TE_ν mode and $P = P_s + P_f + P_c$. It is reminded that $\phi_{fs}^{TE} = \tan^{-1}(\gamma_s/k_{fx})$. It is interesting to specify the percentage of power being carried by the TE_ν mode in each of the three regions. Specifically, from the above equations the following ratios can be defined

$$\frac{P_s}{P} = \frac{1}{h_{eff,\nu}^{TE}} \frac{\cos^2(\phi_{fs}^{TE})}{\gamma_s}, \quad (68)$$

$$\frac{P_f}{P} = \frac{1}{h_{eff,\nu}^{TE}} \left[h + \frac{\sin(2k_{fx}h - 2\phi_{fs}^{TE}) + \sin(2\phi_{fs}^{TE})}{2k_{fx}} \right], \quad (69)$$

$$\frac{P_c}{P} = \frac{1}{h_{eff,\nu}^{TE}} \frac{\cos^2(k_{fx}h - \phi_{fs}^{TE})}{\gamma_c}. \quad (70)$$

For example, the dependence of $h_{eff,\nu}^{TE}$, and of P_s/P , P_f/P , and P_c/P of the film thickness h is shown in Figs. 15 and 16 for the first 10 modes of a slab waveguide.

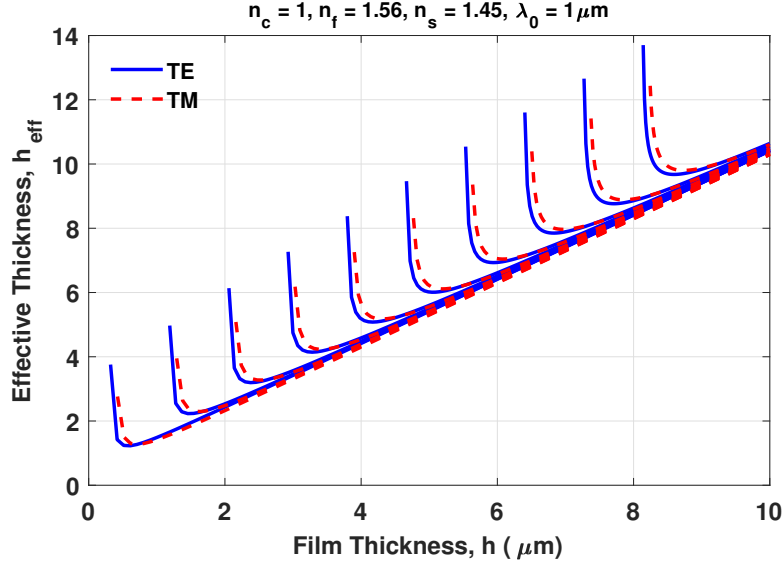


Figure 15: A sample $h_{eff,\nu}$ versus film thickness diagram for a three layer slab waveguide with $n_c = 1.0$, $n_f = 1.56$, $n_s = 1.45$, and $h = 1\mu\text{m}$ for both TE and TM polarizations. The effective thicknesses of the first 10 TE and TM modes are shown.

It is straightforward to show that the average Poynting vector along the x -direction is $P_x = (1/2)\text{Re}\{E_y H_z^*\}$. Specifically, using Eqs. (29) and (26) it is easy to show that $P_x = 0$ as it is expected for guided modes.

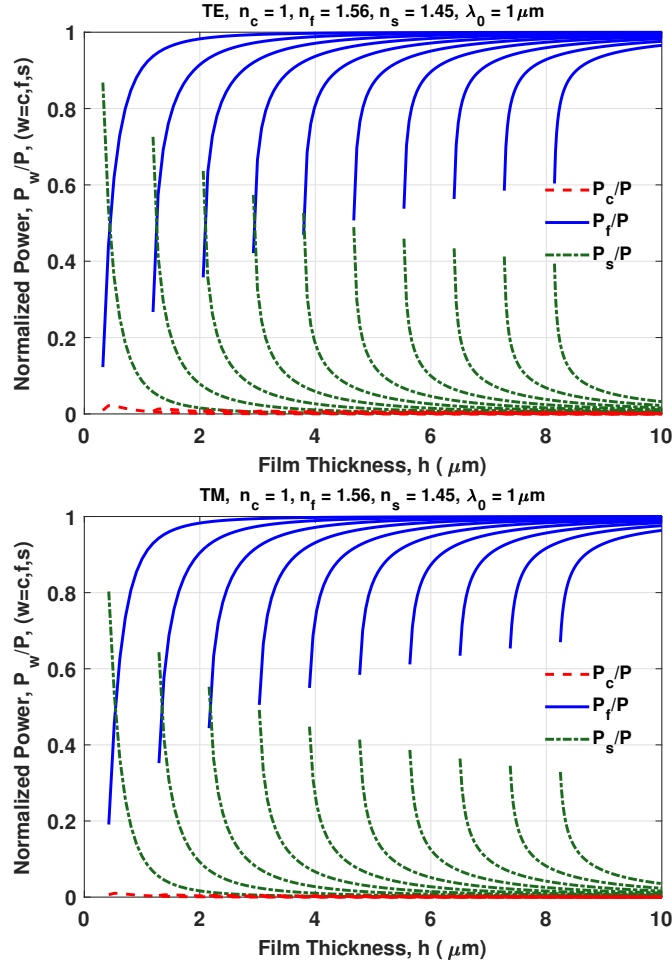


Figure 16: The power percentages in the cover, film and substrate regions as a function of the waveguide film thickness. The slab waveguide parameters are shown on the top of each plot. It is mentioned that the first ten TE and TM modes are included.

10.2 TM Modes

Using the field Eqs. (35) and (31) in the case of TM modes the Poynting vector can be written as $(1/2)\text{Re}\{E_x H_y^*\}$ and the resulting power components become

$$P_s = \frac{N_\nu^{TM}}{4} \sqrt{\frac{\mu_0}{\epsilon_0}} |H_0|^2 \frac{\cos^2(\phi_{fs}^{TM})}{n_s^2 \gamma_s}, \quad (71)$$

$$P_f = \frac{N_\nu^{TM}}{4} \sqrt{\frac{\mu_0}{\epsilon_0}} |H_0|^2 \frac{1}{n_f^2} \left[h + \frac{\sin(2k_{fx}h - 2\phi_{fs}^{TM}) + \sin(2\phi_{fs}^{TM})}{2k_{fx}} \right], \quad (72)$$

$$P_c = \frac{N_\nu^{TM}}{4} \sqrt{\frac{\mu_0}{\epsilon_0}} |H_0|^2 \frac{\cos^2(k_{fx}h - \phi_{fs}^{TM})}{n_c^2 \gamma_c}, \quad (73)$$

$$P = \frac{N_\nu^{TM}}{4} \sqrt{\frac{\mu_0}{\epsilon_0}} |H_0|^2 \frac{1}{n_f^2} \left[h + \frac{1}{q_s \gamma_s} + \frac{1}{q_c \gamma_c} \right] = \frac{N_\nu^{TM}}{4} \sqrt{\frac{\mu_0}{\epsilon_0}} |H_0|^2 \frac{1}{n_f^2} h_{eff,\nu}^{TM}, \quad (74)$$

where $h_{eff,\nu}^{TM}$ is the effective thickness of the TM_ν mode and $P = P_s + P_f + P_c$, and $q_c = (N_\nu^{TM})^2/n_c^2 + (N_\nu^{TM})^2/n_f^2 - 1$, $q_s = (N_\nu^{TM})^2/n_s^2 + (N_\nu^{TM})^2/n_f^2 - 1$. It is reminded that in this case $\phi_{fs}^{TM} = \tan^{-1}[(\gamma_s/n_s^2)/(k_{fx}/n_f^2)]$. It is interesting to specify the percentage of power being carried by the TM_ν mode in each of the three regions. Specifically, from the above equations the following ratios can be defined

$$\frac{P_s}{P} = \frac{n_f^2}{h_{eff,\nu}^{TM}} \frac{\cos^2(\phi_{fs}^{TM})}{n_s^2 \gamma_s}, \quad (75)$$

$$\frac{P_f}{P} = \frac{n_f^2}{h_{eff,\nu}^{TM}} \frac{1}{n_f^2} \left[h + \frac{\sin(2k_{fx}h - 2\phi_{fs}^{TM}) + \sin(2\phi_{fs}^{TM})}{2k_{fx}} \right], \quad (76)$$

$$\frac{P_c}{P} = \frac{n_f^2}{h_{eff,\nu}^{TM}} \frac{\cos^2(k_{fx}h - \phi_{fs}^{TM})}{n_c^2 \gamma_c}. \quad (77)$$

For example, the dependence of $h_{eff,\nu}^{TM}$, and of P_s/P , P_f/P , and P_c/P of the film thickness h is shown in Figs. 15 and 16 for the first 10 modes of a slab waveguide. Similar to the TE modes, it is straightforward to show that the average Poynting vector along the x -direction is $P_x = (1/2)\text{Re}\{-E_z H_y^*\}$. Specifically, using Eqs. (35) and (32) it is easy to show that $P_x = 0$ as it is expected for guided modes.

11. Multi-Layered Slab Waveguides

Multi-layered slab waveguides are of practical interest since they represent realizable semiconductor waveguides that are used in semiconductor lasers. Furthermore, even dielectric waveguides of more than a single film region are used. An example is the case that a buffer layer may be needed to be inserted between the film region and metallic electrodes that are used in electro-optic applications. Also, many times from a fabrication point of view, the resulting waveguides end up with a refractive index profile that is of graded (smooth) variation instead of step (abrupt) variation. Examples of such waveguides are the ones fabricated via ion-exchange, ion-bombardment, or metal-in-diffusion techniques. In latter cases the graded-index waveguides can be approximated by an arbitrary number of layers of constant index thus approximating the smoothly varying refractive index. This technique can also be used to characterized fabricated waveguides. For example, if the mode propagation constants can be measured (which is feasible through a prism-coupling method) the parameters of the fabricated waveguide can be determined by an inverse problem to the one that determines the propagation constants. For all these reasons, the extension of the three region slab waveguide to the general multi-layered case is presented here.

A general multi-layered slab waveguide is shown in Fig. 17. The waveguide is comprised of N film regions (of refractive index n_i and thickness h_i), in addition to the cover and substrate regions, of refractive indices, n_c and n_s , respectively. The $x = 0$ is now selected at the cover-film layer #1 boundary and its direction is towards the substrate. This direction is opposite of the one used in Fig. 1. This is chosen since it matches the coordinate system that was utilized in the computer program that was pre-written to these notes in order to implement the procedure for the multi-layered slab waveguide. All the basic equations are the same with the only care needed in the selection of the signs of the exponential terms of the evanescent waves in cover and substrate regions.

The approach to determine the modes (TE or TM) in the case of a multi-layered slab waveguide is conceptually identical with the one that was followed in Secs. 3.1 and 3.2. For example, if the TE modes should be determined, then the unknown electric field amplitude coefficients are E_c , E_{11} , E_{12} , \dots , E_{i1} , E_{i2} , \dots , E_{N1} , E_{N2} , E_s , as well as the propagation constant β_ν . Therefore the number of unknowns is $2(N + 1)$ excluding β_ν . The number of

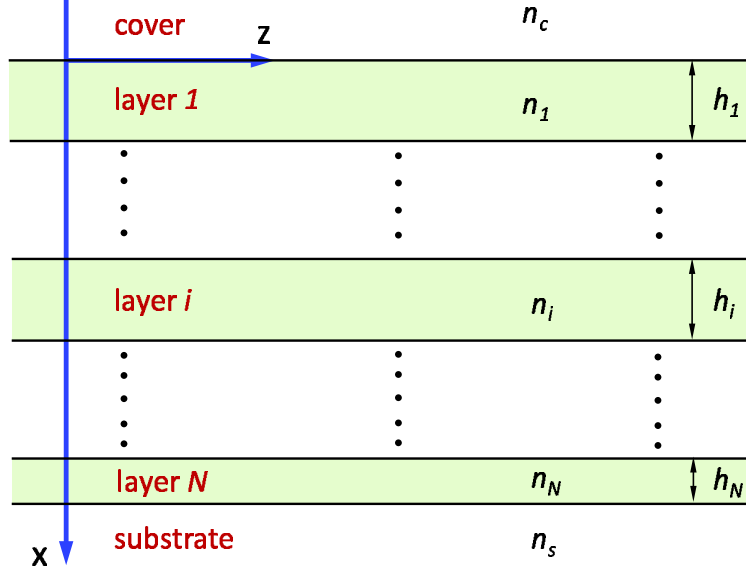


Figure 17: A general multilayered slab waveguide. The waveguide is comprised of N layers, with refractive indices, n_i , and layer thicknesses, h_i ($i = 1, 2, \dots, N$). The cover ($x < 0$) has a refractive index of n_c , whereas the substrate ($x > \sum_1^N h_i$) has a refractive index of n_s .

equations that can be derived from the boundary conditions are $2(N + 1)$. The boundary conditions result in a matrix equation similar to Eq. (27) but of dimension $2(N + 1) \times 2(N + 1)$. The determinant of the coefficient matrix is the dispersion equation and solutions of this in the interval $k_0 n_s < \beta_\nu < k_0 \max_i \{n_i\}$ determine the propagation constants β_ν . However, in this section the *Transfer Matrix* technique as described in Ref. [11–13] will be presented.

11.1 TE Guided Modes

As in the case of the three-layer slab waveguide, the TE modes are characterized by the electromagnetic fields $\{E_y, H_x, H_z\}$. Using the first two of Eqs. (14) and the first of Eqs. (15) the following equations can be written for a region of refractive index n

$$\frac{d}{dx} \begin{bmatrix} E_y \\ H_z \end{bmatrix} = \begin{bmatrix} 0 & -j\omega\mu_0 \\ -j\omega\epsilon + j\frac{\beta^2}{\omega\mu_0} & 0 \end{bmatrix} \begin{bmatrix} E_y \\ H_z \end{bmatrix}, \quad (78)$$

$$H_x = -\frac{\beta}{\omega\mu_0} E_y. \quad (79)$$

The solution of E_y in each region can be summarized as follows

$$\vec{E} = \hat{y} \begin{cases} E_c e^{\gamma_c x} e^{-j\beta z}, & x \leq 0, \\ [E_{i1} e^{-jk_{xi}(x-D_i)} + E_{i2} e^{+jk_{xi}(x-D_i)}] e^{-j\beta z}, & D_i \leq x \leq D_{i+1}, \\ E_s e^{-\gamma_s(x-D_{N+1})} e^{-j\beta z}, & x \geq D_{N+1}, \end{cases} \quad (80)$$

where the distances D_i are given by

$$D_i = \sum_{\ell=1}^{i-1} h_\ell, \quad \text{with } D_1 = 0, \quad i = 1, 2, \dots, N, \quad (81)$$

and $\gamma_c = (\beta^2 - k_0^2 n_c^2)^{1/2}$, $\gamma_s = (\beta^2 - k_0^2 n_s^2)^{1/2}$, $k_{xi} = (k_0^2 n_i^2 - \beta^2)^{1/2}$ ($i = 1, 2, \dots, N$), and E_c , E_s , and E_{i1} , E_{i2} the unknown amplitude constants. The tangential magnetic field can be determined from Eqs. (78) and is given by

$$H_z = \frac{1}{-j\omega\mu_0} \begin{cases} \gamma_c E_c e^{\gamma_c x} e^{-j\beta z}, & x \leq 0, \\ [-jk_{xi} E_{i1} e^{-jk_{xi}(x-D_i)} + jk_{xi} E_{i2} e^{+jk_{xi}(x-D_i)}] e^{-j\beta z}, & D_i \leq x \leq D_{i+1}, \\ -\gamma_s E_s e^{-\gamma_s(x-D_{N+1})} e^{-j\beta z}, & x \geq D_{N+1}. \end{cases} \quad (82)$$

The next step is to use the boundary conditions at $x = 0, D_2, \dots, D_{N+1}$ for continuity of E_y and H_z . The resulting equations are written in the following form

$$\begin{bmatrix} 1 & 1 \\ -jk_{x1} & jk_{x1} \end{bmatrix} \begin{bmatrix} E_{11} \\ E_{12} \end{bmatrix} = \begin{bmatrix} 1 \\ \gamma_c \end{bmatrix} E_c, \quad \text{at } x = 0, \quad (83)$$

At the boundary between the i -th and the $(i+1)$ -th layer the following equations can be derived

$$\begin{bmatrix} e^{-jk_{xi}h_i} & e^{+jk_{xi}h_i} \\ -jk_{xi}e^{-jk_{xi}h_i} & jk_{xi}e^{+jk_{xi}h_i} \end{bmatrix} \begin{bmatrix} E_{i1} \\ E_{i2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -jk_{x,i+1} & jk_{x,i+1} \end{bmatrix} \begin{bmatrix} E_{i+1,1} \\ E_{i+1,2} \end{bmatrix}, \quad \text{at } x = D_{i+1}, \quad (84)$$

whereas at the last boundary at $x = D_{N+1} = \sum_{\ell=1}^N h_\ell$ the following equations are derived

$$\begin{bmatrix} e^{-jk_{xN}h_N} & e^{+jk_{xN}h_N} \\ -jk_{xN}e^{-jk_{xN}h_N} & jk_{xN}e^{+jk_{xN}h_N} \end{bmatrix} \begin{bmatrix} E_{N1} \\ E_{N2} \end{bmatrix} = \begin{bmatrix} 1 \\ -\gamma_s \end{bmatrix} E_s, \quad \text{at } x = D_{N+1}. \quad (85)$$

Now Eq. (83) can be written as

$$\begin{bmatrix} E_{11} \\ E_{12} \end{bmatrix} = \underbrace{\frac{1}{2jk_{x1}} \begin{bmatrix} jk_{x1} & -1 \\ jk_{x1} & 1 \end{bmatrix}}_{\tilde{M}_{1,0}} \begin{bmatrix} 1 \\ \gamma_c \end{bmatrix} E_c, \quad (86)$$

defining the transfer matrix $\tilde{M}_{1,0}$ between regions 0 (cover) and 1 (first film layer). Some special care should be taken in the case that $k_{x1} = 0$ [in this case the electric field solution in region 1 is linear, i.e. $E_{y1}(x) = E_{1,2}(x - D_1) + E_{1,1}$] and $\tilde{M}_{1,0}$ becomes the unity matrix. Boundary condition expressed in Eq. (84) can be written as

$$\begin{bmatrix} E_{i+1,1} \\ E_{i+1,2} \end{bmatrix} = \frac{1}{2} \underbrace{\begin{bmatrix} e^{-jk_{xi}h_i} \left(1 + \frac{k_{xi}}{k_{x,i+1}}\right) & e^{+jk_{xi}h_i} \left(1 - \frac{k_{xi}}{k_{x,i+1}}\right) \\ e^{-jk_{xi}h_i} \left(1 - \frac{k_{xi}}{k_{x,i+1}}\right) & e^{+jk_{xi}h_i} \left(1 + \frac{k_{xi}}{k_{x,i+1}}\right) \end{bmatrix}}_{\tilde{M}_{i+1,i}} \begin{bmatrix} E_{i1} \\ E_{i2} \end{bmatrix}, \quad (87)$$

thus defining the transfer matrix $\tilde{M}_{i+1,i}$ between i -th and $(i+1)$ -th regions. As a special case, when $k_{xi} = 0$ [the electric field solution in region i is linear, i.e. $E_{yi}(x) = E_{i,2}(x - D_i) + E_{i,1}$] and the matrix $\tilde{M}_{i+1,i}$ becomes

$$\tilde{M}_{i+1,i} = \frac{1}{2} \begin{bmatrix} 1 & h_i - \frac{1}{jk_{x,i+1}} \\ 1 & h_i + \frac{1}{jk_{x,i+1}} \end{bmatrix}. \quad (88)$$

Another special case is when $k_{x,i+1} = 0$ [the electric field solution in region $i+1$ is linear, i.e. $E_{y,i+1}(x) = E_{i+1,2}(x - D_{i+1}) + E_{i+1,1}$] and the matrix $\tilde{M}_{i+1,i}$ becomes

$$\tilde{M}_{i+1,i} = \frac{1}{2} \begin{bmatrix} e^{-jk_{xi}h_i} & e^{+jk_{xi}h_i} \\ -jk_{xi}e^{-jk_{xi}h_i} & +jk_{xi}e^{+jk_{xi}h_i} \end{bmatrix}. \quad (89)$$

Finally, Eq. (85) can be written as

$$\begin{bmatrix} 1 \\ -\gamma_s \end{bmatrix} E_s = \underbrace{\begin{bmatrix} e^{-jk_{xN}h_N} & e^{+jk_{xN}h_N} \\ -jk_{xN}e^{-jk_{xN}h_N} & jk_{xN}e^{+jk_{xN}h_N} \end{bmatrix}}_{\tilde{M}_{N+1,N}} \begin{bmatrix} E_{N1} \\ E_{N2} \end{bmatrix}, \quad (90)$$

thus, defining the last transfer matrix $\tilde{M}_{N+1,N}$ between regions N and substrate. In the case that $k_{xN} = 0$ [the electric field solution in region N is linear, i.e. $E_{y,N}(x) = E_{N,2}(x - D_N) + E_{N,1}$] and the matrix $\tilde{M}_{N+1,N}$ becomes

$$\tilde{M}_{N+1,N} = \frac{1}{2} \begin{bmatrix} 1 & h_i \\ 0 & 1 \end{bmatrix}. \quad (91)$$

Using the previous equations all amplitudes E_{i1} and E_{i2} can be eliminated as follows

$$\begin{bmatrix} 1 \\ -\gamma_s \end{bmatrix} E_s = \underbrace{\tilde{M}_{N+1,N} \tilde{M}_{N,N-1} \cdots \tilde{M}_{2,1} \tilde{M}_{1,0}}_{\tilde{M}} \begin{bmatrix} 1 \\ \gamma_c \end{bmatrix} E_c = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \begin{bmatrix} 1 \\ \gamma_c \end{bmatrix} E_c. \quad (92)$$

In the previous equation $\tilde{M} = \prod_{\ell=0}^N \tilde{M}_{N+1-\ell, N-\ell}$ is the total transfer matrix between the cover and the substrate regions. Equation (92) results in

$$\underbrace{\begin{bmatrix} m_{11} + m_{12}\gamma_c & -1 \\ m_{21} + m_{22}\gamma_c & \gamma_s \end{bmatrix}}_{\tilde{\mathcal{A}}_{TE}(\beta^2)} \begin{bmatrix} E_c \\ E_s \end{bmatrix} = 0 \implies \det \left\{ \tilde{\mathcal{A}}_{TE}(\beta^2) \right\} = 0, \quad (93)$$

and equivalently

$$\det \left\{ \tilde{\mathcal{A}}_{TE}(\beta^2) \right\} = 0 \implies \gamma_s m_{11} + \gamma_c m_{22} + \gamma_c \gamma_s m_{12} + m_{21} = 0. \quad (94)$$

Solutions, β_ν , of the last equation for $k_0 n_s < \beta_\nu < k_0 \max_i \{n_i\}$ specify the TE_ν modes. Using Eqs. (86)-(90) all the unknown amplitude coefficients can be determined using a free parameter (for example E_s or E_c). For an example multi-layered waveguide (for $N = 3$) the corresponding TE modes electric field profiles are shown in Fig. 18.

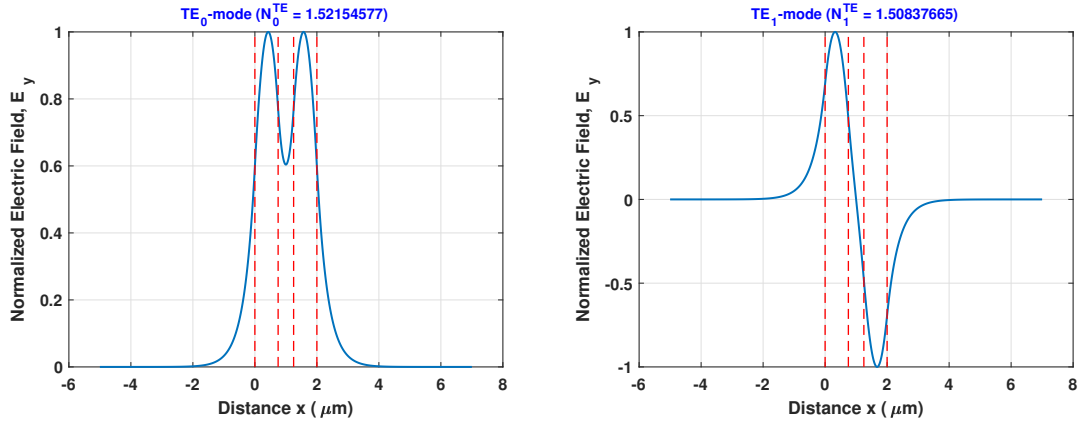


Figure 18: The TE_ν ($\nu = 0, 1$) modes electric field profiles of a sample multi-layered waveguide are shown. The waveguide is comprised of three film layers and its parameters are $n_c = 1.45$, $n_1 = 1.56$, $n_2 = 1.45$, $n_3 = 1.56$, $n_s = 1.45$, $\lambda_0 = 1.0\mu\text{m}$, $h_1 = 0.75\mu\text{m}$, $h_2 = 0.5\mu\text{m}$, and $h_3 = 0.75\mu\text{m}$. The effective indices $N_\nu^{TE} = \beta_\nu/k_0$ are shown on each plot. The dashed lines specify the boundaries between differing regions.

11.2 TM Guided Modes

As in the case of the three-layer slab waveguide, the TM modes are characterized by the electromagnetic fields $\{H_y, E_x, E_z\}$. Using the last two of Eqs. (14) and the second of Eqs. (15) the following equations can be written for a region of refractive index n

$$\frac{d}{dx} \begin{bmatrix} H_y \\ E_z \end{bmatrix} = \begin{bmatrix} 0 & j\omega\epsilon \\ j\omega\mu_0 - j\frac{\beta^2}{\omega\epsilon} & 0 \end{bmatrix} \begin{bmatrix} H_y \\ E_z \end{bmatrix}, \quad (95)$$

$$E_x = \frac{\beta}{\omega\epsilon} H_y. \quad (96)$$

The solution of H_y in each region can be summarized as follows

$$\vec{H} = \hat{y} \begin{cases} H_c e^{\gamma_c x} e^{-j\beta z}, & x \leq 0, \\ [H_{i1} e^{-jk_{xi}(x-D_i)} + H_{i2} e^{+jk_{xi}(x-D_i)}] e^{-j\beta z}, & D_i \leq x \leq D_{i+1}, \\ H_s e^{-\gamma_s(x-D_{N+1})} e^{-j\beta z}, & x \geq D_{N+1}, \end{cases} \quad (97)$$

where the distances D_i are given by Eq. (81), and H_c , H_s , and H_{i1} , H_{i2} the unknown amplitude constants. The tangential magnetic field can be determined from Eqs. (95) and is given by

$$E_z = \frac{1}{j\omega\epsilon_0} \begin{cases} \frac{\gamma_c}{n_c^2} H_c e^{\gamma_c x} e^{-j\beta z}, & x \leq 0, \\ \left[-j \frac{k_{xi}}{n_i^2} H_{i1} e^{-jk_{xi}(x-D_i)} + j \frac{k_{xi}}{n_i^2} H_{i2} e^{+jk_{xi}(x-D_i)} \right] e^{-j\beta z}, & D_i \leq x \leq D_{i+1}, \\ -\frac{\gamma_s}{n_s^2} H_s e^{-\gamma_s(x-D_{N+1})} e^{-j\beta z}, & x \geq D_{N+1}. \end{cases} \quad (98)$$

At the boundary between the i -th and the $(i+1)$ -th layer the boundary conditions can be written as

$$\begin{bmatrix} e^{-jk_{xi}h_i} & e^{+jk_{xi}h_i} \\ j \frac{k_{xi}}{n_i^2} e^{-jk_{xi}h_i} & -j \frac{k_{xi}}{n_i^2} e^{+jk_{xi}h_i} \end{bmatrix} \begin{bmatrix} H_{i1} \\ H_{i2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ j \frac{k_{x,i+1}}{n_{i+1}^2} & -j \frac{k_{x,i+1}}{n_{i+1}^2} \end{bmatrix} \begin{bmatrix} H_{i+1,1} \\ H_{i+1,2} \end{bmatrix}, \quad \text{at } x = D_{i+1}. \quad (99)$$

The last equation is the analogous of Eq. (84) with the difference that $k_{xi} \rightarrow -k_{xi}/n_i^2$ and similarly for $k_{x,i+1}$ whereas those terms appear as amplitude terms. The same replacement is valid for the equations analogous to Eqs. (83), (85) where $\gamma_c \rightarrow -\gamma_c/n_c^2$ and $\gamma_s \rightarrow -\gamma_s/n_s^2$ whereas they appear as amplitude terms. Consequently it is not necessary to repeat the procedure for the TM modes. For example the dispersion equation can be written as

$$\det \left\{ \tilde{\mathcal{A}}_{TM}(\beta^2) \right\} = 0 \Rightarrow -\frac{\gamma_s}{n_s^2} m_{11} - \frac{\gamma_c}{n_c^2} m_{22} + \frac{\gamma_c}{n_c^2} \frac{\gamma_s}{n_s^2} m_{12} + m_{21} = 0. \quad (100)$$

The matrix elements m_{ij} ($i, j = 1, 2$) can be determined in the same manner as for the TE modes with the replacements made previously for the terms that appear as amplitude terms (not inside exponentials or sines and cosines). For an example multi-layered waveguide (for $N = 3$) the corresponding TM modes magnetic field profiles are shown in Fig. 19.

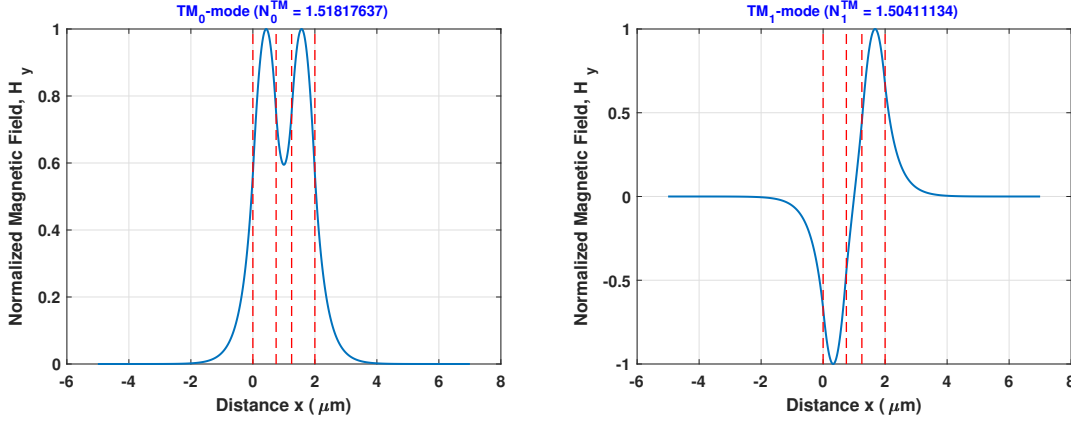


Figure 19: The TM_ν ($\nu = 0, 1$) modes magnetic field profiles of a sample multi-layered waveguide are shown. The waveguide is comprised of three film layers and its parameters are $n_c = 1.45$, $n_1 = 1.56$, $n_2 = 1.45$, $n_3 = 1.56$, $n_s = 1.45$, $\lambda_0 = 1.0\mu\text{m}$, $h_1 = 0.75\mu\text{m}$, $h_2 = 0.5\mu\text{m}$, and $h_3 = 0.75\mu\text{m}$. The effective indices $N_\nu^{TE} = \beta_\nu/k_0$ are shown on each plot. The dashed lines specify the boundaries between differing regions.

11.3 TE and TM Guided Modes using Chen's Approach

In this section the transfer matrix technique as used by Chen *et al.* [17] for the study of multilayer planar waveguides is shortly summarized. For the TE modes the field components for a region of refractive index n can be related by Eq. (78). The solution of E_y in each region can be summarized as follows

$$\vec{E} = \hat{y} \begin{cases} E_c e^{\gamma_c x} e^{-j\beta z}, & x \leq 0, \\ \left\{ E_{i1} \cos[k_{xi}(x - D_i)] + E_{i2} \sin[k_{xi}(x - D_i)] \right\} e^{-j\beta z}, & D_i \leq x \leq D_{i+1}, \\ E_s e^{-\gamma_s(x - D_{N+1})} e^{-j\beta z}, & x \geq D_{N+1}, \end{cases} \quad (101)$$

where the distances D_i are given as before by $D_i = \sum_{\ell=1}^{i-1} h_\ell$, with $D_1 = 0$, $i = 1, 2, \dots, N$, and $\gamma_c = (\beta^2 - k_0^2 n_c^2)^{1/2}$, $\gamma_s = (\beta^2 - k_0^2 n_s^2)^{1/2}$, $k_{xi} = (k_0^2 n_i^2 - \beta^2)^{1/2}$ ($i = 1, 2, \dots, N$), and E_c , E_s , and E_{i1} , E_{i2} the unknown amplitude constants. The tangential magnetic field H_z can easily be determined from Eqs. (78) and (101). Using the continuity of the tangential field components across the layer boundaries it is straightforward to relate the tangential fields $(E_{yi}, dE_{yi}/dx) = (U_i, V_i)$ where $i = 0, 1, \dots, N$ with (U_0, V_0) the tangential fields at the cover/layer-1 interface (U_i, V_i) the tangential fields at the layer- i /layer- $(i+1)$ interface, and (U_N, V_N) the tangential fields at the layer- N /substrate interface. It can be shown [13, 17] that

$$\begin{bmatrix} U_0 \\ V_0 \end{bmatrix} = \underbrace{\tilde{M}_1 \tilde{M}_2 \cdots \tilde{M}_{N-1} \tilde{M}_N}_{\tilde{M}} \begin{bmatrix} U_N \\ V_N \end{bmatrix} = \begin{bmatrix} m_{11} & m_{21} \\ m_{21} & m_{22} \end{bmatrix} \begin{bmatrix} U_N \\ V_N \end{bmatrix}. \quad (102)$$

In the previous equation \tilde{M} is the total transfer matrix between the cover and the substrate regions where

$$\tilde{M}_i = \begin{bmatrix} \cos(k_{xi}h_i) & -\frac{\sin(k_{xi}h_i)}{k_{xi}} \\ k_{xi} \sin(k_{xi}h_i) & \cos(k_{xi}h_i) \end{bmatrix} \quad \text{if } k_{xi} \neq 0, \quad (103)$$

$$\tilde{M}_i = \begin{bmatrix} 1 & -h_i \\ 0 & 1 \end{bmatrix}, \quad \text{if } k_{xi} = 0, \quad (104)$$

where $i = 1, 2, \dots, N$. For nontrivial solutions of Eq. (102) the following dispersion equation can be easily derived

$$F_{TE}(\beta) = \gamma_s m_{22} + \gamma_c m_{11} - \gamma_c \gamma_s m_{12} - m_{21} = 0. \quad (105)$$

In the case of the TM modes the tangential fields are $[H_{yi}, (1/n_i^2)dH_{yi}/dx] = (U_i, V_i)$ with $i = 0, 1, 2, \dots, N$ similar to the TE case. With these tangential fields, it is straightforward to show that the dispersion equation can be derived by a simple replacement of $k_{xi} \rightarrow +k_{xi}/n_i^2$ (only in amplitude terms of the previous equations and not inside phase terms), $\gamma_s \rightarrow +\gamma_s/n_s^2$, and $\gamma_c \rightarrow +\gamma_c/n_c^2$. For completeness in the case of $k_{xi} = 0$ (linear field solution inside layer- i) the element $-h_i$ of matrix \tilde{M}_i appearing in Eq. (104) is replaced by $-n_i^2 h_i$. Then the resulting dispersion equation is

$$F_{TM}(\beta) = \frac{\gamma_s}{n_s^2} m_{11} + \frac{\gamma_c}{n_c^2} m_{22} - \frac{\gamma_c}{n_c^2} \frac{\gamma_s}{n_s^2} m_{12} - m_{21} = 0. \quad (106)$$

It is worth mentioning that both functions F_{TE} and F_{TM} are actually functions of β^2 which implies that if $+\beta$ satisfies Eq. (105) or (106) so is $-\beta$ which corresponds to the backward in z propagating mode.

12. Finite-Difference Frequency-Domain Method

In section the finite-difference frequency-domain (FDFD) method will be applied to the analysis of general isotropic slab waveguides. The wave Helmholtz's equations can be written in the following vector form

$$\nabla^2 \vec{E} + \vec{\nabla} \left[\frac{\vec{\nabla} \varepsilon}{\varepsilon} \cdot \vec{E} \right] + k_0^2 \varepsilon \vec{E} = 0, \quad (107)$$

$$\nabla^2 \vec{H} + \left[\frac{\vec{\nabla} \varepsilon}{\varepsilon} \right] \times (\vec{\nabla} \times \vec{H}) + k_0^2 \varepsilon \vec{H} = 0, \quad (108)$$

where ε is the relative permittivity of the region. In the case of *TE* modes solutions the electric field is of the form $\vec{E} = \hat{y} E_y(x) \exp(-j\beta z)$ and Eq. (107) can be simplified in the form

$$\frac{d^2 E_y}{dx^2} + [k_0^2 n^2(x) - \beta^2] E_y = 0, \quad (109)$$

since the $\vec{\nabla} \varepsilon$ term is along the x direction whereas \vec{E} is along the y direction, and $n^2(x) = \varepsilon(x)$. Similarly, in the case of *TM* modes solutions, the magnetic field is of the form $\vec{H} = \hat{y} H_y(x) \exp(-j\beta z)$ and Eq. (108) can be simplified in the form

$$\frac{d^2 H_y}{dx^2} - \frac{1}{n^2(x)} \frac{dn^2(x)}{dx} \frac{dH_y}{dx} + [k_0^2 n^2(x) - \beta^2] H_y = 0, \quad \text{or} \quad (110)$$

$$n^2(x) \frac{d}{dx} \left(\frac{1}{n^2(x)} \frac{dH_y}{dx} \right) + [k_0^2 n^2(x) - \beta^2] H_y = 0, \quad (111)$$

where for the *TM* case the corresponding equation is more complicated than the *TE* one.

In the finite-difference frequency-domain (FDFD) method the derivatives are approximated using central differences. For example, if $U = E_y$ or H_y , U can be expanded into a Taylor series around a point x . These leads to the following equations

$$U(x + \Delta x) \simeq U(x) + \frac{dU}{dx} \frac{\Delta x}{1!} + \frac{d^2 U}{dx^2} \frac{(\Delta x)^2}{2!} + \frac{d^3 U}{dx^3} \frac{(\Delta x)^3}{3!} + \frac{d^4 U}{dx^4} \frac{(\Delta x)^4}{4!}, \quad (112)$$

$$U(x - \Delta x) \simeq U(x) - \frac{dU}{dx} \frac{\Delta x}{1!} + \frac{d^2 U}{dx^2} \frac{(\Delta x)^2}{2!} - \frac{d^3 U}{dx^3} \frac{(\Delta x)^3}{3!} + \frac{d^4 U}{dx^4} \frac{(\Delta x)^4}{4!}, \quad (113)$$

and adding the last two equations the following equation can be derived

$$U(x + \Delta x) + U(x - \Delta x) = 2U(x) + \frac{d^2 U}{dx^2} (\Delta x)^2 + O[(\Delta x)^4], \quad (114)$$

where the $O[(\Delta x)^4]$ is the term that depends on the forth power of Δx and can be neglected if Δx is small (i.e. $|\Delta x| \ll 1$). Now the discretization along the x axis is considered. For example

the continuous x axis is replaced by the points $x_i = j\Delta x$ where $(i = 0, 1, \dots, M, M+1)$. Then if $x = x_i$, $x + \Delta x = x_{i+1}$, and $x - \Delta x = x_{i-1}$, $U(x_i) = U_i$, $U(x_{i+1}) = U_{i+1}$, and $U(x_{i-1}) = U_{i-1}$, and $n^2(x_i) = n_i^2$. The geometry is summarized in Fig. 20.

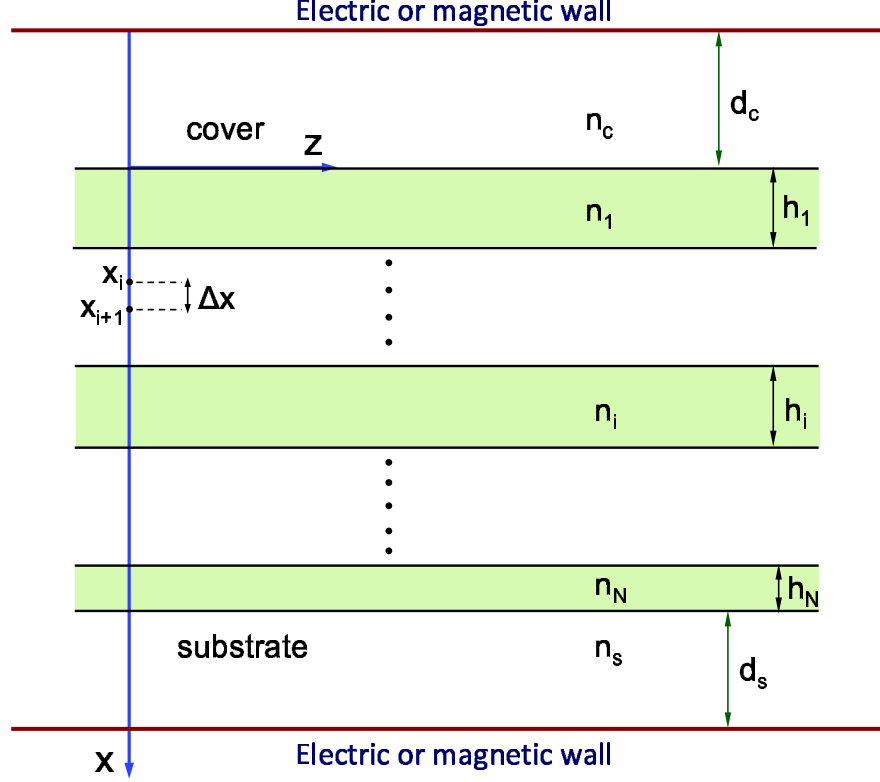


Figure 20: The one dimensional (along the x axis) computational window for the application of the finite-difference frequency-domain (FDFD) in general slab waveguides. It is assumed a multi-layered slab waveguide without thus affecting the generality of the method. The top and bottom lines comprise either electric or magnetic walls (for TE and TM modes respectively) where the electric or magnetic fields are forced to zero. The distances d_c and d_s adjust the total width of the computational window. The grid along the x axis is comprised of points x_i , $i = 1, 2, \dots, M$.

Then from the last equation the second derivative can be approximated (central differences) as follows

$$\frac{d^2 U}{dx^2} = \frac{U_{i-1} - 2U_i + U_{i+1}}{(\Delta x)^2}. \quad (115)$$

Using the approximation for the second derivative in Eq. (109) it can be transformed in the following discretized form

$$\frac{U_{i-1} - 2U_i + U_{i+1}}{(\Delta x)^2} + k_0^2 n_i^2 U_i = \beta^2 U_i, \quad i = 0, 1, \dots, M+1. \quad (116)$$

It is straightforward to make the transformation $X = k_0 x$ and re-write the previous equation in the form

$$\frac{U_{i-1} - 2U_i + U_{i+1}}{(\Delta X)^2} + n_i^2 U_i = N^2 U_i, \quad i = 0, 1, \dots, M+1. \quad (117)$$

where $\Delta X = k_0 \Delta x$ and $N = \beta/k_0$ is the effective index. Applying the last equation for all points x_i and forcing the field to be zero at the upper and lower electric walls (i.e, $U_0 = U_{M+1} = 0$) it is straightforward to form the following system of equations

$$\begin{bmatrix} a_1 & b & 0 & \dots & 0 & 0 \\ b & a_2 & b & 0 & \dots & 0 \\ 0 & b & a_3 & b & \dots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & \dots & b & a_{M-1} & b \\ 0 & \dots & \dots & 0 & b & a_M \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ \vdots \\ U_{M-1} \\ U_M \end{bmatrix} = N^2 \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ \vdots \\ U_{M-1} \\ U_M \end{bmatrix}, \quad (118)$$

where the coefficients $a_i = n_i^2 - [2/(\Delta X)^2]$ for $i = 1, 2, \dots, M$ and $b = 1/(\Delta X)^2$. The last equation is an eigenvalue/eigenvector equation of the form $Av = \lambda v$. The eigenvalues of interest are the ones for which $n_s < N < \max\{n(x)\}$ and the corresponding eigenvectors are the ones that specify the profile of the mode. The above matrix is a tridiagonal matrix which is characteristic of the one-dimensional finite-difference implementation.

In order to apply the same approach for the case of *TM* polarization, Eq. (111), the more complicated first term has to be expressed in discretized form. Following the procedure explained in Ref. [51] it can be shown that

$$n^2(x) \frac{d}{dx} \left(\frac{1}{n^2(x)} \frac{dU}{dx} \right) = \frac{1}{(\Delta X)^2} \frac{2n_i^2}{n_{i+1}^2 + n_i^2} (U_{i+1} - U_i) - \frac{1}{(\Delta X)^2} \frac{2n_i^2}{n_i^2 + n_{i-1}^2} (U_i - U_{i-1}). \quad (119)$$

Inserting the last approximation into Eq. (111) the following discretized wave equation is obtained

$$\begin{aligned} \frac{1}{(\Delta X)^2} \frac{2n_i^2}{n_i^2 + n_{i-1}^2} U_{i-1} &+ \left[n_i^2 - \frac{1}{(\Delta X)^2} \left(\frac{2n_i^2}{n_{i+1}^2 + n_i^2} + \frac{2n_i^2}{n_i^2 + n_{i-1}^2} \right) \right] U_i \\ &+ \frac{1}{(\Delta X)^2} \frac{2n_i^2}{n_{i+1}^2 + n_i^2} U_{i+1} = N^2 U_i, \quad i = 1, 2, \dots, M. \end{aligned} \quad (120)$$

Writing the set of equations for each i ($i = 1, 2, \dots, M$) the following eigenvalue/eigenvector

problem can be formed

$$\begin{bmatrix} a_1 & c_1 & 0 & \dots & 0 & 0 \\ b_2 & a_2 & c_2 & 0 & \dots & 0 \\ 0 & b_3 & a_3 & c_3 & \dots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & \dots & b_{M-1} & a_{M-1} & c_{M-1} \\ 0 & \dots & \dots & 0 & b_M & a_M \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ \vdots \\ U_{M-1} \\ U_M \end{bmatrix} = N^2 \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ \vdots \\ U_{M-1} \\ U_M \end{bmatrix}, \quad (121)$$

where now the matrix elements are given by

$$a_i = n_i^2 - \frac{1}{(\Delta X)^2} \left(\frac{2n_i^2}{n_{i+1}^2 + n_i^2} + \frac{2n_i^2}{n_i^2 + n_{i-1}^2} \right), \quad (i = 1, 2, \dots, M), \quad (122)$$

$$b_i = \frac{1}{(\Delta X)^2} \frac{2n_i^2}{n_i^2 + n_{i-1}^2}, \quad (i = 1, 2, \dots, M) \quad (123)$$

$$c_i = \frac{1}{(\Delta X)^2} \frac{2n_i^2}{n_{i+1}^2 + n_i^2}, \quad (i = 1, 2, \dots, M). \quad (124)$$

An example case is presented next. A simple three-region slab waveguide is selected from semiconductor materials. The parameters are $n_c = 1.0$, $n_f = 3.4$, $n_s = 3.1$, $h = 1.0\mu\text{m}$, and the freespace wavelength is $\lambda_0 = 1.3\mu\text{m}$. The distances from the electric walls are $d_c = 0.5\mu\text{m}$ and $d_s = 1.0\mu\text{m}$. The grid spacing Δx is selected as fraction of the freespace wavelength, i.e., $\Delta x = \lambda_0/L$ where L an integer. The waveguide with these parameters supports two *TE* and two *TM* modes. The results are summarized in Table 1. It is mentioned that the important parameter for convergence is the number of grid points per minimum wavelength as this is measured inside the material. Therefore, for the example case here the minimum wavelength is inside the film region $\lambda_{min} = \lambda_0/3.4$. Consequently, the smallest Δx from the table corresponds to $\lambda_0/200 = \lambda_{min}/58.8$, i.e. in the denser grid results shown there are at least 58 grid points per minimum wavelength. The resulting profiles of the supported modes that can be extracted as the corresponding eigenvectors from Eqs. (118), (121), are shown in Fig. 21.

12.1 FDFD Method Based on Yee's Cell

In this section the Yee's cell [52] is going to be utilized for the determination of the effective propagation constants (effective indices) and the corresponding field profiles for both TE and TM modes for multilayered isotropic slab waveguides. The approach described by Rumpf [53] will be applied appropriately modified in this case. It is assumed that all the materials are non-magnetic ($\mu_r \simeq 1$).

Table 1: Effective Indices Calculated by FDFD

Δx	TE Modes		TM Modes	
	N_0^{TE}	N_1^{TE}	N_0^{TM}	N_1^{TM}
$\lambda_0/20$	3.3585913	3.2362825	3.3518412	3.2129633
$\lambda_0/40$	3.3579455	3.2333453	3.3515178	3.2109972
$\lambda_0/60$	3.3578155	3.2327610	3.3514548	3.2106177
$\lambda_0/80$	3.3577736	3.2325729	3.3514346	3.2104963
$\lambda_0/100$	3.3577538	3.2324843	3.3514251	3.2104392
$\lambda_0/120$	3.3577424	3.2324333	3.3514196	3.2104064
$\lambda_0/150$	3.3577336	3.2323938	3.3514154	3.2103809
$\lambda_0/200$	3.3577268	3.2323632	3.3514215	3.2103612
Exact	3.3577180	3.2323308	3.3514080	3.2103532

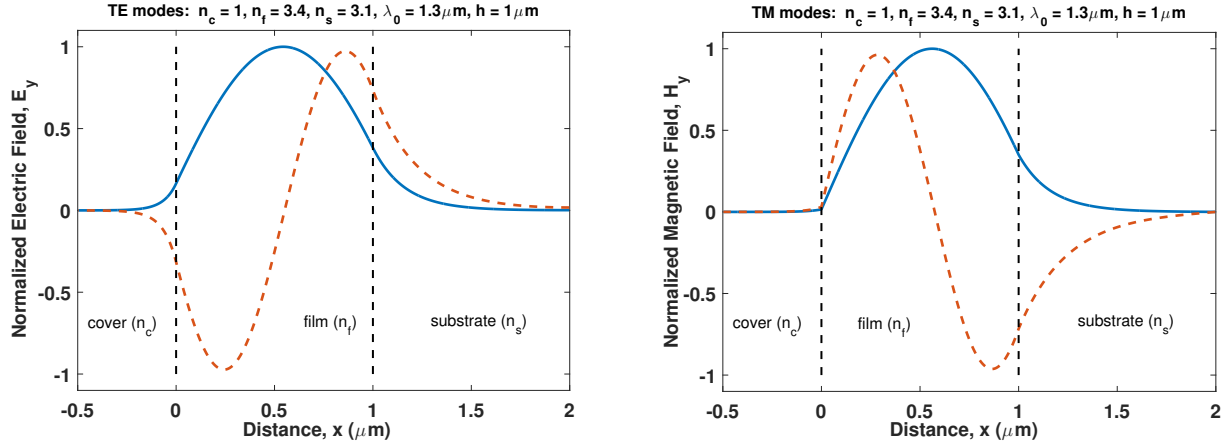


Figure 21: Example electric and magnetic field profiles determined by the FDFD method. This is a case of a slab waveguide of $n_c = 1.0$, $n_f = 3.4$, $n_s = 3.1$, $h = 1.0\mu\text{m}$, and $\lambda_0 = 1.3\mu\text{m}$. The computational window distances $d_c = 0.5\mu\text{m}$ and $d_s = 1.0\mu\text{m}$ were used. In these plots the x -axis matches the one used in Fig. 1. The value of $\Delta x = \lambda_0/200$ is used in these plots. The values of the effective indices are given in Table 1.

Starting from the TE modes determination, it is reminded that the non-zero field components are (E_y, H_x, H_z) where each field component is a function of (x, z) of the form $E_y(z, x) = E_y(x) \exp(-j\beta z) = E_y(x) \exp(-jk_0 N z)$ (as an example), where $E_y(x)$ is the mode profile (for the electric field - similar notation was used all over this chapter) and β , N are the effective propagation constant and the effective refractive index of the TE mode that is sought, and

k_0 is the freespace wavenumber. Using the normalization of the magnetic field components of the form $H'_w = -jZ_0 H_w$ (according to Rumpf [53]) where Z_0 is the free space impedance, the Maxwell's equations can be written in the form

$$-\frac{\partial E_y}{\partial z} = k_0 H'_x \implies -\frac{\partial E_y}{\partial z'} = H'_x, \quad (125)$$

$$+\frac{\partial E_y}{\partial x} = k_0 H'_z \implies +\frac{\partial E_y}{\partial x'} = H'_z, \quad (126)$$

$$-\frac{\partial H'_z}{\partial x} + \frac{\partial H'_x}{\partial z} = k_0 \epsilon_r(x) E_y \implies -\frac{\partial H'_z}{\partial x'} + \frac{\partial H'_x}{\partial z'} = \epsilon_r(x') E_y, \quad (127)$$

where $x' = k_0 x$ and $z' = k_0 z$ are normalized coordinates.

At this point the space domain discretization is done according to Yee's scheme as shown in Fig. 22. The fields are discretized in the coordinate system shown as $E_y(z_i, x_j) = E_y^{i,j}$ taking into account the staggering between various field components.

$$-\frac{E_y^{i+1,j} - E_y^{i,j}}{\Delta z'} = H'_x{}^{i+1/2,j} \implies jNE_y^j = H'_x{}^j e^{-jN\Delta z'/2}, \quad (128)$$

$$\frac{E_y^{i,j+1} - E_y^{i,j}}{\Delta x'} = H'_z{}^{i,j+1/2} \implies \frac{E_y^{j+1} - E_y^j}{\Delta x'} = H'_z{}^{j+1/2}, \quad (129)$$

$$\begin{aligned} -\frac{H'_z{}^{i,j+1/2} - H'_z{}^{i,j-1/2}}{\Delta x'} + \frac{H'_x{}^{i+1/2,j} - H'_x{}^{i-1/2,j}}{\Delta z'} &= \epsilon_r^j E_y^{i,j} \implies \\ -\frac{H'_z{}^{j+1/2} - H'_z{}^{j-1/2}}{\Delta x'} - jNH'_x{}^j &= \epsilon_r^j E_y^j. \end{aligned} \quad (130)$$

When $\Delta z' \rightarrow 0$ the 2D Yee's grid collapses to the compact 1D Yee's cell as it is shown also in the right column (b figures) of Fig. 22. Therefore, the positions of E_y^j and $H'_x{}^j$ coincide (these field components become collocated) while the $H'_z{}^{j+1/2}$ are shifted by $\Delta x'/2$ (forming a staggered 1D grid). Now let's define the vectors $\tilde{E}_y = [E_y^1, E_y^2, \dots, E_y^M]^T$, $\tilde{H}'_x = [H'_x{}^1, H'_x{}^2, \dots, H'_x{}^M]^T$, and $\tilde{H}'_z = [H'_z{}^{1+1/2}, H'_z{}^{2+1/2}, \dots, H'_z{}^{M+1/2}]^T$. Now let's define the difference operators according to Rumpf [53] as follows

$$\tilde{D}_{x'}^E = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 \\ 0 & 0 & -1 & 1 & \dots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & -1 & 1 \\ 0 & \dots & \dots & 0 & 0 & -1 \end{bmatrix}, \quad \tilde{D}_{x'}^H = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ -1 & 1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 1 & 0 & \dots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & -1 & 1 & 0 \\ 0 & \dots & \dots & 0 & -1 & 1 \end{bmatrix}. \quad (131)$$

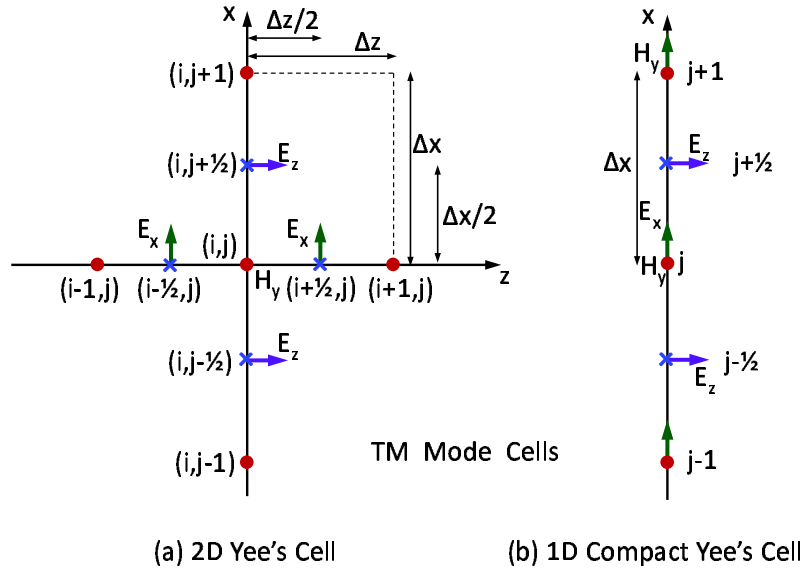
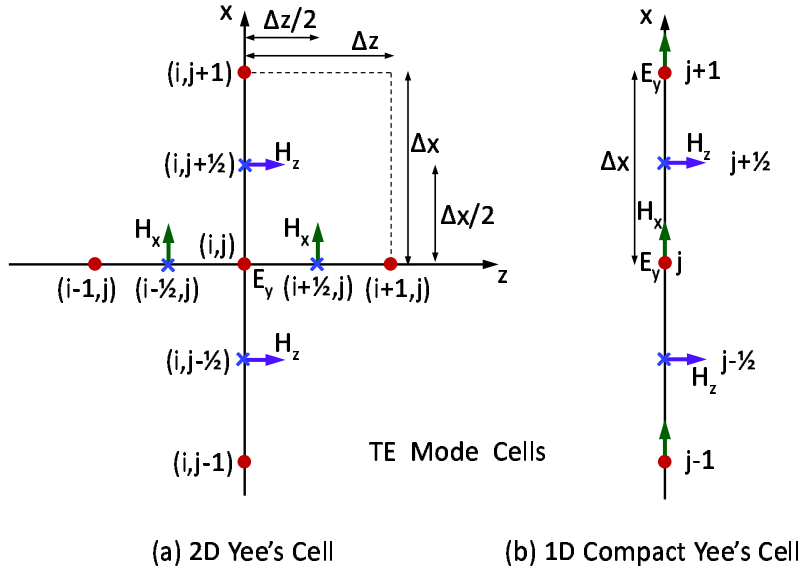


Figure 22: A 2D Yee's cell and its 1D compact counterpart as it applies to the geometry of the slab waveguide for both TE (top) and TM (bottom) modes. The point (z_i, x_j) is represented as (i, j) .

Using the previous definitions, equations (128–130) can be written in the following matrix form:

$$jN\tilde{E}_y = \tilde{H}'_x, \quad (132)$$

$$\frac{1}{\Delta x'} \tilde{D}_{x'}^E \tilde{E}_y = \tilde{H}'_z, \quad (133)$$

$$-\frac{1}{\Delta x'} \tilde{D}_{x'}^H \tilde{H}'_z - jN\tilde{H}'_x = \tilde{\epsilon}_r \tilde{E}_y. \quad (134)$$

where $\tilde{\epsilon}_r$ is a diagonal matrix that contains the relative permittivity, i.e. $\tilde{\epsilon}_r^j = \epsilon_r(x_j)$. Eliminating the magnetic field components from the last equations it is straightforward to obtain the following eigenvalue/eigenvector equation

$$\left\{ \tilde{\epsilon}_r + \frac{1}{(\Delta x')^2} \tilde{D}_{x'}^H \tilde{D}_{x'}^E \right\} \tilde{E}_y = N^2 \tilde{E}_y, \quad (135)$$

where N , and \tilde{E}_y are the eigenvalues and the corresponding eigenvectors (electric field profiles) of the TE modes, respectively.

Observing Eq. (130) there might be a dilemma of what relative permittivity to use especially in the case that the node is at an abrupt permittivity discontinuity. The only thing that is certain is that E_y electric field is continuous at the permittivity discontinuity boundary. Based on this, the following Taylor expansions can be made:

$$\begin{aligned} D_y(x + \Delta x) &= D_y(x) + \Delta x \frac{\partial D_y}{\partial x} = \epsilon_r(x) E_y(x) + \Delta x \frac{\partial D_y}{\partial x}, \\ D_y(x - \Delta x) &= D_y(x) - \Delta x \frac{\partial D_y}{\partial x} = \epsilon_r(x) E_y(x) - \Delta x \frac{\partial D_y}{\partial x}. \end{aligned}$$

Adding up the last two equations the following could be obtained

$$\begin{aligned} \epsilon_r(x) E_y(x) &= \frac{1}{2} [D_y(x + \Delta x) + D_y(x - \Delta x)] = \\ &= \frac{1}{2} [\epsilon(x + \Delta x) E_y(x + \Delta x) + \epsilon_r(x - \Delta x) E_y(x - \Delta x)] = \\ &= \frac{1}{2} [\epsilon(x + \Delta x) + \epsilon_r(x - \Delta x)] E_y(x), \end{aligned}$$

where the continuity of E_y was utilized, i.e. $E_y(x + \Delta x) = E_y(x - \Delta x) = E_y(x)$ for Δx being small. The previous averaging should be inserted into the eigenvalue/eigenvector equation (135) in the form of changing the relative permittivity of the boundary nodes as

$$\epsilon_r^j = \frac{1}{2} [\epsilon_r^{j-1} + \epsilon_r^{j+1}]. \quad (136)$$

As an example case, the same single film layer semiconductor waveguide that was presented in the previous section is used. This three-region slab waveguide has the parameters $n_c = 1.0$,

$n_f = 3.4$, $n_s = 3.1$, $h = 1.0\mu\text{m}$, and the freespace wavelength is $\lambda_0 = 1.3\mu\text{m}$. The distances from the electric walls are $d_c = 0.5\mu\text{m}$ and $d_s = 1.0\mu\text{m}$. The results for the effective indices of the two supported TE modes are shown in Table 2 and are compared to the previous solutions of the Helmholtz equation with the collocated fields. It is observed that with any of the three methods the results are very close. The bottom half of the table shows the error which is defined as $Error(\%) = 100[N_{exact} - N(\Delta x)]/N_{exact}$. It is observed that the Yee's cell with averaging at the boundary nodes gives slightly better results for the practical case of $\lambda_0/80 \simeq \lambda_{min}/20$ where $\lambda_{min} = \lambda_0/n_f$.

In the case of TM modes the non-zero field components are (H_y, E_x, E_z) where again each field component is of the form $H_y(z, x) = H_y(x) \exp(-j\beta z) = H_y(x) \exp(-jk_0 N z)$ (as an example), where $H_y(x)$ is the mode profile (for the magnetic field) and β , N are the effective propagation constant and the effective refractive index of the TM mode that is sought, and k_0 is the freespace wavenumber. Using again the normalization of the magnetic field components of the form $H'_w = -jZ_0 H_w$ the Maxwell's equations can be written in the form

$$-\frac{\partial H'_y}{\partial z} = k_0 \epsilon_r E_x \implies -\frac{\partial H'_y}{\partial z'} = \epsilon_r E_x, \quad (137)$$

$$\frac{\partial H'_y}{\partial x} = k_0 \epsilon_r E_z \implies \frac{\partial H'_y}{\partial x'} = \epsilon_r E_z, \quad (138)$$

$$\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = k_0 H'_y \implies \frac{\partial E_x}{\partial z'} - \frac{\partial E_z}{\partial x'} = H'_y, \quad (139)$$

where $x' = k_0 x$ and $z' = k_0 z$ are normalized coordinates.

At this point the space domain discretization is done according to Yee's scheme as shown in Fig. 22. The fields are discretized in the coordinate system shown as $H'_y(z_i, x_j) = H'_y{}^{i,j}$ taking into account the staggering between various field components.

$$-\frac{H'_y{}^{i,j} - H'_y{}^{i-1,j}}{\Delta z'} = \epsilon_r^{i-1/2,j} E_x^{i-1/2,j} \implies j N H'_y{}^j = \epsilon_{rxx}^j E_x^j e^{+jN\Delta z'/2}, \quad (140)$$

$$\frac{H'_y{}^{i,j} - H'_y{}^{i,j-1}}{\Delta x'} = \epsilon_r^{i,j-1/2} E_z^{i,j-1/2} \implies \frac{H'_y{}^j - H'_y{}^{j-1}}{\Delta x'} = \epsilon_{rzz}^{j-1/2} E_z^{j-1/2}, \quad (141)$$

$$-\frac{E_x^{i+1/2,j} - E_x^{i-1/2,j}}{\Delta z'} + \frac{E_z^{i,j+1/2} - E_z^{i,j-1/2}}{\Delta x'} = H'_y{}^{i,j} \implies -j N E_x^j - \frac{E_z^{j+1/2} - E_z^{j-1/2}}{\Delta x'} = H'_y{}^j, \quad (142)$$

where $\epsilon_{rxx}^j = \epsilon_r(x_j)$ and $\epsilon_{rzz}^{j-1/2} = \epsilon_r(x_{j-1/2})$. When $\Delta z' \rightarrow 0$ the 2D Yee's grid collapses to the compact 1D Yee's cell as it is shown also in the right column (b figures) of Fig. 22. Therefore,

Table 2: Effective Indices (of TE Modes) Calculated by FDFD either using Helmholtz Equation (collocated fields) or using Yee's Cell (staggered fields)

TE Modes						
Helmholtz			Yee's Cell		Yee's Cell	
Δx			No Averaging		Averaging	
	N_0^{TE}	N_1^{TE}	N_0^{TE}	N_1^{TE}	N_0^{TE}	N_1^{TE}
$\lambda_0/20$	3.3585913	3.2362825	3.3585914	3.2362911	3.3570617	3.2310872
$\lambda_0/40$	3.3579455	3.2333453	3.3579314	3.2332928	3.3575479	3.2319896
$\lambda_0/60$	3.3578155	3.2327610	3.3578157	3.2327737	3.3576393	3.2321748
$\lambda_0/80$	3.3577736	3.2325729	3.3577720	3.2325782	3.3576744	3.2322471
$\lambda_0/100$	3.3577538	3.2324843	3.3577531	3.2324938	3.3576897	3.2322788
$\lambda_0/120$	3.3577424	3.2324333	3.3577426	3.2324471	3.3576981	3.2322964
$\lambda_0/150$	3.3577336	3.2323938	3.3577338	3.2324078	3.3577053	3.2323113
$\lambda_0/200$	3.3577268	3.2323632	3.3577268	3.2323769	3.3577109	3.2323231
$\lambda_0/500$	3.3577193	3.2323297	3.3577195	3.2323443	3.3577169	3.2323357
$\lambda_0/750$	3.3577185	3.2323262	3.3577187	3.2323409	3.3577176	3.2323371
$\lambda_0/1000$	3.3577183	3.2323250	3.3577184	3.2323397	3.3577178	3.2323376
Exact	3.3577180	3.2323308	3.3577180	3.2323308	3.3577180	3.2323308

TE Modes						
Helmholtz			Yee's Cell		Yee's Cell	
Δx			No Averaging		Averaging	
	N_0^{TE}	N_1^{TE}	N_0^{TE}	N_1^{TE}	N_0^{TE}	N_1^{TE}
	Error (%)	Error (%)	Error (%)	Error (%)	Error (%)	Error (%)
$\lambda_0/20$	-0.0260	-0.1223	-0.0260	-0.1225	0.0195	0.0385
$\lambda_0/40$	-0.0068	-0.0314	-0.0064	-0.0298	0.0051	0.0106
$\lambda_0/60$	-0.0029	-0.0133	-0.0029	-0.0137	0.0023	0.0048
$\lambda_0/80$	-0.0017	-0.0075	-0.0016	-0.0077	0.0013	0.0026
$\lambda_0/100$	-0.0011	-0.0047	-0.0010	-0.0050	0.0008	0.0016
$\lambda_0/120$	-0.0007	-0.0032	-0.0007	-0.0036	0.0006	0.0011
$\lambda_0/150$	-0.0005	-0.0019	-0.0005	-0.0024	0.0004	0.0006
$\lambda_0/200$	-0.0003	-0.0010	-0.0003	-0.0014	0.0002	0.0002
$\lambda_0/500$	0.0000	0.0000	0.0000	-0.0004	0.0000	-0.0002
$\lambda_0/750$	0.0000	0.0001	0.0000	-0.0003	0.0000	-0.0002
$\lambda_0/1000$	0.0000	0.0002	0.0000	-0.0003	0.0000	-0.0002

the positions of H_y^j and E_x^j coincide (these field components become collocated) while the $E_z^{j-1/2}$ are shifted by $-\Delta x'/2$ (forming a staggered 1D grid). Now let's define the vectors $\tilde{H}_y' = [H_y'^1, H_y'^2, \dots, H_y'^M]^T$, $\tilde{E}_x = [E_x^1, E_x^2, \dots, E_x^M]^T$, and $\tilde{E}_z = [E_z^{1-1/2}, E_z^{2-1/2}, \dots, E_z^{M-1/2}]^T$. Using the previous definitions, equations (140–142) can be written in the following matrix form:

$$jN\tilde{H}_y' = \tilde{\epsilon}_{rxx}\tilde{E}_x, \quad (143)$$

$$\frac{1}{\Delta x'}\tilde{D}_{x'}^H\tilde{H}_y' = \tilde{\epsilon}_{rzz}\tilde{E}_z, \quad (144)$$

$$-jN\tilde{E}_x - \frac{1}{\Delta x'}\tilde{D}_{x'}^E\tilde{E}_z = \tilde{H}_y'. \quad (145)$$

where $\tilde{\epsilon}_{rxx}$ and $\tilde{\epsilon}_{rzz}$ are diagonal matrices that contain the relative permittivity, i.e. $\tilde{\epsilon}_{rxx}^j = \epsilon_r(x_j)$ and $\tilde{\epsilon}_{rzz}^j = \epsilon_r(x_{j+1/2})$. Eliminating the electric field components from the last equations it is straightforward to obtain the following eigenvalue/eigenvector equation,

$$\left\{ \tilde{\epsilon}_{rxx} + \frac{1}{(\Delta x')^2} \tilde{\epsilon}_{rxx}\tilde{D}_{x'}^E\tilde{\epsilon}_{rzz}^{-1}\tilde{D}_{x'}^H \right\} \tilde{H}_y = N^2\tilde{H}_y, \quad (146)$$

where N , and \tilde{H}_y are the eigenvalues and the corresponding eigenvectors (magnetic field profiles) of the TM modes, respectively.

Observing Eq. (140) there might be a dilemma of what relative permittivity to use especially in the case that the node is at an abrupt permittivity discontinuity (similarly to the TE mode case). The only thing that is certain in this case, is that D_x electric displacement field is continuous across the permittivity boundary (since there are no free charges). Based on this, the following Taylor expansions can be made:

$$\begin{aligned} \frac{D_x(x + \Delta x)}{\epsilon_r(x + \Delta x)} &= E_x(x + \Delta x) = E_x(x) + \Delta x \frac{\partial E_x}{\partial x}, \\ \frac{D_x(x - \Delta x)}{\epsilon_r(x - \Delta x)} &= E_x(x - \Delta x) = E_x(x) - \Delta x \frac{\partial E_x}{\partial x}. \end{aligned}$$

The last two equations can be manipulated as following in order to obtain the most appropriate expression for the relative permittivity that will be used in Eq. (140).

$$\begin{aligned} \frac{1}{2} \left[\frac{D_x(x + \Delta x)}{\epsilon_r(x + \Delta x)} + \frac{D_x(x - \Delta x)}{\epsilon_r(x - \Delta x)} \right] &= E_x(x), \\ D_x(x) \frac{1}{2} \left[\frac{1}{\epsilon_r(x + \Delta x)} + \frac{1}{\epsilon_r(x - \Delta x)} \right] &= E_x(x), \\ \frac{D_x(x)}{\epsilon_{r,eq}(x)} &= E_x(x), \\ D_x(x) &= \epsilon_{r,eq}(x)E_x(x), \\ \text{where } \epsilon_{r,eq}(x) &= \left\{ \frac{1}{2} \left[\frac{1}{\epsilon_r(x + \Delta x)} + \frac{1}{\epsilon_r(x - \Delta x)} \right] \right\}^{-1}. \end{aligned}$$

where used of the boundary condition $D_x(x + \Delta x) = D_x(x - \Delta x) = D_x(x)$ when Δx is small. The above averaging (harmonic average) should be inserted into the eigenvalue/eigenvector equation (146) in the form of changing the relative permittivity of the boundary nodes of $\tilde{\epsilon}_r$ as

$$\epsilon_{rxx}^j = \left\{ \frac{1}{2} \left[\frac{1}{\epsilon_{rxx}^{j-1}} + \frac{1}{\epsilon_{rxx}^{j+1}} \right] \right\}^{-1}. \quad (147)$$

Using the same example as before the above procedure is implemented for the determination of the TM modes. The results are tabulated in Table 3. Again it is observed that the Yee's cell with the harmonic averaging gives slightly better results than the solution of Helmholtz's equation using the Kawano scheme described in the previous subsection. The Yee's cell without the averaging performs consistently worse than the other two schemes.

12.1.1. Average-Scheme Modification for Arbitrary Boundary Location

The previously presented average scheme for the FDFD method for slab waveguides is valid only for the special case where the permittivity discontinuity plane boundary passes through an electric field node (E_y node for TE modes) [Eq. (136)] or a magnetic field node (H_y node for TM modes) [Eq. (147)]. In this section let's consider the case that the permittivity discontinuity boundary crosses a Yee's cell at an arbitrary location. This is shown schematically for the case of TE and TM modes in Figs. 23a and 23b, respectively.

Table 3: Effective Indices (of TM Modes) Calculated by FDFD either using Helmholtz Equation (collocated fields) or using Yee’s Cell (staggered fields)

TM Modes						
Δx	Helmholtz Kawano [51]		Yee’s Cell No Averaging		Yee’s Cell Averaging	
	N_0^{TM}	N_1^{TM}	N_0^{TM}	N_1^{TM}	N_0^{TM}	N_1^{TM}
$\lambda_0/20$	3.3518412	3.2129633	3.3496125	3.2041094	3.3512255	3.2110965
$\lambda_0/40$	3.3515178	3.2109972	3.3504428	3.2067131	3.3513622	3.2105083
$\lambda_0/60$	3.3514548	3.2106177	3.3507380	3.2077721	3.3513869	3.2104143
$\lambda_0/80$	3.3514346	3.2104963	3.3509033	3.2083858	3.3513963	3.2103793
$\lambda_0/100$	3.3514251	3.2104392	3.3509982	3.2087442	3.3514004	3.2103641
$\lambda_0/120$	3.3514196	3.2104064	3.3510631	3.2089916	3.3514026	3.2103558
$\lambda_0/150$	3.3514154	3.2103809	3.3511305	3.2092505	3.3514045	3.2103487
$\lambda_0/200$	3.3514215	3.2103612	3.3511995	3.2095179	3.3514060	3.2103430
$\lambda_0/500$	3.3514086	3.2103396	3.3513237	3.2100031	3.3514076	3.2103368
$\lambda_0/750$	3.3514082	3.2103373	3.3513516	3.2101129	3.3514078	3.2103361
$\lambda_0/1000$	3.3514081	3.2103364	3.3513656	3.2101682	3.3514078	3.2103358
Exact	3.3514080	3.2103532	3.3514080	3.2103532	3.3514080	3.2103532

TM Modes						
Δx	Helmholtz Kawano [51]		Yee’s Cell No Averaging		Yee’s Cell Averaging	
	N_0^{TM} Error (%)	N_1^{TM} Error (%)	N_0^{TM} Error (%)	N_1^{TM} Error (%)	N_0^{TM} Error (%)	N_1^{TM} Error (%)
$\lambda_0/20$	−0.0129	−0.0813	0.0536	0.1945	0.0054	−0.0232
$\lambda_0/40$	−0.0033	−0.0201	0.0297	0.1134	0.0014	−0.0048
$\lambda_0/60$	−0.0014	−0.0082	0.0200	0.0804	0.0006	−0.0019
$\lambda_0/80$	−0.0008	−0.0045	0.0153	0.0613	0.0003	−0.0008
$\lambda_0/100$	−0.0005	−0.0027	0.0124	0.0501	0.0002	−0.0003
$\lambda_0/120$	−0.0003	−0.0017	0.0103	0.0424	0.0002	−0.0001
$\lambda_0/150$	−0.0002	−0.0009	0.0083	0.0343	0.0001	0.0001
$\lambda_0/200$	−0.0004	−0.0002	0.0063	0.0260	0.0001	0.0003
$\lambda_0/500$	0.0000	0.0004	0.0025	0.0109	0.0000	0.0005
$\lambda_0/750$	0.0000	0.0005	0.0017	0.0075	0.0000	0.0005
$\lambda_0/1000$	0.0000	0.0005	0.0013	0.0058	0.0000	0.0005

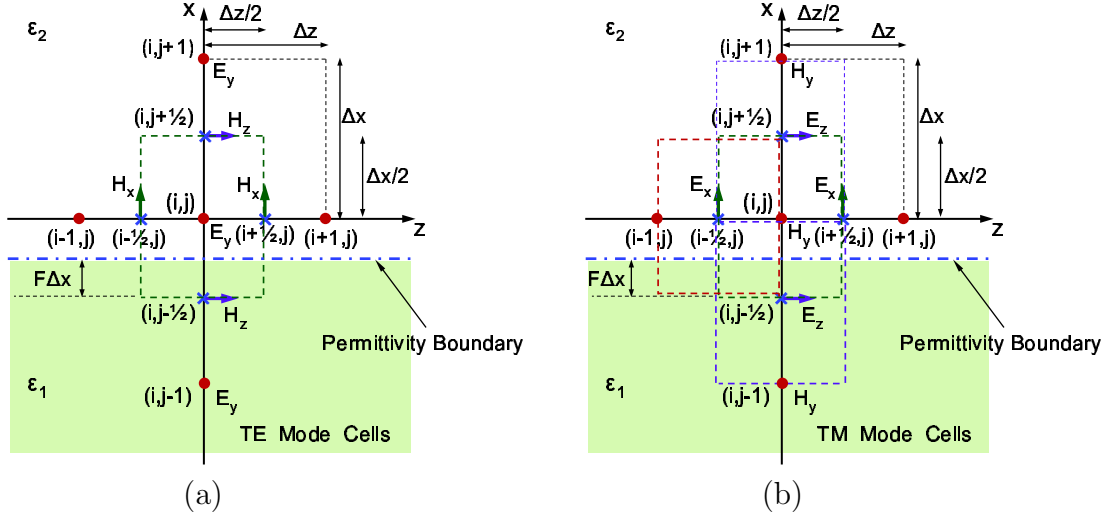


Figure 23: (a) A 2D Yee's cell as it applies to the geometry of the slab waveguide for TE modes. The point (z_i, x_j) is represented as (i, j) . The location of the boundary is shown with dashed-dotted line. It is assumed that the grid size along the z direction tends to zero ($\Delta z \rightarrow 0$). For the sake of simplicity the magnetic fields and the grid lengths are shown as unprimed. (b) Same as (a) but for the case of TM modes.

In order to determine the averaging scheme in this case the procedure developed by Hwang and Cangellaris [54] will be utilized. For the TE modes, let's apply the Ampere's law (in time-harmonic form) in its integral form, for the area S specified by the cell centered at the $E_y^{i,j}$ and the contour C defined by the perimeter of the cell (see Fig. 23a).

$$\begin{aligned}
 j\omega \iint_S \vec{D} \cdot d\vec{S} &= \oint_C \vec{H} \cdot d\vec{\ell} \implies \\
 j\omega\epsilon_0 \int_{x_j-\Delta x/2}^{x_j+\Delta x/2} \int_{z_i-\Delta z/2}^{z_i+\Delta z/2} \epsilon_r(x) E_y(z, x) dx dz &= \oint_C [H_x dx + H_z dz] \implies \\
 j\omega\epsilon_0 \int_{x_j-\Delta x/2}^{x_j+\Delta x/2} \epsilon_r(x) E_y^j(x) dx [e^{-jk_0 N z_i} \Delta z] &= - (H_z^{j+1/2} - H_z^{j-1/2}) [e^{-jk_0 N z_i} \Delta z] - \\
 &\quad H_x^j [-jk_0 N \Delta z e^{-jk_0 N z_i}] \Delta x \implies \\
 j\omega\epsilon_0 \int_{x_j-\Delta x/2}^{x_j+\Delta x/2} \epsilon_r(x) E_y^j(x) dx &= - (H_z^{j+1/2} - H_z^{j-1/2}) - H_x^j [-jk_0 N] \Delta x \implies \\
 j\omega\epsilon_0 [F\epsilon_1 + (1-F)\epsilon_2] E_y^j \Delta x &= - (H_z^{j+1/2} - H_z^{j-1/2}) - H_x^j [-jk_0 N] \Delta x \implies \\
 [F\epsilon_1 + (1-F)\epsilon_2] E_y^j = \epsilon_{r,eff}^j E_y^j &= - \frac{H_z^{j+1/2} - H_z^{j-1/2}}{\Delta x'} - jN H_x^j, \quad (148)
 \end{aligned}$$

where the last equation is the same with Eq. (130) with e_r^j replaced by $\epsilon_{r,eff}^j$. In the above derivation of the RHS integral with respect to x variable the continuity of the E_y (tangential

to the boundary) component was used. Therefore, in this arbitrary location of the permittivity discontinuity boundary the corresponding averaging is expressed by the following equation [which is the generalized form Eq. (136)],

$$\epsilon_{r,eff}^j = [F\epsilon_1 + (1 - F)\epsilon_2] = \epsilon_r^j = [F\epsilon_r^{j-1} + (1 - F)\epsilon_r^{j+1}]. \quad (149)$$

The above equation reduces to the Eq. (136) in the case that the boundary passes from the $E_y^{i,j}$ node since in this case $F = 1/2$. Since the materials are non-magnetic no averaging is needed for the other equations since the relative permeability is everywhere unity.

Now let's examine the case of the TM modes. Again the Ampere's law (in time-harmonic form) will be used. Assume the cell that is centered at the point of $E_x^{i-1/2,j}$ and is perpendicular to the xz -plane and apply again the Ampere's law as it was done previously. The following can be easily shown (where in this case S the area of the cell of size $\Delta z \times L_y$ with L_y a unit length in the y direction and C its perimeter)

$$\begin{aligned} j\omega \iint_S \vec{D} \cdot d\vec{S} &= \oint_C \vec{H} \cdot d\vec{\ell} \implies \\ j\omega \int_{L_y} \int_{z_i - \Delta z/2}^{z_i + \Delta z/2} \epsilon_r(x) E_x(x, z) dy dz &= \oint_C H_y dy \implies \\ j\omega \epsilon_0 \int_{z_i - \Delta z/2}^{z_i + \Delta z/2} \epsilon_r(x) E_x(x) e^{-jk_0 N z} dz L_y &= (H_y^{i-1,j} - H_y^{i,j}) L_y = H_y^j e^{-jk_0 N z_i} j k_0 N \Delta z L_y \implies \\ j\omega \epsilon_0 \epsilon_r(x) E_x(x) \left(\frac{e^{-jk_0 N z}}{-jk_0 N} \right) \Big|_{z_i - \Delta z/2}^{z_i + \Delta z/2} &= H_y^j e^{-jk_0 N z_i} j k_0 N \Delta z \implies \\ j\omega \epsilon_0 \epsilon_r(x) E_x(x) &= j k_0 N H_y^j \implies \\ \epsilon_r(x) E_x(x) &= j N H_y^j. \end{aligned} \quad (150)$$

The last equation is similar to Eq. (140). However, due to the existence of the permittivity boundary there is some ambiguity in the selection of the term $\epsilon_r(x) E_x(x)$. For example the surface S in Ampere's law can be any surface that ends up on the contour C . Therefore x could be selected anywhere in the cell. For this reason it is reasonable to assume that since D_x is continuous at the permittivity boundary to take an average value of $E_x(x)$ and an effective relative permittivity term. The weighted average of the the electric field at the center of this

cell is

$$\begin{aligned}
E_x^j &= F E_x(x_b^-) + (1 - F) E_x(x_b^+) = D_x^j \left[\frac{F}{\epsilon_{r1}} + \frac{1 - F}{\epsilon_{r2}} \right] \\
&= \frac{D_x^j}{\epsilon_{r,eff}^j} \quad \text{where} \\
\epsilon_{r,eff} &= \left[\frac{F}{\epsilon_{r1}} + \frac{1 - F}{\epsilon_{r2}} \right]^{-1} = \epsilon_{rxx}^j,
\end{aligned} \tag{151}$$

and x_b is the position of the permittivity discontinuity boundary while D_x^j is the approximately constant electric flux density within the cell used. The same result can be obtained if the cell centered at $H_y^{i,j}$ is used and the Faraday's law is used with S the area of the cell ($\Delta x \times \Delta z$) and contour C the perimeter of the cell. The above defined effective permittivity is the generalization of the average permittivity defined in Eq. (147). Of course these two expressions coincide in the case of the permittivity boundary passing through the $H_y^{i,j}$ node ($F = 1/2$).

In the TM modes case there is a need to apply the Ampere's law to another cell to generate the equivalent of Eq. (141). In this case the cell centered at $E_z^{i,j-1/2}$ is utilized with its area being perpendicular to xz plane and its dimensions of $\Delta x \times L_y$.

$$\begin{aligned}
j\omega \iint_S \vec{D} \cdot d\vec{S} &= \oint_C \vec{H} \cdot d\vec{\ell} \implies \\
j\omega \int_{L_y} \int_{x_j - \Delta x}^{x_j} \epsilon_r(x) E_z(x, z) dy dx &= \oint_C H_y dy \implies \\
j\omega \epsilon_0 \int_{x_j - \Delta x}^{x_j} \epsilon_r(x) E_z(x) e^{-jk_0 N z} dx L_y &= (H_y^{i,j} - H_y^{i,j-1}) L_y = (H_y^j - H_y^{j-1}) e^{-jk_0 N z_i} L_y \implies \\
j\omega \epsilon_0 \int_{x_j - \Delta x}^{x_j} \epsilon_r(x) E_z(x) dx &= (H_y^j - H_y^{j-1}) \implies \\
j\omega \epsilon_0 [F_1 \epsilon_{r1} + (1 - F_1) \epsilon_{r2}] E_z^{j-1/2} \Delta x &= (H_y^j - H_y^{j-1}) \implies \\
[F_1 \epsilon_{r1} + (1 - F_1) \epsilon_{r2}] E_z^{j-1/2} &= \epsilon_{rzz}^{j-1/2} E_z^{j-1/2} = \frac{H_y'^j - H_y'^{j-1}}{\Delta x'}, \quad \text{where} \\
\epsilon_{r,eff} &= F_1 \epsilon_{r1} + (1 - F_1) \epsilon_{r2} = \epsilon_{rzz}^{j-1/2},
\end{aligned} \tag{152}$$

where the term $F_1 = F + 1/2$ (if $F \leq 1/2$). In the case that $F > 1/2$ then the cell which is centered on $E_z^{i,j+1/2}$ is affected. Summarizing the new effective permittivity term to be used in Eq. (141) is given by

$$\epsilon_{r,eff} = \begin{cases} F_1 \epsilon_{r1} + (1 - F_1) \epsilon_{r2} = \epsilon_{rzz}^{j-1/2}, & F_1 = F + 1/2, \quad \text{for } F \leq 1/2, \\ F_2 \epsilon_{r1} + (1 - F_2) \epsilon_{r2} = \epsilon_{rzz}^{j+1/2}, & F_2 = F - 1/2, \quad \text{for } F > 1/2. \end{cases} \tag{153}$$

In the case that the permittivity boundary passes through the $H_y^{i,j}$ node ($F = 1/2$) then $\epsilon_{r,eff} = \epsilon_{r1}$ as it was used in the previous section. I.e., in this case the relative permittivity of Eq. (141) does not require any averaging.

When a single film layer slab waveguide is analyzed, there is always possible to select Δx in such a manner as to force either electric field nodes (for TE modes) or magnetic field nodes (for TM modes) to be located exactly on the two permittivity discontinuities boundaries. This simplifies significantly the averaging scheme which is as the one presented in the previous section. However, when there are more film layers, in general it is not possible to have field nodes located at all permittivity boundaries. As a comment, it can be stated, that a similar procedure could be used in the case that magnetic materials are used. In the latter case similar appropriate average relative permeabilities can be defined.

13. Graded-Index Slab Waveguides

As it was mentioned earlier, waveguide fabrication techniques such as metal-in-diffusion, ion bombardment, or ion migration result in waveguides with graded variation of the refractive index instead of the step-index profile that was studied up to this point. An example graded-index waveguide is shown in Fig. 24. A more general refractive index profile of a graded-index waveguide is shown in Fig. 25. Starting from the left-hand side, there is the cover region of refractive index n_c , then there is the substrate top region that is affected by the fabrication process, of refractive index $n_s(x)$, and finally, deep into the substrate (i.e. for large enough values of x), is the unaffected from the fabrication process part of the substrate with refractive index n_s . After, analyzing multi-layered slab waveguides any resulting graded index profile could potentially be approximated by a stair-step approximation and could be analyzed according to the theory developed in Sec. 11. However, in this section a short summary of an approximate technique will be presented, even if its utility at the present day computer era is questionable. The method has the initials WKB (from the initials of the researchers who first proposed it, Wentzel, Kramers, and Brillouin) and was originated in the area of quantum mechanics.

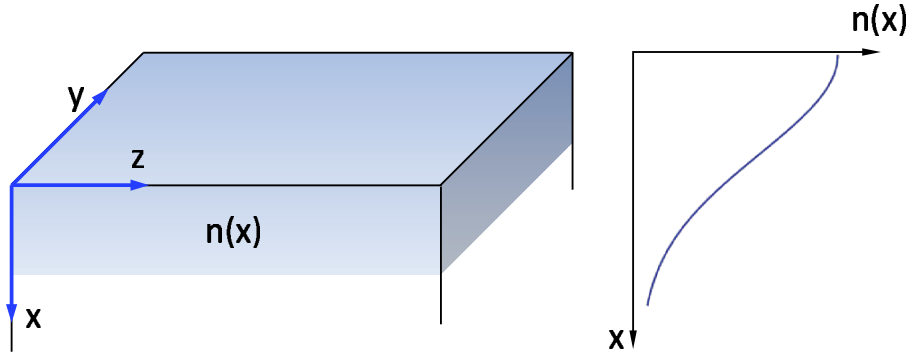


Figure 24: An example graded-index slab waveguide with an exponential-like refractive index profile. The dimension along the y -axis is much larger than the dimension along the x -axis so the light confinement is along the x direction and its propagation along the z direction.

Before presenting the basics of WKB method it is useful to start from the Maxwell's equations once more but now assuming that $\epsilon = \epsilon(x)$ for the graded-index waveguide. Assuming that the materials are linear, nonmagnetic, and isotropic, the Helmholtz's equation is written

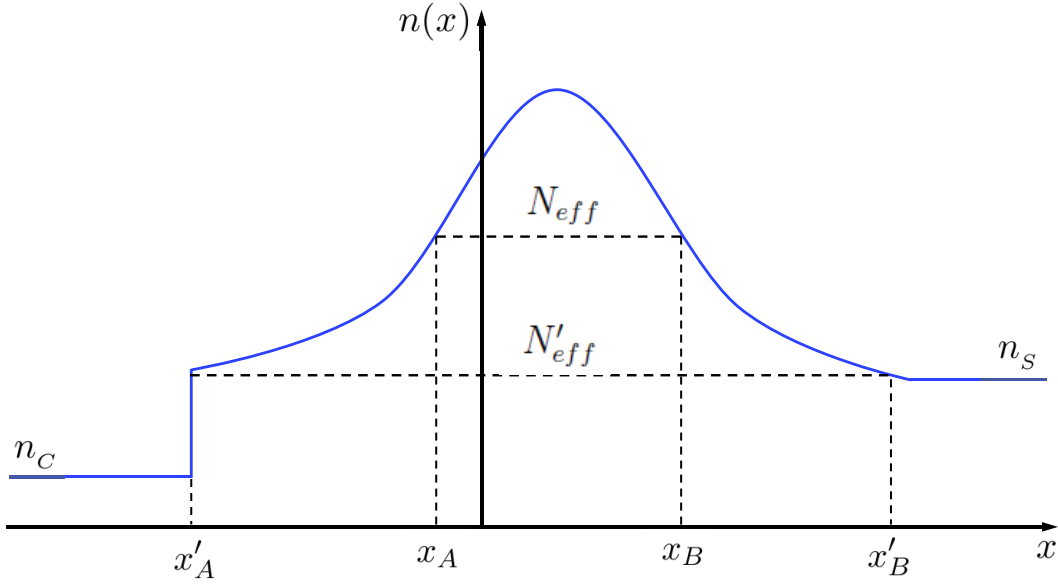


Figure 25: An example refractive index profile of a graded-index waveguide. For a mode of effective index N_{eff} the turning points x_A and x_B are shown. For a mode of effective index $N'_{eff} < n(x = x'_A)$ the turning points are x'_A (which is at the cover/graded-index-layer boundary) and x'_B could vary within the graded-index region.

as

$$\vec{\nabla} \left(\vec{\nabla} \cdot \vec{E} \right) - \nabla^2 \vec{E} = k_0^2 n^2(x) \vec{E}. \quad (154)$$

In order to simplify the solution it is now assumed that TE polarized solutions are sought. Therefore, $\vec{E} = \hat{y} E_y(x) \exp(-j\beta z)$. The y polarized electric field is perpendicular to the gradient of the relative permittivity variation, i.e. $\vec{\nabla} n^2 \cdot \vec{E} = 0$. Using the Gauss's law $\vec{\nabla} \cdot \vec{D} = 0 \Rightarrow \vec{\nabla} n^2 \cdot \vec{E} + n^2 \vec{\nabla} \cdot \vec{E} = 0 \Rightarrow \vec{\nabla} \cdot \vec{E} = 0$. Consequently the divergence of the electric field is zero, and the previous Helmholtz's equation simplifies to the form

$$\frac{d^2 E_y}{dx^2} + [k_0^2 n^2(x) - \beta^2] E_y(x) = 0. \quad (155)$$

The last equation is a second-order ordinary differential equation with varying coefficients. Usually, there are only very specific relative permittivity, $n^2(x)$, profiles for which analytical solutions are feasible. For example analytical solutions can be found in the case of parabolic, linear, and exponential profile [44]. In general, though there are no analytical solutions of this equation. As it was mentioned earlier, one approximate approach is the WKB method. In this

method a trial solution for E_y is selected of the form $\psi = \psi_0 \exp[jk_0 S(x)]$. Expanding $S(x)$ in the form

$$S(x) = S_0(x) + \frac{1}{k_0} S_1(x) + \frac{1}{k_0^2} S_2(x) + \cdots, \quad (156)$$

results in a trial solution of the form

$$\psi(x) = \psi_0 \exp \left[jk_0 S_0(x) + jS_1(x) + j\frac{1}{k_0} S_2(x) + \cdots \right], \quad (157)$$

where all higher order terms can be neglected since k_0 is a large number for optical frequencies. Using the previous trial solution in Helmholtz's equation, and retaining terms of the same order in k_0 , the S_0 and S_1 terms are given by

$$S_0(x) = \frac{1}{k_0} \int [k_0^2 n^2(x) - \beta^2]^{1/2} dx, \quad (158)$$

$$S_1(x) = \frac{j}{2} \ln \left| \frac{dS_0}{dx} \right|, \quad (159)$$

and the resulting trial solution $\psi(x)$ has the form

$$E_y(x) = \psi(x) = \begin{cases} \frac{\psi_0}{\sqrt{Q}} \exp \left[\pm j \int Q dx \right], & \text{if } [k_0^2 n^2(x) - \beta^2] = Q^2 > 0, \\ \frac{\psi_0}{\sqrt{P}} \exp \left[\pm \int P dx \right], & \text{if } [k_0^2 n^2(x) - \beta^2] = -P^2 < 0. \end{cases} \quad (160)$$

The above solutions for the example case shown in Fig. 25 take the form [44]

$$E_y(x) = \psi(x) = \begin{cases} \frac{\psi_0}{2[\beta^2 - k_0^2 n^2(x)]^{1/4}} \exp \left[- \int_x^{x_A} [\beta^2 - k_0^2 n^2(x')]^{1/2} dx' \right], & x < x_A, \\ \frac{\psi_0}{[k_0^2 n^2(x) - \beta^2]^{1/4}} \cos \left[\int_{x_A}^x [k_0^2 n^2(x') - \beta^2]^{1/2} dx' - \frac{\pi}{4} \right], & x_A < x < x_B, \\ \frac{\psi_0}{2[\beta^2 - k_0^2 n^2(x)]^{1/4}} \exp \left[- \int_{x_B}^x [\beta^2 - k_0^2 n^2(x')]^{1/2} dx' \right], & x > x_B, \end{cases} \quad (161)$$

where the solution between the turning points is propagating and elsewhere is evanescent. The above solution diverges when $x = x_A$ or $x = x_B$, i.e. at the turning points. Therefore, the above solutions are valid somewhat away from the turning points. Near the turning points special connections formulas are necessary (see Ref. [1]) in order to describe the electric field. However, further details of the WKB will not be pursued here. The interested reader can refer to Refs. [1] and [44]. Instead, a simple approach developed in Ref. [45] will be presented. By

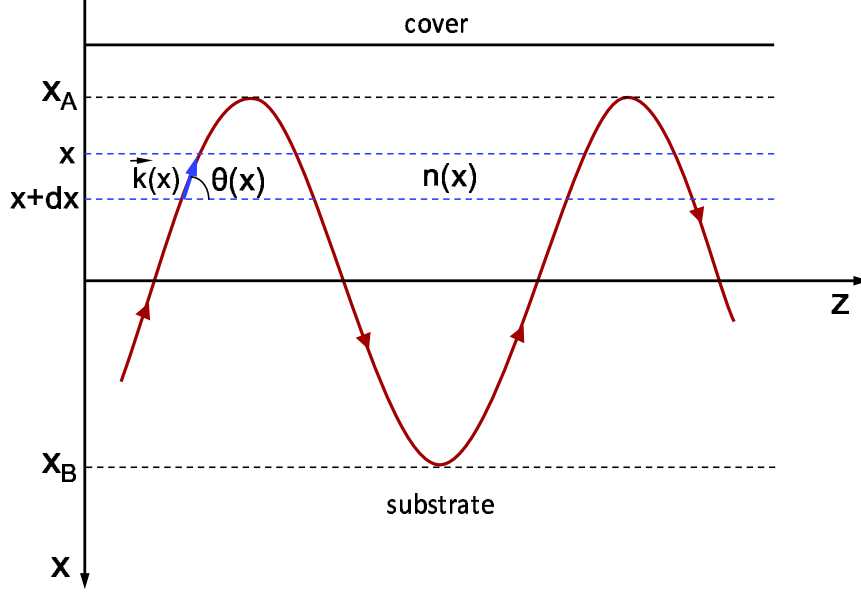


Figure 26: A picture diagram inside a graded-index waveguide for a guided mode. The turning points at x_A and x_B are shown. In addition, between x and $x + dx$ a wavevector $\vec{k}(x)$ is shown where $|\vec{k}(x)| = k_0 n(x)$ as an approximation.

doing so a dispersion equation for the determination of the propagation constant β can be easily developed.

Using the ray diagram shown in Fig. 26 it can be assumed that locally, between x and $x + dx$ the ray “sees” a constant refractive index of $n(x)$. Consequently, the plane wave solution can be used for the ray between x and $x + dx$. This means that the local wavevector $\vec{k}(x)$ satisfies locally the dispersion for a plane wave. Therefore,

$$k_x^2(x) + \beta^2 = k_0^2 n^2(x) = |\vec{k}(x)|^2. \quad (162)$$

Defining again $\beta = k_0 N$, where $N = N_{eff}$ is the effective index of the resulting guided mode, and using the angle $\theta(x)$ as defined in Fig. 26, it can be easily shown that $\theta(x) = \cos^{-1}[N/n(x)]$. Of course $\theta(x)$ takes real values when $n(x) > N = N_{eff}$ as it can be seen in Fig. 26 when $x_A < x < x_B$. However, when $x < x_A$ or $x > x_B$, $\theta(x)$ takes purely imaginary values. At the turning points $\theta(x_A) = \theta(x_B) = 0$, thus the ray becomes parallel to the z axis at the turning points. For a given value of the effective index, the turning points can be evaluated by inverse solution of the equation $n(x) = N \Rightarrow x_w = n^{-1}(N)$ ($w = A$ or B) and $n^{-1}(N)$ specifies the inverse function. Using the analogous of Eq. (4) for this case, the following dispersion condition

can be written

$$-2 \int_{x_A}^{x_B} k_x(x) dx + \phi_A + \phi_B = 2\nu\pi, \quad \nu = 0, \pm 1, \pm 2, \dots, \quad (163)$$

where ϕ_A and ϕ_B are the phase shifts occurring at the turning points x_A and x_B , respectively. These phase shifts can be determined by the use of Fresnel reflection coefficient in the limiting case between $x_A - \Delta x < x < x_A + \Delta x$ (and similarly for x_B) where Δx ($\Delta x > 0$) is a small variation around x_A (or x_B). Using the *TE* polarization Fresnel coefficient for the ray going from $x_A + \Delta x$ to $x_A - \Delta x$ (as seen in Fig. 26) the following can be written

$$r = \frac{k_x(x_A + \Delta x) - k_x(x_A - \Delta x)}{k_x(x_A + \Delta x) + k_x(x_A - \Delta x)} = \frac{1 - \frac{k_x(x_A - \Delta x)}{k_x(x_A + \Delta x)}}{1 + \frac{k_x(x_A - \Delta x)}{k_x(x_A + \Delta x)}}, \quad (164)$$

and using Eq. (162)

$$r = \frac{1 - \frac{[n^2(x_A - \Delta x) - N^2]^{1/2}}{[n^2(x_A + \Delta x) - N^2]^{1/2}}}{1 + \frac{[n^2(x_A - \Delta x) - N^2]^{1/2}}{[n^2(x_A + \Delta x) - N^2]^{1/2}}}. \quad (165)$$

At this point, it is necessary to use a Taylor series expansion for the square of the refractive index and retain up to first-order terms. This results in the following approximations:

$$n^2(x_A - \Delta x) \simeq n^2(x_A) - \frac{dn}{dx} \Delta x = N^2 - \frac{dn}{dx} \Delta x, \quad (166)$$

$$n^2(x_A + \Delta x) \simeq n^2(x_A) + \frac{dn}{dx} \Delta x = N^2 + \frac{dn}{dx} \Delta x. \quad (167)$$

Using the previous approximations the reflection coefficient becomes

$$\begin{aligned} r &= \frac{1 - \left\{ \frac{[n^2(x_A - \Delta x) - N^2]}{[n^2(x_A + \Delta x) - N^2]} \right\}^{1/2}}{1 + \left\{ \frac{[n^2(x_A - \Delta x) - N^2]}{[n^2(x_A + \Delta x) - N^2]} \right\}^{1/2}} \simeq \frac{1 - \left\{ \frac{-(dn/dx)\Delta x}{+(dn/dx)\Delta x} \right\}^{1/2}}{1 + \left\{ \frac{-(dn/dx)\Delta x}{+(dn/dx)\Delta x} \right\}^{1/2}} = \\ &= \frac{1 - j}{1 + j} = 1 \exp\left(-j\frac{\pi}{2}\right) = 1 \exp(j\phi_A). \end{aligned} \quad (168)$$

Therefore, the resulting phase shift $\phi_A = -\pi/2$. Similarly, the phase shift at turnig point x_B is also $\phi_B = -\pi/2$. Replacing the two phase factors in Eq. (163) the dispersion equation for the determination of the effective index for a graded index slab waveguide becomes

$$k_0 \int_{x_A=n^{-1}(N)}^{x_B=n^{-1}(N)} \sqrt{(n^2(x') - N^2)} dx' = \left(\nu + \frac{1}{2}\right) \pi, \quad \nu = 0, 1, 2, \dots. \quad (169)$$

In the latter dispersion equation the unknown effective index N is inside the integrand as well as in the integration limits. Of course the above equation can be solved only numerically. The solution provides the effective index of the mode while the field profile can be approximated by Eqs. (161) (relatively away though from the turning points).

Example 1: The first example case is comprised of a graded index slab waveguide with a relative permittivity profile of the form

$$n^2(x) = \varepsilon_s + \Delta\varepsilon \exp \left[-\frac{(x - x_0)^2}{w_0^2} \right], \quad (170)$$

where it is assumed that the variation is inside the substrate. For $\varepsilon_s = 4.80$, $\Delta\varepsilon = 0.045$, $x_0 = 8\mu\text{m}$, $w_0 = 2\mu\text{m}$, and $\lambda_0 = 0.6328\mu\text{m}$, the relative permittivity variation is shown in Fig. 27 along with a stair-step approximation by 50 layers (in the interval between $x_0 - 4w_0 < x < x_0 + 4w_0$). The effective indices were determined with three methods: (a) using the multi-layered waveguide analysis with the stair-step approximation of the profile (comprised of 50 film layers), (b) using the finite-difference frequency-domain (FDFD) method (presented in Sec. 12), and (c) the WKB method, i.e. the numerical solution of Eq. (169).

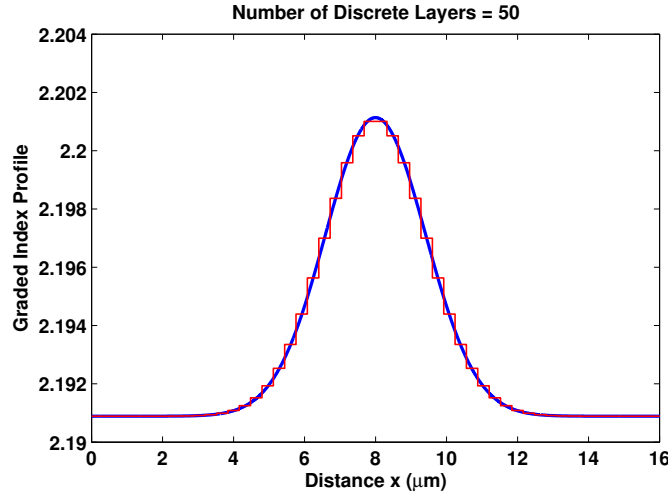


Figure 27: An example graded index waveguide profile of the form $n^2(x) = \varepsilon_s + \Delta\varepsilon \exp[-(x - x_0)^2/w_0^2]$. The parameters are $\varepsilon_s = 4.80$, $\Delta\varepsilon = 0.045$, $x_0 = 8\mu\text{m}$, $w_0 = 2\mu\text{m}$, and $\lambda_0 = 0.6328\mu\text{m}$. A stair-step approximation of the graded index profile is also shown for 50 layers of equal thickness of $h = 8w_0/50$.

The effective indices are shown tabulated in Table 4. It can be seen that the effective indices agree very well for all methods. The multi-layered waveguide analysis determines the solution

Table 4: Effective Indices of Example # 1 Graded Index Waveguide

	Multilayered Waveguide	Finite	Finite	
	Approximation	Difference	Difference	WKB
TE_ν	<i>Number of Layers</i> = 50	$\Delta x = 0.05\mu\text{m}$	$\Delta x = 0.025\mu\text{m}$	Approximation
	N_ν	N_ν	N_ν	N_ν
TE_0	2.198894532	2.198926188	2.198925969	2.198818399
TE_1	2.195047282	2.194992344	2.194991579	2.194884443
TE_2	2.192239107	2.192152744	2.192151661	2.192051926
TE_3	2.190896246	Not Found	Not Found	Not Found

TE_3 which both the FDFD and the WKB methods fail to determine. For the FDFD analysis this is due to the relatively small computational window that is used. Since TE_3 is near cutoff the evanescent tails extend well beyond the area shown in Fig. 27 and larger computational window could be needed for the determination of this mode. The WKB approximation fails to determine this mode since it is near cutoff and the corresponding approximations may not hold or it could be just a numerical issue. Furthermore, the TE_3 mode found by the multilayered approximation of the graded-index profile could be a product of this discretization. The normalized electric field profiles of the calculated TE_ν modes are shown in Fig. 28 when the multi-layered slab waveguide theory is used in conjunction with the stair-step approximation of the graded refractive index profile. Similar graphs are shown in Fig. 29 when the FDFD method is used.

Example 2: In this example the graded relative permittivity profile has the following form

$$n^2(x) = \begin{cases} \varepsilon_s + \Delta\varepsilon \exp\left[-\frac{x^2}{w_0^2}\right], & \text{if } x > 0, \\ n_c^2, & \text{if } x < 0. \end{cases} \quad (171)$$

For $\varepsilon_s = 4.80$, $\Delta\varepsilon = 0.045$, $w_0 = 4\mu\text{m}$, and $\lambda_0 = 0.6328\mu\text{m}$, and $n_c = 1.0$, the relative permittivity variation is shown in Fig. 30 along with a stair-step approximation by 40 layers (in the interval between $0 < x < 5w_0$) and it is shown in Fig. 30. Again, the effective indices are determined with three methods: (a) the multi-layered waveguide analysis with the stair-step approximation of the profile (40 film layers), (b) using the finite-difference frequency-domain (FDFD) method (presented in Sec. 12), and (c) the WKB method. However, for the WKB method, some special care has to be taken since the turning point x_A falls always at the cover-

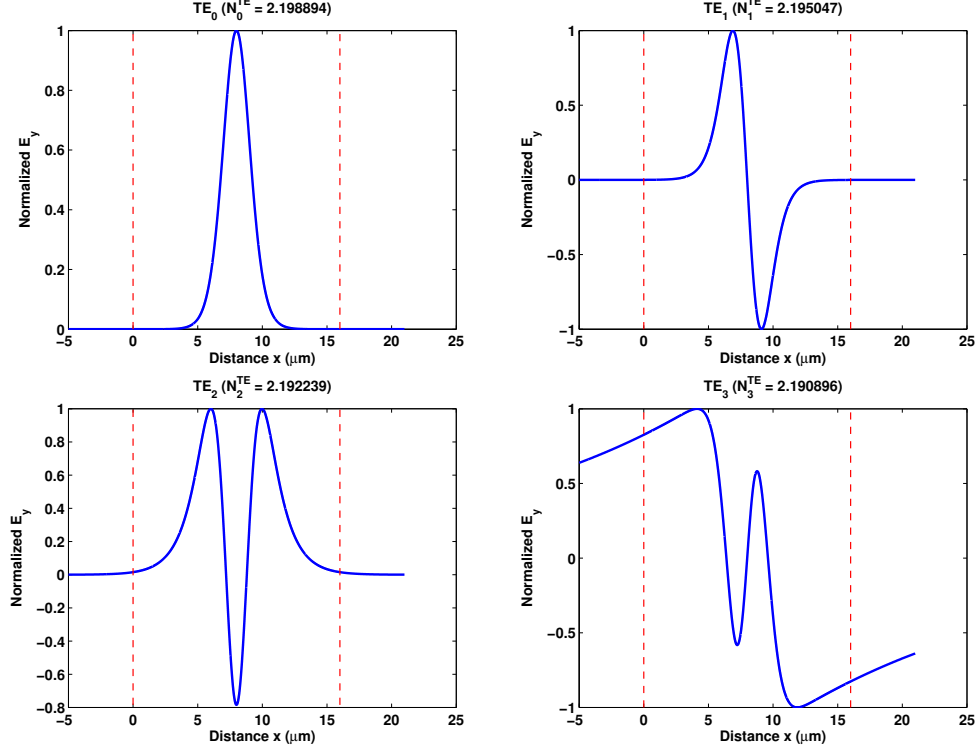


Figure 28: Normalized electric field profiles of TE_0 , TE_1 , TE_2 , and TE_3 modes for the example # 1 graded-index slab waveguide. The profile was calculated using the multi-layered slab waveguide theory. The vertical dashed lines show the cover-first film layer boundary (left dashed line at $x = 0$) and the boundary between the last film layer and substrate region (right dashed line). The boundaries between the film layers due to the stair-step approximation are not shown for simplicity of the figure.

substrate boundary and this means that the phase term ϕ_A should be given by the standard term expressing the phase shift upon total internal reflection between the graded index layer near the edge and the cover index layer. Therefore Eq. (169) has to be modified as

$$k_0 \int_{x_A=0}^{x_B=n^{-1}(N)} \sqrt{(n^2(x') - N^2)} dx' - \tan^{-1} \left\{ \left[\frac{N^2 - n^2(x=0^-)}{n^2(x=0^+) - N^2} \right]^{1/2} \right\} - \frac{\pi}{4} = \nu\pi, \quad \nu = 0, 1, 2, \dots \quad (172)$$

or equivalently

$$k_0 \int_{x_A=0}^{x_B=n^{-1}(N)} \sqrt{(n^2(x') - N^2)} dx' - \tan^{-1} \left\{ \left[\frac{N^2 - n_c^2}{\varepsilon_s + \Delta\varepsilon - N^2} \right]^{1/2} \right\} - \frac{\pi}{4} = \nu\pi, \quad \nu = 0, 1, 2, \dots \quad (173)$$

The effective indices are shown tabulated in Table 5. It can be seen that the effective indices agree very well for all methods. The normalized electric field profiles of the calculated

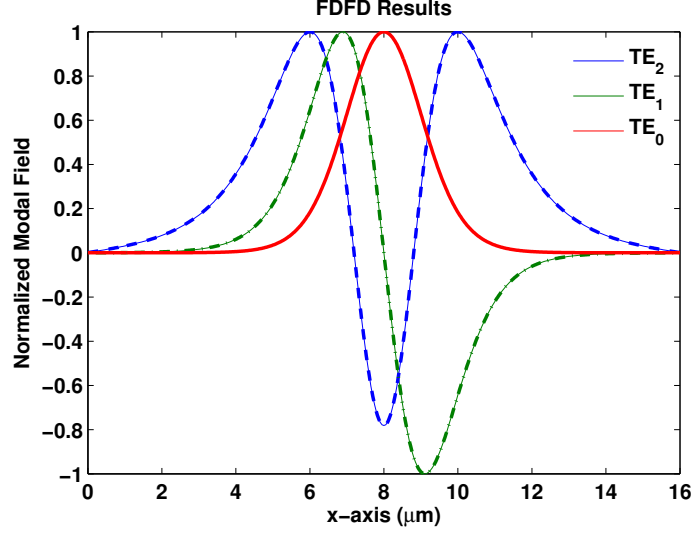


Figure 29: Normalized electric field profiles of TE_0 , TE_1 , and TE_2 modes for the example # 1 graded-index slab waveguide. The profile was calculated using the FDFD method with $\Delta x = 0.025\mu\text{m}$.

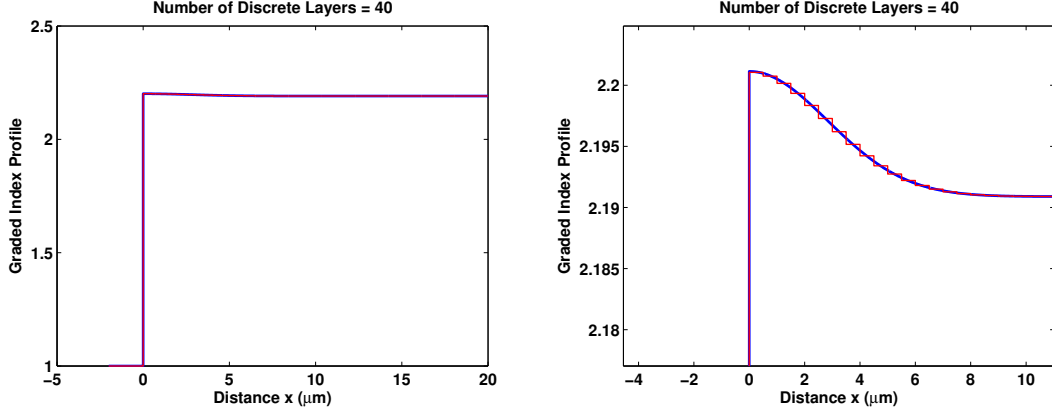


Figure 30: Another example of a graded index waveguide profile of the form $n^2(x) = \varepsilon_s + \Delta\varepsilon \exp[-x^2/w_0^2]$ for $x > 0$, $n^2(x) = n_c^2$ for $x < 0$ and . The parameters are $\varepsilon_s = 4.80$, $\Delta\varepsilon = 0.045$, $n_c = 1.0$, $w_0 = 4\mu\text{m}$, and $\lambda_0 = 0.6328\mu\text{m}$. A stair-step approximation of the graded index profile is also shown for 40 layers of equal thickness of $h = 5w_0/40$. The right plot is a zoomed version of the left in order to show the stair-step approximation.

TE_ν modes are shown in Fig. 31 when the multi-layered slab waveguide theory is used in conjunction with the stair-step approximation of the graded refractive index profile. Similar graphs are shown in Fig. 32 when the FDFD method is used.

Table 5: Effective Indices of Example # 2 Graded Index Waveguide

	Multilayered Waveguide	Finite	Finite	
	Approximation	Difference	Difference	WKB
TE_ν	$Number\ of\ Layers = 40$	$\Delta x = 0.05\mu m$	$\Delta x = 0.025\mu m$	Approximation
	N_ν	N_ν	N_ν	N_ν
TE_0	2.197868744	2.197905044	2.197877837	2.197825988
TE_1	2.194253274	2.194235177	2.194204855	2.194152293
TE_2	2.191721704	2.191680005	2.191658955	2.191615558

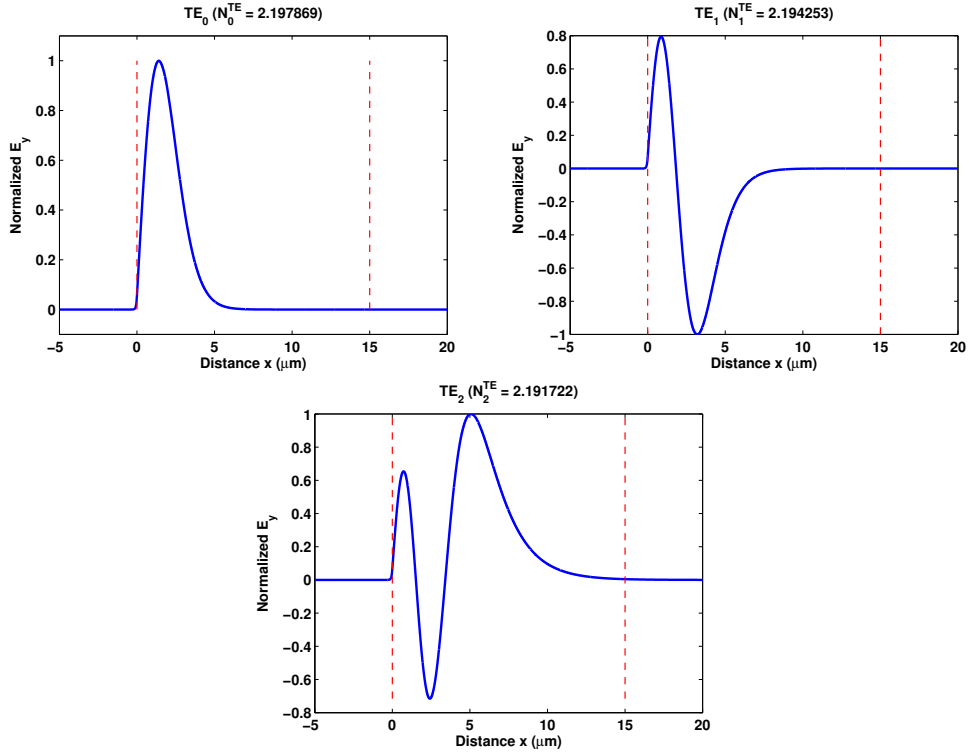


Figure 31: Normalized electric field profiles of TE_0 , TE_1 , and TE_2 modes for the example # 2 graded-index slab waveguide. The profile was calculated using the multi-layered slab waveguide theory (using 40 layers in the stair-step approximation). The vertical dashed lines show the cover-first film layer boundary (left dashed line at $x = 0$) and the boundary between the last film layer and substrate region (right dashed line). The boundaries between the film layers due to the stair-step approximation are not shown for simplicity of the figure.

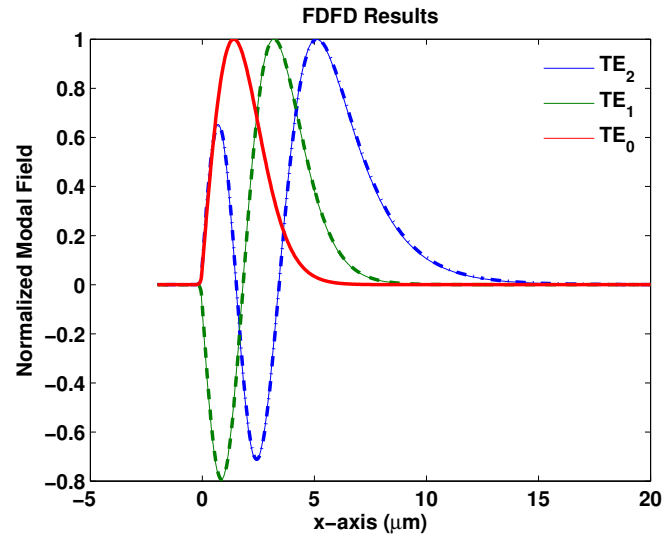


Figure 32: Normalized electric field profiles of TE_0 , TE_1 , and TE_2 modes for the example # 2 graded-index slab waveguide. The profile was calculated using the FDTD method with $\Delta x = 0.025\mu\text{m}$.

14. Lossy, Leaky, and Active Planar Multilayer Waveguides

In this section numerical methods for determining the *complex* propagation constants and their corresponding modal characteristics for planar multilayer waveguides are presented. In the previous lossless waveguides results presented, the real propagation constants were determined using the bisection method [8]. However, the calculation of the complex propagation constants that they can represent lossy, active, or leaky modes, is a lot more complicated. Initially, three similar globally converging, and robust numerical methods for determining zeros of analytic functions will be summarized. Then these methods are applied to the case of the planar multilayer waveguides. The three numerical mathematical methods are (a) the Argument Principal Method (APM) [10], (b) the Abd-ellal, Delves, and Reid (ADR) method [34, 36, 37], and the winding-number based on the analytic function's phase method [35]. In all cases the methods find good estimates of the function's zeros which are subsequently refined using the Müller's iterative method [9] with deflation [8].

14.1 Basic Theory of the Principle Argument Method (APM)

The application of the argument principle method for the determination of the modes of planar multilayer waveguides was first published by Smith *et al.* [14, 15] and Anemogiannis and Glytsis [16]. The argument principle method (APM) [10] relies on the generalized argument principle theorem from complex analysis. This theorem relates the number of zeros of an analytic function to a contour integral, and it has the form [21]

$$\sigma_k = \frac{1}{j2\pi} \oint_C z^k \frac{f'(z)}{f(z)} dz = \sum_{i=1}^{N_z} \zeta_i^k, \quad (174)$$

where $f(z)$ is an analytic function inside and on the simple closed contour C (Fig. 33a) and on which no zeros are located. In addition, ζ_i^k are the $i = 1, \dots, N_z$ zeros of $f(z)$ inside C raised to the k -th power. The quantity σ_k is their summation and $f'(z)$ denotes the derivative of $f(z)$ with respect to z . For the case that $f'(z)$ is not analytically given, it can be calculated numerically via the Cauchy's theorem [21], according to which the first derivative of a function $f(z)$ at the point $z = z_0$, $f'(z_0)$ is given by

$$f'(z_0) = \frac{1}{j2\pi} \oint_D \frac{f(z)}{(z - z_0)^2} dz, \quad (175)$$

where D is any simple path which encloses the point z_0 and $f(z)$ is analytic inside and on D .

The path integral of Eq. (174) evaluated for $k = 0$ gives the number of zeros of $f(z)$ enclosed by contour C . Also for $k = 1, \dots, N_z$, the same integral produces the summations $\sigma_1, \dots, \sigma_n$ which can be used to evaluate the coefficients of a polynomial $S(z)$ of degree n , which has the same roots, $\zeta_1, \dots, \zeta_{N_z}$, as the function $f(z)$ inside C . The approximation polynomial, $S(z)$, can be written as

$$S(z) = \sum_{\ell=0}^{N_z} c_\ell z^\ell, \quad (176)$$

with $c_{N_z} = 1$. The coefficients c_k are given via the Newton recursive formulas [55, 56]

$$(N_z - \ell) c_\ell + \sigma_1 c_{\ell+1} + \sigma_2 c_{\ell+2} + \dots + \sigma_{N_z-\ell} c_{N_z} = 0 \quad \text{for } \ell = N_z - 1, \dots, 0. \quad (177)$$

The polynomial, $S(z)$, can be solved by well known techniques [55], which in general are more efficient and accurate for low-degree polynomials. Therefore the maximum number of zeros enclosed by C must be low, i.e., $N_z \leq 4$, as proposed in [10, 21, 57]. Following the above procedure, the initial difficult problem of finding the zeros of an arbitrary function, $f(z)$, [or the poles of $f^{-1}(z)$], is transformed to the problem of finding the zeros of the polynomial, $S(z)$, for which a variety of reliable numerical methods exist. By suitably dividing the region where the zeros, ζ_i , are located and using Eq. (174) with $k = 0$, the degree of $S(z)$ can be set such that the polynomial root-solver can be more effective.

In practice, the zeros of the polynomial $S(z)$, ζ'_i , do not coincide with the zeros, ζ_i , of the initial function $f(z)$, due to errors introduced by the numerical evaluation of integrals of Eqs. (174) and (175). For this reason, a further refinement must be done on the roots of $S(z)$, applying a zero-finding numerical method [9, 55] with initial conditions ζ'_i .

As an example case the function $f(z) = e^{3z} + 2z \cos z - 1$, from Ref. [37], is considered. The derivative of the function is $f'(z) = 3e^{3z} + 2 \cos z - 2z \sin z$. The contour $ABCD$ is specified in Ref. [37] by the A , B , C , and D (see Fig. 33b caption). In order to determine the zeros of $f(z)$ inside the domain enclosed by $ABCD$ the APM method is applied. Using Eq. (174) for $k = 0$ the number of zeros is determined to be $\sigma_0 = N_z = 6$. Then the terms σ_k ($k = 1, 2, \dots, N_z$) are calculated from Eq. (174) and are shown in Table 6. The zero estimates and the refined by Müller's method zeros are shown in Table 7. It is observed that the zero estimates and the refined zeros are practically identical which means that the integrations were very accurately performed. Furthermore, it can be observed that the σ_k coefficients increase in magnitude as k

increases. Therefore, as it was mentioned previously, it is safer to apply the method in regions that contain $N_z \leq 4$ roots. In the case that $N_z > 4$ the initial contour can be subdivided into smaller contours that each one contains smaller than 4 roots.

Table 6: APM σ_k 's of function $f(z) = e^{3z} + 2z \cos z - 1$ from Ref. [37].

k	σ_k Coefficients
0	$6.00000000000000 - j0.0000000000000000$
1	$2.0467702626400 - j0.0000000000000000$
2	$-14.1574340854587 - j0.0000000000000001$
3	$-84.7999601511681 - j0.0000000000000007$
4	$-30.4653454987963 + j0.0000000000000006$
5	$696.0895474735294 - j0.0000000000000028$
6	$2527.5240087190520 + j0.0000000000000123$

Table 7: APM zeros estimates and refined zeros of function $f(z) = e^{3z} + 2z \cos z - 1$ from Ref. [37].

APM Zeros Estimates	Zeros after Müller's Refinement
$-1.844233953262218 - j0.0000000000000002$	$-1.844233953262213 + j0.0000000000000000$
$0.0000000000000019 + j0.0000000000000007$	$0.0000000000000000 - j0.0000000000000000$
$0.530894930292922 - j1.331791876751122$	$0.530894930292931 - j1.331791876751121$
$0.530894930292921 + j1.331791876751117$	$0.530894930292931 + j1.331791876751121$
$1.414607177658184 - j3.047722062627170$	$1.414607177658184 - j3.047722062627173$
$1.414607177658189 + j3.047722062627174$	$1.414607177658184 + j3.047722062627173$

14.2 Basic Theory of the Abd-ellal, Delves, and Reid (ADR) Method

The algorithm summarized here extracts the zeros of a complex function, $f(z)$, without using knowledge of the function's derivative. The algorithm used in [10], is rather complicated since it uses the multivalued $\log[f(z)]$ function. Another algorithm which was first presented by Abd-ellal, Delves, and Reid, (ADR method) [34], can be applied to meromorphic functions, i.e., complex functions that are analytic in the region enclosed by a closed contour C , except at a finite number of points (poles) inside C . The authors in [34] describe four different algorithms which extract both poles and zeros of a meromorphic function in a complex region, and apply their algorithms inside the unit circle on well behaved (polynomial) functions.

Here, the Abd-ellal *et al.* [34] results are adopted in order to be used for the eigenvalue dispersion equations of optical, microwave, and semiconductor structures. In the case that the function $f(z)$ has N_z zeros inside the complex contour C , a polynomial can be constructed having the same zeros as $f(z)$. The polynomial has the form

$$P_{N_z}(z) = c_0 + c_1 z + c_2 z^2 + \cdots + c_{N_z-1} z^{N_z-1} + z^{N_z}, \quad (178)$$

and its coefficients, c_i , ($i = 0, \dots, N_z - 1$), are obtained after solving the system of linear equations [34].

$$\sum_{j=0}^{N_z-1} c_j G_{r+j} + G_{r+N_z} = 0, \quad \text{for } r = 0, 1, 2, \dots, (N_z - 1). \quad (179)$$

The G_k quantities are given as

$$G_k = \frac{1}{j2\pi} \oint_C \frac{z^k}{f(z)} dz, \quad \text{for } k = 0, 1, 2, \dots, (2N_z - 1). \quad (180)$$

The number of zeros N_z inside the domain enclosed by contour C (see Fig. 33a) is critical in the above process and can be determined according to the argument principle theorem [21] (including possible multiplicities) [35] by

$$N_z = \frac{1}{j2\pi} \oint \frac{f'(z)}{f(z)} dz = \frac{1}{j2\pi} \Delta_C \{\ln[f(z)]\} = \frac{1}{2\pi} \Delta_C \{\arg[f(z)]\}, \quad (181)$$

where $f'(z)$ is the first derivative of $f(z)$ and the contour C is traveled counterclockwise. Therefore, knowing N_z from Eq. (185), the coefficients G_k ($k = 0, 1, 2, \dots, 2N_z - 1$) can be calculated from Eq. (180), and then the polynomial coefficients c_j ($j = 0, 1, 2, \dots, N_z - 1$) can be evaluated from the solution of Eq. (179). Finally, the roots of the polynomial of Eq. (178) can be used as estimates in the Müller's iterative procedure for the final refinement of the function's zeros.

An alternative method for finding the zeros estimates has been proposed by Kravanja *et al.* [36] and Gillan *et al.* [37]. Their approach calculates the G_k coefficients as they are defined

in Eq. (180). Then the following matrices can be defined:

$$\mathbf{H}^< = \begin{bmatrix} G_1 & G_2 & G_3 & \dots & G_{N_z} \\ G_2 & G_3 & G_4 & \dots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ G_{N_z} & \dots & \dots & \dots & G_{2N_z} \end{bmatrix} \quad (182)$$

$$\mathbf{H} = \begin{bmatrix} G_0 & G_1 & G_2 & \dots & G_{N_z-1} \\ G_1 & G_2 & G_3 & \dots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ G_{N_z-1} & \dots & \dots & \dots & G_{2N_z-2} \end{bmatrix} \quad (183)$$

The zeros of the $f(z)$ function can be calculated by the following generalized eigenvalue problem,

$$(\mathbf{H}^< - \zeta' \mathbf{H}) \mathbf{V} = 0, \quad (184)$$

where ζ' correspond to the $\zeta'_1, \dots, \zeta'_{N_z}$ are the estimates of the zeros of the function, and \mathbf{V} are the corresponding wavefunctions of the generalized eigenvalue problem (that are not utilized and therefore are computed).

As an example case, again the function $f(z) = e^{3z} + 2z \cos z - 1$ from Ref. [37], is considered. The contour $ABCD$ is specified in Ref. [37] by the A , B , C , and D (see Fig. 33b caption). In order to determine the number of zeros of $f(z)$ inside the domain enclosed by $ABCD$ the ADR method is applied. Using Eq. (185) the number of zeros is determined to be $N_z = 6$. Then the coefficients G_k ($k = 0, 1, \dots, 2 \times 6 - 1$) are given in Table 8. The corresponding zero estimates, ζ'_i as solutions of the generalized eigenvalue problem [Eq. (184)] along with the final refined by Müller's method are shown in Table 9. It is observed that the estimates are identical to the refined roots and they agree perfectly with the ones specified in Ref. [37]. Also it can be observed that the G_k coefficients increase in magnitude as k increases. Therefore, it is safer to apply the method in regions that contain $N_z \leq 4$ roots. In the case that $N_z > 4$ the initial contour can be subdivided into smaller contours that each one contains smaller than 4 roots as it was proposed for APM in Sec. 14.1.

14.3 Basic Theory of the Derivative-Free Zero-Extraction by Phase-based Enclosure Method (DFZEPE)

The idea of the zero extraction by phase-based enclosure is based on the Ying and Katz [35] winding number evaluation of an analytic function. Assume a function $f(z)$ (where $z = x + jy$)

Table 8: ADR G_k 's of function $f(z) = e^{3z} + 2z \cos z - 1$ from Ref. [37].

k	G_k Coefficients
0	$-0.07519739 - j0.00000000$
1	$0.27503184 + j0.00000000$
2	$-0.90593139 - j0.00000000$
3	$2.07560724 + j0.00000000$
4	$-2.01447904 + j0.00000000$
5	$2.44932797 + j0.00000000$
6	$-22.63060397 + j0.00000000$
7	$14.84101715 + j0.00000000$
8	$99.50440329 + j0.00000000$
9	$456.78080464 - j1.832 \times 10^{-13}$
10	$-483.81132715 - j9.047 \times 10^{-14}$
11	$-5297.84498932 - j2.099 \times 10^{-12}$

Table 9: ADR zeros estimates and refined zeros of function $f(z) = e^{3z} + 2z \cos z - 1$ from Ref. [37].

ADR Zeros Estimates	Zeros after Müller's Refinement
$-1.844233953262213 - j0.0000000000000000$	$-1.844233953262213 + j0.0000000000000000$
$-0.000000000000000017 - j0.000000000000000032$	$0.0000000000000000 - j0.0000000000000000$
$0.530894930292934 - j1.331791876751112$	$0.530894930292930 - j1.331791876751121$
$0.530894930292934 + j1.331791876751058$	$0.530894930292931 + j1.331791876751121$
$1.414607177658188 - j3.047722062627168$	$1.414607177658184 - j3.047722062627173$
$1.414607177658189 + j3.047722062627171$	$1.414607177658184 + j3.047722062627173$

which is analytic in a bounded domain enclosed by the rectangular contour C (defined by $ABCD$) shown in Fig. 33a. According to the argument principle theorem [21] the number of zeros (including possible multiplicities), N_z , of function $f(z)$ within the domain in the interior of contour C is given as [35]

$$N_z = \frac{1}{j2\pi} \oint \frac{f'(z)}{f(z)} dz = \frac{1}{j2\pi} \Delta_C \{\ln[f(z)]\} = \frac{1}{2\pi} \Delta_C \{\arg[f(z)]\}, \quad (185)$$

where $f'(z)$ is the first derivative of $f(z)$ and the contour C is traveled counterclockwise. The range of the argument function “arg” is in the interval $(-\pi, +\pi]$. For a straight section defined by $[z_i, z_{i+1}]$ (along the contour C) that satisfies $|\Delta_{[z_i, z_{i+1}]} \arg[f(z)]| \leq \pi$ the phase accumulation is $\Phi_t|_{z_i}^{z_{i+1}} = \arg[f(z_{i+1})/f(z_i)]$. Therefore, the total length of the contour C can be decomposed into linear segments satisfying the previous condition, by $4M$ points (M points per side of the rectangle) such as $C = \cup_{i=1}^{4M} [z_i, z_{i+1}]$ ($i = 1, 2, \dots, 4M$ with $z_1 = z_{4M+1}$). Then, N_z can be expressed in the form [35]

$$N_z = \frac{1}{2\pi} \sum_{i=1}^{4M} \arg \left[\frac{f(z_{i+1})}{f(z_i)} \right] = \frac{1}{2\pi} \Phi_t, \quad (186)$$

where Φ_t is the total accumulated phase along the closed contour $ABCD$.

The last equation, Eq. (186), is the heart of the Derivative-Free Zero-Extraction by Phase-based Enclosure (DFZEPE) method. Assume that the size of the sub-rectangle that approximates each zero of the function in the complex plane has pre-selected dimensions $w_x \times w_y$ as shown in Fig. 33a where $w_x \ll L_x = x_2 - x_1$ and $w_y \ll L_y = y_2 - y_1$. The w_x and w_y specify the zeros-enclosing sub-rectangles maximum dimensions and are directly related to the accuracy of the estimation of the function's $f(z)$ zeros. For well separated zeros w_x and w_y could be of the order of $10^{-1}L_x$ and $10^{-1}L_y$, respectively. For closely located zeros w_x and w_y could be of the order of $10^{-6}L_x$ and $10^{-6}L_y$ or even smaller. However, it is not necessary to isolate all roots since the Müller's algorithm which employs deflation can detect them inside the $w_x \times w_y$ enclosing rectangle in most cases. In general though for close-spaced roots, isolation of the roots is more reliable. Of course in the isolated roots the deflation procedure is unnecessary. The DFZEPE algorithm is summarized in the following steps:

- Define the number of points M per segment of each rectangle or sub-rectangle for the use of Eq. (186). This number could vary from 50-200 for functions $f(z)$ with well-separated

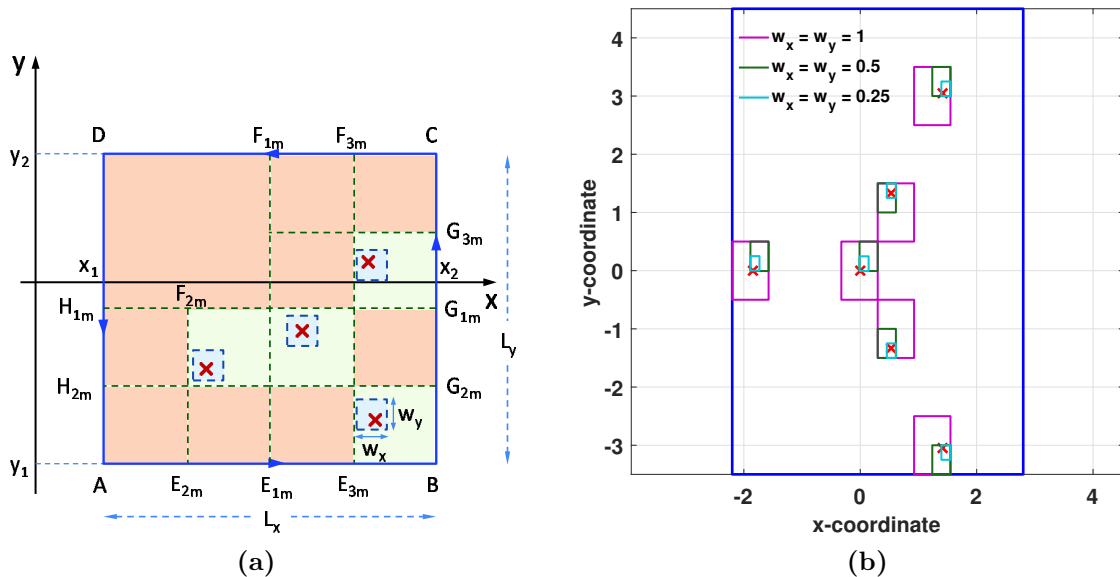


Figure 33: (a) A rectangular bounded domain in the complex plane ($z = x + jy$). The rectangular contour $ABCD$ encloses the domain where the function $f(z)$ is analytic. The “ \times ” correspond to zeros (roots) of $f(z)$ in this domain. The DFZEPE procedure encloses each root of the function by successively shrinking sub-rectangles. The sub-rectangles of dimensions $\leq w_x, w_y$ correspond to the final sub-rectangles that enclose each of the zeros. (b) An example case for $f(z) = e^{3z} + 2z \cos z - 1$ (from Ref. [37]) where the final sub-rectangles are shown for various selections of w_x and w_y for visualization of the DFZEPE process. The “ \times ” correspond to the positions of the zeros as specified in Ref. [37]. The realistic values of w_x and w_y that were used for this example were $w_x = w_y = 0.0025$ but in the scale of the figure these small sub-rectangles will show like dots. The A , B , C , and D points correspond to $x_1 = -2.2$, $x_2 = 2.5$, $y_1 = -3.5$, and $y_2 = 4.5$ as given in Ref. [37].

zeros up to 500-1000 (or even more if root isolation is sought) for functions with very closely located zeros and/or fast phase variation. This can be done in an adaptive manner. For example, one could start with $M = M_0$ and then repeat the process for $M = 2M_0$ and $M = 3M_0$. If the results are identical (number and location of roots) then M_0 is sufficient for all phase accumulation computations. However, it should be pointed out that the selection of M is strongly dependent on the selection of w_x and w_y . Actually, the latter are more critical for the DFZEPE method since large w_x and w_y results in crude two-dimensional bracketing of the roots and this could cause failure of the Müller's refinement algorithm (employing deflation which is needed if the roots are not isolated).

- Subdivide the initial rectangle $ABCD$ into four equal sub-rectangles using the horizontal divider segment $H_{1m}G_{1m} = L_x$ and the vertical divider segment $E_{1m}F_{1m} = L_y$. Now each

of the four equal sub-rectangles has dimensions $0.5L_x \times 0.5L_y$.

- Using Eq. (186) check the number of zeros within each of the newly-formed four sub-rectangles. Check the value of the function along the sub-rectangles boundaries. In case that $|f(z)| \leq \varepsilon_F$ (where ε_F is related to the computer precision and to the function behavior within the complex domain of interest) move divider $H_{1m}G_{1m}$ slightly to the bottom of the midpoint in y -direction and/or divider $E_{1m}F_{1m}$ slightly to the left to the midpoint in the x direction (for each sub-rectangle). In this manner, mitigation of the boundary proximity or even crossing to a zero is achieved. In the case of dividers movement from the midpoints, recalculate the number of zeros of each sub-rectangle.
- In case that the number of zeros of a sub-rectangle is zero disregard it.
- If the number of zeros of a sub-rectangle is greater or equal to one continue subdividing it while $width > w_x$ or $height > w_y$ where $width \simeq 0.5^\ell L_x$ and $height \simeq 0.5^\ell L_y$ denote the dimensions of the sub-rectangle after ℓ subdivisions. For example in Fig. 33a two subdivisions are shown to illustrate the process (see the dividers $E_{2m}F_{2m}$, $H_{2m}G_{2m}$, and $E_{3m}F_{3m}$ which represent the second-level subdivisions ($\ell = 2$). Observe that the second subdivision process occurred only in the sub-rectangles that include zeros inside. Store the four corners of each sub-rectangle and its N_z if $N_z \geq 1$ at the end of this loop and repeat the process for all stored sub-rectangles while the conditions $width > w_x$ or $height > w_y$ hold.
- Use the centers of the converged sub-rectangles that enclose the zeros as the estimates of the zeros of function $f(z)$.
- Use these zero estimates in the Müller iterative algorithm with deflation [8] to obtain refinement of the function's $f(z)$ zeros. Store the final refined zeros. Since the Müller's method requires three initial starting points, if z_e is the zero estimate provided by DFZEPE's enclosure process, the three initial points are z_e and $z_e \pm \min\{w_x, w_y\}/2$.
- End of the algorithm.

As an example case the function $f(z) = e^{3z} + 2z \cos z - 1$ from Ref. [37] is considered. The contour $ABCD$ is specified in Ref. [37] by the A , B , C , and D (see Fig. 33b caption).

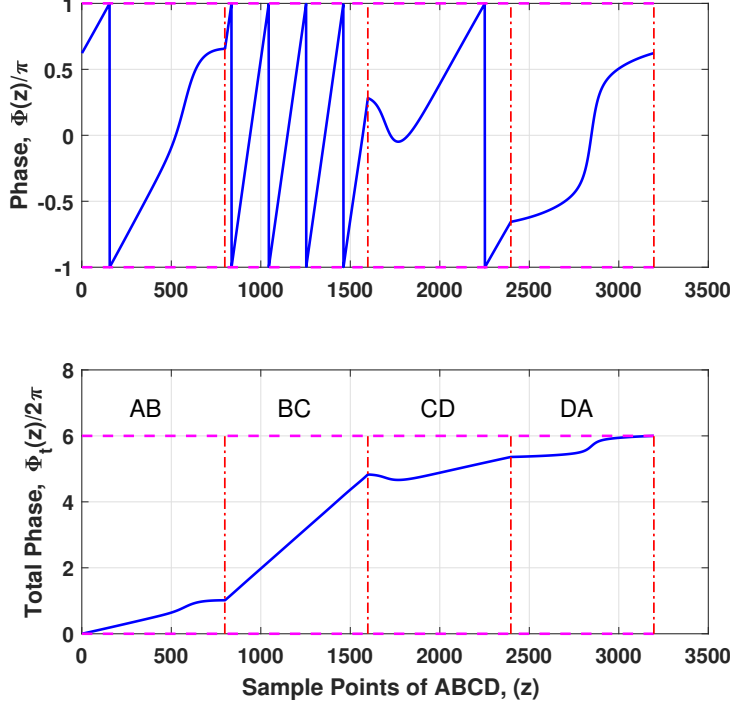


Figure 34: (a) The phase variation around $ABCD$, $\Phi(z) = \arg[f(z)]$, for the function $f(z) = e^{3z} + 2z \cos z - 1$ (from Ref. [37]) and (b) the total phase accumulation $\Phi_t(z) = \sum_{i=1}^I \arg[f(z_{i+1})/f(z_i)]$ where I is an index such $z_I = z$ and $\Phi_t(z_1) = 0$. The AB , BC , CD , and DA show the phase accumulation within each segment of the contour C . Obviously, the total accumulated phase is $6 \times 2\pi$ that reveals the existence of 6 zeros ($N_z = 6$) within the domain enclosed by C .

In order to determine the number of zeros of $f(z)$ inside the domain enclosed by $ABCD$ the DFZEPE method is applied. The $\Phi(z) = \arg[f(z)]$ and the accumulated phase $\Phi_t(z) = \sum_{i=1}^I \arg[f(z_{i+1})/f(z_i)]$ where I is an index such $z_I = z$ are shown in Fig. 34 for demonstration purposes. It is obvious that the function $f(z)$ has six zeros in the specified domain. It is also interesting to mention that the largest phase accumulation occurs on the BC segment of the contour. For illustration purposes the convergence of the subdivision process is shown Fig. 33b for three different selections of $w_x = w_y = 1, 0.5$, and 0.25 . It can be observed how the converged sub-rectangles enclose tighter the zeros of the function as w_x and w_y become smaller. In the actual numerical test $w_x = w_y = 0.0025$, $M = 250$, and $\varepsilon_F = 10^{-11}$ were selected and the zero estimates along with the Müller's refinement are shown in Table 10. The agreement with the results provided by Gillan *et al.* [37] is perfect for the precision shown in the table.

It is worth mentioning that the number of points per contour segment, M , could be selected

Table 10: DFZEPE zeros estimates and refined zeros of function $f(z) = e^{3z} + 2z \cos z - 1$ from Ref. [37].

DFZEPE Zeros Estimates	Zeros after Müller's Refinement
$-1.844165039062500 + j0.000976562500000$	$-1.844233953262213 + j0.000000000000000$
$0.000317382812500 + j0.000260416666667$	$0.000000000000000 - j0.000000000000000$
$0.531323242187500 - j1.331054687500000$	$0.530894930292931 - j1.331791876751121$
$0.531323242187500 + j1.331054687500000$	$0.530894930292931 + j1.331791876751121$
$1.415112304687500 - j3.047851562500000$	$1.414607177658184 - j3.047722062627173$
$1.415112304687500 + j3.047851562500000$	$1.414607177658184 + j3.047722062627173$

smaller and the w_x and w_y can be selected larger if it is not required that the mode indices to be separated by different enclosing sub-rectangles. In such a case the final sub-rectangles could enclose more than a single mode but the Müller's algorithm with deflation can successfully distinguish all the zeros since their number is known. This was tested for modes separated by $\Delta N_{eff,r} \leq 10^{-6}$ with $w_x = w_y = 10^{-3}$ with success in the case of two coupled waveguides that was reported in Ref. [38]. However, it is always safer to isolate by different enclosing rectangles each mode (since for the case of multilayer waveguides there are no multiple zeros). Cases where the final sub-rectangles could enclose more than a single zero could exist in problems where there are zeros with higher than one multiplicities.

14.4 Application of APM, ADR, and DFZEPE methods to Planar Multilayer Waveguides

A general planar multilayer waveguide is shown in Fig. 35a. The waveguide is uniform in the y direction, while the light confinement is in the x direction, and the light propagation is in the z direction. It is comprised of the cover region ($x < 0$) of refractive index n_c , the substrate region ($x > \sum_{i=1}^N h_i$) of refractive index n_s , and N film layers of height h_i , and refractive index n_i (where $i = 1, 2, \dots, N$). The refractive indices could be of the form $n = n_R + jn_I$ where $n_R > 0$ and $n_I = 0$ for lossless layers, $n_R > 0$ and $n_I < 0$ for lossy layers, and $n_R > 0$ and $n_I > 0$ for active layers. Of course this selection of signs is consistent with $\exp[+j(\omega t - kz)]$ (ω is the radial frequency and k the corresponding wavenumber) field dependences for modes traveling in the positive z direction. The freespace wavelength is λ_0 and the freespace wavenumber is $k_0 = 2\pi/\lambda_0$. The transfer matrix technique [11–13] is used in order to find the dispersion

equation for the guided modes of this planar waveguide. The resulting dispersion equation is [Eqs. (105) and (106)]

$$F(\beta) = F(k_0 N_{eff}) = \frac{\gamma_s}{a_s} m_{22} + \frac{\gamma_c}{a_c} m_{11} - \frac{\gamma_c \gamma_s}{a_c a_s} m_{12} - m_{21} = 0, \quad (187)$$

where $a_c = a_s = 1$ for TE polarization, and $a_c = n_c^2$, $a_s = n_s^2$ for TM polarization and m_{ij} ($i, j = 1, 2$), γ_s, γ_c are defined in Sec. 11. The electric and magnetic field phasors are of the form $A(x) \exp(-j\beta z)$. The propagation constant β can be in general complex (for lossy or active waveguides) and can be written in the form $\beta = k_0 N_{eff} = k_0(N_{eff,r} + jN_{eff,i})$ where $N_{eff,r}$ is the real part and $N_{eff,i}$ is the imaginary part of the effective index of a mode, respectively. For lossy waveguides and for propagation in the positive z direction $N_{eff,r} > 0$ and $N_{eff,i} < 0$ while for active waveguides $N_{eff,r} > 0$ and $N_{eff,i} > 0$. Of course, for lossless waveguides, $N_{eff,i} = 0$ for all guided modes. At the branch-cut points $N_{eff} = n_c$ and $N_{eff} = n_s$ it is important to remain in the same Riemann surface when N_{eff} (and correspondingly β) are varying in the complex plane. This can be accomplished by selecting the sign of γ_c and γ_s in such a way in order to satisfy the conditions $\text{Re}\{\gamma_c\} > 0$ and $\text{Re}\{\gamma_s\} > 0$ (where $\text{Re}\{*\}$ denotes real part) that guarantee confinement within the guiding layers.

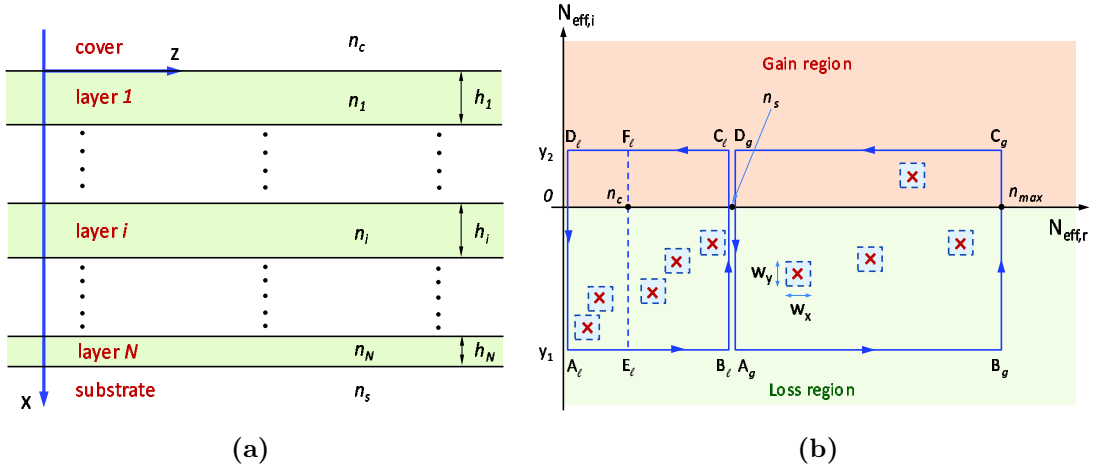


Figure 35: (a) The geometry of a general planar multilayer waveguide. The z -axis is the propagation axis and the x axis is the confinement axis. The i -th layer has a refractive index n_i and a height of h_i . The number of layers is N . The substrate and the cover refractive indices are n_s and n_c , respectively. (b) Zero search in the complex domain for guided (lossless, lossy, or active) and leaky waveguide modes.

The case of leaky modes requires a little more attention. The leaky modes are a compact representation of the radiation field of the waveguide [58, 59]. Therefore, even for lossless

waveguides the leaky modes have complex propagation constants (with $N_{eff,i} < 0$) that represent the radiation of the electromagnetic field into the substrate or into the cover or into both regions. In order to remain in the same Riemann surface when N_{eff} (and β) are varying in the complex plane, the sign of γ_c and γ_s must be selected in such a way in order to satisfy the conditions [16, 20, 33]: (a) $\text{Re}\{\gamma_c\} > 0$ and $\text{Im}\{\gamma_s\} > 0$ (where $\text{Im}\{*\}$ denotes imaginary part) for confinement in the cover region and radiation into the substrate region; (b) $\text{Im}\{\gamma_c\} > 0$ and $\text{Im}\{\gamma_s\} > 0$ for radiation into both the cover and the substrate regions. These conditions are such if $n_s > n_c$. In the case of $n_c > n_s$ the roles of γ_s and γ_c as well as of substrate and cover regions in the previous conditions are interchanged.

The modes of the planar multilayer waveguide can be determined by applying either the APM (Sec. 14.1), the ADR (Sec. 14.2), or the DFZEPE procedure (Sec. 14.3) to the function of Eq. (187). The contours $A_g B_g C_g D_g$ (for guided modes) and $A_\ell B_\ell C_\ell D_\ell$ for leaky modes are shown in Fig. 35b. The right part of $A_\ell B_\ell C_\ell D_\ell$ domain ($E_\ell B_\ell C_\ell F_\ell$) contains leaky modes radiating into the substrate region (assuming $n_s > n_c$) but confined in the cover region, while the left part ($A_\ell E_\ell F_\ell D_\ell$) contains leaky modes radiating into both the substrate and the cover regions. The guided modes of lossy waveguides will lay in the bottom part ($N_{eff,i} < 0$) of the complex plane while the guided modes of active waveguides could lay either in the top part ($N_{eff,i} > 0$) (if they experience gain) or in the bottom part (if they experience loss). Similarly, the leaky modes will lay on the bottom half ($N_{eff,i} < 0$) but for the range $0 < N_{eff,r} < \max\{n_c, n_s\} = n_s$ (since it is assumed that $n_s > n_c$). All these are true for forward mode propagation ($z > 0$). The converged sub-rectangles (with exaggerated size) that enclose the location of the mode effective indices are also shown in Fig. 35b which will be the result of the DFZEPE subdivision process. The APM and ADR methods do not require any systematic zero enclosure process since their main result is the coefficients of an approximating polynomial or the eigenvalues of a generalized eigenvalue problem that involves orthogonal approximating polynomials, respectively. Then the estimated by the centers of the final sub-rectangles mode effective indices are inserted into the Müller's refinement process (with deflation) for the final determination of the mode complex effective indices. Similarly, the APM and ADR methods use their approximating zeros as estimates for the Müller's refinement process. It is true that the approximating zeros of the APM and ADR methods are usually closer than the approximations of the DFZEPE method but after the refinement process the final roots (modal propagation

constants) of all methods are identical. After the effective index of each mode is determined it is straightforward to calculate all modal field components using the transfer matrix technique presented in Sec. 11.

Furthermore, the APM method requires the derivative of the function. The derivative could be calculated numerically using Eq. (175) or can be determined analytically in the case of multilayer planar waveguides [14, 17]. Defining as before $\beta = k_0 N_{eff} = k_0(N_{eff,r} + jN_{eff,i})$ the function $f(\beta) = F(N_{eff})$ in Eq. (187). Then the derivative with respect to N_{eff} of the function is given by

$$\begin{aligned} \frac{dF}{dN_{eff}} = & \frac{1}{a_s} \frac{d\gamma_s}{dN_{eff}} m_{22} + \frac{\gamma_s}{a_s} \frac{dm_{22}}{dN_{eff}} + \frac{1}{a_c} \frac{d\gamma_c}{dN_{eff}} m_{11} + \frac{\gamma_c}{a_c} \frac{dm_{11}}{dN_{eff}} \\ & - \frac{\gamma_s}{a_s a_c} \frac{d\gamma_c}{dN_{eff}} m_{12} - \frac{\gamma_c}{a_s a_c} \frac{d\gamma_s}{dN_{eff}} m_{12} - \frac{\gamma_c \gamma_s}{a_c a_s} \frac{dm_{12}}{dN_{eff}} - \frac{dm_{21}}{dN_{eff}}. \end{aligned} \quad (188)$$

In the above equation the required derivatives are given by

$$\frac{d\gamma_c}{dN_{eff}} = k_0^2 \frac{N_{eff}}{\gamma_c}, \quad (189)$$

$$\frac{d\gamma_s}{dN_{eff}} = k_0^2 \frac{N_{eff}}{\gamma_s}, \quad (190)$$

$$\frac{d\tilde{M}}{dN_{eff}} = \sum_{i=1}^N \left(\frac{d\tilde{M}_i}{dN_{eff}} \prod_{\substack{j=1 \\ j \neq i}}^N \tilde{M}_j \right) \quad (191)$$

$$\frac{d\tilde{M}_i}{dN_{eff}} = k_0^2 N_{eff} \begin{bmatrix} \frac{h_i \sin(k_{xi} h_i)}{k_{xi}} & \frac{a_i h_i \cos(k_{xi} h_i)}{k_{xi}^2} - \frac{a_i \sin(k_{xi} h_i)}{k_{xi}^3} \\ -\frac{\sin(k_{xi} h_i)}{a_i k_{xi}} - \frac{h_i}{a_i} \cos(k_{xi} h_i) & \frac{h_i \sin(k_{xi} h_i)}{k_{xi}} \end{bmatrix}, \quad (192)$$

where $a_i = 1$ for *TE* modes and $a_i = n_i^2$ for *TM* modes. In the special case that a $k_{xi} = 0$ the corresponding \tilde{M}_i matrix is treated as constant in the previous equations and its derivative is not needed.

14.5 Results

Only waveguide structures that have been published in the literature will be tested to demonstrate the effectiveness of the APM, ADR, and DFZEPE methods. All these methods produced identical results. However, in this section the results were obtained using the DFZEPE method which is the simplest and does not require either the derivative of the function or any complex integrations. Specifically, the following cases will be tested: (a) A case of a lossless multilayer

waveguide [11, 16, 17, 38] where both guided and leaky modes will be determined. (b) A case of a slightly-lossy multilayer waveguide [11, 16, 17, 38] where the guided modes and leaky modes will be determined. (c) A case of an active waveguide [2, 38] from doped semiconductor materials where guided modes and sample leaky modes will be determined, and (d) A case of an ARROW waveguide [17, 38] where only leaky modes exist. The results of the DFZEPE method for the effective modal indices in all cases are identical to the published results within the numerical accuracy that is provided in the literature. Therefore, a deviation from the published results in terms of a percentage difference is meaningless. All the results were calculated on an Intel(R) Core(TM) i7-7700, 3.60 GHz processor. The computation times varied from 4-20 sec for most cases. The most time-demanding calculations were for the ARROW waveguide where the computation times were between 160-400 sec (results of Table 14).

14.5.1. Lossless Dielectric Waveguide

This multilayer lossless waveguide was presented initially in Ref. [11]. It is comprised of 4 film layers ($N = 4$) and its parameters are given in the caption of Table 11. To apply the DFZEPE method for the guided modes the following parameters were selected: $x_1 = 1.501 \gtrsim n_s$, $x_2 = 1.659 \lesssim n_{max} = \max_i \{n_i\}$, $y_1 = -0.25$, $y_2 = 0.20$, $w_x = w_y = 0.005$, $M = 250$, and $\varepsilon_F = 10^{-11}$ (this parameter checks the proximity of zero to contour segments and it is discussed in Sec. 14.3). The effective indices of the guided TE and TM modes are shown in Table 11. Observe that the imaginary part of the effective index of each guided mode is zero as it is expected since the waveguide is lossless.

In the same table are also shown the effective indices of TE and TM leaky modes that are radiating into the substrate but are confined in the cover. For these leaky modes the parameters of the DFZEPE method are $x_1 = 1.001 \gtrsim n_c$, $x_2 = 1.499 \lesssim n_s$, $y_1 = -0.25$, $y_2 = 0.20$, $w_x = w_y = 0.005$, $M = 250$, and $\varepsilon_F = 10^{-11}$. There are five TE and five TM leaky modes that radiate into the substrate only. If $x_1 = 0.001 \gtrsim 0$ and $y_1 = -5.0$ the DFZEPE method detects additional leaky modes that radiate into both the substrate and the cover regions. For latter case the DFZEPE method identified 25 additional TE and 26 additional TM leaky modes that radiate in both cover and substrate with effective indices $0 < N_{eff,r} < n_c = 1.0$ and $-5 < N_{eff,i} < 0$. Since these leaky modes have much higher losses (compare to the ones that radiate only into the substrate) are not of much interest. However, for completeness all TE and

Table 11: Effective indices of guided and leaky modes of a multilayer lossless waveguide from Ref. [11]
 $\lambda_0 = 0.6328 \mu\text{m}$, $n_c = 1.0$, $n_s = 1.45$, $n_1 = 1.66$, $n_2 = 1.53$, $n_3 = 1.60$, $n_4 = 1.66$, $h_1 = h_2 = h_3 = h_4 = 0.5 \mu\text{m}$.

Mode	$N_{eff} = N_{eff,r} + jN_{eff,i}$	Mode	$N_{eff} = N_{eff,r} + jN_{eff,i}$
Guided Modes			
TE_0	$1.62272868 + j0$	TM_0	$1.62003132 + j0$
TE_1	$1.60527569 + j0$	TM_1	$1.59478848 + j0$
TE_2	$1.55713615 + j0$	TM_2	$1.55498069 + j0$
TE_3	$1.50358711 + j0$	TM_3	$1.50181780 + j0$
Leaky Modes - Substrate Radiating			
TE_4	$1.46185664 - j0.00715587$	TM_4	$1.45153498 - j0.01192359$
TE_5	$1.38248922 - j0.01816588$	TM_5	$1.37066437 - j0.03014206$
TE_6	$1.28136443 - j0.03587739$	TM_6	$1.27373706 - j0.05679177$
TE_7	$1.14231446 - j0.05287607$	TM_7	$1.15731285 - j0.08757849$
TE_8	$1.00303702 - j0.07077094$	TM_8	$1.03695026 - j0.10307808$
Leaky Modes - Substrate/Cover Radiating			
TE_9	$0.80402477 - j0.15549191$	TM_9	$0.96341519 - j0.165250319488834$
TE_{10}	$0.49261437 - j0.33590355$	TM_{10}	$0.76239325 - j0.222733601859893$
TE_{11}	$0.29877905 - j0.69942867$	TM_{11}	$0.46058337 - j0.370232916081077$
TE_{12}	$0.25212085 - j1.00504264$	TM_{12}	$0.24771086 - j0.718279099245514$
TE_{13}	$0.25050946 - j4.95922057$	TM_{13}	$0.18839165 - j1.013610354108087$
TE_{14}	$0.24991236 - j4.79387973$	TM_{14}	$0.14364341 - j1.274982619497127$
TE_{15}	$0.24586267 - j4.62848156$	TM_{15}	$0.12685382 - j1.524936733594305$
TE_{16}	$0.24161573 - j4.45910352$	TM_{16}	$0.11859313 - j1.949587977521766$
TE_{17}	$0.24026651 - j4.28845411$	TM_{17}	$0.11837505 - j1.753757137213375$
TE_{18}	$0.23986335 - j4.12047459$	TM_{18}	$0.10254993 - j2.145238444255885$
TE_{19}	$0.23543581 - j3.95230034$	TM_{19}	$0.10232727 - j2.720059502187552$
TE_{20}	$0.23077652 - j3.77917269$	TM_{20}	$0.09894013 - j2.349610858911572$
TE_{21}	$0.22948257 - j3.43177029$	TM_{21}	$0.09740231 - j2.546721039915246$
TE_{22}	$0.22944624 - j3.60410128$	TM_{22}	$0.09539809 - j3.435254140965898$
TE_{23}	$0.22470439 - j3.25890969$	TM_{23}	$0.09169923 - j4.123746135758958$
TE_{24}	$0.22207063 - j1.26968361$	TM_{24}	$0.09054857 - j2.896630870912667$
TE_{25}	$0.21976462 - j2.71626746$	TM_{25}	$0.08946336 - j4.796978313580165$
TE_{26}	$0.21968734 - j3.07943141$	TM_{26}	$0.08924637 - j3.085873731780236$
TE_{27}	$0.21863277 - j2.89664559$	TM_{27}	$0.08914771 - j3.271127601096534$
TE_{28}	$0.21504559 - j1.94511532$	TM_{28}	$0.08509431 - j3.603072946747228$
TE_{29}	$0.21495125 - j2.53439786$	TM_{29}	$0.08488961 - j3.964394978515887$
TE_{30}	$0.21238769 - j1.74120004$	TM_{30}	$0.08452455 - j3.785071780722443$
TE_{31}	$0.21178594 - j1.51632839$	TM_{31}	$0.08236169 - j4.640441344235671$
TE_{32}	$0.21039529 - j2.14368451$	TM_{32}	$0.08209538 - j4.286677523338038$
TE_{33}	$0.21009037 - j2.34244177$	TM_{33}	$0.08181951 - j4.464558548926473$
		TM_{34}	$0.08025401 - j4.956855738939457$

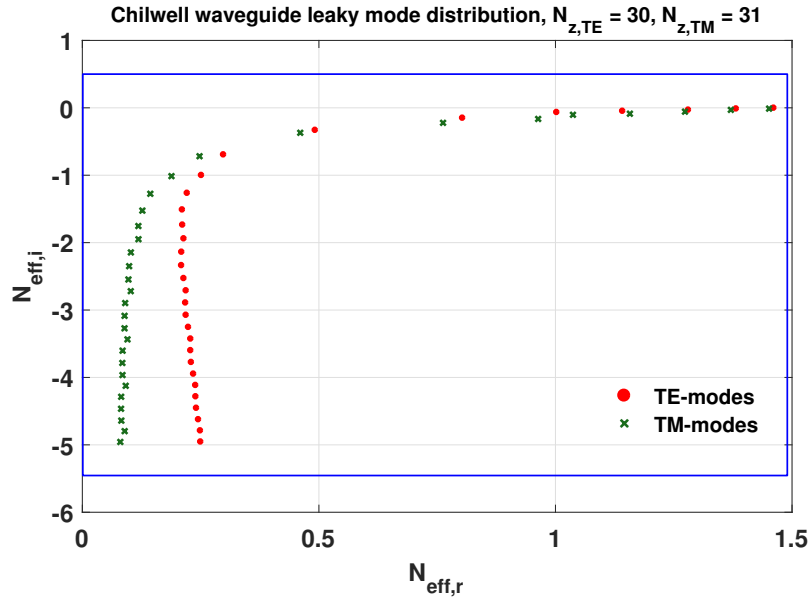


Figure 36: The leaky mode distribution for TE and TM modes for the Chilwell [11] waveguide of Table 11. There are 29 TE leaky modes and 30 TM modes. The parameters for the DFZEPE method where $x_1 = 0.001$, $x_2 = 1.499$, $y_1 = -0.25$ and $y_2 = 0.20$.

TM leaky modes (radiating into both cover and substrate) are shown in Table 11. A subset of these leaky modes (with $N_{eff,i} > -3.70$) has been published in Ref. [25]. If $y_1 < -5$ many more leaky modes are detected but since they all have high radiation losses may not be a good representation of the radiation field [58, 59] and therefore are not of much interest. A graphical representation of all the leaky modes of Table 11 are shown in Fig. 36. Furthermore, the electric field profile of the first leaky (substrate-radiating) TE (TE_4) mode, and the magnetic field profile of the first leaky (substrate-radiating) TM (TM_4) mode, are shown in Fig. 37. The radiation of these modes into the substrate is apparent. The increasing amplitude of the fields in the substrate region denotes the radiation of the modes and is a characteristic of the leaky mode approximation (implying that power from previous z positions of the mode is arriving at higher points into the substrate).

14.5.2. Low-Loss Dielectric Waveguide

This is the same multilayer waveguide that was presented in the previous section (Sec. 14.5.1) with the addition of some small loss in two of its film layers [11]. The parameters of the structure are given in the caption of Table 12. In order to apply the DFZEPE method for the guided

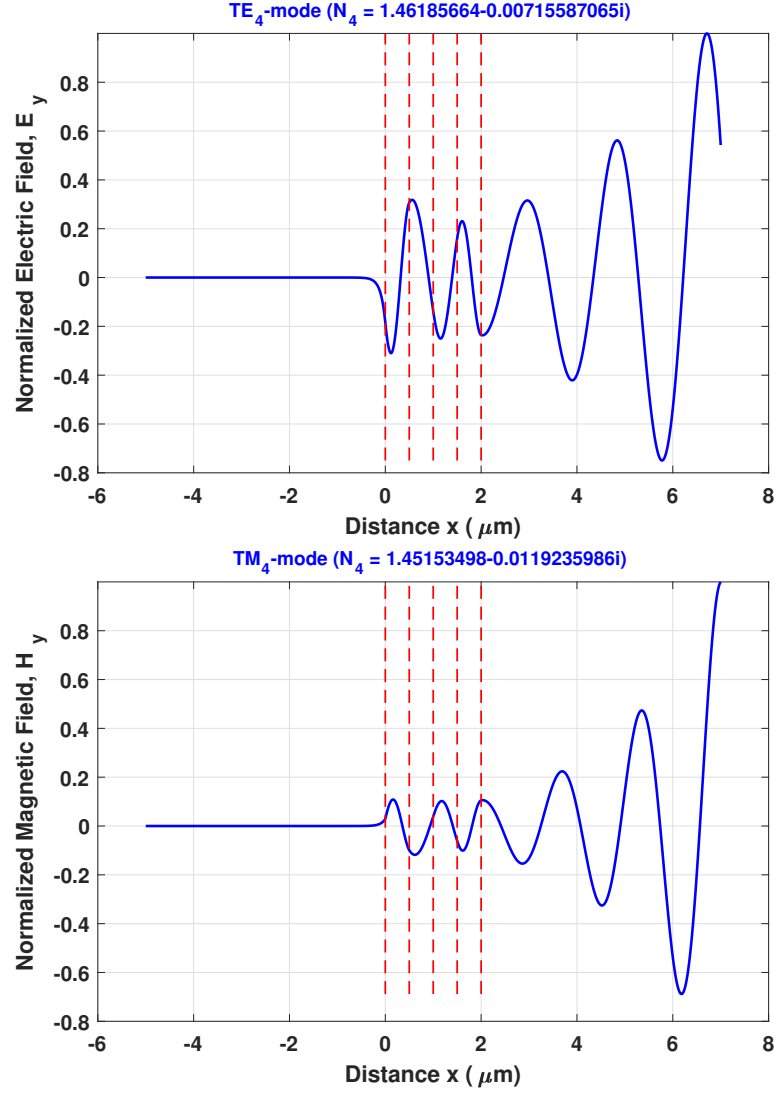


Figure 37: The electric field profile of the first leaky (substrate-radiating) TE (TE_4) mode, and the magnetic field profile of the first leaky (substrate-radiating) TM (TM_4) mode, for the Chilwell [11] waveguide of Table 11. Observe the radiation into the substrate region as well as the increasing amplitude of the radiation field.

modes the same parameters with the ones used in Sec. 14.5.1 were used. The effective indices of the guided TE and TM modes are shown also in Table 12. Observe that now all modes have negative imaginary part of their effective index due to the lossy layers of the waveguide.

Table 12: Effective indices of guided and leaky modes of a multilayer small-loss waveguide from Ref. [11] $\lambda_0 = 0.6328 \mu\text{m}$, $n_c = 1.0$, $n_s = 1.45$, $n_1 = 1.66 - j1.66 \times 10^{-4}$, $n_2 = 1.53 - j1.53 \times 10^{-4}$, $n_3 = 1.60$, $n_4 = 1.66$, $h_1 = h_2 = h_3 = h_4 = 0.5 \mu\text{m}$.

Mode	$N_{eff} = N_{eff,r} + jN_{eff,i} \times 10^{+4}$	Mode	$N_{eff} = N_{eff,r} + jN_{eff,i} \times 10^{+4}$
Guided Modes			
TE_0	$1.62272868 - j0.00673727$	TM_0	$1.62003131 - j0.00892759$
TE_1	$1.60527569 - j1.66244285$	TM_1	$1.59478847 - j1.65565266$
TE_2	$1.55713612 - j0.20880097$	TM_2	$1.55498066 - j0.23704828$
TE_3	$1.50358696 - j0.55032495$	TM_3	$1.50181764 - j0.42530043$
Leaky Modes - Substrate Radiating			
TE_4	$1.46185448 - j0.00726710$	TM_4	$1.45153751 - j0.01202887$
TE_5	$1.38249997 - j0.01827662$	TM_5	$1.37068384 - j0.03024261$
TE_6	$1.28137151 - j0.03596266$	TM_6	$1.27375077 - j0.05687731$
TE_7	$1.14233026 - j0.05299360$	TM_7	$1.15732794 - j0.08766890$
TE_8	$1.00303470 - j0.07087449$	TM_8	$1.03694118 - j0.10316486$

Even though for this lossy structure leaky modes were not reported in the literature, they are included here for completeness. Specifically, the leaky modes that radiate only into the substrate are shown in Table 12. It can be observed that the real parts of the leaky mode effective indices are very close to the ones appearing for the lossless waveguide in Table 11. However, their imaginary parts of their effective indices are in magnitude slightly higher since they represent both the loss due to absorption in the lossy layers as well as the loss due to radiation into the substrate.

14.5.3. Active Waveguide

Semiconductor materials with loss and gain are utilized for optical amplifiers or lasers. In this example case the structure proposed by Visser *et al.* [2] based on InP and InGaAsP system is analyzed using the DFZEPE method. The same case is also reported in Ref. [38]. It is comprised of 4 film layers with either loss or gain. The structure's parameters are defined in the caption of Table 13. To apply the DFZEPE method for the guided modes the following parameters were

selected: $x_1 = 3.17 \gtrsim n_s$, $x_2 = 3.59 \lesssim \max_i[\text{Re}\{n_i\}]$, $y_1 = -0.25$, $y_2 = 0.20$, $w_x = w_y = 0.005$, $M = 500$, and $\varepsilon_F = 10^{-11}$. The results are summarized in Table 13 for both TE and TM guided modes. Observe that TE_0 and TM_1 modes have positive imaginary part of their effective index which imply gain due to the active layer. These modes are mainly concentrated in the active layer (corresponding to InGaAsP, $n_3 = 3.60 + j0.002$, and $h_3 = 0.15 \mu\text{m}$). On the other hand, TM_0 mode has a negative imaginary part which imply loss (this is actually a plasmon mode mainly concentrated at the interface between the thin metal layer (corresponds to a thin gold layer with $n_1 = 0.18 - j10.2$ and thickness of $h_1 = 0.04 \mu\text{m}$) and the second film layer (which corresponds to InP with $n_2 = 3.16 - j0.0001$ and $h_2 = 1.0 \mu\text{m}$). These modes normalized electric and magnetic fields profiles are shown in Fig. 38.

Table 13: Effective indices of guided modes and first leaky mode of a multilayer active waveguide from Ref. [2,38] $\lambda_0 = 1.30 \mu\text{m}$, $n_c = 1.0$, $n_s = 3.16$, $n_1 = 0.18 - j10.2$, $n_2 = 3.16 - j0.0001$, $n_3 = 3.6 + j0.002$, $n_4 = 3.16 - j0.0001$, $h_1 = 0.04 \mu\text{m}$, $h_2 = 1.0 \mu\text{m}$, $h_3 = 0.15 \mu\text{m}$, $h_4 = 3 \mu\text{m}$.

Mode	$N_{eff} = N_{eff,r} + jN_{eff,i} \times 10^{+4}$		Mode	$N_{eff} = N_{eff,r} + jN_{eff,i} \times 10^{+4}$	
Guided Modes					
TE_0	$3.28088001 + j$	9.13918191	TM_0	$3.33449848 - j$	75.18872326
			TM_1	$3.24809848 + j$	5.46307013
First Leaky Mode					
TE_1	$3.13650356 - j$	376.20259075	TM_2	$3.13622674 - j$	377.75840427

For completeness a search for leaky modes into the substrate was performed. Using $x_1 = 1.001 \gtrsim n_c$, $x_2 = 3.159 \lesssim n_s$, while retaining the rest of the parameters of DFZEPE the same as in the case of guided modes 12 TE leaky modes and 11 TM leaky modes were determined using the specified values of y_1 and y_2 . All these leaky modes radiate into the substrate and they all have negative imaginary parts of their effective indices revealing radiation loss. For demonstration purposes, only the first leaky mode for each polarization is shown in Table 13.

14.5.4. Leaky Modes of an ARROW Waveguide

Antiresonant reflecting optical waveguides (ARROW) utilize antiresonant (multi-interference) reflection as the guiding mechanism instead of total internal reflection [5,6]. The structure under investigation was also used by Chen *et al.* [17] and Semwal and Rastogi [38]. It is

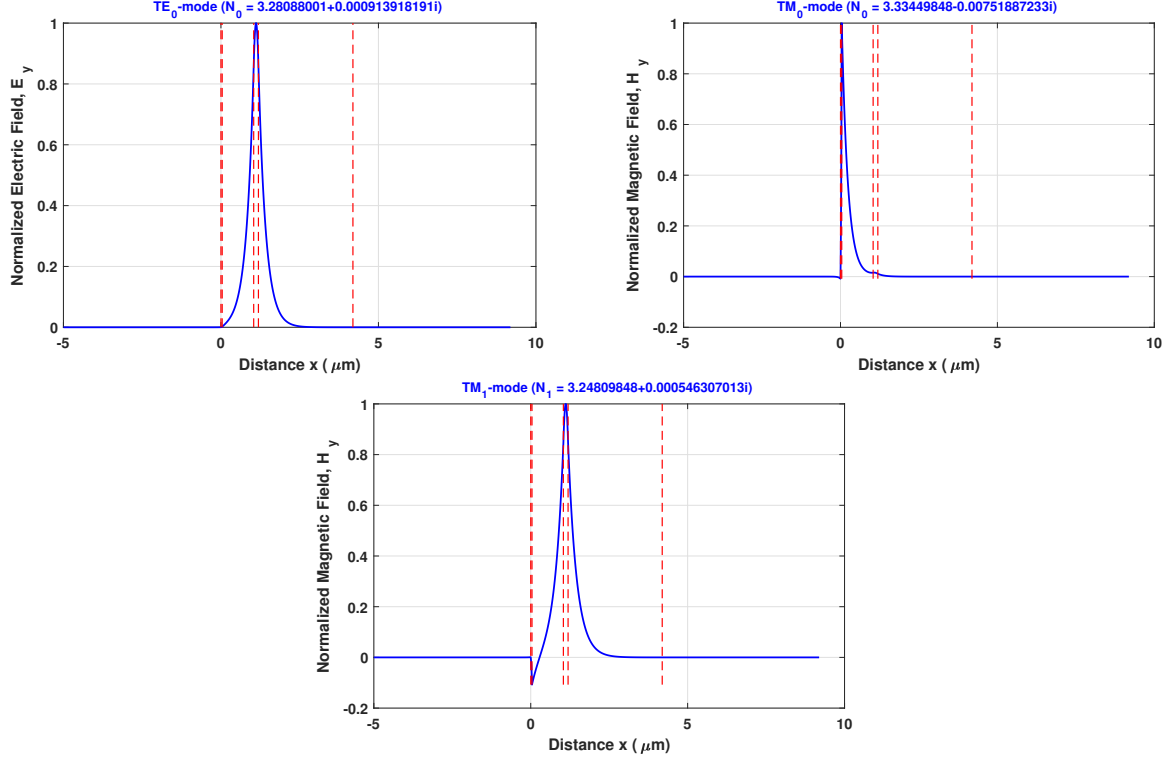


Figure 38: The active and lossy modes TE_0 , TM_0 , and TM_1 , for the Semwal [38] active waveguide of Table 13. Observe that the active modes are confined in the active layer (TE_0 and TM_1), while the lossy mode (plasmon mode) TM_0 is confined in the metal-semiconductor interface. The thin metal layer of $n_1 = 0.18 - j10.2$ and $h_1 = 0.04 \mu\text{m}$ is represented by the thicker dashed line on the left part of the figures.

comprised of $N = 9$ film layers with interchanging high and low refractive indices on top of a high index substrate (silicon). The structure parameters are defined in the caption of Table 14. Only leaky modes are supported by this structure. The DFZEPE method parameters were: $x_1 = 1.4501$, $x_2 = 1.499 \lesssim \max_i[\text{Re}\{n_i\}]$, $y_1 = -0.25$, $y_2 = 0.20$, $w_x = 2.5 \times 10^{-6}$, $w_y = 5 \times 10^{-4}$, $M = 500 - 1000$, and $\varepsilon_F = 10^{-4}$. The value of ε_F was much larger than previous cases since the dispersion function reaches values of the order of 10^{44} in the domain of interest. Therefore, ε_F represents a drop of the function of almost 48 orders. Furthermore, the large number of M is necessary due to the large phase variation that is observed in the vertical segments of the $ABCD$ contour. For demonstration, the phase variation of this ARROW waveguide is shown in Fig. 39.

Table 14: Effective indices of leaky modes of a multilayer ARROW waveguide from Ref. [38] $\lambda_0 = 0.6328 \mu\text{m}$, $n_c = 1.0$, $n_s = 3.50$, $n_1 = 1.46$, $n_2 = 1.50$, $n_3 = 1.46$, $n_4 = 1.50$, $n_5 = 1.46$, $n_6 = 1.50$, $n_7 = 1.46$, $n_8 = 1.50$, $n_9 = 1.46$, $h_1 = 2.00 \mu\text{m}$, $h_2 = 0.448 \mu\text{m}$, $h_3 = 4.00 \mu\text{m}$, $h_4 = 0.448 \mu\text{m}$, $h_5 = 2.00 \mu\text{m}$, $h_6 = 0.448 \mu\text{m}$, $h_7 = 4.00 \mu\text{m}$, $h_8 = 0.448 \mu\text{m}$, $h_9 = 2.00 \mu\text{m}$.

Mode	$N_{eff} = N_{eff,r} + jN_{eff,i} \times 10^{+4}$	Mode	$N_{eff} = N_{eff,r} + jN_{eff,i} \times 10^{+4}$
TE_0	$1.473925808 - j0.000000801$	TM_0	$1.473275805 - j0.000005809$
TE_1	$1.473697976 - j0.000017405$	TM_1	$1.473027205 - j0.032900856$
TE_2	$1.473696644 - j0.005452261$	TM_2	$1.473026854 - j0.000035036$
TE_3	$1.473459693 - j0.000001142$	TM_3	$1.472767027 - j0.000008508$
TE_4	$1.457920191 - j0.007106241$	TM_4	$1.457925423 - j0.045880488$
TE_5	$1.457791244 - j0.009053396$	TM_5	$1.457782773 - j0.057163274$
TE_6	$1.453780369 - j0.114698816$	TM_6	$1.453795448 - j0.645756672$
TE_7	$1.453045406 - j0.420121480$	TM_7	$1.452928429 - j2.555862981$
TE_8	$1.451864807 - j0.693651857$	TM_8	$1.451781628 - j4.567101184$
TE_9	$1.450269491 - j0.732515868$	TM_9	$1.450247659 - j4.357488809$

The results are summarized in Table 14 for both TE and TM guided modes. The leaky modes of the ARROW waveguides are numerically harder to find due to the close proximity of the zeros. In addition for the accurate computation of the accumulated phase Φ_t around each of the sub-rectangles a much higher number of points per segment M was used. The accumulated phase increases fast in the BC part of the rectangle contours while decreases fast in the DC parts of the contours (see Fig. 39). This means that it is critical to count all of the phase jumps in order to accurately determine the accumulated phase. This is the main reason for the

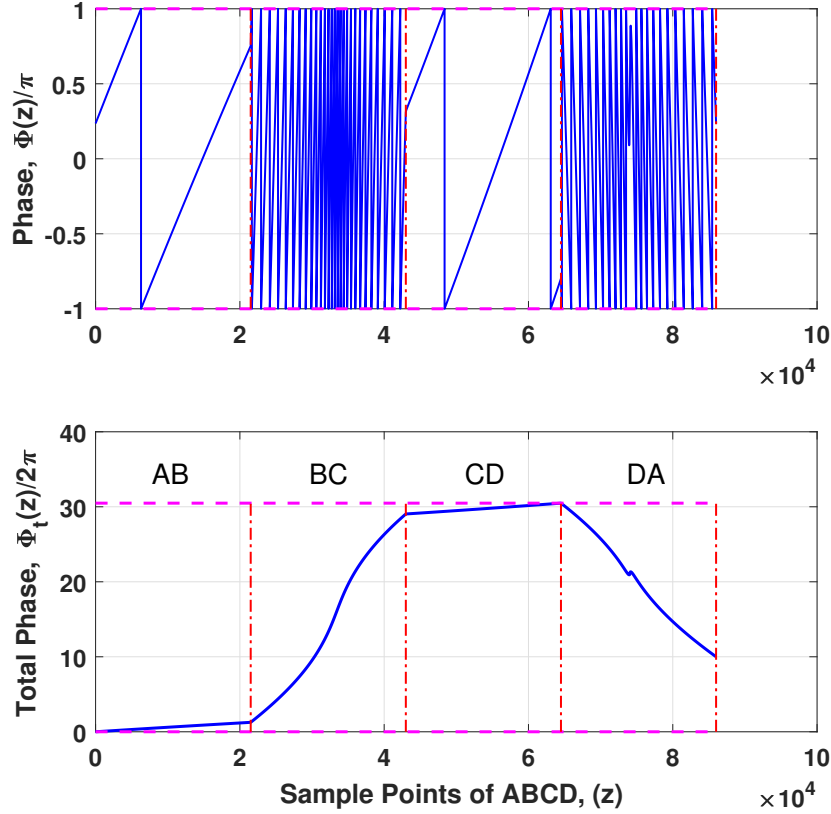


Figure 39: (a) The phase variation around $ABCD$, $\Phi(z) = \arg[f(z)]$, for the function of the ARROW waveguide and $x_1 = 1.4501$, $x_2 = 1.4999$, $y_1 = 0.20$, $y_2 = -0.25$, $M = 20000$ and (b) the total phase accumulation $\Phi_t(z) = \sum_{i=1}^I \arg[f(z_{i+1})/f(z_i)]$ where I is an index such $z_I = z$ and $\Phi_t(z_1) = 0$. The AB , BC , CD , and DA show the phase accumulation within each segment of the contour C . Obviously, the total accumulated phase is $10 \times 2\pi$ that reveals the existence of 10 leaky modes ($N_z = 10$) within the domain enclosed by C . Observe the very fast variation along the vertical segments of the contour.

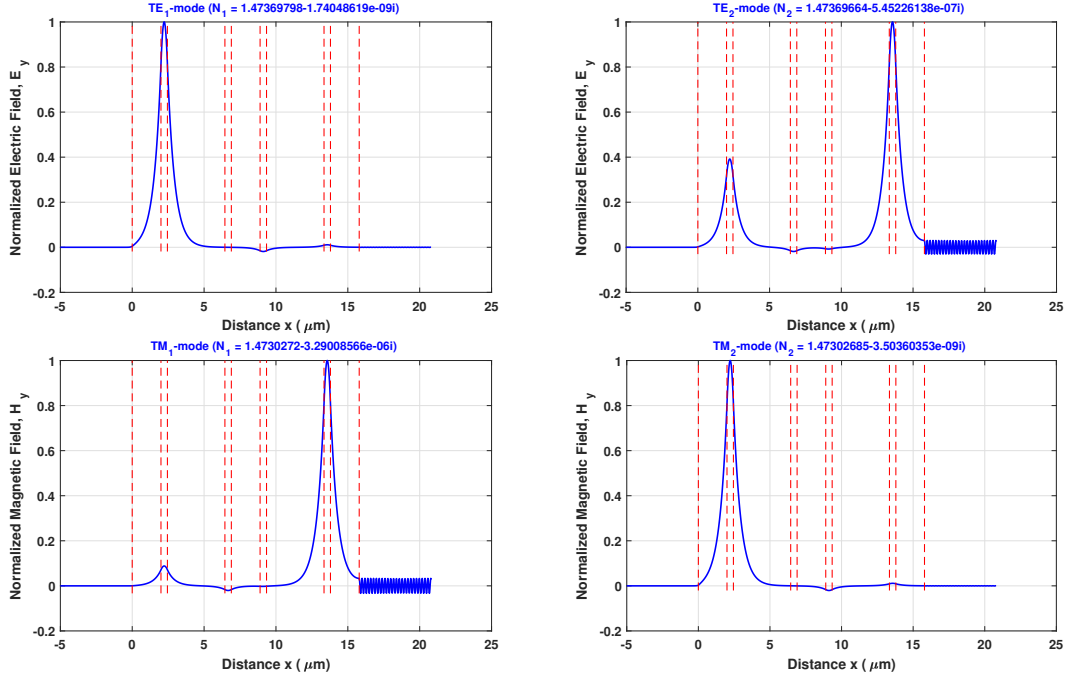


Figure 40: The close in effective index leaky modes TE_1 and TE_2 , as well TM_1 , and TM_2 , for the ARROW waveguide [38] of Table 14. Observe that the leaky modes that closer to the substrate radiate more and have larger(in magnitude) $N_{eff,i}$.

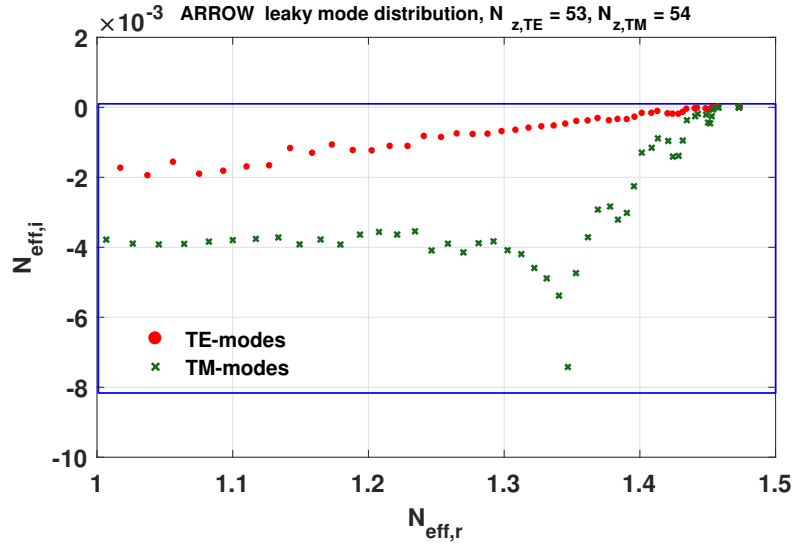


Figure 41: The leaky (substrate-radiating) mode distribution for TE and TM modes for the ARROW waveguide of Table 14. There are 53 TE leaky modes and 54 TM modes (the first ten of TE and TM are shown in the table). There are no leaky modes in the $1.5 - 3.5 N_{eff,r}$ interval. The parameters for the DFZEPE method where $x_1 = 1.001$, $x_2 = 3.499$, $y_1 = -0.25$ and $y_2 = 0.20$.

large number of points needed. From Table 14, for both polarizations, there are two clusters of modes. One that is near 1.45 (6 leaky modes) and one around 1.47 (4 leaky modes). Because the accumulated phase around a contour can accurately determine the number of modes the process of finding closely spaced modes can be performed in parts of the domain by changing appropriately the DFZEPE parameters. It is mentioned that only the 1.45 leaky mode cluster is published in Ref. [17] and only a subset of them (3 out the 6 modes from 1.45 cluster and 1 out of 4 modes from 1.47 cluster) in Ref. [38]. However, the proposed DFZEPE method can reliably determine all the leaky modes of the ARROW waveguide. For example, the total number of substrate-radiating leaky modes in the interval of $N_{eff,r}$ from 1 – 3.5 is 53 for TE and 54 for TM polarization, respectively. It is interesting to note that there are no leaky modes found in the interval $N_{eff,r} \in (1.5 - 3.5)$. Only the first ten are given with their effective indices in Table 14. However, the distribution of all the substrate-radiating leaky modes is shown in Fig. 41. For illustration purposes the normalized electric and magnetic fields profiles of the leaky modes TE_1 and TE_2 , as well as TM_1 , and TM_2 , are shown in Fig. 40, respectively. It is interesting to note that the modes that are closer to the substrate experience higher radiation losses as it is naturally expected due to their proximity to the high index silicon substrate.

References

- [1] A. W. Snyder and J. D. Love, *Optical Waveguide Theory*. London: Chapman & Hall, 1983.
- [2] T. D. Visser, H. Blok, and D. Lenstra, “Modal analysis of a planar waveguide with gain and losses,” *IEEE J. Quantum Electron.*, vol. 31, pp. 1803–1810, Oct. 1995.
- [3] M. Principe, G. Castaldi, M. Consales, A. Cusano, and V. Galdi, “Supersymmetry-inspired non-Hermitian optical couplers,” *Sci. Rep. UK*, vol. 5, pp. 8568:1–7, Feb. 2015.
- [4] E. G. Turitsyna, I. V. Shadrivov, and Y. S. Kivshar, “Guided modes in non-Hermitian optical waveguides,” *Phys. Rev. A*, vol. 96, p. 033824, Sept. 2017.
- [5] M. A. Duguay, Y. Kokubun, T. L. Koch, and L. Pfeiffer, “Antiresonant reflecting optical waveguides in SiO₂-Si multilayer structures,” *Appl. Phys. Lett.*, vol. 49, pp. 13–15, July 1986.
- [6] T. Baba and Y. Kokubun, “Dispersion and radiation loss characteristics of antiresonant reflecting optical waveguides - numerical results and analytical expressions,” *IEEE J. Quantum Electron.*, vol. 28, pp. 1689–1700, July 1992.
- [7] J. F. Traub, *Iterative Methods for the Solution of Equations*. Englewood Cliffs, NJ: Prentice-Hall, 2005.
- [8] S. D. Conte and C. de Boor, *Elementary Numerical Analysis: An Algorithmic Approach*, ch. The solution of nonlinear equations, pp. 72–127. International Series in Pure and Applied Mathematics, 3rd Ed., New York: McGraw-Hill Book Company, 1980.
- [9] D. E. Müller, “A method for solving algebraic equations using an automatic computer,” *Mathematical Tables and Other Aids to Computation*, vol. 10, no. 56, pp. 208–215, 1956.
- [10] L. M. Delves and J. N. Lyness, “A numerical method for locating the zeros of an analytic function,” *Math. Comput.*, vol. 21, pp. 543–560, 1967.
- [11] J. Chilwell and I. Hodhkinson, “Thin-films field-transfer matrix theory of planar multilayer waveguides and reflection from prism-loaded waveguide,” *J. Opt. Soc. Amer. A*, vol. 1, pp. 742–753, July 1984.

- [12] L. M. Walpita, “Solutions for planar optical waveguide equations by selecting zero elements in a characteristic matrix,” *J. Opt. Soc. Amer. A*, vol. 2, pp. 595–602, Apr. 1985.
- [13] T. Tamir, *Guided-Wave Optoelectronics*. Springer series in Electronics and Photonics 26, New York: Springer-Verlag, 2 ed., 1988.
- [14] R. E. Smith, S. N. Haude-Walter, and G. W. Forbes, “Numerical determination of planar waveguide modes using the analyticity of the dispersion equation,” *Opt. Lett.*, vol. 16, pp. 1316–1318, Sept. 1991.
- [15] R. E. Smith, S. N. Haude-Walter, and G. W. Forbes, “Mode determination for planar waveguide using the four-sheeted dispersion relation,” *IEEE J. Quantum Electron.*, vol. 28, pp. 1520–1526, June 1992.
- [16] E. Anemogiannis and E. N. Glytsis, “Multilayer waveguides: Efficient numerical analysis of general structures,” *J. Lightwave Technol.*, vol. 10, pp. 1344–1351, Oct. 1992.
- [17] C. Chen, P. Berini, D. Feng, S. Tanev, and V. P. Tzolov, “Efficient and accurate numerical analysis of multilayer planar optical waveguides in lossy anisotropic media,” *Opt. Express*, vol. 7, pp. 260–272, Oct. 2000.
- [18] M.-S. Kwon and S.-Y. Shin, “Simple and fast numerical analysis of multilayer waveguide modes,” *Opt. Commun.*, vol. 233, no. 1, pp. 119–126, 2004.
- [19] M.-S. Kwon, “A numerically stable analysis method for complex multilayer waveguides based on modified transfer-matrix equations,” *J. Lightwave Technol.*, vol. 27, pp. 4407–4414, Oct. 2009.
- [20] K. A. Michalski and M. M. Mustafa, “Numerically stable and reliable computation of electromagnetic modes in multilayered waveguides using the Cauchy integration method with automatic differentiation,” *IEEE Trans. Microwave Theory Tech.*, vol. 66, pp. 3981–3992, Sept. 2018.
- [21] J. D. Brown and R. V. Churchill, *Complex Variables and Applications*, ch. 4 and 7. New York: McGraw-Hill Education, 9th ed., 2014.

- [22] A. Ghatak, K. Thyagarajan, and M. Shenoy, "Numerical analysis of planar optical waveguides using matrix approach," *J. Lightwave Technol.*, vol. 5, pp. 660–667, May 1987.
- [23] E. Anemogiannis, E. N. Glytsis, and T. K. Gaylord, "Determination of guided and leaky modes in lossless and lossy planar multilayer optical waveguides: Reflection pole method and wavevector density method," *J. Lightwave Technol.*, vol. 17, pp. 929–941, May 1999.
- [24] K. Mehrany, S. Khorasani, and B. Rashidian, "Variational approach for extraction of eigenmodes in layered waveguides," *J. Opt. Soc. Amer. B*, vol. 19, pp. 1978–1981, Sept. 2002.
- [25] J. Petracek and K. Singh, "Determination of leaky modes in planar multilayer waveguides," *IEEE Phot. Technol. Lett.*, vol. 14, pp. 810–812, June 2002.
- [26] K. H. Schlereth and M. Tacke, "The complex propagation constant of multilayer waveguides: an algorithm for a personal computer," *IEEE J. Quantum Electron.*, vol. 26, pp. 627–630, Apr. 1990.
- [27] G. P. Zouros, "CCOMP: An efficient algorithm for complex roots computation of determinantal equations," *Comp. Phys. Comm.*, vol. 222, pp. 339–350, 2018.
- [28] M. Koshiba, H. Kumagami, and M. Suzuki, "Finite-element solution of planar arbitrarily anisotropic diffused optical waveguides," *J. Lightwave Technol.*, vol. 3, pp. 773–778, Aug. 1985.
- [29] M. Koshiba and H. Kumagami, "Theoretical study off silicon-clad planar diffused optical waveguides," *IEE Proc., Pt. J*, vol. 134, pp. 333–338, Dec. 1987.
- [30] E. Schweig and W. B. Bridges, "Computer analysis of dielectric waveguides: A finite-difference method," *IEEE Trans. Microwave Theory Tech.*, vol. 32, pp. 531–541, May 1984.
- [31] C. M. Arft and A. Knoesen, "An efficient finite-difference frequency-domain method including thin layers," *Microw. Opt. Techn. Lett.*, vol. 43, pp. 40–44, Oct. 2004.
- [32] S. D. Gedney, "An anisotropic perfectly matched layer absorbing media for the truncation of FDTD lattices," *IEEE Trans. Antennas & Prop.*, vol. 44, pp. 1630–1639, Dec. 1996.

- [33] E. Anemogiannis, E. N. Glytsis, and T. K. Gaylord, "Efficient solution of eigenvalue equations of optical waveguiding structures," *J. Lightwave Technol.*, vol. 12, pp. 2080–2084, Dec. 1994.
- [34] L. F. Abd-ellal, L. M. Delves, and J. K. Reid, "A numerical method for locating zeros and poles of a meromorphic function," in *Numerical Methods for Nonlinear Algebraic Equations* (P. Rabinowitz, ed.), ch. 3, pp. 47–59, London: Gordon and Breach, Science Publishers Ltd. 1970.
- [35] X. Ying and I. N. Katz, "A reliable argument principle algorithm to find the number of zeros of an analytic function in a bounded domain," *Numer. Math.*, vol. 53, pp. 143–163, Jan. 1988.
- [36] P. Kravanja and M. van Barel, "A derivative-free algorithm for computing zeros of analytic functions," *Computing*, vol. 63, pp. 69–91, 1999.
- [37] C. J. Gillan, A. Schuchinsky, and I. Spence, "Computing zeros of analytic functions in the complex plane without using derivatives," *Comp. Phys. Comm.*, vol. 175, pp. 304–313, Aug. 2006.
- [38] G. Semwal and V. Rastogi, "Derivative free method for computing modes of multilayer planar waveguides," *Opt. Quant. Electron.*, vol. 46, pp. 423–438, Mar. 2014.
- [39] E. N. Glytsis and E. Anemogiannis, "A simple derivative-free method for zero-extraction by phase-based enclosure for determination of complex propagation constants in planar multilayer waveguides," *Appl. Opt.*, pp. xxx–xxx, Dec. 2018.
- [40] X. Ying and I. N. Katz, "A simple reliable solver for all the roots of a nonlinear function in a given domain," *Computing*, vol. 41, pp. 317–333, Dec. 1989.
- [41] M. Dellnitz, O. Schutze, and Q. Zheng, "Locating all the zeros of an analytic function in one complex variable," *J. Comput. Appl. Math.*, vol. 138, no. 2, pp. 325–333, 2002.
- [42] P. Kowalczyk and W. Marynowski, "Efficient complex root tracing algorithm for propagation and radiation problems," *IEEE Trans. Antennas & Prop.*, vol. 65, pp. 2540–2546, May 2017.

- [43] P. Kowalczyk, “Complex root finding algorithm based on Delaunay triangulation,” vol. 41, pp. 19:1–19:13, June 2015.
- [44] M. J. Adams, *An introduction to Optical Waveguides*. Chichester: John Wiley & Sons, Inc., 1981.
- [45] D. L. Lee, *Electromagnetic Principles of Integrated Optics*. New York: John Wiley & Sons, Inc., 1986.
- [46] R. G. Hunsperger, *Integrated Optics: Theory and Technology*. Optical Sciences, New York: Springer-Verlag, 2009.
- [47] D. Marcuse, *Theory of Dielectric Optical Waveguides*. San Diego: Academic Press Inc., 1991.
- [48] C. Vassallo, *Optical Waveguide Concepts*. Optical Wave Sciences and Technology, New York: Elsevier Science Publishing Company, Inc., 1991.
- [49] A. B. Buckman, *Guide-Wave Photonics*. New York: Sauders College Publishing, 1992.
- [50] K. Okamoto, *Fundamentals of Optical Waveguides*. Burlington, MA: Academic Press as Inprint of Elsevier Inc., 2 ed., 2006.
- [51] K. Kawano and T. Kitoh, *Introduction to Optical Waveguide Analysis*. New York: John Wiley & Sons, 2001.
- [52] K. S. Yee, “Numerical solution of initial boundary value problems involving Maxwell’s equations in isotropic media,” *IEEE Trans. Antennas & Prop.*, vol. 14, pp. 302–307, May 1966.
- [53] R. C. Rumpf, “Simple implementation of arbitrary shaped total-field/scattered-field regions in finite-difference frequency-domain,” *Progress in Electromagnetics Research B*, vol. 36, pp. 221–248, 2012.
- [54] K.-P. Hwang and A. C. Cangellaris, “Effective permittivities for second-order accurate FDTD equations at dielectric interfaces,” *IEEE Microw. Wirel. Compon. Lett.*, vol. 11, pp. 158–160, Apr. 2001.

- [55] A. S. Householder, *The Numerical Treatment of a Single Nonlinear Equation*. International series in pure and applied mathematics, McGraw-Hill, 1970.
- [56] L. C. Botten, M. S. Craig, and R. C. McPhedran, “Complex zeros of analytic functions,” *Comp. Phys. Comm.*, vol. 29, no. 3, pp. 245–259, 1983.
- [57] L. C. Botten, M. S. Craig, and R. C. McPhedran, “Highly conducting lamellar diffraction gratings,” *Opt. Acta*, vol. 28, no. 8, pp. 1103–1106, 1981.
- [58] S.-L. Lee, Y. Chung, L. A. Coldren, and N. Dagli, “On leaky mode approximations for modal expansion in multilayer open waveguides,” *IEEE J. Quantum Electron.*, vol. 31, pp. 1790–1802, Oct. 1995.
- [59] R. E. Smith and S. N. Haude-Walter, “Leaky guiding in nontransparent waveguides,” *J. Opt. Soc. Amer. A*, vol. 12, pp. 715–724, Apr. 1995.