

Understanding leaky modes: slab waveguide revisited

Jonathan Hu* and Curtis R. Menyuk

University of Maryland Baltimore County, TRC 205A, 5200 Westland Boulevard,
Baltimore, Maryland 21227, USA

*Corresponding author: hu1@umbc.edu

Received July 25, 2008; revised November 5, 2008; accepted November 6, 2008;
posted November 6, 2008 (Doc. ID 99380); published January 29, 2009

Computational methods for determining the complex propagation constants of leaky waveguide modes have become so powerful and so readily available that it is possible to use these methods with little understanding of what they are calculating. We compare different computational methods for calculating the propagation constants of the leaky modes, focusing on the relatively simple context of a W-type slab waveguide. In a lossless medium with infinite transverse extent, a direct determination of the leaky mode by using mode matching is compared with complete mode decomposition. The mode matching method is analogous to the multipole method in two dimensions. We then compare these results with a simple finite-difference scheme in a transverse region with absorbing boundaries that is analogous to finite-difference or finite-element methods in two dimensions. While the physical meaning of the leaky modes in these different solution methods is different, they all predict a nearly identical evolution for an initial, nearly confined mode profile over a limited spatial region and a limited distance. Finally, we demonstrate that a waveguide that uses bandgap confinement with a central defect produces analogous results.

OCIS codes: 000.4430, 230.7400.

1. Introduction.	60
2. Wave Equation and Its Solutions in Slab Waveguides.	63
3. Nonleaky Waveguides.	69
3.1. Uniform Medium.	69
3.2. Three-Layer Waveguide.	72
4. W-Type Waveguide.	76
4.1. Leaky Mode Analysis.	77
4.2. Perturbation Analysis.	79
4.3. Physical Explanation of Exponential Decay.	79
4.4. Radiation Mode Decomposition.	80
4.5. Comparison and Analysis.	83
5. Bandgap Waveguide.	86
5.1. Eigenvalue Equation.	86
5.2. Alternative Solution Procedures.	89

6. Waveguide with Absorbing Layers.	92
6.1. Optimization of the Absorbing Layer.	93
6.2. Mode Decomposition and the Leaky Mode.	95
7. Answers to the Introductory Questions.	99
Appendix A: How We Did It and Software Link.	102
A.1. How We Did It.	102
A.2. Software Link.	103
A.3. For More Information.	103
Acknowledgments.	103
References.	103

Understanding leaky modes: slab waveguide revisited

Jonathan Hu and Curtis R. Menyuk

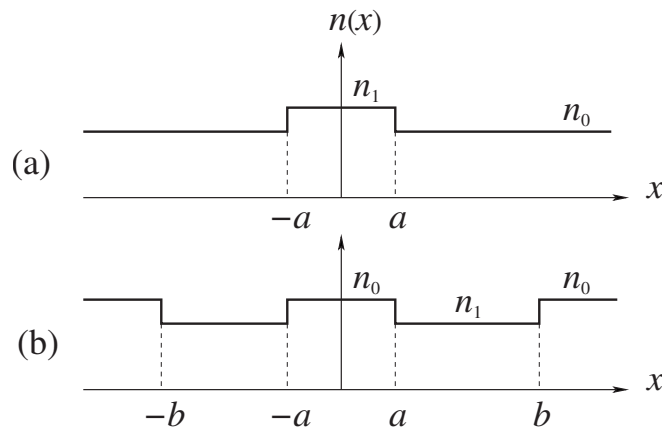
1. Introduction

In the 1960s and 1970s, the optical modes in solid-state waveguides were the subject of intensive study [1–4]. The waveguides that could be made then were fairly simple; a higher-index material was typically embedded in a lower-index material with a slab or rectangular waveguide in the case of semiconductor waveguides and a step or graded-index waveguide in the case of optical fibers. In the past decade, it has become possible to make highly complex optical waveguides in both semiconductors and glass, and, as a consequence, the study of optical waveguides has undergone a renaissance. A number of highly sophisticated computational algorithms have been developed to find the modes and their propagation constants inside these waveguides, including the finite-element method [5–7], the finite-difference method [6,8], the multipole method [9,10], the Galerkin method [11], and the plane-wave method [12,13]. In general, the modes in these modern waveguides may be evanescent or leaky even when there are no material losses. The finite-difference method, the finite-element method, and the multipole method allow one to determine the mode profiles and attenuation. Commercial finite-difference and finite-element solvers have in practice become so sophisticated and at the same time so easy to use that it is possible to determine mode profiles and their complex propagation constants with little understanding of what a complex propagation constant really means.

There is in fact a serious conceptual issue with leaky modes. Strictly speaking, they are not modes of the infinite lossless transverse waveguide. In principle, given any initial transverse profile at the beginning of a waveguide, one can decompose that profile into a finite number of guided modes and a continuum of radiation modes. These modes then propagate without attenuation in a lossless medium, and any attenuation of an initial profile that occurs must be due to spreading of the radiation modes. So, what is it that the finite-element, the multipole method, or other mode solvers are finding?

Leaky modes appear when an initial profile is nearly guided. In the three-layer slab waveguide shown in Fig. 1(a), in which $n_1 > n_0$, at least one guided mode exists. By contrast, in the five-layer W-type slab waveguide, shown in Fig. 1(b), with $n_1 < n_0$, only continuum radiation modes still exist. Nonetheless, when the width of the lower-index region becomes large, an optical beam is observed that looks much like the guided mode that would exist if the outer, higher-index layer were not present. This beam attenuates exponentially

Figure 1



The refractive index profile for (a) a three-layer waveguide and (b) a W-type waveguide.

with a rate that decreases rapidly as the width of the lower-index layer increases. How do we mathematically relate the radiation modes that compose this nearly guided waveguide to the guided mode that exists in the absence of the outer higher-index layers? Why is the attenuation in the W-type waveguide exponential?

Exponential attenuation is not the only possibility. If we send light into a medium with a single index of refraction n_0 , then the light diffracts and its intensity diminishes algebraically, rather than exponentially. This algebraic decay occurs because the light is not even partially confined. When the light comes from a point source, the algebraic decay is proportional to the square of the inverse distance from the source, since the light spreads in both transverse dimensions. When the light source is extended in one dimension, then the algebraic decay is proportional to the inverse of the distance, since light spreads in only one dimension. In a purely diffractive medium, the mode solvers will not find a single mode. In leaky mode waveguides, most of the energy will eventually leak out from the nearly guided waveguide, and the light will then diffract. When does diffraction dominate and the exponential decay of the leaky modes become algebraic?

A further issue with leaky modes is that at sufficiently large transverse distances from the mode center they increase exponentially. This behavior is clearly unphysical in the limit $x \rightarrow \pm\infty$. What is its origin? Is an exponential increase away from the transverse center ever visible?

The waveguides shown in Fig. 1 are quite simple. Is the behavior in more complex waveguides, in particular bandgap or photonic crystal waveguides, qualitatively similar? Can bandgap waveguides be studied by using the same methods as in the case of the simpler W-type waveguides?

A related issue is that one normally applies the finite-difference or the finite-element method by using a finite lossless region that is surrounded by absorbing layers or regions. The purpose of the absorbing regions is to avoid reflections and simulate outward-going boundary conditions. Thus, inside the lossless region, the solution approximates the solution with the same initial

conditions that one would find in a lossless medium of infinite transverse extent. However, the problem that one is solving, even in the limit as the discretization grows finer, is in fact fundamentally different from the problem of a lossless medium of infinite transverse extent and a lossless medium. Are the leaky modes real modes in this system? Do they still grow exponentially in the direction transverse to the propagation, at least in the lossless regions? Is there an analog to diffraction? What is the mode decomposition in this case? Finally—and perhaps most important—is the behavior that is predicted in this case the same as in the case of a lossless medium of infinite extent?

The answers to these questions may be found scattered throughout the technical literature, but not together and not in a form that is easily accessible to newcomers to the study of optical waveguides. The discussions rely on the steepest descent method for evaluating integrals and other asymptotic methods, whose physical meaning may not always be clear. In this introductory tutorial, our goal is to answer all these questions and to show the relationship among them. We will present the mathematical basis for these answers, but, at the same time, we will present pictures and animations whose goal is to illuminate what is happening both mathematically and physically.

We will focus here on simple one-dimensional slab waveguides, although two-dimensional waveguides are more important in practice. All the questions that we posed can be answered within the context of one-dimensional waveguides. No new concepts appear in two dimensions, although the mathematical complexity increases. Moreover, it is far simpler to illustrate the phenomena through pictures and animations, since one less dimension must be kept. Further discussion of the mathematical methods in two dimensions may be found in the textbooks by Marcuse [14] and by Snyder and Love [15].

For historical interest, we note that leaky modes were first described in the context of microwave waveguides, and both the textbooks by Marcuse and by Snyder and Love refer to a classic textbook by Collins [16], who refers in turn to reports by Barone [17,18] that appeared in 1956 and 1958. These reports are part of a series of reports from the Brooklyn Polytechnic Institute in which the leaky mode concept was first described. The first archival article reference is a brief report by Marcuvitz [19], who noted the close analogy to quantum-mechanical tunneling. He stated that leaky modes are not members of a complete set of orthogonal basis functions. He noted that this solution to the wave equation gives field representation in a center range with a complex propagation constant, but that the field becomes infinite at the infinite transverse spatial limit. A series of papers by Tamir, Oliner, and their colleagues, begun shortly thereafter, summarizes and categorizes the possible behavior of a leaky mode [20–25]. The first measurement of a leaky mode was carried out in 1961 by Cassedy and Cohn [26]. They confirmed the existence of a leaky wave due to a line current source above a grounded dielectric slab [26]. Hall and Yeh presented both theory and experiment for heteroepitaxial deposition of ZnS or ZnSe on GaSe, which is a three-layer waveguide [27]. In their experiment, the index of the substrate is higher than that of the center layer. Hence, only leaky modes can exist in this waveguide; no guided mode can exist. The same refractive index variation with position also occurs in waveguides with GaAlAs layers on GaAs substrates [28]. In 1975, the W-type, or depressed-cladding, slab waveguide was analyzed by Suematsu and Furuya [29]. At about the same time, the theory for W-type fiber waveguide

was developed by Kawakami and Nishida [30,31]. The original leaky wave analysis for a cylindrical waveguide was carried out in 1969 [32,33], which is 13 years after the leaky wave theory for a slab waveguide was first presented [17]. The attenuation coefficient of leaky modes has been obtained by solving the appropriate eigenvalue equation [27,30,31,33–35], by using Poynting’s vector theorem [36–39], and by using ray optics [39–42]. For the multilayer slab waveguide, the mode-matching method has been extensively used to find leaky modes [43–45].

We do not discuss the ray optics method, and, given its historical importance, this omission requires comment. Historically, ray optics was used in the analysis of optical waveguides to find reasonably accurate analytical approximations to Maxwell’s equations in contexts where exact analytical solutions could not be obtained. This approximation in principle requires the wavelength of the light in the waveguide to be small compared with the waveguide’s dimensions. What “small” means in this context is difficult to precisely define, and the attempt is rarely made. When the leakage is due to partial refraction, as in the case of a mode just below cutoff, ray optics can provide a reasonably accurate estimate of the leakage [15], along with a physical picture that some find compelling [15,46]. However, when the leakage is due to tunneling, as in the slab waveguide examples presented in this tutorial, there is no refraction, and in principle no leakage in the limit of small wavelengths. So, ray optics cannot be used without additional, *ad hoc* assumptions. Moreover, ray optics, in contrast to mode-matching methods and finite-difference or finite-element methods, is ill suited for use in computational mode solvers. That is particularly the case in the complex geometries that are becoming increasingly common. As a consequence, it has been little used in recent years for waveguide analysis.

The rest of this paper is organized as follows: in Section 2, we show the wave equation and its solutions in slab waveguides. Section 3 shows wave propagation and asymptotic analysis in nonleaky waveguides for a uniform medium and a three-layer slab waveguide. Section 4 shows the analysis for a W-type slab waveguide. Section 5 shows the analysis for a bandgap slab waveguide. Section 6 shows the analysis for a W-type slab waveguide with absorbing layers. In Section 7, we provide answers to the introductory questions.

2. Wave Equation and Its Solutions in Slab Waveguides

When polarization effects can be ignored, so that the electric field vector can be represented by a single component of the vector, then the field in optical waveguides can be approximately described by the scalar Helmholtz equation [6,47]. When we restrict our study to one transverse dimension, as shown in Fig. 1, this equation becomes

$$\frac{\partial^2 A(z,x)}{\partial z^2} + \frac{\partial^2 A(z,x)}{\partial x^2} + k_0^2 n^2(x) A(z,x) = 0, \quad (1)$$

where z and x denote the dimensions along and transverse to the waveguide, $A(z,x)$ is the complex electric field, normalized so that $|A(z,x)|^2$ is the power per unit length, k_0 is the propagation constant in vacuum, which is equal to the

ratio of the angular frequency to the speed of light ω/c , and $n(x)$ is the index of refraction. Here, we will assume that $n(x)$ is purely real, so that the waveguide has no material losses. Throughout this paper, we will refer to power per unit length simply as power. We will also use the negative carrier frequency convention in which all fields vary proportional to $\exp(-i\omega t)$, where $i = \sqrt{-1}$. While this convention is common among physicists, electrical engineers typically use the positive carrier frequency convention in which all fields vary proportional to $\exp(j\omega t)$, where $j = \sqrt{-1}$. The optics literature is split, and one occasionally finds the positive carrier frequency convention along with the use of $i = \sqrt{-1}$, as in [15]. The reader of the optics literature must always check which convention is being used, since the authors are often not explicit.

While Eq. (1) is only approximate in general, it becomes exact when the electric field only has a y component, as in the case of a TE mode. Moreover, the basic behavior of leaky modes remains unchanged when the full-vector Maxwell's equations are used. Likewise, the basic behavior is unchanged when, instead of one transverse dimension, two transverse dimensions are considered. If we search for solutions to Eq. (1) of the form

$$A(z,x) = E(x)\exp(i\beta z - i\omega t), \quad (2)$$

we find that $E(x)$ obeys the equation

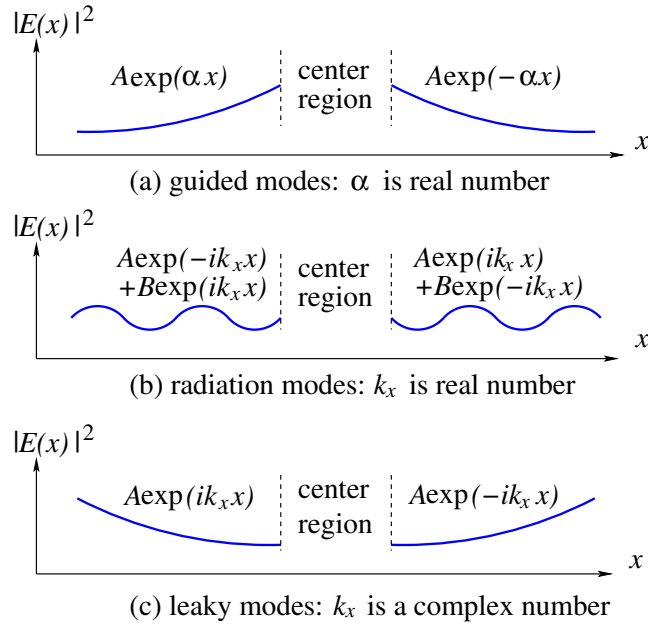
$$\frac{d^2 E(x)}{dx^2} + [k_0^2 n^2(x) - \beta^2] E(x) = 0, \quad (3)$$

where the eigenvalue β corresponds to the propagation constant in the z direction.

In slab waveguides with the property that $n(x)$ is equal to some constant value n_0 when $|x|$ is larger than some value x_0 , there are three qualitatively different types of solution that can appear, as shown in Fig. 2. Beyond $|x| = x_0$, where $n = n_0$, the field $E(x)$ must be expandable in the form $E(x) = E_0 \exp(ik_x x)$, where $k_x = \pm(k_0^2 n_0^2 - \beta^2)^{1/2}$. First, as shown in Fig. 2(a), if β is real and $k_0^2 n_0^2 - \beta^2 < 0$, then $k_x = \pm i\alpha$ is purely imaginary and the field decays exponentially when $|x| > a$ as $x \rightarrow \pm\infty$. These solutions are guided mode solutions. Second, as shown in Fig. 2(b), if β is real and $k_0^2 n_0^2 - \beta^2 > 0$, then k_x is purely real, and the field oscillates when $|x| > b$ as $x \rightarrow \pm\infty$. These solutions are radiation mode solutions. Only a finite number of β values, which may equal zero, may be found for which $k_0^2 n_0^2 - \beta^2 < 0$ and for which Eq. (3) has guided mode solutions. By contrast, radiation mode solutions may be found for any value for which $k_0^2 n_0^2 - \beta^2 > 0$ [48].

The guided modes and the radiation modes constitute a complete set, by which we mean that any physically reasonable initial condition $A(x, z=0)$ can be expanded as a superposition of guided modes and radiation modes [48,49]. Leaky mode solutions of the sort shown in Fig. 2(c) are not part of this complete set. Equation (3) is self-adjoint, which implies that it is possible to choose both the guided mode solutions and the radiation mode solutions that make up the complete set so that they are all real. In all the examples that we will consider, the index of refraction $n(x)$ is symmetric, i.e., $n(x) = n(-x)$. The asymmetric case is more complicated, but the basic behavior is unchanged

Figure 2



Comparison of (a) guided modes, (b) radiation modes, and (c) leaky modes in a one-dimensional waveguide. The solid curves show the mode power outside a center region, which depends on the details of the waveguide index variation.

[14,15]. In the symmetric case, it is possible to choose the solutions to Eq. (3) so that they are even or odd, and we may write

$$A(z=0,x) = \sum_{l=1}^N \tilde{A}_l E_l(x) + \frac{1}{\pi} \int_0^{\infty} \tilde{A}_e(k_x) E_e(k_x,x) dk_x + \frac{1}{\pi} \int_0^{\infty} \tilde{A}_o(k_x) E_o(k_x,x) dk_x, \quad (4)$$

where N is the number of guided mode solutions, which may equal zero, the $E_l(x)$ are the guided mode solutions to Eq. (3), the $E_e(k_x,x)$ are the even radiation modes, and the $E_o(k_x,x)$ are the odd radiation modes. The $E_l(x)$, $E_e(k_x,x)$, and $E_o(k_x,x)$ are all mutually orthogonal. If in addition we choose them to be orthonormal and real, so that

$$\int_{-\infty}^{\infty} E_l(x) E_m(x) dx = \delta_{lm}, \quad (5a)$$

$$\int_{-\infty}^{\infty} E_e(k_x,x) E_e(k'_x,x) dx = \pi \delta(k_x - k'_x), \quad (5b)$$

$$\int_{-\infty}^{\infty} E_o(k_x,x) E_o(k'_x,x) dx = \pi \delta(k_x - k'_x), \quad (5c)$$

then we find that

$$\tilde{A}_l = \int_{-\infty}^{\infty} A(z=0,x)E_l(x)dx, \quad (6a)$$

$$\tilde{A}_e(k_x) = \int_{-\infty}^{\infty} A(z=0,x)E_e(k_x,x)dx, \quad (6b)$$

$$\tilde{A}_o(k_x) = \int_{-\infty}^{\infty} A(z=0,x)E_o(k_x,x)dx. \quad (6c)$$

Equations (6b) and (6c) are analogous to cosine and sine transforms. It is possible and often useful to define an analog to the Fourier transform by writing $\tilde{A}(k_x) = \tilde{A}_e(k_x) - i\tilde{A}_o(k_x)$, $E(k_x,x) = E_e(k_x,x) + iE_o(k_x,x)$, $\tilde{A}(-k_x) = \tilde{A}^*(k_x)$, and $E(-k_x,x) = E^*(k_x,x)$. We now find that Eq. (4) becomes

$$A(z=0,x) = \sum_{l=1}^N \tilde{A}_l E_l(x) + \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{A}(k_x) E(k_x,x) dk_x, \quad (7)$$

and Eqs. (6b) and (6c) become

$$\tilde{A}(k_x) = \int_{-\infty}^{\infty} A(z=0,x)E^*(k_x,x)dx. \quad (8)$$

We note that $E_e(k_x,x) \propto \cos[k_x x + \varphi_e(k_x)]$ and $E_o(k_x,x) \propto \sin[k_x x + \varphi_o(k_x)]$ when $x > |x_0|$, with $E_e(k_x,x) = E_e(k_x,-x)$ and $E_o(k_x,x) = -E_o(k_x,-x)$ when $x < |x_0|$, where in general $\varphi_e(k_x) \neq \varphi_o(k_x)$, and we recall that x_0 is the value of $|x|$ beyond which $n(x) = n_0$. As a consequence $E(k_x,x)$ contains components proportional to both $\exp(ik_x x)$ and $\exp(-ik_x x)$ as $x \rightarrow \pm\infty$. If the initial field consists of purely forward-going waves in the z direction, as would be the case for a beam that is externally injected into a waveguide, then we may write the solution for all z as

$$A(z,x) = \sum_{l=1}^N \tilde{A}_l E_l(x) \exp(i\beta_l z) + \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{A}(k_x) E(k_x,x) \exp[i\beta(k_x)z] dk_x, \quad (9)$$

where we recall that $\beta(k_x) = (k_0^2 n_0^2 - k_x^2)^{1/2}$. When $k_x^2 > k_0^2 n_0^2$, $\beta(k_x) = i(k_x^2 - k_0^2 n_0^2)^{1/2}$, and these contributions to the solution are purely decaying. The energy in these components will be reflected backward.

While this formalism is completely general and allows one to determine in principle the evolution of a wave along a waveguide starting from any initial condition, the inclusion of radiation modes in the analysis is usually difficult, since one must integrate over the continuum of modes, and an exact solution is rarely available. When the problem is discretized for a numerical solution, the number of modes that must be included in the analysis is typically quite large. Moreover, this formalism often fails to capture the essence of the physics. The contribution from the continuous spectrum of radiation modes can be replaced approximately by a summation of discrete modes, which are called leaky modes [50]. If we consider for example the three-layer waveguide shown in Fig. 1(a), then it will have at least one guided mode. By contrast, if we consider the W-type waveguide shown in Fig. 1(b), then it has no guided mode solutions, and any solution must be expressible in terms of the radiation modes. At the same time, it is intuitively clear that when $b \gg a$, the W-type waveguide must have solutions that closely resemble the guided mode

solution to the waveguide in Fig. 1(a), and indeed such solutions can be observed to propagate with a slow exponential decay when an initial condition corresponding to the guided mode in the three-layer waveguide is launched in the W-type waveguide.

In fact, one can find solutions to Eq. (3) for the W-type waveguide that have nearly the same profiles in some range around $x=0$ as the profiles that we observed computationally when an initial condition that corresponds to the guided mode solution in the three-layer waveguide is launched in the W-type waveguide. These leaky mode solutions have complex propagation constants β , and the attenuation rate corresponds to what is observed computationally. However, the leaky mode solutions have the uncomfortable property that they grow exponentially as $x \rightarrow \pm\infty$, as shown schematically in Fig. 2(c), so that beyond some range of $|x|$ they no longer resemble the computationally observed solution. We will show below that this exponential growth is required by flux conservation. While these leaky modes can be very useful in practice, and, as noted in Section 1, they are often what is found by computational mode solvers, they are not normalizable and thus are not part of a complete set of basis functions on the infinite line in the usual sense. While mathematical work is ongoing to explore some unusual senses in which leaky modes are part of a complete set on the infinite line [46], this work is in its early stages and will not be discussed here.

In this discussion, we have assumed thus far that the index of refraction $n(x)$ becomes equal to the same value n_0 as $x \rightarrow \pm\infty$. In a slab waveguide, it is possible for these limits to be different. In that case, instead of just the three possibilities shown in Fig. 2, additional possibilities appear [20,22]. As a consequence, the complete mode decomposition, shown in Eq. (7), becomes more complicated. However, the basic behavior is unchanged by this added complexity. Any mode that grows exponentially as $x \rightarrow +\infty$ or as $x \rightarrow -\infty$ cannot be part of a complete set of basis functions, since they do not satisfy the boundary conditions.

In Sections 3–5 of this tutorial, we will be presenting computational solutions of Eq. (4) or Eq. (9). In numerical calculations, one has a finite spatial window. A boundary condition should be enforced that ensures that the discretized equations remain self-adjoint [48] and that leads to an orthogonal set of basis functions. In a uniform medium with periodic boundary conditions, this orthogonal set will be uniformly spaced in k_x and will be doubly degenerate, allowing the use of complex exponents as a basis set [48,51]. In effect, we are discretizing Eq. (9). However, in more complex waveguides, like the three-layer waveguide that we will consider in Section 3, the W-type waveguide that we will consider in Section 4, and the bandgap waveguides that we will consider in Section 5, it is no longer possible to choose the even and odd modes with self-adjoint boundary conditions so that they are degenerate. Hence, we must in effect discretize Eq. (4) with different choices of the k_x values for the even and the odd modes. In all cases except for the uniform medium, we use Neumann boundary conditions, which means that the derivatives of $E(k_x, x)$ are zero at the ends of the transverse spatial window. In all of our computational examples, we pick initial conditions that are initially symmetric around $x=0$, so that $\tilde{A}_o(k_x)=0$, and we will only need to determine $\tilde{A}_e(k_x)$ at the allowed values of k_x for the even modes. We keep anywhere from 1000 to 20,000 k_x modes, and the boundaries of our spatial window, x

$= \pm B$, are chosen so that B is hundreds to thousand of times larger than the initial beam width. We checked these values in every simulation to ensure that they are large enough, so that all our figures are unaffected by numerical errors. The number of modes and the spatial window must be large to ensure that we have good spatial resolution, while at the same time the field remains equal to zero at the spatial boundary over the entire propagation. We choose $\lambda = 1 \mu\text{m}$ in all cases except for the bandgap waveguides in Section 5. One can obtain results for different wavelengths by scaling the wavelength, the waveguide dimensions, and the mode fields, since Maxwell's equations are scale invariant [52].

When the medium becomes lossy beyond some value in $|x|$ as $x \rightarrow \pm\infty$, as is always assumed in computational mode solvers based on the finite-difference or finite-element methods, then the decomposition of an initial condition into a complete set of modes changes in important ways. First, all modes decay exponentially as $x \rightarrow \pm\infty$, and there are a countably infinite number of discrete modes. Hence, any initial condition may be written in the form

$$A(z=0, x) = \sum_{l=1}^{\infty} \tilde{A}_l(x) E_l(x), \quad (10)$$

so that

$$A(z, x) = \sum_{l=1}^{\infty} \tilde{A}_l(x) E_l(x) \exp(i\beta_l z). \quad (11)$$

Radiation modes and leaky modes are no longer present. However, when guided modes exist in the lossless waveguide, there are modes in the discrete set of Eq. (10) that closely resemble the guided modes. Likewise, when leaky modes exist, there are modes in this discrete set that closely resemble the leaky modes up to the values of $|x|$ where the waveguide becomes lossy. A second major difference from the lossless case is that Eq. (3) is no longer self-adjoint. As a consequence, both $E_l(x)$ and β_l become complex, and $E_l^*(x)$ is not a solution of Eq. (3). However, Eq. (3) is symmetric, by which we mean that

$$\int_{-\infty}^{\infty} f(x) \left\{ \frac{d^2}{dx^2} + [k_0^2 n(x) - \beta^2] \right\} g(x) dx = \int_{-\infty}^{\infty} g(x) \left\{ \frac{d^2}{dx^2} + [k_0^2 n(x) - \beta^2] \right\} f(x) dx \quad (12)$$

for any $f(x)$ and $g(x)$ for which the integrals exist. As a consequence, the $E_l(x)$ are self-dual, and we may determine the \tilde{A}_l by using the relation

$$\tilde{A}_l = \int_{-\infty}^{\infty} A(z=0, x) E_l(x) dx. \quad (13)$$

The appropriate normalization condition is

$$\int_{-\infty}^{\infty} E_l^2(x) dx = 1, \quad (14)$$

which sets the overall phase as well as the amplitude of $E_l(x)$. Because $E_l(x)$ is complex, one might worry that the integral in Eq. (14) could equal zero, in which case $E_l(x)$ would not be normalizable. However, that cannot happen, at

least for any finite discretization, because Eq. (3) is symmetric and hence normal [53].

In the remainder of the paper, we will elucidate the relationship between the leaky modes and the modes that form a complete basis set in the lossless guide, as described in Eqs. (4)–(9). We will also elucidate the relationship between the leaky modes and the modes in a lossy waveguide, as described in Eqs. (10), (11), (13), and (14). While these two descriptions are quite different, we will find that they yield nearly identical behavior for nearly confined light over some range of $|x|$ and $|z|$. We will also elucidate the relationship between the leaky modes and the modes that are found by computational mode solvers.

3. Nonleaky Waveguides

In this section, we consider two nonleaky waveguides that we will compare with the leaky waveguides. The first is a uniform waveguide in which simple diffraction occurs, and the intensity decays algebraically rather than exponentially as $z \rightarrow \infty$. The second is the three-layer waveguide shown in Fig. 1(a).

3.1. Uniform Medium

In this case, we only have radiation modes, and they are the standard Fourier modes. Referring to Eq. (9), we find $N=0$ and $E(k_x, x) = \exp(ik_x x)$. It follows that at any point z along the waveguide we may write

$$A(z, x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{A}(k_x) \exp[ik_x x + i\beta(k_x)z] dk_x, \quad (15)$$

where $\tilde{A}(k_x) = \int_{-\infty}^{\infty} A(z=0, x) \exp(-ik_x x) dx$.

In most cases, it is not possible to find an analytical solution to Eq. (15). An important exception is when the beam is initially Gaussian distributed so that $A(z=0, x) = A_0 \exp(-x^2/2w^2)$, where w is the initial beam width. In this case, we find

$$\tilde{A}(k_x) = A_0 \int_{-\infty}^{\infty} \exp(-x^2/2w^2) \exp(-ik_x x) dx = \sqrt{2\pi} w A_0 \exp(-k_x^2 w^2/2). \quad (16)$$

In the paraxial approximation, which often holds in optical waveguides, we may assume that $k_0 n_0 \gg k_x$, so that $\beta(k_x) \approx k_0 n_0 [1 - (k_x^2/2k_0^2 n_0^2)]$, where n_0 is the index of the uniform waveguide. Equation (9) then becomes

$$\begin{aligned} A(z, x) &= \frac{w A_0 \exp(ik_0 n_0 z)}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp \left[-\frac{1}{2} \left(w^2 + \frac{iz}{k_0 n_0} \right) k_x^2 + ik_x x \right] dk_x \\ &= \frac{w A_0 \exp(ik_0 n_0 z)}{(w^2 + iz/k_0 n_0)^{1/2}} \exp \left[\frac{-x^2}{2(w^2 + iz/k_0 n_0)} \right]. \end{aligned} \quad (17)$$

We see that the field remains Gaussian distributed, but the argument becomes complex, corresponding to the development of a curved phase front. The

intensity per unit length $|A(z,x)|^2$ becomes

$$|A(z,x)|^2 = \frac{w^2 A_0^2}{(w^4 + z^2/k_0^2 n_0^2)^{1/2}} \exp\left[\frac{-w^2 x^2}{(w^4 + z^2/k_0^2 n_0^2)}\right]. \quad (18)$$

When $z \gg k_0 n_0 w^2 = 2\pi w^2 n_0 / \lambda$, where λ is the vacuum wavelength, we find that the on-axis intensity per unit length decreases as z^{-1} and the beam width spreads proportional to z .

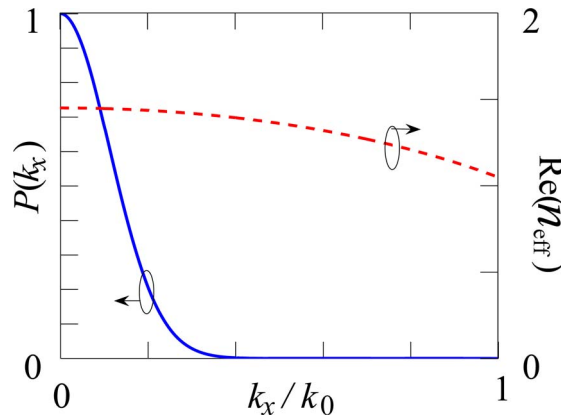
As a special case of the Gaussian beam, we may consider the case $w = \lambda$, corresponding to a relatively narrow beam. In Fig. 3, we plot $P(k_x) = |\tilde{A}(k_x)|^2 / \max[|\tilde{A}(k_x)|^2]$. As k_x increases, $P(k_x)$ decreases, and when it falls to 0.01, we find that $\beta(k_x) / \beta(k_x = 0) = 0.94$, so that the propagation is nearly paraxial. We also show for reference the real part of the effective index $\text{Re}(n_{\text{eff}})$ as a function of k_x . The effective index, $n_{\text{eff}} = \beta / k_0$, equals the ratio of the propagation constant to the wavenumber k_0 . In Fig. 4, we show a movie of the beam as it propagates through the medium. We solve Eq. (9) after discretization with the Gaussian input beam profile, $A(z=0, x) = \exp(-x^2/2w^2)$, setting $w = \lambda$. We then multiply $A(z, x)$ by $\exp(-i\omega t)$ and allow ωt to increase, where ω is the angular frequency of the input wave. We show the real part of the field in Fig. 4. We use a spatial transverse limit of $B = 500\lambda$, and 1000 k_x modes. As noted in Section 2, the k_x modes are uniformly spaced in these cases. We set $n_0 = 1.45$, which is a typical value for silica waveguides.

The same qualitative features are present with almost any initial beam shape when z becomes large, as may be shown by using the method of stationary phase [54]. We may rewrite Eq. (9) in the form

$$A(z,x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{A}(k_x) \exp[ik_x x + i\beta(k_x)z] dk_x = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{A}(k_x) \exp(i\varphi) dk_x. \quad (19)$$

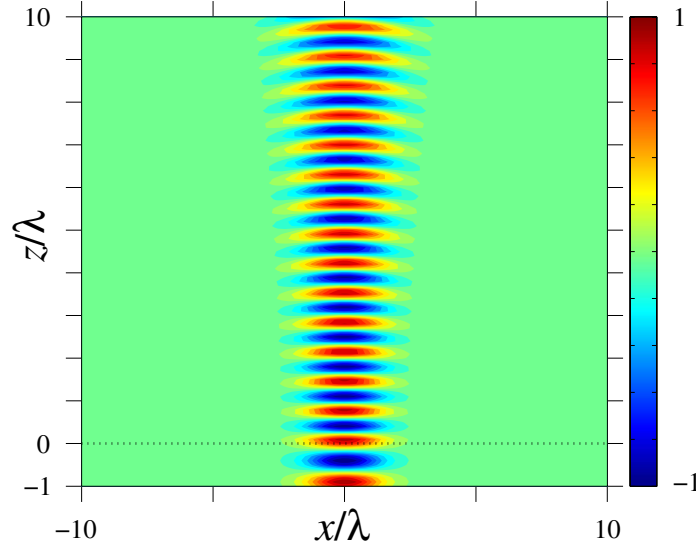
Expanding $\varphi = k_x x + \beta(k_x)z$ about the point $k_x = k_s$, we find $\varphi = \varphi^{(0)} + \varphi^{(1)}(k_x - k_s) + (1/2)\varphi^{(2)}(k_x - k_s)^2 + \text{higher-order terms}$, where $\varphi^{(0)} = k_s x + \beta(k_s)z$, $\varphi^{(1)} = x$

Figure 3



The normalized spectral power density $P(k_x) = |\tilde{A}(k_x)|^2 / \max[|\tilde{A}(k_x)|^2]$ and the real part of n_{eff} as a function of k_x .

Figure 4



Wave propagation in a uniform medium. Light is injected into a uniform medium at $z=0$. The movie ([Media 1](#)) shows the real part of the electric field.

$+\beta'(k_s)z$, $\varphi^{(2)}=\beta''(k_s)z$. We have written $\beta'(k_s)=d\beta/dk_x$ at $k_x=k_s$ and $\beta''(k_s)=d^2\beta/d^2k_x$ at $k_x=k_s$. From the expansion $\beta(k_x)=(k_0^2n_0^2-k_x^2)^{1/2}$, we find $\beta'(k_x)=-k_x/\beta(k_x)$. The stationary phase point satisfies the condition $0=\varphi^{(1)}=x-k_s z/\beta(k_s)$. It follows that $k_s=[x/(x^2+z^2)^{1/2}]k_0n_0$ and $\beta(k_s)=[z/(x^2+z^2)^{1/2}]k_0n_0$. Using this result, we find that $\varphi^{(0)}=(1+x^2/z^2)^{1/2}k_0n_0z$ and $\varphi^{(2)}=-(1+x^2/z^2)^{3/2}(z/k_0n_0)$. Equation (19) now becomes

$$A(z,x)=\frac{1}{2\pi}\exp\left[i\left(1+\frac{x^2}{z^2}\right)^{1/2}k_0n_0z\right]\times\int_{-\infty}^{\infty}\tilde{A}(k_x)\exp\left[-\frac{i}{2}\left(1+\frac{x^2}{z^2}\right)^{3/2}\frac{z}{k_0n_0}(k_x-k_s)^2\right]dk_x. \quad (20)$$

As z becomes larger, the oscillations about the stationary phase point $k_x=k_s$ become increasingly more rapid; so, to lowest order in an expansion in increasing powers of z , we may replace $\tilde{A}(k_x)$ with $\tilde{A}(k_s)$, and Eq. (20) becomes

$$A(z,x)=\frac{1}{2\pi}\exp\left[i\left(1+\frac{x^2}{z^2}\right)^{1/2}k_0n_0z\right]\tilde{A}\left[\frac{x}{(x^2+z^2)^{1/2}}k_0n_0\right]\times\int_{-\infty}^{\infty}\exp\left[-\frac{i}{2}\left(1+\frac{x^2}{z^2}\right)^{3/2}\frac{z}{k_0n_0}(k_x-k_s)^2\right]dk_x\\ =\frac{1-i}{2\sqrt{\pi}}\left(1+\frac{x^2}{z^2}\right)^{-3/4}\left(\frac{k_0n_0}{z}\right)^{1/2}\exp\left[i\left(1+\frac{x^2}{z^2}\right)^{1/2}k_0n_0z\right]\times\tilde{A}\left[\frac{x}{(x^2+z^2)^{1/2}}k_0n_0\right]. \quad (21)$$

In the paraxial limit, which is of greatest practical interest, Eq. (21) reduces to

$$A(z,x) = \frac{1-i}{2\sqrt{\pi}} \left(\frac{k_0 n_0}{z} \right)^{1/2} \exp(ik_0 n_0 z) \tilde{A} \left(\frac{x}{z}, -k_0 n_0 \right), \quad (22)$$

and the intensity per unit length becomes

$$|A(z,x)|^2 = \frac{1}{2\pi} \frac{k_0 n_0}{z} \left| \tilde{A} \left(\frac{x}{z}, -k_0 n_0 \right) \right|^2. \quad (23)$$

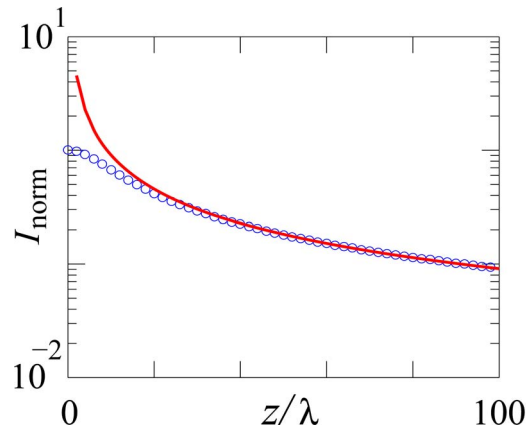
Hence, any initial condition ultimately resembles its initial Fourier transform, diminishes proportional to z^{-1} for any fixed ratio x/z , and has a width that is proportional to z .

Equation (20) is the first term in an asymptotic expansion of $A(z,x)$ in powers of $z^{-1/2}$. Corrections will be small as long as $z \gg k_0 n_0 w^2$, where w is the initial beam width. This limit, referred to as the Fresnel limit, is rapidly reached in most applications. In the case of the Gaussian beam, this limit can be obtained directly from the complex solution that is shown in Eqs. (17) and (18). In Fig. 5, we show $I_{\text{norm}} = |A(z,x=0)|^2 / |A(z=0,x=0)|^2$, the on-axis normalized intensity per unit length, for both the complete and the first-order asymptotic solutions in the case of the Gaussian beam. At the point $z/\lambda = 20$, the two solutions are nearly indistinguishable.

3.2. Three-Layer Waveguide

A simple waveguide that has guided modes as well as radiation modes is the three-layer symmetric waveguide, shown in Fig. 1(a). As long as $n_1 > n_0$, this waveguide has at least one guided mode [47]. We note that since the waveguide is symmetric, all modes must be even or odd, or must be the superposition of two degenerate mode in the case of radiation modes, where there is one even and one odd mode for each allowed value of β .

Figure 5



$I_{\text{norm}} = |A(z,x=0)|^2 / |A(z=0,x=0)|^2$ as a function of z/λ for a Gaussian beam. Blue circles represent the complete solution, while the red solid curve represents the lowest-order asymptotic approximation.

From the wave equation, Eq. (3), we find that the even guided modes are expressible as

$$E(x) = \begin{cases} C_1 \cos(k_c x) & |x| \leq a \\ C_0 \exp(-\alpha|x|) & a \leq |x| \end{cases}, \quad (24)$$

where $\alpha = [\beta^2(\alpha) - k_0^2 n_0^2]^{1/2}$ and $k_c = (k_0^2 n_1^2 - \beta^2)^{1/2} = [k_0^2(n_1^2 - n_0^2) - \alpha^2]^{1/2}$. Consequently, the guided modes must satisfy the condition $k_c^2 < k_0^2(n_1^2 - n_0^2)$. Any mode that satisfies Eq. (3) and its derivative must be continuous at the slab interfaces, from which we infer

$$C_1 \cos(k_c a) = C_0 \exp(-\alpha a), \quad (25a)$$

$$k_c C_1 \sin(k_c a) = \alpha C_0 \exp(-\alpha a), \quad (25b)$$

which yields the dispersion relation

$$k_c \tan(k_c a) = \alpha = [(k_0^2(n_1^2 - n_0^2) - k_c^2)^{1/2}]. \quad (26)$$

This transcendental relation always has at least one solution and will have at least two if $k_0 a(n_1^2 - n_0^2)^{1/2} > \pi$. The orthonormality condition, Eq. (5), becomes [47]

$$1 = \int_{-\infty}^{\infty} E^2(x) dx = C_1^2 \left[a + \frac{\sin(2k_c a)}{2k_c} + \frac{\cos^2(k_c a)}{\alpha} \right]. \quad (27)$$

The discussion for the odd guided modes is similar. When $x > 0$, we find

$$E(x) = \begin{cases} C_1 \sin(k_c x) & 0 \leq x \leq a \\ C_0 \exp(-\alpha x) & a \leq x \end{cases}, \quad (28)$$

and $E(x) = -E(-x)$ when $x < 0$. The dispersion relation becomes

$$-k_c \cot(k_c a) = \alpha = [k_0^2(n_1^2 - n_0^2) - k_c^2]^{1/2}, \quad (29)$$

which will have at least one solution as long as $k_0 a(n_1^2 - n_0^2)^{1/2} > \pi/2$. Conversely, there is only one even guided mode when $k_0 a(n_1^2 - n_0^2)^{1/2} < \pi/2$. Quite generally, it follows from Sturm–Liouville theory that the propagation numbers for the even and odd modes are interleaved and are nondegenerate [51].

When $x > 0$, the even radiation mode may be written as

$$E_e(k_x, x) = \begin{cases} C_{1e} \cos(k_c x) & 0 \leq x \leq a \\ C_{0e} \cos(k_x x + \varphi_e) & a \leq x \end{cases}, \quad (30)$$

and odd radiation mode solutions may be written as

$$E_o(k_x, x) = \begin{cases} C_{1o} \sin(k_c x) & 0 \leq x \leq a \\ C_{0o} \sin(k_x x + \varphi_o) & a \leq x \end{cases}, \quad (31)$$

with $E_e(k_x, x) = E_e(k_x, -x)$ and $E_o(k_x, x) = -E_o(k_x, -x)$ when $x < 0$. Any value of k_x is allowed, and we find $\beta(k_x) = (k_0^2 n_0^2 - k_x^2)^{1/2}$ and $k_c(k_x) = [k_0^2(n_1^2 - n_0^2) + k_x^2]^{1/2}$. Taking the combination $E(k_x, x) = E_e(k_x, x) + iE_o(k_x, x)$ at each value of k_x , we may write

$$E(k_x, x) = \begin{cases} C_1 \exp(ik_c x) & 0 \leq x \leq a \\ C_0 \exp(ik_x x) + D_0 \exp(-ik_x x) & a \leq x \end{cases}, \quad (32)$$

where $C_1 = C_{1e} = C_{1o}$, $C_0 = [C_{0e} \exp(i\varphi_e) + C_{0o} \exp(i\varphi_o)]/2$, and $D_0 = [D_{0e} \exp(-i\varphi_e) - C_{0o} \exp(-i\varphi_o)]/2$ when $x > 0$ and $E(k_x, x) = E^*(k_x, -x)$ when $x < 0$. From the continuity of $E(k_x, x)$ and its derivative at $x = a$, we find

$$C_1 \exp(ik_c a) = C_0 \exp(ik_x a) + D_0 \exp(-ik_x a), \quad (33a)$$

$$ik_c C_1 \exp(ik_c a) = ik_x C_0 \exp(ik_x a) - ik_x D_0 \exp(-ik_x a), \quad (33b)$$

which determines the ratios D_0/C_1 and C_0/C_1 . From the orthogonality condition, Eq. (5), we find $|C_0|^2 + |D_0|^2 = 1$. The solution given in Eq. (32) is like the solutions $E(k_x, x) = \exp(ik_x x)$ that we found in the case of the uniform medium. However, there is one important difference. When $k_x > 0$, the solution in the uniform medium propagates rightward, which means that it is purely outgoing as $x \rightarrow +\infty$ and is purely incoming as $x \rightarrow -\infty$. The opposite holds when $k_x < 0$. By contrast, the solution in Eq. (32) has both incoming and outgoing components as $x \rightarrow \pm\infty$.

We now consider as an example a medium in which $n_1 = 1.45$ and $n_0 = 0.96n_1 = 1.39$. The width in the center region is chosen so that $k_0 a (h_1^2 - n_0^2)^{1/2} = 1$, which implies $a = 0.39\lambda$. There is only one guided mode [47], and we consider a Gaussian input beam $A(z=0, x) = A_0 \exp(-x^2/2a^2)$, where a is half the width of the central layer, as shown in Fig. 1. In this example, 70% of the initial power is in the guided mode, and the rest is in the radiation modes. We discretize the radiation contribution to Eq. (4) using Neumann boundary conditions, as discussed in Section 2. For even modes, we insert the ratios D_0/C_1 and C_0/C_1 that we obtained from Eq. (33) into Eq. (32). We set $dE/dx = 0$ at $x = \pm B$, and we then obtain

$$\tan[k_x(B-a)] + \frac{k_c}{k_x} \tan(k_c a) = 0. \quad (34)$$

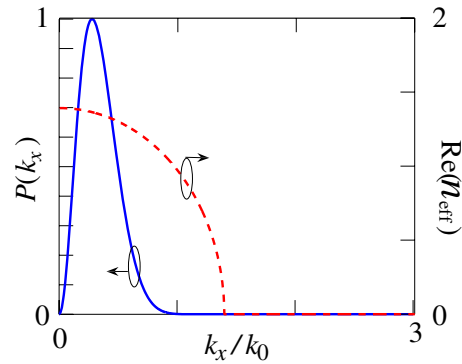
We note for completeness that the corresponding dispersion relation for the odd modes is

$$\tan[k_x(B-a)] - \frac{k_c}{k_x} \cot(k_c a) = 0, \quad (35)$$

although we will not need the odd modes in this example, since the initial condition is symmetric. We keep 2000 k_x modes, and we set $B = 500a$. The function $\tilde{A}(k_x) = \tilde{A}_e(k_x)$ is purely real in this case, and we show $|\tilde{A}(k_x)|^2$ normalized to its maximum for $k_x > 0$ in Fig. 6. We also show for reference the real part of the effective index $\text{Re}(n_{\text{eff}})$ as a function of k_x . Note that n_{eff} is a purely real number when $k_x < n_0 k_0$, and n_{eff} becomes a purely imaginary number when $k_x > n_0 k_0$. As in the case of the uniform guide, the radiation components for which $k_x > n_0 k_0$ do not propagate. They are exponentially damped, and their energy will be reflected back.

In Fig. 7, we show a movie of the wave propagation in the z direction. The waveguide parameters are the same as in Fig. 6. We keep 2000 k_x modes, and we set $B = 1000a$. As expected, the Gaussian input profile separates into a guided-wave component that propagates without diminishing and a diffractive

Figure 6

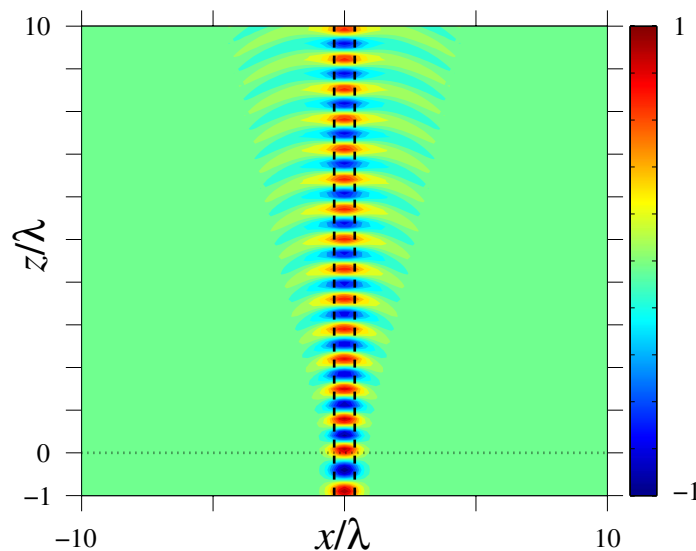


The normalized spectral power density $P(k_x) = |\tilde{A}(k_x)|^2 / \max[|\tilde{A}(k_x)|^2]$ and the $\text{Re}(n_{\text{eff}})$ as a function of k_x .

component that diminishes algebraically, not exponentially. The behavior of the diffractive component in the three-layer waveguide is qualitatively similar to the behavior of a beam in the uniform medium.

We find, as in the case of the uniform medium, that the solution spreads proportional to z and that the intensity per unit length diminishes proportional to z^{-1} . We may now once again use the method of stationary phase to obtain the contribution of the radiation modes in the Fresnel limit when $z \gg k_0 n_0 w^2$ for the three-layer waveguide. Focusing on the radiation field contribution A_{rad} to the total field, we rewrite this field at z in the form

Figure 7



Wave propagation in a three-layer waveguide. The light is injected into the waveguide at $z=0$. The movie ([Media 2](#)) shows the real part of the electric field. The black dashed lines indicate $x = \pm a$.

$$\begin{aligned}
A_{\text{rad}}(z,x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{A}(k_x) E(k_x,x) \exp[i\beta(k_x)z] dk_x \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{A}(k_x) E(k_x,x) \exp(i\varphi) dk_x.
\end{aligned} \tag{36}$$

There is no analytical solution of this equation. Hence, for $x=0$, we proceed by writing

$$A_{\text{rad}}(z,x=0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{A}(k_x) E(k_x,0) \exp(i\varphi) dk_x. \tag{37}$$

Expanding $\varphi = \beta(k_x)z$ about the point $k_x = k_s$, we find $\varphi = \varphi^{(0)} + \varphi^{(1)}(k_x - k_s) + (1/2)\varphi^{(2)}(k_x - k_s)^2 + \text{higher-order terms}$, where $\varphi^{(0)} = \beta(k_s)z$, $\varphi^{(1)} = \beta'(k_s)z$, $\varphi^{(2)} = \beta''(k_s)z$, and the stationary phase point satisfies the condition $0 = \varphi^{(1)} = -k_s z / \beta(k_s)$. It follows that $k_s = 0$ and $\beta(k_s) = k_0 n_0$. Using this result, we find that $\varphi^{(0)} = k_0 n_0 z$ and $\varphi^{(2)} = -(z/k_0 n_0)$. We also write $\tilde{A}(k_x)$ and $E(k_x, 0)$ as a Taylor series, from which we obtain $\tilde{A}(k_x) = \tilde{A}(k_s) + \tilde{A}'(k_x)(k_x - k_s) + \text{higher-order terms}$ and $E(k_x, 0) = E(k_s, 0) + E'(k_s, 0)(k_x - k_s) + \text{higher-order terms}$. Both $\tilde{A}(k_x)$ and $E(k_x, 0)$ are zero at $k_x = 0$. Equation (37) now becomes

$$\begin{aligned}
A_{\text{rad}}(z,x=0) &= \frac{1}{2\pi} \exp[ik_0 n_0 z] \int_{-\infty}^{\infty} \tilde{A}'(0) E'(0,0) k_x^2 \exp\left(-\frac{i}{2} \frac{z}{k_0 n_0} k_x^2\right) dk_x \\
&= \frac{1+i}{2\sqrt{\pi}} \exp(ik_0 n_0 z) \tilde{A}'(0) E'(0,0) \left(\frac{k_0 n_0}{z}\right)^{3/2}.
\end{aligned} \tag{38}$$

The power then becomes

$$|A_{\text{rad}}(z,x=0)|^2 = \frac{1}{2\pi} \left(\frac{k_0 n_0}{z}\right)^3 |\tilde{A}'(0) E'(0,0)|^2. \tag{39}$$

Hence, the power at $x=0$ diminishes proportional to z^{-3} . At other values of x/z , the power at $\tilde{A}(k_s) \neq 0$, and the power will diminish proportional to z^{-1} at large z , just as is the case in the uniform medium.

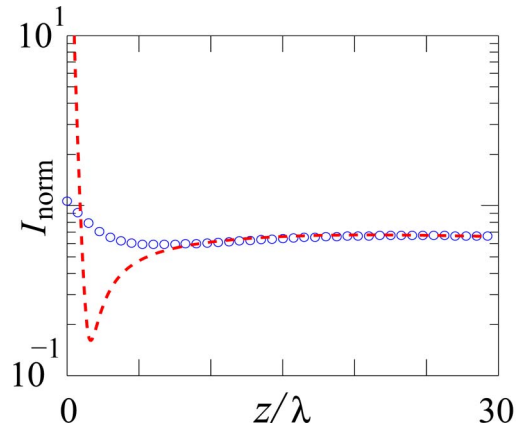
In Fig. 8, we show $I_{\text{norm}} = |A(z,x=0)|^2 / |A(z=0,x=0)|^2$ as a function of z/λ . The blue circles represent the power calculated by solving the propagation equation, Eq. (4), by using complete decomposition. The red dashed curve represents the sum of the fundamental mode propagation and the steepest descent estimated according to Eq. (39). The numerical integration and asymptotic analysis agree at $z > 10\lambda$.

The steepest descent approach that we have used to evaluate the radiation integrals is a powerful mathematical technique that will allow us in the following sections to separate the leaky mode contributions to the radiation integrals from the diffractive contributions.

4. W-Type Waveguide

We now turn to consideration of what is perhaps the simplest waveguide that has leaky modes—the W-type waveguide whose profile is shown in Fig. 1(b). A

Figure 8



$I_{\text{norm}} = |A(z, x=0)|^2 / |A(z=0, x=0)|^2$ as a function of z/λ for a Gaussian beam. The blue circles represent the power calculated by solving the propagation equation, Eq. (4), by using the complete decomposition, while the red dashed curve represents the sum of the asymptotic approximation from Eq. (39) and the guided mode contribution.

solution to Eq. (3) is purely outgoing when $x > b$ as $x \rightarrow +\infty$ if $E(x) \propto \exp(ik_x^+ x)$, where $\text{Re}(k_x^+) > 0$ and is purely incoming if $\text{Re}(k_x^+) < 0$. A solution to Eq. (3) is purely outgoing when $x < -b$ as $x \rightarrow -\infty$ if $E(x) \propto \exp(ik_x^- x)$, where $\text{Re}(k_x^-) < 0$. A leaky mode is defined as a solution to Eq. (3) that has no incoming components as $x \rightarrow \pm\infty$ and has an outgoing component either as $x \rightarrow +\infty$ or $x \rightarrow -\infty$ or both. In an asymmetric guide, it is possible for a leaky mode to be guided on one side and outgoing on the other, but in a symmetric guide, like that shown in Fig. 1(b), it must be purely outgoing as $x \rightarrow \pm\infty$. Just as a waveguide may not have guided mode solutions, it may not have leaky mode solutions. For example, the uniform waveguide that we considered in Subsection 3.1 only has solutions that are incoming as $x \rightarrow -\infty$ if they are outgoing as $x \rightarrow +\infty$ and vice versa, since $E(k_x, x) = \exp(ik_x x)$. When leaky mode solutions exist, their growth rate as $x \rightarrow \pm\infty$ must be small in order for them to be of practical interest.

4.1. Leaky Mode Analysis

We focus on the index profiles shown in Fig. 1(b) for which $n(x) = n_0$ for $|x| \leq a$, $n(x) = n_1$, $a \leq |x| \leq b$, and $n(x) = n_0$ for $b < |x|$. We will search for a solution that when $x > 0$ may be written as

$$E(x) = \begin{cases} C \cos k_x x & 0 \leq x \leq a \\ \frac{C \cos(k_x a)}{\cosh(\alpha a + \varphi)} \cosh(\alpha x + \varphi) & a \leq x \leq b \\ \frac{C \cos(k_x a) \cosh(\alpha b + \varphi)}{\cosh(\alpha a + \varphi)} \exp[ik_x(x - b)] & b \leq x \end{cases}, \quad (40)$$

where $k_x = (k_0^2 n_0^2 - \beta^2)^{1/2}$ and $\alpha = (\beta^2 - k_0^2 n_1^2)^{1/2}$. When $x < 0$, we set $E(-x) = E(x)$, so that the solution is even. We also demand $\text{Re}(k_x) > 0$, so that the solution

is outgoing as $x \rightarrow \pm\infty$. By matching the x derivatives of the fields at $x=a$ and $x=b$, we obtain

$$k_x \tan(k_x a) = -\alpha \tanh(\alpha a + \varphi), \quad (41a)$$

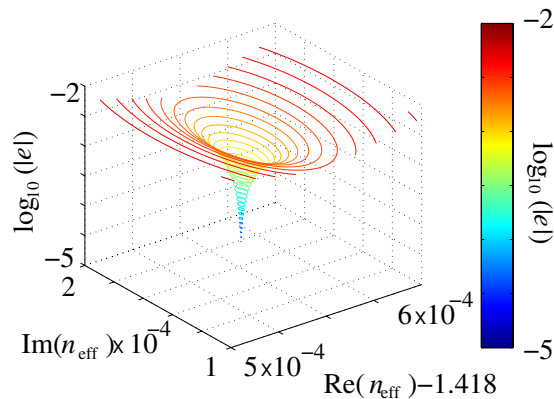
$$\alpha \tanh(\alpha b + \varphi) = ik_x. \quad (41b)$$

Eliminating the constant φ , we obtain the dispersion relation for the W-type waveguide [55],

$$\tan(k_x a) = \frac{\alpha}{k_x} \tanh \left[\tanh^{-1} \left(-i \frac{k_x}{\alpha} \right) + \alpha(b-a) \right]. \quad (42)$$

We now consider a specific numerical example that is closely related to the three-layer waveguide that we considered in Subsection 3.1. In this example, we set $n_0 = 1.45$ and $n_1 = 0.96n_0 = 1.39$, which is the same ratio as in the three-layer waveguide, except that the roles of n_0 and n_1 are interchanged, because it is now the higher-index material that is present when $x \rightarrow \infty$. The width in the center region is chosen so that $k_0 a (n_0^2 - n_1^2)^{1/2} = 1$ and $b/a = 5$, so that $a = 0.39\lambda$ and $b = 1.96\lambda$. With these choices, the waveguide is the same up to $|x| \leq b$ as in the case of the three-layer waveguide. In Fig. 9, we show a logarithmic plot of the absolute value of e , the difference between the left and the right sides of Eq. (42), as a function of real and imaginary parts of n_{eff} . There is evidently a pole in the plot, corresponding to a resonance of Eq. (42). To find its location accurately, we first find the derivative along the real axis and then determine its derivative in the imaginary direction using the Cauchy–Riemann relation for e [56]. We then use the Newton–Raphson method to find the point where $e \rightarrow 0$. In the case considered here, we obtain $n_{\text{eff}} = \beta/k_0 = 1.4185997 + 1.577 \times 10^{-4}i$. We have written the answer to eight significant figures, because the imaginary part is 4 orders of magnitude smaller than the real part. It is typical in practical problems for the imaginary part of β to be much smaller than the real part. However, obtaining this level of accuracy is not a problem in a modern-day 32 bit computer that has approximately 15 digits of accuracy for a double precision number.

Figure 9



Logarithm of the magnitude of the difference between the left- and the right-hand side of Eq. (42).

4.2. Perturbation Analysis

We now analyze this same problem by using perturbation theory. From a practical standpoint, it becomes increasingly difficult to obtain the imaginary part of β accurately as b/a increases, since the imaginary part of β rapidly decreases, and, when close to 15 digits of accuracy are needed, one will typically have large computational round-off errors. At the same time, this limit is precisely the one in which perturbation theory is expected to work well and can substantially decrease the computational time required to find $\text{Im}(\beta)$ [31]. From a conceptual standpoint, the leaky modes in a W-type waveguide, shown in Fig. 1(b), are expected to be a slight modification of the corresponding guided modes in the corresponding three-layer waveguide, shown in Fig. 1(a). Perturbation theory allows us to make this connection directly. We begin by writing $\beta \simeq \beta_0 + \Delta\beta$, where β_0 is the propagation constant for the three-layer waveguide. We then have

$$k_x = [(k_0 n_1)^2 - (\beta_0 + \Delta\beta)^2]^{1/2} \simeq k_{x0} - \beta_0 \Delta\beta / k_{x0}, \quad (43a)$$

$$\alpha = [(\beta_0 + \Delta\beta)^2 - (k_0 n_0)^2]^{1/2} \simeq \alpha_0 + \beta_0 \Delta\beta / \alpha_0, \quad (43b)$$

where k_{x0} and α_0 are the solution for the corresponding three-layer waveguide. We may now substitute these expressions into the dispersion relation, Eq. (42), and carry out a Taylor expansion in powers of $\Delta\beta$, keeping only the zero-order and first-order term. Solving for $\Delta\beta$, we obtain

$$\Delta\beta = \frac{2 \exp[-2\alpha_0(b-a)][ik_{x0}^2 \tan(k_{x0}a) + \alpha_0^2]}{\beta_0 M_W}, \quad (44)$$

where

$$\begin{aligned} M_W = & 2 + (\alpha_0/k_{x0} - k_{x0}/\alpha_0)[i + \tan(k_{x0}a)] + a(\alpha_0 - ik_{x0})[1 + \tan^2(k_{x0}a)] \\ & - 2i \tan(k_{x0}a) + \{4\alpha_0(b-a) - 4 + 2iak_{x0}[1 + \tan^2(k_{x0}a)] + 4i \tan(k_{x0}a) \\ & + 4i \tan(k_{x0}a)(b-a)k_{x0}/\alpha_0\} \exp[-2\alpha_0(b-a)]. \end{aligned} \quad (45)$$

We now find that $n_{\text{eff}} = \beta/k_0 = 1.4185997 + 1.567 \times 10^{-4}i$, and we see that the direct computation of $\text{Im}(\beta)$ and the results from perturbation theory agree to three decimal places.

4.3. Physical Explanation of Exponential Decay

The exponential decay in z implies that there is an exponential growth in $|x|$. Mathematically, Eqs. (43) and (44) imply that a positive imaginary component change in β , corresponding to decay in z , implies a negative imaginary change in k_x and growth in $|x|$. From a physical standpoint, we may understand this growth as a consequence of flux conservation.

The Helmholtz equation, Eq. (1), has a conserved time-averaged flux that may be written as

$$\mathbf{F}(z, x) = (1/2i\omega)[A^*(z, x) \nabla A(z, x) - A(z, x) \nabla A^*(z, x)], \quad (46)$$

where $\nabla(\cdot) = \hat{z}\partial/\partial z + \hat{x}\partial/\partial x$ is the transverse gradient operator. Using Eq. (1), we find that $\nabla \cdot \mathbf{F} = 0$. When $A(z, x)$ corresponds to a TE wave, then \mathbf{F} is

proportional to the Poynting flux. If $A(z, x) = E(x)\exp(i\beta z)$ is a waveguide mode, then F_z , the z component of \mathbf{F} , becomes $F_z = [\text{Re}(\beta) / \omega] |A(x)|^2$, which is positive definite. Referring to Fig. 10, flux conservation implies that the inward flux must equal the outward when integrated over the sides of the rectangles defined by the z values z_1 and z_2 and by any two values of x , for example, $x = \pm A$ or $x = \pm B$. When $x = \pm A$, the difference between the flux entering at $z = z_1$ and leaving at $z = z_2$ is proportional to $[1 - \exp(-2 \text{Im}(\beta)|z_1 - z_2|)] \int_{-A}^A |E(x)|^2 dx$, and when $x = \pm B$, the difference is proportional to $[1 - \exp(-2 \text{Im}(\beta)|z_1 - z_2|)] \int_{-B}^B |E(x)|^2 dx$. Since the second integral is larger than the first, the flux that exits from the sides at $|x| = \pm B$ must be larger than the flux that exits from the sides at $|x| = \pm A$, as long as the waves that exit from the sides are purely outgoing, as is the case for the leaky mode at sufficiently large $|x|$. This increase is only possible if $|E(x)|$ increases as well.

4.4. Radiation Mode Decomposition

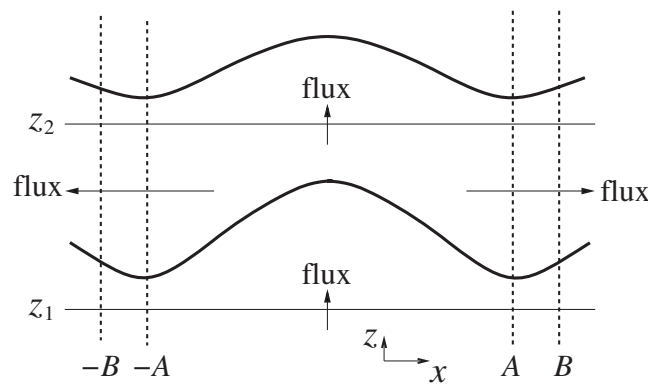
There are no guided modes in the W-type waveguide. As a consequence, it must be possible to express any square-integrable initial condition as a superposition of radiation modes, just as in the case of the uniform waveguide. At the same time, when $b \gg a$, we expect the behavior to resemble the behavior in the three-layer waveguide and, in particular, when the initial profile is close to the guided mode, as was the case with the Gaussian profile that we showed in Subsection 3.2, we expect the evolution to closely resemble that of the leaky mode in the central waveguide region. In this subsection, we will show how a superposition of the radiation modes leads to leaky behavior in a W-type waveguide.

When $x > 0$, the radiation modes may be written as [57]

$$E(x) = \begin{cases} C_2 \exp(ik_x x) & 0 \leq x \leq a \\ C_1 \exp(-\alpha x) + D_1 \exp(\alpha x) & a \leq x \leq b \\ C_0 \exp(-ik_x x) + D_0 \exp(ik_x x) & b \leq x \end{cases}, \quad (47)$$

where $\beta(k_x) = (k_0^2 n_0^2 - k_x^2)^{1/2}$ and $\alpha(k_x) = [k_0^2 (n_0^2 - n_1^2) - k_x^2]^{1/2}$. When $x < 0$, we find $E(k_x, x) = E^*(k_x, -x)$, and k_x can have any real value. Note that in contrast to

Figure 10



Schematic illustration of the flux flow.

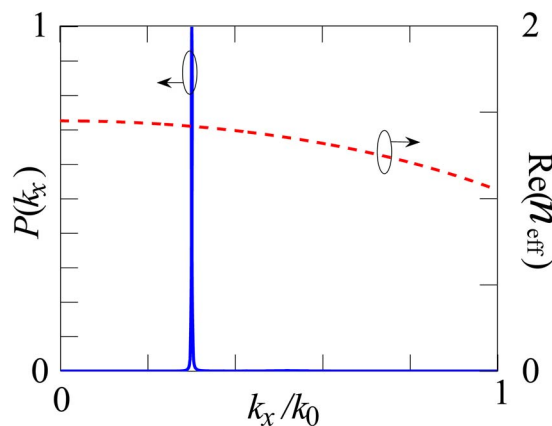
the leaky mode *ansatz* in Eq. (40), the radiation modes include both incoming and outgoing waves. Equation (47) has both incoming and outgoing components at $|x| > b$, in contrast to Eq. (47), which only has outgoing component at $|x| > b$. We may find the dispersion relation by matching $E(k_x, x)$ and its derivatives at $x=a$ and $x=b$, and the orthonormality condition becomes $|C_0|^2 + |D_0|^2 = 1$. The dispersion relation in this case is complicated, and we do not show it here.

In Fig. 11, we show the normalized coefficient $P(k_x) = |\tilde{A}(k_x)|^2 / \max[|\tilde{A}(k_x)|^2]$ and $\text{Re}(n_{\text{eff}})$ for the same Gaussian input beam that we considered in Subsection 3.2 in which $A(z=0, x) = \exp(-x^2/2a^2)$, where a is half of the center layer width, as shown in Fig. 1. We calculated $\tilde{A}(k_x)$ computationally, starting from the expression

$$\tilde{A}(k_x) = \int_{-\infty}^{\infty} A(z=0, x) E^*(k_x, x) dx, \quad (48)$$

using the decomposition procedure described in Section 2 with 2000 k_x modes and with $B=2000a$. We find that $P(k_x)$ is sharply peaked around a value of k_x that we denote k_r . This resonant behavior of $\tilde{A}(k_x)$ differs sharply from that of the uniform waveguide or the analogous three-layer waveguide, in which $\tilde{A}(k_x)$ varied smoothly. There is a smoothly varying portion of $\tilde{A}(k_x)$, in which $k_x^2 > k_0^2(n_0^2 - n_1^2)$, that corresponds to a diffractive contribution, analogous to the diffractive contribution in the three-layer waveguide. However, we find that as a result of the resonant behavior of $\tilde{A}(k_x)$, the integral will become dominated at an early stage of evolution by the behavior near $k_x = k_r$, rather than the stationary phase points of $\beta(k_x)$. There will also be a Lorentzian peak when $k_x = -k_r$, in addition to the Lorentzian peak at $k_x = k_r$. A close examination of $\tilde{A}(k_x)E(k_x, x=0)$ shows that the Lorentzian peaks that dominate its behavior have a Lorentzian (single-pole) shape, so that we may write to good approximation

Figure 11



The normalized spectral power density $P(k_x) = |\tilde{A}(k_x)|^2 / \max[|\tilde{A}(k_x)|^2]$ and the real part of effective index $\text{Re}(n_{\text{eff}})$ as a function of k_x/k_0 .

$$\tilde{A}(k_x)E(k_x, x=0) = -\frac{ik_r\tilde{A}(k_r)E(k_r, x=0)}{k_x - k_r - ik_i} + \frac{ik_i\tilde{A}(-k_r)E(-k_r, x=0)}{k_x + k_r + ik_i}. \quad (49)$$

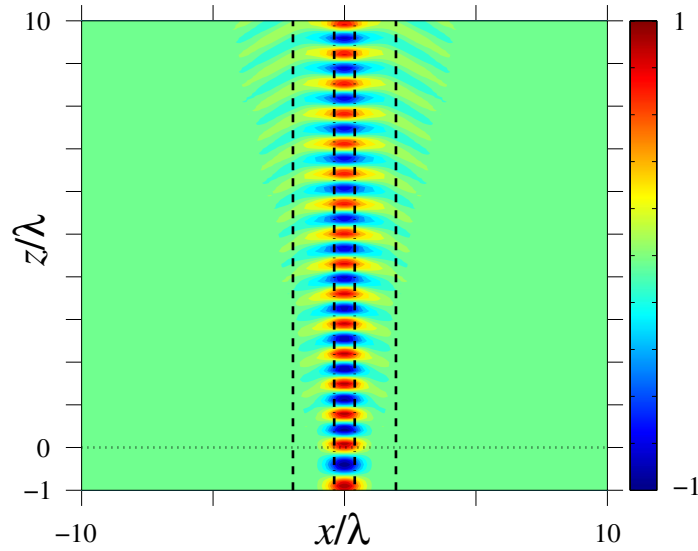
Lorentzian peaks like the ones in Eq. (49) *always* appear when leaky modes are present! These Lorentzian peaks and the corresponding poles in the complex k_x domain are the signature of a leaky mode. Just as $-i/(k_x - k_r - ik_i) + i/(k_x + k_r + ik_i)$ is the Fourier transform in the ordinary sense of $\exp[(ik_r - k_i)|x|]$ when $k_i > 0$, it is also the transform of this function in a generalized sense when $k_i < 0$, in which the contour of integration over k_x is deformed so that $[\text{Im}(k_x) - k_i]x < 0$ as $x \rightarrow +\infty$ and $[\text{Im}(k_x) + k_i]x < 0$ as $x \rightarrow -\infty$ [58]. Hence, we find that k_r corresponds to the oscillation wavenumber of the leaky mode when $|x| > b$ and is nearly equal to k_c in the corresponding waveguide, while $|k_i|$ corresponds to the leaky mode's transverse exponentiation factor when $|x| > b$. From Eq. (49), it also follows that $|k_i|$ corresponds to the width of the Lorentzian peaks in the k_x domain. When $b - a$ increases, we find that $|k_i|$ rapidly decreases, and the Lorentzian peaks tend toward δ functions, while k_r becomes exactly equal to k_c . In general, it must be the case that $|k_i| < k_r$ in order for a leaky mode to propagate for an observable distance. That is the case in the example discussed here, and we thus find the following equation by using the residue theorem [56]:

$$\begin{aligned} A(z, x=0) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{A}(k_x)E(k_x, x=0) \exp[i\beta(k_x)z] dk_x \\ &\simeq \frac{1}{2} k_i \tilde{A}(k_r)E(k_r, x=0) \exp[i\beta_r z - \beta_i z] + \text{c.c.}, \end{aligned} \quad (50)$$

where $\beta_r = \text{Re}[\beta(k_x = k_r + ik_i)] \simeq \beta(k_r)$ and $\beta_i = \text{Im}[\beta(k_x = k_r + ik_i)] \simeq -k_r k_i / \sqrt{k_0^2 n_0^2 - k_r^2} = k_i \partial \beta / \partial k_x|_{k_x = k_r}$. We find $\text{Re}(n_{\text{eff}}) = \text{Re}(\beta_r / k_0) = 1.4185984$ and $\text{Im}(n_{\text{eff}}) = \text{Im}(\beta_i / k_0) = 1.588 \times 10^{-4}$, which are close to the values that we found from the leaky mode analysis. Equation (50) shows the intuitively expected result that $A(z, x=0)$ is approximately proportional to $\tilde{A}(k_r)$, which is the overlap integral of $A(z=0, x)$ with the mode at $k_x = k_r$ that corresponds closely to the guided mode in the three-layer waveguide. We find that $\tilde{A}(k_x)$ and $E(k_x, x=0)$ have the same shape around $k_x = k_r$. Both are sharply peaked. Finally, $A(z, x=0)$ has an exponential decay rate in z that is inversely proportional to the width of the resonance in $\tilde{A}(k_x)$. In Fig. 12, we show a movie of the evolution of a Gaussian input beam in a W-type waveguide with $A(z=0, x) = \exp(-x^2/2a^2)$, just as in the three-layer waveguide simulation in Fig. 7. The waveguide parameters are the same that we used in Fig. 9, i.e., $n_0 = 1.45$, $n_1 = 0.96n_0$, $k_0 a(n_0^2 - n_1^2)^{1/2} = 1$, and $b/a = 5$. We keep 1000 k_x modes, and we set $B = 2000a$. After an initial transient in which a portion of the initial beam rapidly diffracts, the beam settles down into the shape of the leaky mode in the central region of waveguide $|x| < b$, after which the gradual exponential loss is visible.

This discussion assumes that the change in $\beta(k_x)z$ is much less than one over the Lorentzian linewidth. This assumption will eventually break down as z increases, and the beam evolution will no longer be exponential. We will discuss this point shortly.

Figure 12

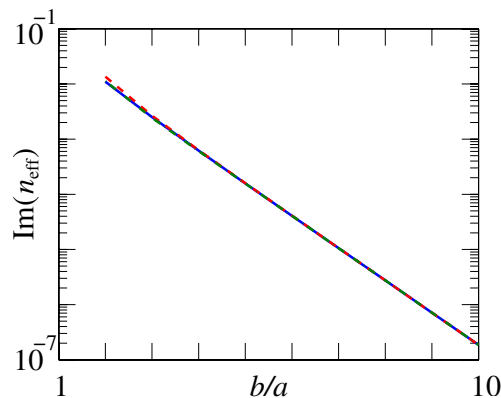


Wave propagation in a W-type waveguide. Light is injected into the W-type waveguide at $z=0$. The black dashed-dotted lines and black dashed lines indicate $x=\pm a$ and $x=\pm b$, respectively. The movie ([Media 3](#)) shows the real part of the electric field.

4.5. Comparison and Analysis

Figure 13 shows $\text{Im}(n_{\text{eff}})$ as a function of b/a . As the ratio b/a increases, all the methods that we have used for calculating β_l —the direct determination of the leaky mode solution, the perturbation method, and the determination from the radiation mode solutions—yield nearly identical results. The linear falloff on a logarithmic plot indicates that the imaginary part of n_{eff} decreases

Figure 13



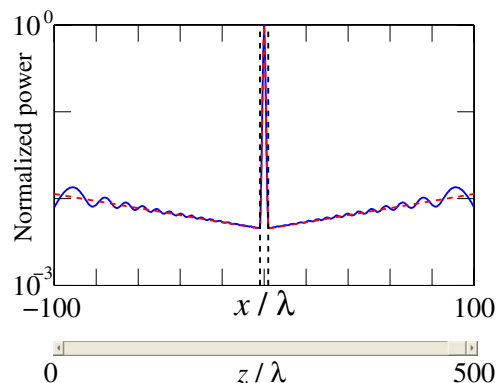
$\text{Im}(n_{\text{eff}})$ as a function of b/a . The blue solid curve, green dashed-dotted curve, and red dashed curve represent the leakage loss calculated from the direct determination of the leaky mode solution, the perturbation method, and the determination from the radiation mode solution, respectively.

exponentially as b/a increases. This falloff is expected, since the magnitude of the nearly guided mode decreases exponentially before reaching the interfaces at $x = \pm b$.

Figure 14 shows a movie of the transverse mode for the Gaussian input beam that we considered in Fig. 7 as it propagates along the W-type waveguide. We keep 5000 k_x modes, and we set $B = 8000a$. The waveguide parameters are the same as in Figs. 9 and 12. The red dashed curve indicates the leaky mode solution. The mode profiles are normalized to 1 at $x = 0$, so that the attenuation as z increases is not visible. The two solutions overlap in the center, showing that the mode preserves its shape. The exponential increase as $x \rightarrow \pm\infty$ that is expected for a leaky mode is also apparent in both solutions. However, the dynamical solution has the following characteristics: first, a portion of the beam that is mismatched to the leaky mode in the central region of the waveguide rapidly diffracts. Second, as the beam propagates along the z direction, the beam's power gradually increases in the cladding region of the waveguide, $|x| > b$, and resembles the leaky mode profile over larger values of $|x|$. Hence, the power that is radiated from the core is actually not lost, since it is simply redistributed from the core into the cladding region. Third, the dynamic solution has a front in $\pm x$ beyond which it rapidly tends to zero. The oscillations in Fig. 14 are due to phase interference from different points along the initial x profile, analogous to the oscillations that are observed in Fraunhofer diffraction from a single slit.

We noted earlier that as z becomes large, we expect the exponential decay to cease. Mathematically, this effect will occur when the change in $\beta(k_x)z$ over the bandwidth of the Lorentzian peak becomes large. The stationary phase point of $\beta(k_x)$ will once again dominate the solution, and the wave should exhibit algebraic decay that is consistent with diffraction, rather than the exponential decay that is characteristic of leakage. This long-term diffraction should not be confused with the early-term diffraction that is observed immediately after a beam is launched into the waveguide. In contrast to the early-term diffraction

Figure 14

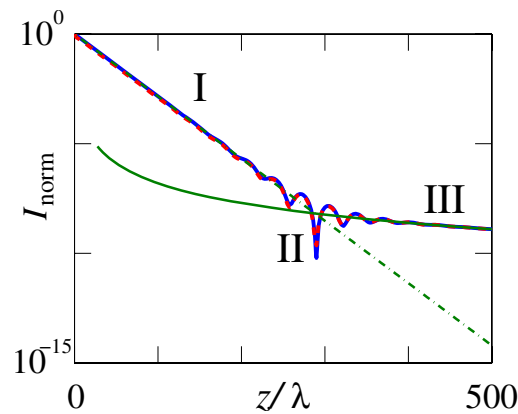


Movie (Media 4) of the transverse mode evolution as the mode propagates along a W-type slab waveguide. The red dashed curve and blue solid curve represent the transverse mode power from the leaky mode and the actual profile that is found by integrating Eq. (4). The black dashed lines indicate $x = \pm b$.

that is due to the portions of the spectrum $\tilde{A}(k_x)$ in which $k_x^2 > k_0^2(n_0^2 - n_1^2)$, the long-term diffraction is due to the portion of the spectrum $\tilde{A}(k_x)$ in which $k_x^2 < k_0^2(n_0^2 - n_1^2)$. This portion of the spectrum includes $k_x = \pm k_r$, so that the components of $\tilde{A}(k_x)$ that lead to long-term diffraction also act as a background, canceling out the initial exponential growth in the $\pm x$ directions. During the period of evolution when the leaky mode dominates, the phases of the continuum background become increasingly mismatched, increasingly revealing the leaky mode at larger values of $|x|$, as we observed in Fig. 14.

In Fig. 15, we show $I_{\text{norm}} = |A(z, x=0)|^2 / |A(z=0, x=0)|^2$ as a function of z/λ . The red dashed curve shows the power of the field using numerical integration. The whole curve can be separated into three regions. In region I, we find the expected exponential decay. In region III, we find the expected algebraic decay, which is proportional to z^{-3} at $x=0$, according to Eq. (39). In the transition region between, oscillations are visible, as the contributions due to diffraction and leakage interfere either constructively or destructively. The green dashed-dotted and green solid curves show, respectively, the steepest descent analysis for the evolution at $x=0$ and the leaky mode evolution. They agree in the appropriate limits with the exact evolution. We note that we have set $b/a = 2.5$ instead of $b/a = 5$, as in our previous examples. The algebraic decay is difficult to observe, and the larger value of $b/a = 5$ implies a low leakage loss and a low power at the transition point where the decay rate changes from exponential to algebraic. The power at that point is so low that round-off errors made it impossible for us to observe. In principle, if the input field were perfectly matched to the leaky mode there would not be a transition to algebraic decay, but since perfect matching is impossible given the infinite extent of the leaky mode, the transition will always occur.

Figure 15



$I_{\text{norm}} = |A(z, x=0)|^2 / |A(z=0, x=0)|^2$ as a function of z/λ for a Gaussian beam with $b/a = 2.5$. The red dashed curve shows the power of the field obtained by using numerical integration. The green dashed-dotted and green solid curves show, respectively, the steepest descent analysis for the evolution at $x = 0$ and the leaky mode evolution. The blue solid curve shows the I_{norm} that is calculated by summing the fields from the steepest descent analysis that were used to produce I and II.

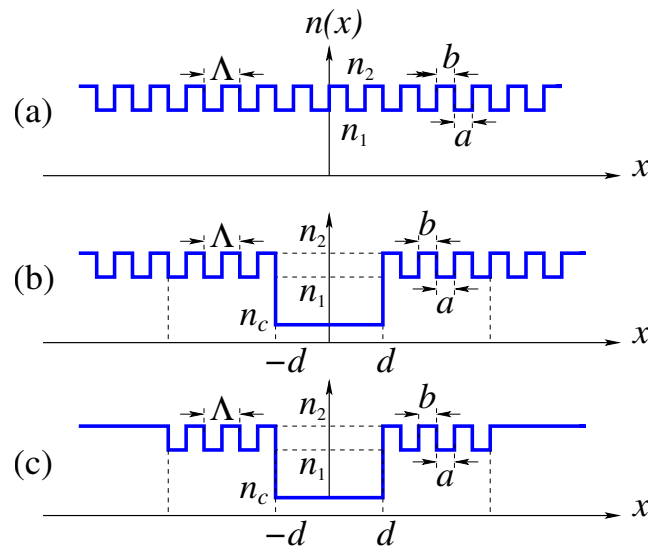
5. Bandgap Waveguide

In the slab waveguides that we have studied thus far, the index of refraction in the guiding regions is always higher than in the immediately surrounding regions. However, it is advantageous in some cases to be able to guide waves in a region of lower index of refraction. For example, by filling a region with air, one can greatly lower the nonlinearity. One can confine modes in a lower-index region by using the bandgap effect [12,59,60]. In a periodically varying medium, like the one shown in Fig. 16(a), frequency bands where the light cannot propagate through the periodically varying medium are referred to as bandgaps. By creating a defect in the periodic structure, as shown in Fig. 16(b), and launching light at a frequency that is in the bandgap of the periodic structure, one can confine light inside the defect, even when the index of refraction is smaller than in the surrounding regions. In practice, however the periodic variations have a finite extent, as shown in Fig. 16(c). In this case, the waveguide is leaky. We will show in this section that the leaky modes in these bandgap waveguides behave much like the leaky modes in the W-type waveguides. In optical fibers, an analogous approach has been widely used to confine light to a low-index core. This guidance is referred to as photonic bandgap guidance or capillary guidance [46].

5.1. Eigenvalue Equation

We begin by considering an infinitely periodic structure, of which Fig. 16(a) shows one example. Because Eq. (3) is a second-order ordinary differential equation, it must be possible to write any solution $E(x)$ as a superposition of two independent solutions $E_1(x)$ and $E_2(x)$. In particular, since the equation is

Figure 16



The refractive index profile for (a) an infinite periodic structure, (b) an infinite periodic structure with a center defect, and (c) a leaky bandgap waveguide.

periodic with period Λ , it must be possible to write $E_1(x+\Lambda)=AE_1(x)+BE_2(x)$ and $E_2(x+\Lambda)=CE_1(x)+DE_2(x)$, where $A, B, C,$ and D are constants that depend on the details of the periodic structure. We may write these two relations in matrix form as

$$\begin{pmatrix} E_1(x+\Lambda) \\ E_2(x+\Lambda) \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} E_1(x) \\ E_2(x) \end{pmatrix}, \quad (51)$$

where $A, B, C,$ and D are four constants that depend on the details of the periodic variation with one important constraint. Since Eq. (3) has no first derivative terms, its Wronskian will be constant [51]. That implies in turn that $AD-BC=1$. To determine whether propagating solutions to Eq. (3) exist at a given value of β , we first search for a particular set of solutions $E_+(x)$ and $E_-(x)$ that have the property $E_+(x)=\exp(iK_+x)u_+(x)$ and $E_-(x)=\exp(iK_-x)u_-(x)$, where $u_+(x)$ and $u_-(x)$ are strictly periodic in Λ . The existence of $u_{\pm}(x)$ is guaranteed by the Bloch–Floquet theorem, which holds for any periodic structure. The solutions $E_{\pm}(x)$ satisfy the conditions $E_{\pm}(x+\Lambda)=\lambda_{\pm}E_{\pm}(x)$, where $\lambda_{\pm}=\exp(iK_{\pm}\Lambda)$. We may find λ_{\pm} and hence K_{\pm} and $u_{\pm}(x)$ by solving the eigenvalue problem

$$\begin{vmatrix} A-\lambda & B \\ C & D-\lambda \end{vmatrix} = 0 = \lambda^2 - \lambda(A+D) + 1. \quad (52)$$

We now infer

$$\lambda_{\pm} = \frac{A+D}{2} \pm \left[\left(\frac{A+D}{2} \right)^2 - 1 \right]^{1/2}. \quad (53)$$

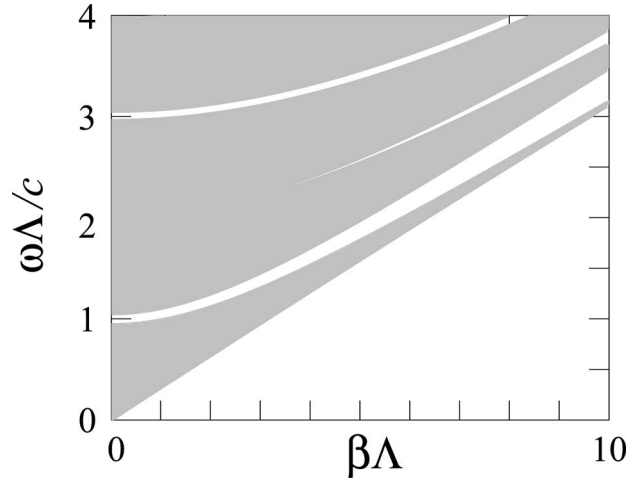
From the condition $AD-BC=1$, it follows that $\lambda_+\lambda_-=1$, so that $K_+=-K_-$, and both are purely real or purely imaginary. If $(A+D)^2/4 < 1$, then K_{\pm} are real, and waves can propagate along x , which is required in order for an initial beam that is introduced at $z=0$ to propagate in the $+z$ direction. Otherwise, there is no propagation, and an initial beam attenuates. The case shown schematically in Fig. 16(a) in which there are just two different indices has been extensively analyzed in the literature [59,61]. We consider an example from [59], in which $\lambda=1.15 \mu\text{m}$, $n_1=2.89$, $n_2=3.38$, $n_c=1$, and $a=b=0.1 \mu\text{m}$, which applies to the cladding region in the waveguide for a gas laser. In Fig. 17, we show the band structure as a function of normalized frequency and propagation constant. The dark areas are the allowed bands, where $(A+D)^2/4 < 1$.

We now consider the case of a defect, shown in Fig. 16(b). The lowest-order defect mode will be even, and we write

$$E(x) = \begin{cases} C_1 \cos(k_c x) & 0 \leq x \leq d \\ C_0 u_+(x) \exp(iK_+ x) + D_0 u_-(x) \exp(-iK_- x) & d \leq x \end{cases}, \quad (54)$$

where $k_c=(k_0^2 n_c^2 - \beta^2)^{1/2}$, with the condition that $E(x)=E(-x)$ when $x < 0$. Since we are only interested in confined modes within the bandgap, we may assume $K_+=i\alpha$, where α is real, and $K_-=-i\alpha$. As a consequence, the contribution $D_0 u_-(x) \exp(-iK_- x) = D_0 u_-(x) \exp(\alpha x)$ grows exponentially as $x \rightarrow +\infty$, and we must set $D_0=0$. Thus, Eq. (54) becomes

Figure 17



Band structure as a function of normalized frequency and propagation constant. The dark areas are the allowed bands.

$$E(x) = \begin{cases} C_1 \cos(k_c x) & 0 \leq x \leq d \\ C_0 u_+(x) \exp(-\alpha x) & d \leq x \end{cases} \quad (55)$$

Matching the function and its derivative at the boundary, we obtain

$$C_1 \cos(k_c d) = C_0 u_+(d) \exp(-\alpha d), \quad (56a)$$

$$-C_1 k_c \sin(k_c d) = C_0 [u'_+(d) - \alpha u_+(d)] \exp(-\alpha d). \quad (56b)$$

Eliminating constants C_1 and C_0 , we find the dispersion relation

$$-k_c \tan(k_c d) = \frac{u'_+(d) - \alpha u_+(d)}{u_+(d)}, \quad (57)$$

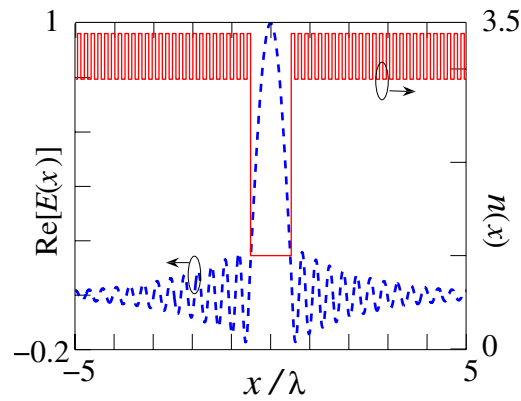
from which the allowed value of β or n_{eff} may be determined. For the example that we are considering, in which $n_1=2.89$, $n_2=3.38$, $n_c=1$, $a=b=0.1 \mu\text{m}$, $d=6a$, and $\lambda=11.5a$, we find $n_{\text{eff}}=\beta/k_0=0.89295$. We show a picture of $n(x)$ and real part of $E(x)$ in Fig. 18.

Finally, we turn to consideration of the leaky mode in the waveguide shown in Fig. 16(c). In this case, Eq. (55) becomes

$$E(x) = \begin{cases} C_2 \cos(k_c x) & 0 < x \leq d \\ C_1 u_+(x) \exp(-\alpha x) + D_1 u_-(x) \exp(\alpha x) & d \leq x \leq d + M\Lambda \\ C_0 \exp(iK_x x) & d + M\Lambda \leq x \end{cases}, \quad (58)$$

where M is the number of periods in the intermediate region, and we set $E(-x)=E(x)$, where $x < 0$. We have $K_x=(k_0^2 n_2^2 - \beta^2)^{1/2}$. We may solve for the propagation constant β and n_{eff} by using exactly the same mode-matching technique that we used for the W-type waveguides described in Subsection 4.1. In this case, we find $n_{\text{eff}}=\beta/k_0=0.89085 + 1.543 \times 10^{-3}i$ when the number of periods $M=10$.

Figure 18



Refractive index and real part of $E(x)$ as a function of x .

5.2. Alternative Solution Procedures

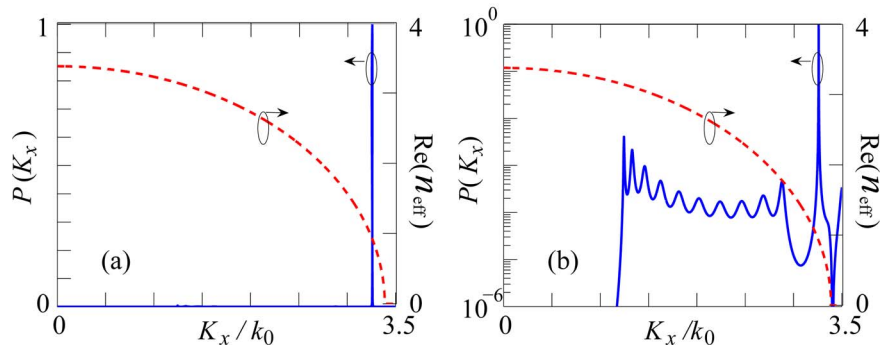
In our study of the W-type waveguide, we considered two alternative solution procedures. In the first, we perturbed around the nonleaky solution, assuming that the leakage is small. In the second, we used complete mode decomposition, which consists of only radiation modes. We showed that both these procedures yield nearly the same answer as a direct determination of the leaky mode solution. The same is true for the bandgap leaky modes.

The perturbation analysis proceeds exactly as in the case of the W-type waveguide. We write $\beta = \beta_0 + \Delta\beta$, substitute this expression into the dispersion relation for the leaky mode, expand in powers of $\Delta\beta$, and keep the zeroth- and first-order contributions. We thus obtain an expression of the form $M_0(\beta_0) + M_1(\beta_0)\Delta\beta = 0$. Since we are using β_0 from the solution to the nonleaky waveguide, $M_0(\beta_0)$ is only due to the difference between the leaky and nonleaky waveguides and will be small. Using this approach on our example system, we find $n_{\text{eff}} = (\beta_0 + \Delta\beta)/k_0 = 0.89097 + 1.275 \times 10^{-4}i$. The imaginary part is within 20% of what we found using a direct solution in the case with 10 periods in each cladding region.

The analysis using radiation modes again proceeds by analogy to what we found with the W-type waveguide. In Fig. 19 we show the $\text{Re}(n_{\text{eff}})$ and the normalized coefficient $P(K_x) = |\tilde{A}(K_x)|^2 / \max[|\tilde{A}(K_x)|^2]$ as a function of K_x/k_0 for our example. We keep 2000 k_x modes, and we set $B = 1000\Lambda$. We consider a Gaussian input beam $A(z=0, x) = A_0 \exp(-x^2/2d^2)$, where d is half the width of the center region. A sharp Lorentzian peak in $|\tilde{A}(K_x)|^2$ is visible, just as in the case of the W-type waveguide. Using Eq. (50) once more to find n_{eff} , we obtain $n_{\text{eff}} = 0.89084 + 1.540 \times 10^{-4}i$, which agrees well with the result that we obtained by using the direct method. We note that there are small peaks located between $k_x/k_0 = 1$ and $k_x/k_0 = 3$. The small peaks are not visible in Fig. 19(a) on a linear scale, but are visible in Fig. 19(b) on a logarithmic scale. These are modes that are confined inside the high-index portions of the bandgap regions by the neighboring lower-index portions. Their mode effective indices are between n_1 and n_2 .

Figure 20 shows a movie of wave propagation in the leaky bandgap waveguide. Figure 21 shows a movie of the transverse mode evolution as the mode

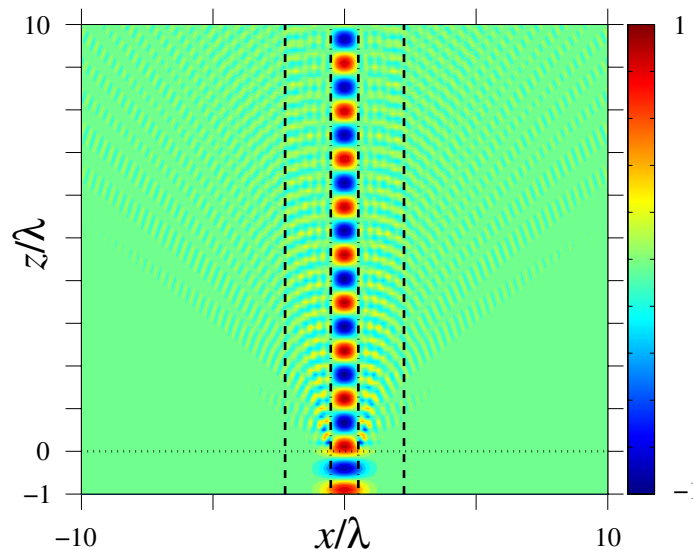
Figure 19



The real part of n_{eff} and the normalized coefficient $P(K_x) = |\tilde{A}(K_x)|^2 / \max[|\tilde{A}(K_x)|^2]$ with (a) a linear scale and (b) a logarithmic scale as a function of K_x .

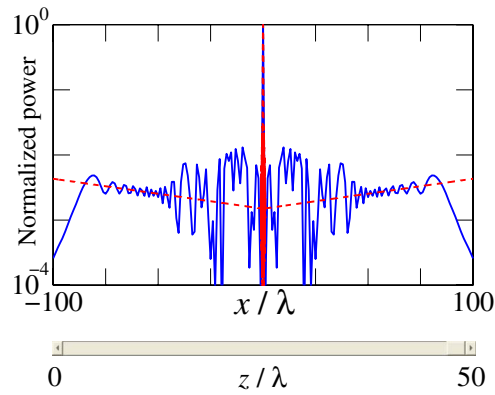
propagates along the bandgap slab waveguide with ten periods. The red dashed curve and blue solid curve show the transverse leaky mode power and the computational solution of Eq. (4), respectively. The power in both curves is normalized to 1 at $x=0$. As the mode propagates in the $+z$ direction, the mode gradually fills in the power in the outside region of the waveguide. Again, the power that is radiated from the core is not lost; it is simply redistributed from the core into the outside region. Note that the solid blue curve has an irregular shape around the center region. The irregular shape is caused by the small

Figure 20



Wave propagation in a leaky bandgap waveguide. A beam is injected into the waveguide at $z=0$. The movie (Media 5) shows the real part of the electric field. The black dashed-dotted lines and black dashed lines indicate $x = \pm d$ and $x = \pm(d + M\lambda)$, respectively.

Figure 21

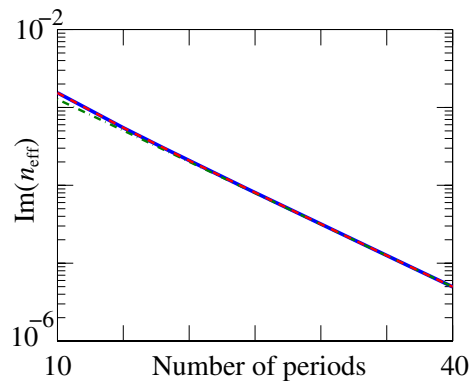


Movie ([Media 6](#)) of the transverse mode evolution as it propagates along a leaky bandgap slab waveguide. The red dashed curve and blue solid curve represent the power of the leaky mode and the computational solution of Eq. (4). The boundary lines are not shown in this figure, since they are very close to the center, as shown in Fig. 20.

peaks shown in Fig. 19(b). If we remove these small peaks in the coefficient function $P(K_x)$, then the irregular fluctuations in the blue solid curve go away.

In Fig. 22 we show $\text{Im}(n_{\text{eff}})$ for all three mode-matching methods, for our example system, as we allow the number of periods to grow. Beyond 20 periods, disagreement among them is slight.

Figure 22



$\text{Im}(n_{\text{eff}})$ as a function of the number of periods. The blue solid curve, green dashed-dotted curve, and red dashed curve represent the leakage loss calculated from the direct determination of the leaky mode solution, the perturbation method, and the determination from the radiation mode solution, respectively.

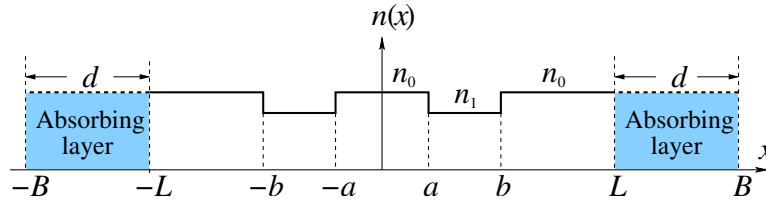
6. Waveguide with Absorbing Layers

Up to this point, we have solved for the modes in the waveguide by using the fact that in a region in which the index of refraction is constant, the solution can be written exactly as a sum of exponents or a sum of cosines and sines. Matching the solutions and their derivatives across the boundaries of the regions with different indices yields a matrix equation whose solution produces the propagation constant. In two dimensions, the solution can no longer be written as a sum of exponents, but when each of the regions of constant index of refraction has a circular profile, a closely analogous method based on Bessel functions can be developed and is referred to as the multipole method [9,10]. When all the indices are real, so that the Helmholtz equation, Eq. (3), is self-adjoint, there is a complete mode decomposition that consists of some finite number of guided modes and a continuum of radiation modes. Leaky modes are not part of this complete set, although they can usefully approximate the behavior of nearly guided modes, as we have discussed in detail.

While this approach is the basis for all analytical studies of optical waveguides, it is neither the most useful, nor the most widespread computational approach. Most computational approaches are based on finite-element or finite-difference discretizations of Maxwell's equation [5,6,62]. While the discretization in two transverse dimensions is far from trivial to implement, highly robust commercial software is available from several different vendors. The finite-difference and finite-element approaches are highly flexible, since they can deal with arbitrary geometries.

To simulate outgoing boundary conditions, the finite-difference and finite-element approaches are usually implemented with an absorbing layer at the simulation boundary in which the index of refraction is complex [5]. The goal is to obtain a solution that reproduces as closely as possible, in a limited spatial region, the solution with a lossless medium of infinite extent. That will be possible only if the absorbing layer is far from the initial beam. However, even given this constraint, the mathematical consequences of adding this absorbing layer are profound. First, all the modes are confined within a finite region, so that the mode decomposition always consists of a countable number of modes as the spatial discretization becomes increasingly fine, as discussed in Section 2. Second, the equations that describe the wave propagation are no longer self-adjoint, and the modes can no longer be chosen so that they are purely real. The propagation constants are in general complex. Third, we have found that when we approximate the W-type waveguide with a lossless region, surrounded by a lossy region when $|x| > L$, as shown in Fig. 23, there is *always* a leaky mode that is part of the complete mode decomposition. That is true even when L is very large, and the maxima of the mode's exponential tails are actually larger than the mode's central peak. As $L \rightarrow \infty$, the mode decomposition does not appear to converge to the mode decomposition for the self-adjoint problem, where there is no loss in the waveguide material that we considered in Section 4. This behavior can be considered a generalization of the result that even with self-adjoint equations, the behavior depends upon the choice of the boundary conditions [48]. Thus, it is important to understand when and why the solutions to this problem can be expected to reproduce the solutions to the problem of a lossless medium of infinite extent.

Figure 23



Refractive index profile for a W-type waveguide with absorbing layers.

In practice, the details of the absorbing layer implementation can have a significant impact on the propagation constant of the leaky mode. In order to minimize reflections from the absorbing layers, it has become common to use perfectly matched layers—a technique that was first introduced in finite-difference time domain simulations [63,64]. An ideal perfectly matched layer is a layer that gives no reflection at any frequency and angle. However, it is important to recognize that once they are spatially discretized, the perfectly matched layers are no longer perfectly matched, and reflections from them do occur [64]. In our prior studies of two-dimensional waveguides, we have found that shifting the location of the absorbing layer slightly can make a difference in the propagation constant that far exceeds the round-off or discretization error.

In the remainder of this section, we discuss the optimization of the absorbing layer and then discuss the complete mode decomposition and the role of the leaky mode for a simple finite-difference method with absorbing boundary layers.

6.1. Optimization of the Absorbing Layer

Figure 23 shows the structure of a symmetric W-type waveguide with absorbing layers. We have chosen not to implement perfectly matched layers in order to simplify the discussion and because they offer little or no advantage relative to a straightforward absorber for the discretization that we use. We use a simple finite-difference method, so that Eq. (3) becomes

$$\frac{1}{\Delta^2}[E(x_{k-1}) - 2E(x_k) + E(x_{k+1}))] + [k_0^2 n^2(x_k) - \beta^2]E(x_k) = 0, \quad (59)$$

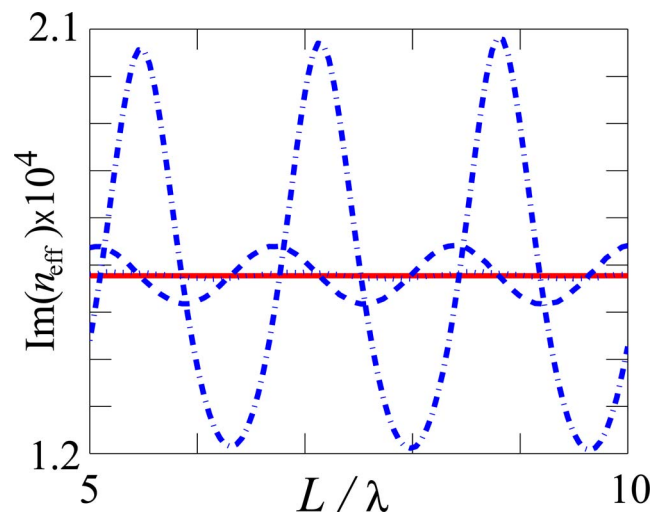
where $x_k = -B + (k-1)\Delta$ and $\Delta = 2B/(N-1)$ is the discretization width. The index k varies from 1 to N , where the simulation window extends from $-B$ to B . The index of refraction $n(x_k)$ is real when $|x_k| \leq L$ and is complex when $|x_k| > L$. The problem of solving Eq. (59) is thus a matter of finding the N eigenvalues β of the $N \times N$ matrix in Eq. (59), where for convenience we have applied Dirichlet boundary conditions, setting $E_0 = E_{N+1} = 0$. This choice of boundary condition does affect the mode decomposition, but has no significant effect on the leaky mode.

The permittivity of the absorbing layer is

$$\varepsilon = n_0^2 \left[1 + i \left(\frac{|x| - L}{d} \right)^2 s' \right], \quad (60)$$

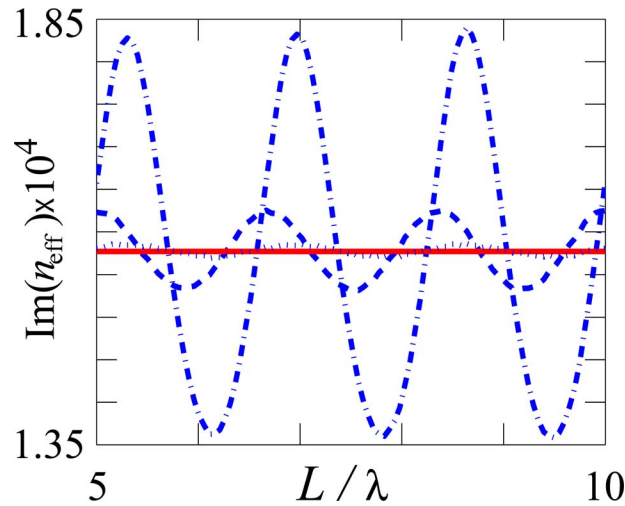
where s' is a parameter that we choose to minimize the reflections. Figure 24 shows $\text{Im}(n_{\text{eff}}) = \text{Im}(\beta)/k_0$ as a function of L/λ , where β is the propagation constant of the leaky mode that we obtain from the solution of Eq. (59). We have chosen $b/a=5$, $n_1=0.96n_0$, $n_0=1.45$, and $ka(n_0^2 - n_1^2)^{1/2}=1$, which are the same parameters as in Subsection 4.1. We have set $N=10^5$. The blue dashed-dotted curve, dashed curve, and dotted curve show the results with normalized absorbing layer widths $d/\lambda=5, 10, 15$, respectively. We have set $s'=1$ in all three cases. The red solid line shows the results from Eq. (42). We see that $\text{Im}(n_{\text{eff}})$ varies sinusoidally, and its magnitude diminishes rapidly as d/λ increases. These oscillations are due primarily to reflections from the boundary at $x=\pm B$. The boundary produces strong reflection when the width of the absorbing layer is too small. In Fig. 25, we show the results when $s'=5, 2, 1$. The red solid line shows the result from the eigenvalue of Eq. (42). The normalized width of absorbing layer d/λ is set to be 15 for all three cases. A large value of s' leads once again to large oscillations, primarily due to reflections from within the absorbing layer, close to $|x|=L$. In Fig. 26, we look in more detail at the behavior of $\text{Im}(n_{\text{eff}})$ as s' varies from 10^{-2} to 10. At each value of s' , we calculated the average value of $\text{Im}(n_{\text{eff}})$ and the standard deviation for 100 evenly spaced values of L/λ as we allow it to vary from 5 to 10. When $s' < 0.04$, the reflection from the boundary at $\pm B$ causes a large error and hence a large standard deviation. When $s' > 1$, the reflection at the absorbing layer cause a large error and a large standard deviation. Any value between $s'=0.04$ and $s'=1.0$ yields nearly the same average, which is about 1.578×10^{-4} , but, for a single computation, it is better to use $0.04 < s' < 0.1$ so that the standard deviation is low.

Figure 24



$\text{Im}(n_{\text{eff}}) = \text{Im}(\beta)/k_0$ as a function of L/λ . The blue dashed-dotted curve, dashed curve, and dotted curve show the results with normalized absorbing layer widths $d/\lambda=5, 10, 15$, respectively. The red solid line shows the results from Eq. (42).

Figure 25



$\text{Im}(n_{\text{eff}}) = \text{Im}(\beta) / k_0$ as a function of L/λ . The blue dashed–dotted curve, dashed curve, and dotted curve show the results with $s' = 5, 2, 1$, respectively. The red solid line shows the results from Eq. (42).

In general, it is useful to use absorbing layers with different values of L and then to average the results for $\text{Im}(n_{\text{eff}})$.

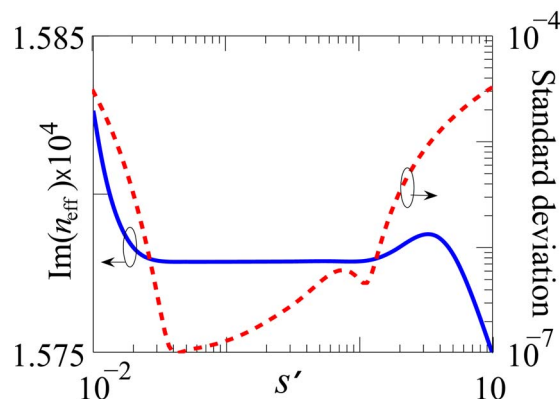
6.2. Mode Decomposition and the Leaky Mode

Equation (59) has the form of a matrix eigenvalue problem

$$(\mathbf{M} - \beta^2 \mathbf{I}) \mathbf{E} = 0, \quad (61)$$

where \mathbf{I} is the identity matrix, and, letting $\delta_{j,k}$ denote the Kronecker delta function, the matrix element

Figure 26



Average value of $\text{Im}(n_{\text{eff}})$ and the standard deviation for 100 evenly spaced values of L/λ as we allow it to vary from 5 to 10.

$$M_{j,k} = \frac{\delta_{j-1,k} - 2\delta_{j,k} + \delta_{j+1,k}}{\Delta^2} + k_0^2 n_k^2 \delta_{j,k} \quad (62)$$

is the j, k th element of the matrix \mathbf{M} . The vector \mathbf{E} is a column vector, whose k th element is $E(x_k)$. We note that \mathbf{M} is a symmetric matrix, which automatically implies that it is normal, i.e., $\mathbf{M}\mathbf{M}^\dagger = \mathbf{M}^\dagger\mathbf{M}$. That in turn implies that its eigenvectors form a complete set [65], so that we may write

$$A(x_k) = \sum_{l=1}^N \tilde{A}_l E_l(x_k), \quad (63)$$

where $E_l(x_k)$ denotes the k th element of the column vector \mathbf{E}_l , and the \mathbf{E}_l are solutions of the right eigenvalue equation

$$(\mathbf{M} - \beta^2 \mathbf{I})\mathbf{E}_l = 0. \quad (64)$$

Equation (63) is the discretized version of Eq. (10) in Section 2. In principle, we may find \tilde{A}_l by defining left eigenvectors that satisfy

$$\mathbf{F}_l^T (\mathbf{M} - \beta^2 \mathbf{I}) = 0, \quad (65)$$

where \mathbf{F}_l is a column vector and \mathbf{F}_l^T is the corresponding row vector, and noting that

$$\tilde{A}_l = \sum_{j=1}^N A(x_j) F_j(x_k). \quad (66)$$

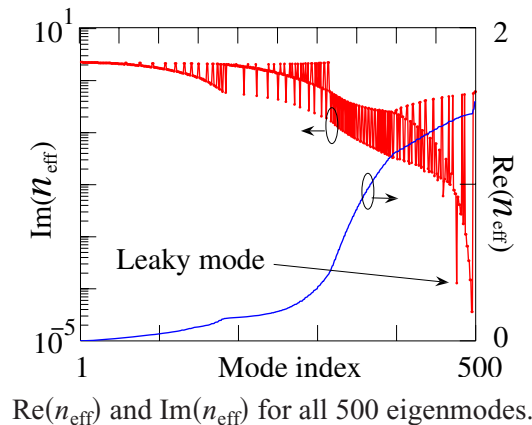
Because the matrix \mathbf{M} is symmetric, we find $F_l(x_k) = E_l(x_k)$. To determine the field at any x_l and any z , we use the expression

$$A(z, x_k) = \sum_{l=1}^N \tilde{A}_l E_l(x_k) \exp(i\beta_l z), \quad (67)$$

where we stress that the β_l are in general complex. Equation (67) is the analog to Eq. (4) for the finite-difference method and is the discretized version of Eq. (11).

In Fig. 27, we show the $\text{Re}(n_{\text{eff}})$ and $\text{Im}(n_{\text{eff}})$. We set $N=500$ in this case. We

Figure 27

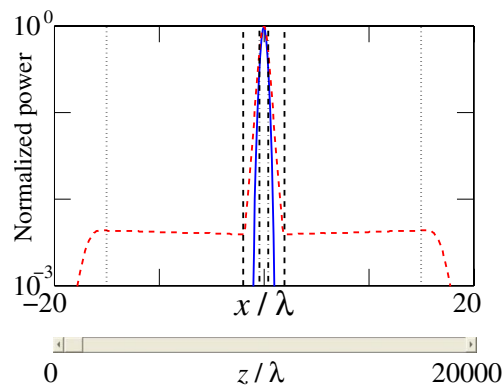


$\text{Re}(n_{\text{eff}})$ and $\text{Im}(n_{\text{eff}})$ for all 500 eigenmodes.

use the same parameters as in Subsection 4.4 with $L/\lambda=15$, $(B-L)/\lambda=15$, and $s'=1$. We arrange the indices from small $\text{Re}(n_{\text{eff}})$ to large $\text{Re}(n_{\text{eff}})$, which corresponds roughly to arranging the indices from large $|k_x|$ to small $|k_x|$. In this case, mode number 476 corresponds to the leaky mode. Its imaginary part is substantially lower than its neighbors. It is the modes with small indices, corresponding to $k_x^2 > k_0^2(n_0^2 - n_1^2)$, that contribute to the early term diffraction. All of these modes are more lossy than the leaky mode. There is one mode with a lower $\text{Im}(n_{\text{eff}})$ than the leaky mode, which is mode number 498. This mode is the lowest-order cavity mode that is confined between the absorbing layer and the layer with low index of refraction n_1 . In general, the loss of the cavity modes is low because their transverse derivatives are small, implying from Eq. (46) that they have little outward flux into the absorbing region. Modes 499 and 500 are cladding modes that are located almost entirely in the absorbing regions and have high loss.

In Fig. 28, we show a movie of the transverse normalized power of the same Gaussian input beam that we considered in Subsection 4.4, as it propagates along the waveguide with the same parameters as in Subsection 4.1 and with $L/\lambda=15$, $(B-L)/\lambda=15$, and $s'=1$. We also show the power of the leaky mode as the red dashed curve. We have normalized the peak of the mode power profiles to 1. There are several different propagation regions. First, the field leaks into the outer region of the W-type waveguide in which $b < |x| < L$ and almost exactly reproduces the profile of the leaky mode, although small oscillations are visible for $2 < x/\lambda < 20$, just as shown in Fig. 14. However, when $z/\lambda > 10,000$, the lowest-order cavity mode becomes the dominant mode. The distance at which it dominates is determined by the magnitude of its small, but finite overlap with the initial beam. Since this cavity mode is located primarily in the region $b < |x| < L$, the profile of the power changes significantly once this cavity mode becomes dominant. The lowest-order cavity mode eventually dominates because it has a lower value of $\text{Im}(n_{\text{eff}})$ than any other mode, including the leaky mode, and because it has a small, but nonzero overlap with the initial beam. It is possible in principle to launch an initial

Figure 28



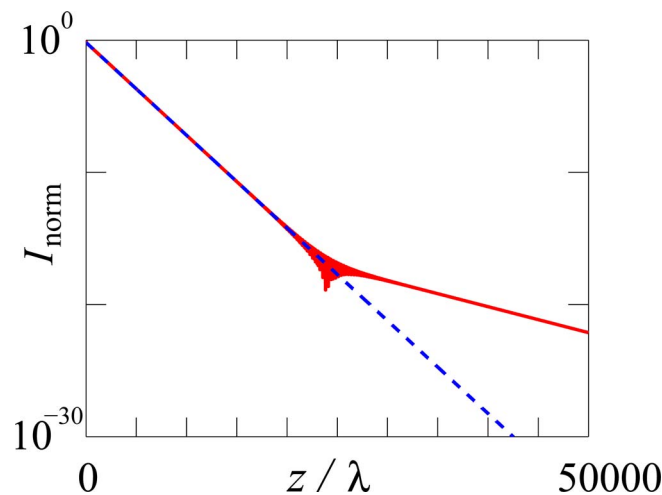
Movie (Media 7) of the transverse normalized power of the same initial Gaussian beam that we considered in Subsection 4.4. We also show the power of the leaky mode as a red dashed curve. We have normalized the peak of the mode power profiles to 1. The black dashed-dotted lines, black dashed line, and black dotted lines indicate $x = \pm a$, $x = \pm b$, and $x = \pm L$, respectively.

beam that has no overlap with this cavity mode or, in fact, is a pure leaky mode. However, one cannot do so when the initial beam is localized in the center of the simulation window, as is always required in a realistic simulation. In practice, any initial beam will be composed of many modes, of which the leaky mode is just one. Figure 29 shows $I_{\text{norm}} = |A(z, x=0)|^2 / |A(z=0, x=0)|^2$ as a function of z/λ for a Gaussian input beam. The beam and waveguide parameters are the same as in Fig. 27. Owing to the large attenuation, normalization is needed to view these results. The red solid curve shows the result keeping all 500 modes, while the blue dashed line shows the result keeping only the leaky mode. This figure is analogous to Fig. 15, in which long-term diffraction ultimately dominates the evolution. In this case, however, the attenuation at large z is still exponential, but the attenuation occurs at a slower rate than in the case of the leaky mode, because the energy ultimately resides in the lowest-order cavity mode, rather than a continuum of radiation modes.

In Fig. 30, we show a slide show for the input wave (blue solid curves) and its decomposition into the eigenmodes (red dashed curves). Starting with the leaky mode, which has index number 476, we add the other eigenmodes with their coefficients \tilde{A}_l . We see that the other modes ultimately cancel the tails of the leaky mode, so that the mode profile represents the Gaussian input beam. Since these modes attenuate more rapidly than the leaky mode, the leaky mode is ultimately revealed as the wave propagates along the waveguide. This behavior is different from the infinite guide, where no modes attenuate and the leaky mode is revealed only in a finite region surrounding the center by diffraction of the continuum of radiation modes. However, the behavior in the central region is still the same in both cases up to the point where the large- z behavior begins to dominate the evolution.

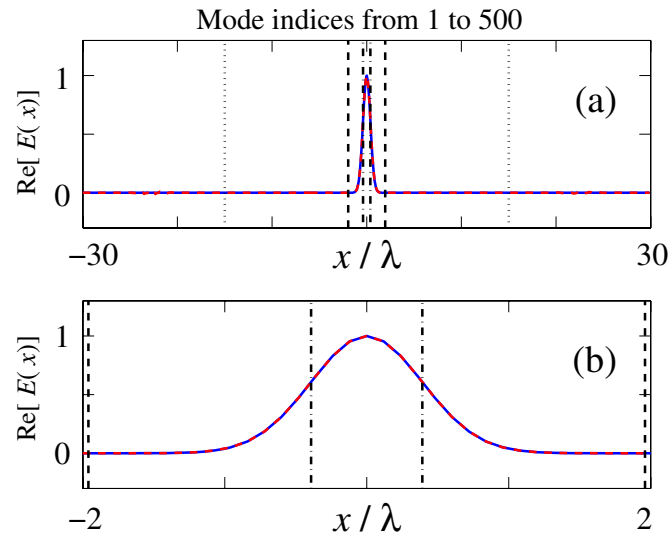
In the example that we considered here, the leaky mode does not have large exponential tails. However, we have examined the behavior as L and B become large. As long as d is large enough and s' is not too big, so that reflections

Figure 29



$I_{\text{norm}} = |A(z, x=0)|^2 / |A(z=0, x=0)|^2$ as a function of z/λ for a Gaussian beam. The red solid curve shows the result keeping all 500 modes, while the blue dashed curve shows the result keeping only the leaky mode.

Figure 30



Slide show ([Media 8](#)) for the input wave (blue solid curves) and its decomposition into the eigenmodes (red dashed curves). In (b), we show the central region from (a). The black dashed-dotted lines, black dashed line, and black dotted lines indicate $x = \pm a$, $x = \pm b$, and $x = \pm L$, respectively.

from the absorbing regions are avoided, the qualitative behavior is unchanged. There is always a leaky mode that dominates the evolution until z becomes large, at which point one or more cavity modes dominate the evolution. We have observed the following points: with b/a kept fixed, as L increases, so do the maxima of the exponential tails of the leaky modes. Ultimately these tails become bigger at their maxima than the central peak of the mode. Nonetheless, the leaky mode is a real mode of this waveguide system. Additionally, as L increases, the number of cavity modes with less attenuation than the leaky mode increases. However, the lowest-order cavity mode always has the lowest loss and ultimately dominates the evolution.

7. Answers to the Introductory Questions

There are two basic types of computational method that are used to find optical waveguide modes. One type is mode-matching methods, in which one uses exact analytical solutions in regions where the index of refraction is constant and matches the solutions and their derivatives at the boundaries. In the case of slab waveguides, one uses exponential functions. The other type is finite-difference and finite-element methods.

When using mode-matching methods, it is usual to specify the problem in a lossless waveguide so that the index of refraction is real at all points in the space. In this case, all the waveguide modes are lossless and have real propagation constants. If the index of refraction becomes equal to a constant value at some finite distance from the origin, then in general there is a continuum of radiation modes and a finite number of discrete guided modes. Any physically reasonable initial profile can be decomposed into a generalized sum over

the waveguide modes, in which the sum includes an integral over the radiation modes. One can then determine the beam profile at any subsequent point in the waveguide by multiplying the amplitude of each mode by an appropriate exponential factor and then resumming the modes.

This problem is equivalent mathematically to the quantum-mechanical problem of a particle that is confined in a potential well, and the set of facts just stated is often repeated in elementary textbooks on quantum mechanics and waveguide theory. The proof is, however, far from trivial [48]. It begins with a finite region in space with boundary conditions that ensure that the eigenvalue problem—given in our case by Eq. (3)—remains self-adjoint. In this case, one has a countably infinite set of modes. As one lets the boundary tend toward infinity, most of the modes coalesce into a continuum, leaving a finite set of guided modes. This continuum will always contain both outgoing and incoming waves. That occurs because self-adjoint boundary conditions on a finite region of space always lead to reflections, so that as the boundary tends toward infinity, there are always waves propagating both outward and inward. Because of its importance in quantum mechanics, as well as optical waveguides, the theory of self-adjoint eigenvalue equations has been extensively studied.

Given the complete mode decomposition into a continuum of radiation modes and a finite number of guided modes, the light evolution in the three-slab waveguide of Fig. 1(a) is not difficult to understand. The portion of the beam profile that couples into the guided modes remains confined, while the portion that couples into the radiation modes diffracts.

By contrast, the evolution of the light in the W-type waveguide, in which only a continuum of radiation modes is present, is not so easy to understand. The key point is that the continuum contains sharp Lorentzian peaks at values of k_x that nearly equal the transverse wavenumbers of the central portion of the guided modes of the corresponding three-slab waveguide. Immediately after a beam is injected into the waveguide, there is a transient stage in which a portion of the energy rapidly diffracts, in close analogy to the behavior in a three-slab waveguide. However, Lorentzian peaks in the continuum quickly dominate the early evolution, leading to leaky modes whose damping rates are given by the widths of the Lorentzian peaks. The amplitudes corresponding to all the transverse wavenumbers in a Lorentzian peak are initially in phase, but when the propagation distance becomes long enough so that these amplitudes are out of phase, then long-term diffraction dominates. This long-term diffraction, which corresponds to the dissolution of the leaky mode and occurs outside the central region of the waveguide, should not be confused with the early term diffraction, during which the initial beam settles down into the shape of the leaky mode in the central region of the waveguide. In effect, the continuum can be decomposed into three contributions—a smooth portion at large transverse wavenumbers that leads to the rapid initial diffraction, Lorentzian peaks that correspond to the leaky modes, and a smooth background to the Lorentzian peaks that compensates for the exponential growth at large transverse distances and leads to diffractive radiation at large propagation distances.

While leaky modes can be understood as a consequence of the Lorentzian peaks in the continuum, that does not explain the evolution of their exponential growth transverse to the direction of propagation. We have found that this exponential growth is real within a limited transverse distance surrounding the

origin, and this transverse distance continues to grow as long as the leaky mode dominates over the long-term diffraction. Using conservation of flux, we have shown that any mode that decreases exponentially as it propagates must increase exponentially transverse to the direction of propagation once the transverse dimensions are large enough that the index of refraction equals its final constant value. However, exponential growth that extends to infinity is unphysical. There have been attempts in the literature to bypass this problem by claiming that at some transverse distance the material becomes lossy or interacts with air, and the problem goes away. However, the protective coating or air is often very far from the region of interest, so that it is easier both mathematically and computationally to treat the medium as though the indices of refraction at the boundary of the region of interest extended to infinity. Invoking a protective coating or air does not help us understand how the appearance of leaky modes in the mathematical formulation of the problem as a lossless waveguide can provide useful numbers or even be consistent with physical reality.

The Lorentzian peaks in the continuum, if separated from the continuum background lead to solutions that grow exponentially in the transverse dimension. The background cancels this exponential growth. As the light propagates, the different components of the background evolve so that they are increasingly out of phase, revealing the exponential growth. This trend continues until the components of the Lorentzian peaks evolve so that they are also out of phase. Thereafter, long-term diffraction dominates the evolution.

Remarkably, the basic behavior that we have just described remains the same in far more complex systems than the three-slab and W-type waveguides that were the focus of most of our discussion. We demonstrated this point in detail for the important case of bandgap waveguides, but it remains true when the transverse structures become two-dimensional.

When using finite-difference or finite-element methods, it is usual to surround the region of interest with a lossy region, whose purpose is to absorb outgoing waves. For the class of problems that we are considering here, in which the index of refraction reaches a constant value at some finite distance from the origin, the absorbing boundary would typically be placed shortly beyond the distances at which the constant value is reached. The wave flux is strictly outward in the lossless problem at transverse distances that are both beyond the initial beam width and beyond the points at which the index of refraction reaches its final value. Hence, one expects physically that surrounding the region of interest with an absorbing layer will produce the same behavior within the region of interest as does a lossless guide that extends to infinity. Some sort of absorber is needed because any lossless boundary conditions in a finite spatial region will produce reflections, which very visibly changes the behavior in the region of interest!

Because of their flexibility and the ease relative to other methods with which they can be numerically implemented, finite-difference and finite-element methods are the methods of choice in geometries with any significant complexity. As a consequence, considerable effort has gone into developing algorithms for absorbing layers that reflect as little as possible, while using as small a number of node points as possible [5]. The absorbing layer must be optimized for each geometry, and we showed how this optimization procedure

works in the case of a simple finite-difference scheme and a simple absorber for the W-type waveguide.

As long as the absorbing layer has been optimized to produce negligible reflections and the number of points is sufficiently large on a well-chosen mesh, then the finite-difference and finite-element methods will produce results for the evolution of an initially localized beam that agrees in the region of interest with the results of the mode-matching methods for some finite propagation distance. Moreover, when we are solving for the modes, the damping rates for the leaky modes will agree. Not surprisingly, mode-matching methods produce results different from those for the finite-difference and finite-element methods when the energy that is outside the region of interest becomes large.

While the finite-difference and finite-element methods for the quantities of interest may produce the same results as do the mode-matching methods, it is important to recognize that the mathematical problem has profoundly changed, and so has the mode decomposition. Since the problem formulation is no longer self-adjoint, one can no longer use the mathematical apparatus that was developed for self-adjoint problems. Indeed, there is no guarantee in general that the modes constitute a complete set, although we showed that the decomposition was complete for the simple case that we considered. Hence, the completeness must be verified on a case-by-case basis [65].

For the simple finite-difference algorithm that we considered, we observed the following differences from the mode-matching method: (1) The mode decomposition consists of a finite number of discrete modes. That will be the case for *any* finite-difference or finite-element method. (2) The leaky mode is a real mode of the system. (3) The transverse exponential behavior is revealed when modes that have larger loss than the leaky mode attenuate. There is no real diffraction, although the evolution reproduces the early term diffraction. (4) There are one or more cladding modes that have less loss than the leaky mode. These cladding modes decay exponentially, but at a slower rate than the leaky mode. Algebraic decay associated with long-term diffraction is not observed.

It might seem surprising at first that the mode decomposition used with mode-matching methods and the mode decomposition used with finite-difference or finite-element methods should yield the same result for the quantities of interest when the decompositions differ so profoundly. However, it is a reflection of a deep result that can be found throughout physics. A problem can often be formulated in two different ways. As long as both formulations are correct, they must yield the same results. Often these formulations lead to different, complementary physical pictures of the phenomenon being studied. That is the case here.

Appendix A: How We Did It and Software Link

A.1. How We Did It

There are many ways to generate animations. We proceeded as follows: for each animation, we generated 10–100 frames in postscript (PS) format, labeled sequentially as fig1.ps, fig2.ps, fig3.ps, etc. Then we used the following Linux command to transform files in PS format to files in GIF format:

```
mogrify -format gif fig*.ps
```

After we generated a series of GIF files, we linked all the GIF files to build the animation, using the commercial software GIF Construction Set Professional (<http://www.mindworkshop.com/alchemy/gifcon.html>). Both MOV files and GIF files can be generated. The schematic diagrams in the tutorial paper, Figs. 1, 2, 10, 16, and 23, were not created by using MATLAB. These figures were created by using xfig (<http://www.xfig.org/>).

A.2. Software Link

Download the data archive ([Media 9](#)) and unpack it to a directory. Please see the Terms of Use on the *Advances in Optics and Photonics* website. To use the code, simply download it and run it from MATLAB. To generate a particular figure, type “FigureNumber(*N*),” where *N* is the index of the figure or movie that you want to generate. For questions or problems regarding the code, contact Jonathan Hu.

A.3. For More Information

The code UndStdLeakyMode was developed by Jonathan Hu in the Computational Photonics Laboratory at UMBC. More information about the laboratory and Jonathan Hu, as well as additional useful software, may be found at our Web site (<http://www.umbc.edu/photonics/software>).

Acknowledgments

This work has been supported in part by the Naval Research Laboratory, and we have benefitted from our interactions with scientists there who are working on the subject of holey and photonic crystal fibers, including I. D. Aggarwal, T. F. Carruthers, P. Falkenstein, E. J. Friebele, B. L. Justus, J. S. Sanghera, L. B. Shaw, T. F. Taunay, and B. M. Wright. We also thank M. de Sterke for his encouragement during the writing of this tutorial introduction. Finally, we thank B. Kuhlmeiy for a detailed critique of an earlier version of this manuscript.

We dedicate this paper to D. Marcuse, with whom C. R. Menyuk had the privilege of working for several years and whose published work has greatly benefited J. Hu.

References

1. S. E. Miller, “Integrated optics: an introduction,” *Bell Syst. Tech. J.* **48**, 2059–2069 (1969).
2. E. A. Marcatili, “Dielectric rectangular waveguide and directional coupler for integrated optics,” *Bell Syst. Tech. J.* **48**, 2071–2102 (1969).
3. E. Snitzer, “Cylindrical dielectric waveguide modes,” *J. Opt. Soc. Am.* **51**, 491–198 (1961).
4. D. Gloge, “Weakly guiding fibers,” *Appl. Opt.* **10**, 2252–2258 (1971).
5. J. Jin, *The Finite Element Method in Electromagnetics*, 2nd ed. (Wiley, 2002).
6. K. Kawano and T. Kitoh, *Introduction to Optical Waveguide Analysis*

Solving Maxwell's Equations and the Schrödinger Equation (Wiley, 2001).

7. M. Koshiba and Y. Tsuji, "Curvilinear hybrid edge/nodal elements with triangular shape for guided-wave problems," *J. Lightwave Technol.* **18**, 737–743 (2000).
8. M. N. O. Sadiku, *Numerical Techniques in Electromagnetics*, 2nd ed. (CRC Press, 2001).
9. T. P. White, B. T. Kuhlmeiy, R. C. McPhedran, D. Maystre, G. Renversez, C. M. de Sterke, and L. C. Botten, "Multipole method for microstructured optical fibers. I. Formulation," *J. Opt. Soc. Am. B* **19**, 2322–2330 (2002).
10. B. T. Kuhlmeiy, T. P. White, G. Renversez, D. Maystre, L. C. Botten, C. M. de Sterke, and R. C. McPhedran, "Multipole method for microstructured optical fibers. II. Implementation and results," *J. Opt. Soc. Am. B* **19**, 2331–2340 (2002).
11. D. Marcuse, "Solution of the vector wave equation for general dielectric waveguides by the Galerkin method," *IEEE J. Quantum Electron.* **28**, 459–465 (1992).
12. S. Johnson and J. Joannopoulos, "Block-iterative frequency-domain methods for Maxwell's equations in a planewave basis," *Opt. Express* **8**, 173–190 (2001).
13. J. M. Lourtioz, H. Benisty, V. Berger, J.-M. Gérard, D. Maystre, and A. Tchebnokov, *Photonic Crystals: towards Nanoscale Photonic Devices* (Springer, 2005).
14. D. Marcuse, *Theory of Dielectric Optical Waveguides* (Academic, 1991).
15. A. W. Snyder and J. D. Love, *Optical Waveguide Theory* (Kluwer Academic, 1983).
16. R. E. Collin, *Field Theory of Guided Waves*, 2nd ed. (IEEE, 1991).
17. S. Barone, "Leaky wave contributions to the field of a line source above a dielectric slab," Report R-532-546, PIB-462 (Microwave Research Institute, Polytechnic Institute of Brooklyn, Nov. 26, 1956).
18. S. Barone and A. Hessel, "Leaky wave contributions to the field of a line source above a dielectric slab—part II," Report R-698-58, PIB-626 (Microwave Research Institute, Polytechnic Institute of Brooklyn, Dec. 1958).
19. N. Marcuvitz, "On field representations in terms of leaky modes or eigenmodes," *IRE Trans. Antennas Propag.* **4**, 192–194 (1956).
20. T. Tamir and A. A. Oliner, "Guided complex waves. Part 1: fields at an interface," *Proc. IEEE* **110**, 310–324 (1963).
21. T. Tamir and A. A. Oliner, "Guided complex waves. Part 2: relation to radiation pattern," *Proc. IEEE* **110**, 325–334 (1963).
22. C. W. Hsue and T. Tamir, "Evolution of transverse-electric surface and leaky waves guided by an asymmetric layer configuration," *J. Opt. Soc. Am. A* **1**, 923–931 (1984).
23. T. Tamir and F. Y. Kou, "Varieties of leaky waves and their excitation along multilayered structures," *IEEE J. Quantum Electron.* **22**, 544–551 (1986).
24. S. T. Peng and A. A. Oliner, "Guidance and leakage properties of a class of open dielectric waveguides. I. Mathematical formulations," *IEEE Trans. Microwave Theory Tech.* **MTT-29**, 843–855 (1981).
25. A. A. Oliner, S. T. Peng, T. I. Hsu, and A. Sanchez, "Guidance and leakage properties of a class of open dielectric waveguides. II. New physical

- effects,” *IEEE Trans. Microwave Theory Tech.* **MTT-29**, 855–869 (1981).
26. E. S. Cassedy and M. Cohn, “On the existence of leaky waves due to a line source above a grounded dielectric slab,” *IRE Trans. Microwave Theory Tech.* **9**, 243–247 (1961).
 27. D. B. Hall and C. Yeh, “Leaky waves in a heteroepitaxial film,” *J. Appl. Phys.* **44**, 2271–2274 (1973).
 28. H. Haus and D. Miller, “Attenuation of cutoff modes and leaky modes of dielectric slab structures,” *IEEE J. Quantum Electron.* **22**, 310–318 (1986).
 29. Y. Suematsu and K. Furuya, “Quasi-guided modes and related radiation losses in optical dielectric waveguides with external higher index surroundings,” *IEEE Trans. Microwave Theory Tech.* **23**, 170–175 (1975).
 30. S. Kawakami and S. Nishida, “Characteristics of a doubly clad optical fiber with a low-index inner cladding,” *IEEE J. Quantum Electron.* **10**, 879–887 (1974).
 31. S. Kawakami and S. Nishida, “Perturbation theory of a doubly clad optical fiber with a low-index inner cladding,” *IEEE J. Quantum Electron.* **11**, 130–138 (1975).
 32. J. Arnbak, “Leaky waves on a dielectric rod,” *Electron. Lett.* **5**, 41–42 (1969).
 33. J. R. James, “Leaky waves on a dielectric rod,” *Electron. Lett.* **5**, 252–254 (1969).
 34. J. Burke, “Propagation constants of resonant waves on homogeneous, isotropic slab waveguides,” *Appl. Opt.* **9**, 2444–2452 (1970).
 35. V. V. Shevchenko, “On the behavior of wave numbers beyond the critical value for waves in dielectric waveguides (media with losses),” *Radiophys. Quantum Electron.* **15**, 194–200 (1972).
 36. A. W. Snyder and D. J. Mitchell, “Ray attenuation in lossless dielectric structures,” *J. Opt. Soc. Am.* **64**, 956–963 (1974).
 37. A. W. Snyder and D. J. Mitchell, “Leaky mode analysis of circular optical waveguides,” *Opto-electronics (London)* **6**, 287–296 (1974).
 38. M. Maeda and S. Yamada, “Leaky modes on W-fibers: mode structure and attenuation,” *Appl. Opt.* **16**, 2198–2203 (1977).
 39. A. W. Snyder, “Leaky-ray theory of optical waveguides of circular cross section,” *Appl. Phys. (N.Y.)* **4**, 273–298 (1974).
 40. A. W. Snyder and D. J. Mitchell, “Leaky rays on circular optical fibers,” *J. Opt. Soc. Am.* **64**, 599–607 (1974).
 41. A. W. Snyder, D. J. Mitchell, and C. Pask, “Failure of geometric optics for analysis of circular optical fibers,” *J. Opt. Soc. Am.* **64**, 608–614 (1974).
 42. J. D. Love and C. Winkler, “Attenuation and tunneling coefficients for leaky rays in multilayered optical waveguides,” *J. Opt. Soc. Am.* **67**, 1627–1633 (1977).
 43. J. T. Chilwell and I. J. Hodgkinson, “Thin-films field-transfer matrix theory of planar multilayer waveguides and reflection from prism-loaded waveguides,” *J. Opt. Soc. Am. A* **1**, 742–753 (1984).
 44. L. Torner, F. Canal, and J. Hernandez-Marco, “Leaky modes in multilayer uniaxial optical waveguides,” *Appl. Opt.* **29**, 2805–2814 (1990).
 45. J. Petraček and K. Singh, “Determination of leaky modes in planar multilayer waveguides,” *IEEE Photon. Technol. Lett.* **14**, 810–812 (2002).
 46. F. Zolla, G. Renversez, A. Nicolet, B. Kuhlmeiy, S. Guenneau, and D. Felbacq, *Foundations Of Photonic Crystal Fibres* (Imperial College Press, 2005).

47. K. Okamoto, *Fundamentals of Optical Waveguides* (Academic, 2000), Chap. 2.
48. E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations* (McGraw-Hill, 1984).
49. W. Klaus and W. R. Leeb, "Transient fields in the input coupling region of optical single-mode waveguides," *Opt. Express* **15**, 11808–11826 (2007).
50. 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.* **31**, 1790–1802 (1995).
51. G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, 5th ed. (Academic, 2001).
52. J. D. Joannopoulos, S. G. Johnson, J. N. Winn, R. D. Meade, *Photonic Crystals: Molding the Flow of Light*, 2nd ed. (Princeton U. Press, 2008).
53. K. Hoffman and R. Kunze, *Linear Algebra*, 2nd ed. (Prentice Hall, 1971), Chap. 8.5.
54. P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, 1953), Chap. 4.6.
55. M. J. Adams, *An Introduction to Optical Waveguides* (Wiley, 1981), Chap. 2.6.
56. R. V. Churchill and J. W. Brown, *Complex Variables and Applications*, 5th ed. (McGraw-Hill, 1990), Chap. 2.
57. D. Marcuse, *Light Transmission Optics* (Van Nostrand Reinhold, 1972), Chap. 8.4.
58. R. N. Bracewell, *The Fourier Transform and Its Applications*, 3rd ed. (McGraw-Hill, 1999).
59. P. Yeh, A. Yariv, and C.-S. Hong, "Electromagnetic propagation in periodic stratified media. I. General theory," *J. Opt. Soc. Am.* **67**, 423–438 (1977).
60. J. Hu and C. R. Menyuk, "Leakage loss and bandgap analysis in air-core photonic bandgap fiber for nonsilica glasses," *Opt. Express* **15**, 339–349 (2007).
61. A. Yariv and P. Yeh, *Optical Waves in Crystals* (Wiley, 1984), Chap. 11.10.
62. S. Guo, F. Wu, S. Albin, H. Tai, and R. Rogowski, "Loss and dispersion analysis of microstructured fibers by finite-difference method," *Opt. Express* **12**, 3341–3352 (2004).
63. J. P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comput. Phys.* **114**, 110–117 (1994).
64. A. Taflove and S. C. Hagness, *Computational Electrodynamics*, 2nd ed. (Artech House, 2000).
65. G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. (Johns Hopkins U. Press, 1996).