

improving liquid crystal displays, and other products, such as various optoelectronic components, cosmetics, and "hot" and "cold" mirrors for architectural and automotive windows.

8.14 Problems

- 8.1 Prove the reflectance and transmittance formulas (8.4.6) in FTIR.
- 8.2 *Computer Experiment—FTIR.* Reproduce the results and graphs of Figures 8.4.3–8.4.5.
- 8.3 *Computer Experiment—Surface Plasmon Resonance.* Reproduce the results and graphs of Figures 8.5.3–8.5.7.
- 8.4 Working with the electric and magnetic fields across an negative-index slab given by Eqs. (8.6.1) and (8.6.2), derive the reflection and transmission responses of the slab given in (8.6.8).
- 8.5 *Computer Experiment—Perfect Lens.* Study the sensitivity of the perfect lens property to the deviations from the ideal values of $\epsilon = -\epsilon_0$ and $\mu = -\mu_0$, and to the presence of losses by reproducing the results and graphs of Figures 8.6.3 and 8.6.4. You will need to implement the computational algorithm listed on page 329.
- 8.6 *Computer Experiment—Antireflection Coatings.* Reproduce the results and graphs of Figures 8.7.1–8.7.3.
- 8.7 *Computer Experiment—Omnidirectional Dielectric Mirrors.* Reproduce the results and graphs of Figures 8.8.2–8.8.10.
- 8.8 Derive the generalized Snell's laws given in Eq. (8.10.10). Moreover, derive the Brewster angle expressions given in Eqs. (8.11.4) and (8.11.5).
- 8.9 *Computer Experiment—Brewster angles.* Study the variety of possible Brewster angles and reproduce the results and graphs of Example 8.11.1.
- 8.10 *Computer Experiment—Multilayer Birefringent Structures.* Reproduce the results and graphs of Figures 8.13.1–8.13.2.

Waveguides

Waveguides are used to transfer electromagnetic power efficiently from one point in space to another. Some common guiding structures are shown in the figure below. These include the typical coaxial cable, the two-wire and microstrip transmission lines, hollow conducting waveguides, and optical fibers.

In practice, the choice of structure is dictated by: (a) the desired operating frequency band, (b) the amount of power to be transferred, and (c) the amount of transmission losses that can be tolerated.

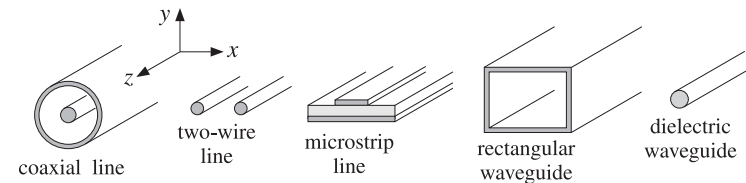


Fig. 9.0.1 Typical waveguiding structures.

Coaxial cables are widely used to connect RF components. Their operation is practical for frequencies below 3 GHz. Above that the losses are too excessive. For example, the attenuation might be 3 dB per 100 m at 100 MHz, but 10 dB/100 m at 1 GHz, and 50 dB/100 m at 10 GHz. Their power rating is typically of the order of one kilowatt at 100 MHz, but only 200 W at 2 GHz, being limited primarily because of the heating of the coaxial conductors and of the dielectric between the conductors (dielectric voltage breakdown is usually a secondary factor.) However, special short-length coaxial cables do exist that operate in the 40 GHz range.

Another issue is the single-mode operation of the line. At higher frequencies, in order to prevent higher modes from being launched, the diameters of the coaxial conductors must be reduced, diminishing the amount of power that can be transmitted.

Two-wire lines are not used at microwave frequencies because they are not shielded and can radiate. One typical use is for connecting indoor antennas to TV sets. Microstrip lines are used widely in microwave integrated circuits.

Rectangular waveguides are used routinely to transfer large amounts of microwave power at frequencies greater than 3 GHz. For example at 5 GHz, the transmitted power might be one megawatt and the attenuation only 4 dB/100 m.

Optical fibers operate at optical and infrared frequencies, allowing a very wide bandwidth. Their losses are very low, typically, 0.2 dB/km. The transmitted power is of the order of milliwatts.

9.1 Longitudinal-Transverse Decompositions

In a waveguiding system, we are looking for solutions of Maxwell's equations that are propagating along the guiding direction (the z direction) and are confined in the near vicinity of the guiding structure. Thus, the electric and magnetic fields are assumed to have the form:

$$\begin{aligned} \mathbf{E}(x, y, z, t) &= \mathbf{E}(x, y) e^{j\omega t - j\beta z} \\ \mathbf{H}(x, y, z, t) &= \mathbf{H}(x, y) e^{j\omega t - j\beta z} \end{aligned} \quad (9.1.1)$$

where β is the propagation wavenumber along the guide direction. The corresponding wavelength, called the *guide wavelength*, is denoted by $\lambda_g = 2\pi/\beta$.

The precise relationship between ω and β depends on the type of waveguiding structure and the particular propagating mode. Because the fields are confined in the transverse directions (the x, y directions,) they cannot be uniform (except in very simple structures) and will have a non-trivial dependence on the transverse coordinates x and y . Next, we derive the equations for the phasor amplitudes $\mathbf{E}(x, y)$ and $\mathbf{H}(x, y)$.

Because of the preferential role played by the guiding direction z , it proves convenient to decompose Maxwell's equations into components that are *longitudinal*, that is, along the z -direction, and components that are *transverse*, along the x, y directions. Thus, we decompose:

$$\mathbf{E}(x, y) = \underbrace{\hat{\mathbf{x}} E_x(x, y) + \hat{\mathbf{y}} E_y(x, y)}_{\text{transverse}} + \underbrace{\hat{\mathbf{z}} E_z(x, y)}_{\text{longitudinal}} \equiv \mathbf{E}_T(x, y) + \hat{\mathbf{z}} E_z(x, y) \quad (9.1.2)$$

In a similar fashion we may decompose the gradient operator:

$$\nabla = \underbrace{\hat{\mathbf{x}} \partial_x + \hat{\mathbf{y}} \partial_y}_{\text{transverse}} + \hat{\mathbf{z}} \partial_z = \nabla_T + \hat{\mathbf{z}} \partial_z = \nabla_T - j\beta \hat{\mathbf{z}} \quad (9.1.3)$$

where we made the replacement $\partial_z \rightarrow -j\beta$ because of the assumed z -dependence. Introducing these decompositions into the source-free Maxwell's equations we have:

$$\begin{aligned} \nabla \times \mathbf{E} &= -j\omega\mu\mathbf{H} & (\nabla_T - j\beta\hat{\mathbf{z}}) \times (\mathbf{E}_T + \hat{\mathbf{z}} E_z) &= -j\omega\mu(\mathbf{H}_T + \hat{\mathbf{z}} H_z) \\ \nabla \times \mathbf{H} &= j\omega\epsilon\mathbf{E} & (\nabla_T - j\beta\hat{\mathbf{z}}) \times (\mathbf{H}_T + \hat{\mathbf{z}} H_z) &= j\omega\epsilon(\mathbf{E}_T + \hat{\mathbf{z}} E_z) \\ \nabla \cdot \mathbf{E} &= 0 & (\nabla_T - j\beta\hat{\mathbf{z}}) \cdot (\mathbf{E}_T + \hat{\mathbf{z}} E_z) &= 0 \\ \nabla \cdot \mathbf{H} &= 0 & (\nabla_T - j\beta\hat{\mathbf{z}}) \cdot (\mathbf{H}_T + \hat{\mathbf{z}} H_z) &= 0 \end{aligned} \quad (9.1.4)$$

where ϵ, μ denote the permittivities of the medium in which the fields propagate, for example, the medium between the coaxial conductors in a coaxial cable, or the medium within the hollow rectangular waveguide. This medium is assumed to be lossless for now.

We note that $\hat{\mathbf{z}} \cdot \hat{\mathbf{z}} = 1$, $\hat{\mathbf{z}} \times \hat{\mathbf{z}} = 0$, $\hat{\mathbf{z}} \cdot \mathbf{E}_T = 0$, $\hat{\mathbf{z}} \cdot \nabla_T E_z = 0$ and that $\hat{\mathbf{z}} \times \mathbf{E}_T$ and $\hat{\mathbf{z}} \times \nabla_T E_z$ are transverse while $\nabla_T \times \mathbf{E}_T$ is longitudinal. Indeed, we have:

$$\begin{aligned} \hat{\mathbf{z}} \times \mathbf{E}_T &= \hat{\mathbf{z}} \times (\hat{\mathbf{x}} E_x + \hat{\mathbf{y}} E_y) = \hat{\mathbf{y}} E_x - \hat{\mathbf{x}} E_y \\ \nabla_T \times \mathbf{E}_T &= (\hat{\mathbf{x}} \partial_x + \hat{\mathbf{y}} \partial_y) \times (\hat{\mathbf{x}} E_x + \hat{\mathbf{y}} E_y) = \hat{\mathbf{z}} (\partial_x E_y - \partial_y E_x) \end{aligned}$$

Using these properties and equating longitudinal and transverse parts in the two sides of Eq. (9.1.4), we obtain the equivalent set of Maxwell equations:

$$\begin{aligned} \nabla_T E_z \times \hat{\mathbf{z}} - j\beta \hat{\mathbf{z}} \times \mathbf{E}_T &= -j\omega\mu\mathbf{H}_T \\ \nabla_T H_z \times \hat{\mathbf{z}} - j\beta \hat{\mathbf{z}} \times \mathbf{H}_T &= j\omega\epsilon\mathbf{E}_T \\ \nabla_T \times \mathbf{E}_T + j\omega\mu \hat{\mathbf{z}} H_z &= 0 \\ \nabla_T \times \mathbf{H}_T - j\omega\epsilon \hat{\mathbf{z}} E_z &= 0 \\ \nabla_T \cdot \mathbf{E}_T - j\beta E_z &= 0 \\ \nabla_T \cdot \mathbf{H}_T - j\beta H_z &= 0 \end{aligned} \quad (9.1.5)$$

Depending on whether both, one, or none of the longitudinal components are zero, we may classify the solutions as transverse electric and magnetic (TEM), transverse electric (TE), transverse magnetic (TM), or hybrid:

$E_z = 0, H_z = 0,$	TEM modes
$E_z = 0, H_z \neq 0,$	TE or H modes
$E_z \neq 0, H_z = 0,$	TM or E modes
$E_z \neq 0, H_z \neq 0,$	hybrid or HE or EH modes

In the case of TEM modes, which are the dominant modes in two-conductor transmission lines such as the coaxial cable, the fields are purely transverse and the solution of Eq. (9.1.5) reduces to an equivalent two-dimensional electrostatic problem. We will discuss this case later on.

In all other cases, at least one of the longitudinal fields E_z, H_z is non-zero. It is then possible to express the transverse field components $\mathbf{E}_T, \mathbf{H}_T$ in terms of the longitudinal ones, E_z, H_z .

Forming the cross-product of the second of equations (9.1.5) with $\hat{\mathbf{z}}$ and using the BAC-CAB vector identity, $\hat{\mathbf{z}} \times (\hat{\mathbf{z}} \times \mathbf{H}_T) = \hat{\mathbf{z}} (\hat{\mathbf{z}} \cdot \mathbf{H}_T) - \mathbf{H}_T (\hat{\mathbf{z}} \cdot \hat{\mathbf{z}}) = -\mathbf{H}_T$, and similarly, $\hat{\mathbf{z}} \times (\nabla_T H_z \times \hat{\mathbf{z}}) = \nabla_T H_z$, we obtain:

$$\nabla_T H_z + j\beta \mathbf{H}_T = j\omega\epsilon \hat{\mathbf{z}} \times \mathbf{E}_T$$

Thus, the first two of (9.1.5) may be thought of as a linear system of two equations in the two unknowns $\hat{\mathbf{z}} \times \mathbf{E}_T$ and $\mathbf{H}_T$, that is,

$$\begin{aligned} \beta \hat{\mathbf{z}} \times \mathbf{E}_T - \omega\mu\mathbf{H}_T &= j\hat{\mathbf{z}} \times \nabla_T E_z \\ \omega\epsilon \hat{\mathbf{z}} \times \mathbf{E}_T - \beta \mathbf{H}_T &= -j\nabla_T H_z \end{aligned} \quad (9.1.6)$$

The solution of this system is:

$$\begin{aligned}\hat{\mathbf{z}} \times \mathbf{E}_T &= -\frac{j\beta}{k_c^2} \hat{\mathbf{z}} \times \nabla_T E_z - \frac{j\omega\mu}{k_c^2} \nabla_T H_z \\ \mathbf{H}_T &= -\frac{j\omega\epsilon}{k_c^2} \hat{\mathbf{z}} \times \nabla_T E_z - \frac{j\beta}{k_c^2} \nabla_T H_z\end{aligned}\quad (9.1.7)$$

where we defined the so-called *cutoff wavenumber* k_c by:

$$k_c^2 = \omega^2 \epsilon \mu - \beta^2 = \frac{\omega^2}{c^2} - \beta^2 = k^2 - \beta^2 \quad (\text{cutoff wavenumber}) \quad (9.1.8)$$

The quantity $k = \omega/c = \omega\sqrt{\epsilon\mu}$ is the wavenumber a uniform plane wave would have in the propagation medium ϵ, μ .

Although k_c^2 stands for the difference $\omega^2 \epsilon \mu - \beta^2$, it turns out that the boundary conditions for each waveguide type force k_c^2 to take on certain values, which can be positive, negative, or zero, and characterize the propagating modes. For example, in a dielectric waveguide k_c^2 is positive inside the guide and negative outside it; in a hollow conducting waveguide k_c^2 takes on certain quantized positive values; in a TEM line, k_c^2 is zero. Some related definitions are the *cutoff frequency* and the *cutoff wavelength* defined as follows:

$$\omega_c = ck_c, \quad \lambda_c = \frac{2\pi}{k_c} \quad (\text{cutoff frequency and wavelength}) \quad (9.1.9)$$

We can then express β in terms of ω and ω_c , or ω in terms of β and ω_c . Taking the positive square roots of Eq. (9.1.8), we have:

$$\beta = \frac{1}{c} \sqrt{\omega^2 - \omega_c^2} = \frac{\omega}{c} \sqrt{1 - \frac{\omega_c^2}{\omega^2}} \quad \text{and} \quad \omega = \sqrt{\omega_c^2 + \beta^2 c^2} \quad (9.1.10)$$

Often, Eq. (9.1.10) is expressed in terms of the wavelengths $\lambda = 2\pi/k = 2\pi c/\omega$, $\lambda_c = 2\pi/k_c$, and $\lambda_g = 2\pi/\beta$. It follows from $k^2 = k_c^2 + \beta^2$ that

$$\frac{1}{\lambda^2} = \frac{1}{\lambda_c^2} + \frac{1}{\lambda_g^2} \Rightarrow \lambda_g = \frac{\lambda}{\sqrt{1 - \frac{\lambda^2}{\lambda_c^2}}} \quad (9.1.11)$$

Note that λ is related to the free-space wavelength $\lambda_0 = 2\pi c_0/\omega = c_0/f$ by the refractive index of the dielectric material $\lambda = \lambda_0/n$.

It is convenient at this point to introduce the *transverse impedances* for the TE and TM modes by the definitions:

$$\eta_{TE} = \frac{\omega\mu}{\beta} = \eta \frac{\omega}{\beta c}, \quad \eta_{TM} = \frac{\beta}{\omega\epsilon} = \eta \frac{\beta c}{\omega} \quad (\text{TE and TM impedances}) \quad (9.1.12)$$

where the medium impedance is $\eta = \sqrt{\mu/\epsilon}$, so that $\eta/c = \mu$ and $\eta c = 1/\epsilon$. We note the properties:

$$\eta_{TE} \eta_{TM} = \eta^2, \quad \frac{\eta_{TE}}{\eta_{TM}} = \frac{\omega^2}{\beta^2 c^2} \quad (9.1.13)$$

Because $\beta c/\omega = \sqrt{1 - \omega_c^2/\omega^2}$, we can write also:

$$\eta_{TE} = \frac{\eta}{\sqrt{1 - \frac{\omega_c^2}{\omega^2}}}, \quad \eta_{TM} = \eta \sqrt{1 - \frac{\omega_c^2}{\omega^2}} \quad (9.1.14)$$

With these definitions, we may rewrite Eq. (9.1.7) as follows:

$$\begin{aligned}\hat{\mathbf{z}} \times \mathbf{E}_T &= -\frac{j\beta}{k_c^2} (\hat{\mathbf{z}} \times \nabla_T E_z + \eta_{TE} \nabla_T H_z) \\ \mathbf{H}_T &= -\frac{j\beta}{k_c^2} \left(\frac{1}{\eta_{TM}} \hat{\mathbf{z}} \times \nabla_T E_z + \nabla_T H_z \right)\end{aligned}\quad (9.1.15)$$

Using the result $\hat{\mathbf{z}} \times (\hat{\mathbf{z}} \times \mathbf{E}_T) = -\mathbf{E}_T$, we solve for $\mathbf{E}_T$ and $\mathbf{H}_T$:

$$\begin{aligned}\mathbf{E}_T &= -\frac{j\beta}{k_c^2} (\nabla_T E_z - \eta_{TE} \hat{\mathbf{z}} \times \nabla_T H_z) \\ \mathbf{H}_T &= -\frac{j\beta}{k_c^2} (\nabla_T H_z + \frac{1}{\eta_{TM}} \hat{\mathbf{z}} \times \nabla_T E_z)\end{aligned}\quad (\text{transverse fields}) \quad (9.1.16)$$

An alternative and useful way of writing these equations is to form the following linear combinations, which are equivalent to Eq. (9.1.6):

$$\begin{aligned}\mathbf{H}_T - \frac{1}{\eta_{TM}} \hat{\mathbf{z}} \times \mathbf{E}_T &= \frac{j}{\beta} \nabla_T H_z \\ \mathbf{E}_T - \eta_{TE} \mathbf{H}_T \times \hat{\mathbf{z}} &= \frac{j}{\beta} \nabla_T E_z\end{aligned}\quad (9.1.17)$$

So far we only used the first two of Maxwell's equations (9.1.5) and expressed $\mathbf{E}_T, \mathbf{H}_T$ in terms of E_z, H_z . Using (9.1.16), it is easily shown that the left-hand sides of the remaining four of Eqs. (9.1.5) take the forms:

$$\begin{aligned}\nabla_T \times \mathbf{E}_T + j\omega\mu \hat{\mathbf{z}} H_z &= \frac{j\omega\mu}{k_c^2} \hat{\mathbf{z}} (\nabla_T^2 H_z + k_c^2 H_z) \\ \nabla_T \times \mathbf{H}_T - j\omega\epsilon \hat{\mathbf{z}} E_z &= -\frac{j\omega\epsilon}{k_c^2} \hat{\mathbf{z}} (\nabla_T^2 E_z + k_c^2 E_z) \\ \nabla_T \cdot \mathbf{E}_T - j\beta E_z &= -\frac{j\beta}{k_c^2} (\nabla_T^2 E_z + k_c^2 E_z) \\ \nabla_T \cdot \mathbf{H}_T - j\beta H_z &= -\frac{j\beta}{k_c^2} (\nabla_T^2 H_z + k_c^2 H_z)\end{aligned}$$

where ∇_T^2 is the two-dimensional Laplacian operator:

$$\nabla_T^2 = \nabla_T \cdot \nabla_T = \partial_x^2 + \partial_y^2 \quad (9.1.18)$$

and we used the vectorial identities $\nabla_T \times \nabla_T E_z = 0$, $\nabla_T \times (\hat{z} \times \nabla_T H_z) = \hat{z} \nabla_T^2 H_z$, and $\nabla_T \cdot (\hat{z} \times \nabla_T H_z) = 0$.

It follows that in order to satisfy all of the last four of Maxwell's equations (9.1.5), it is necessary that the longitudinal fields $E_z(x, y)$, $H_z(x, y)$ satisfy the two-dimensional Helmholtz equations:

$$\begin{cases} \nabla_T^2 E_z + k_c^2 E_z = 0 \\ \nabla_T^2 H_z + k_c^2 H_z = 0 \end{cases} \quad (\text{Helmholtz equations}) \quad (9.1.19)$$

These equations are to be solved subject to the appropriate *boundary conditions* for each waveguide type. Once, the fields E_z, H_z are known, the transverse fields E_T, H_T are computed from Eq. (9.1.16), resulting in a complete solution of Maxwell's equations for the guiding structure. To get the full x, y, z, t dependence of the propagating fields, the above solutions must be multiplied by the factor $e^{j\omega t - j\beta z}$.

The cross-sections of practical waveguiding systems have either cartesian or cylindrical symmetry, such as the rectangular waveguide or the coaxial cable. Below, we summarize the form of the above solutions in the two types of coordinate systems.

Cartesian Coordinates

The cartesian component version of Eqs. (9.1.16) and (9.1.19) is straightforward. Using the identity $\hat{z} \times \nabla_T H_z = \hat{y} \partial_x H_z - \hat{x} \partial_y H_z$, we obtain for the longitudinal components:

$$\begin{cases} (\partial_x^2 + \partial_y^2) E_z + k_c^2 E_z = 0 \\ (\partial_x^2 + \partial_y^2) H_z + k_c^2 H_z = 0 \end{cases} \quad (9.1.20)$$

Eq. (9.1.16) becomes for the transverse components:

$$\begin{cases} E_x = -\frac{j\beta}{k_c^2} (\partial_x E_z + \eta_{TE} \partial_y H_z) \\ E_y = -\frac{j\beta}{k_c^2} (\partial_y E_z - \eta_{TE} \partial_x H_z) \end{cases}, \quad \begin{cases} H_x = -\frac{j\beta}{k_c^2} (\partial_x H_z - \frac{1}{\eta_{TM}} \partial_y E_z) \\ H_y = -\frac{j\beta}{k_c^2} (\partial_y H_z + \frac{1}{\eta_{TM}} \partial_x E_z) \end{cases} \quad (9.1.21)$$

Cylindrical Coordinates

The relationship between cartesian and cylindrical coordinates is shown in Fig. 9.1.1. From the triangle in the figure, we have $x = \rho \cos \phi$ and $y = \rho \sin \phi$. The transverse gradient and Laplace operator are in cylindrical coordinates:

$$\nabla_T = \hat{\rho} \frac{\partial}{\partial \rho} + \hat{\phi} \frac{1}{\rho} \frac{\partial}{\partial \phi}, \quad \nabla_T^2 = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2}{\partial \phi^2} \quad (9.1.22)$$

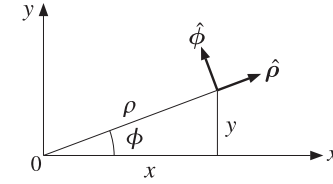


Fig. 9.1.1 Cylindrical coordinates.

The Helmholtz equations (9.1.19) now read:

$$\begin{cases} \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial E_z}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 E_z}{\partial \phi^2} + k_c^2 E_z = 0 \\ \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial H_z}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 H_z}{\partial \phi^2} + k_c^2 H_z = 0 \end{cases} \quad (9.1.23)$$

Noting that $\hat{z} \times \hat{\rho} = \hat{\phi}$ and $\hat{z} \times \hat{\phi} = -\hat{\rho}$, we obtain:

$$\hat{z} \times \nabla_T H_z = \hat{\phi} (\partial_\rho H_z) - \hat{\rho} \left(\frac{1}{\rho} \partial_\phi H_z \right)$$

The decomposition of a transverse vector is $E_T = \hat{\rho} E_\rho + \hat{\phi} E_\phi$. The cylindrical coordinates version of (9.1.16) are:

$$\begin{cases} E_\rho = -\frac{j\beta}{k_c^2} (\partial_\rho E_z - \eta_{TE} \frac{1}{\rho} \partial_\phi H_z) \\ E_\phi = -\frac{j\beta}{k_c^2} (\frac{1}{\rho} \partial_\phi E_z + \eta_{TE} \partial_\rho H_z) \end{cases}, \quad \begin{cases} H_\rho = -\frac{j\beta}{k_c^2} (\partial_\rho H_z + \frac{1}{\eta_{TM} \rho} \partial_\phi E_z) \\ H_\phi = -\frac{j\beta}{k_c^2} (\frac{1}{\rho} \partial_\phi H_z - \frac{1}{\eta_{TM}} \partial_\rho E_z) \end{cases} \quad (9.1.24)$$

For either coordinate system, the equations for H_T may be obtained from those of E_T by a so-called *duality transformation*, that is, making the substitutions:

$$E \rightarrow H, \quad H \rightarrow -E, \quad \epsilon \rightarrow \mu, \quad \mu \rightarrow \epsilon \quad (\text{duality transformation}) \quad (9.1.25)$$

These imply that $\eta \rightarrow \eta^{-1}$ and $\eta_{TE} \rightarrow \eta_{TM}^{-1}$. Duality is discussed in greater detail in Sec. 18.2.

9.2 Power Transfer and Attenuation

With the field solutions at hand, one can determine the amount of power transmitted along the guide, as well as the transmission losses. The total power carried by the fields along the guide direction is obtained by integrating the z -component of the Poynting vector over the cross-sectional area of the guide:

$$P_T = \int_S \mathcal{P}_z dS, \quad \text{where} \quad \mathcal{P}_z = \frac{1}{2} \operatorname{Re}(\mathbf{E} \times \mathbf{H}^*) \cdot \hat{\mathbf{z}} \quad (9.2.1)$$

It is easily verified that only the transverse components of the fields contribute to the power flow, that is, $\mathcal{P}_z$ can be written in the form:

$$\mathcal{P}_z = \frac{1}{2} \operatorname{Re}(\mathbf{E}_T \times \mathbf{H}_T^*) \cdot \hat{\mathbf{z}} \quad (9.2.2)$$

For waveguides with conducting walls, the transmission losses are due primarily to ohmic losses in (a) the conductors and (b) the dielectric medium filling the space between the conductors and in which the fields propagate. In dielectric waveguides, the losses are due to absorption and scattering by imperfections.

The transmission losses can be quantified by replacing the propagation wavenumber β by its complex-valued version $\beta_c = \beta - j\alpha$, where α is the attenuation constant. The z -dependence of all the field components is replaced by:

$$e^{-j\beta z} \rightarrow e^{-j\beta_c z} = e^{-(\alpha + j\beta)z} = e^{-\alpha z} e^{-j\beta z} \quad (9.2.3)$$

The quantity α is the sum of the attenuation constants arising from the various loss mechanisms. For example, if α_d and α_c are the attenuations due to the ohmic losses in the dielectric and in the conducting walls, then

$$\alpha = \alpha_d + \alpha_c \quad (9.2.4)$$

The ohmic losses in the dielectric can be characterized either by its loss tangent, say $\tan \delta$, or by its conductivity σ_d —the two being related by $\sigma_d = \omega \epsilon \tan \delta$. More generally, the effective dielectric constant of the medium may have a negative imaginary part ϵ_I that includes both conductive and polarization losses, $\epsilon(\omega) = \epsilon - j\epsilon_I$, with $\epsilon_I = \epsilon \tan \delta$. Then, the corresponding complex-valued wavenumber β_c is obtained by the replacement:

$$\beta = \sqrt{\omega^2 \mu \epsilon - k_c^2} \rightarrow \beta_c = \sqrt{\omega^2 \mu \epsilon(\omega) - k_c^2}$$

For weakly lossy dielectrics ($\epsilon_I \ll \epsilon$), we may make the approximation:

$$\beta_c = \sqrt{\omega^2 \mu (\epsilon - j\epsilon_I) - k_c^2} = \sqrt{\beta^2 - j\omega^2 \mu \epsilon_I} = \beta \sqrt{1 - j \frac{\omega^2 \mu \epsilon_I}{\beta^2}} \simeq \beta - j \frac{\omega^2 \mu \epsilon_I}{2\beta}$$

Resulting in the attenuation constant, after setting $\mu \epsilon = 1/c^2$ and $\beta c/\omega = \sqrt{1 - \omega_c^2/\omega^2}$,

$$\alpha_d = \frac{\omega^2 \mu \epsilon_I}{2\beta} = \frac{1}{2} \frac{\omega^2 \mu \epsilon}{\beta} \tan \delta = \frac{\omega \tan \delta}{2c \sqrt{1 - \omega_c^2/\omega^2}} \quad (\text{dielectric losses}) \quad (9.2.5)$$

The conductor losses are more complicated to calculate. In practice, the following approximate procedure is adequate. First, the fields are determined on the assumption that the conductors are perfect.

Second, the magnetic fields on the conductor surfaces are determined and the corresponding induced surface currents are calculated by $\mathbf{J}_s = \hat{\mathbf{n}} \times \mathbf{H}$, where $\hat{\mathbf{n}}$ is the outward normal to the conductor.

Third, the ohmic losses per unit conductor area are calculated by Eq. (2.8.7). Figure 9.2.1 shows such an infinitesimal conductor area $dA = dl dz$, where dl is along the cross-sectional periphery of the conductor. Applying Eq. (2.8.7) to this area, we have:

$$\frac{dP_{\text{loss}}}{dA} = \frac{dP_{\text{loss}}}{dl dz} = \frac{1}{2} R_s |\mathbf{J}_s|^2 \quad (9.2.6)$$

where R_s is the surface resistance of the conductor given by Eq. (2.8.4),

$$R_s = \sqrt{\frac{\omega \mu}{2\sigma}} = \eta \sqrt{\frac{\omega \epsilon}{2\sigma}} = \frac{1}{\sigma \delta}, \quad \delta = \sqrt{\frac{2}{\omega \mu \sigma}} = \text{skin depth} \quad (9.2.7)$$

Integrating Eq. (9.2.6) around the periphery of the conductor gives the power loss per unit z -length due to that conductor. Adding similar terms for all the other conductors gives the total power loss per unit z -length:

$$P'_{\text{loss}} = \frac{dP_{\text{loss}}}{dz} = \oint_{C_a} \frac{1}{2} R_s |\mathbf{J}_s|^2 dl + \oint_{C_b} \frac{1}{2} R_s |\mathbf{J}_s|^2 dl \quad (9.2.8)$$

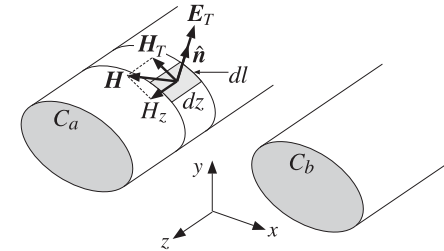


Fig. 9.2.1 Conductor surface absorbs power from the propagating fields.

where C_a and C_b indicate the peripheries of the conductors. Finally, the corresponding attenuation coefficient is calculated from Eq. (2.6.22):

$$\alpha_c = \frac{P'_{\text{loss}}}{2P_T} \quad (\text{conductor losses}) \quad (9.2.9)$$

Equations (9.2.1)–(9.2.9) provide a systematic methodology by which to calculate the transmitted power and attenuation losses in waveguides. We will apply it to several examples later on. Eq. (9.2.9) applies also to the dielectric losses so that in general P'_{loss} arises from two parts, one due to the dielectric and one due to the conducting walls,

$$\alpha = \frac{P'_{\text{loss}}}{2P_T} = \frac{P'_{\text{diel}} + P'_{\text{cond}}}{2P_T} = \alpha_d + \alpha_c \quad (\text{attenuation constant}) \quad (9.2.10)$$

Eq. (9.2.5) for α_d can also be derived directly from Eq. (9.2.10) by applying it separately to the TE and TM modes. We recall from Eq. (1.9.6) that the losses per unit volume in a dielectric medium, arising from both a conduction and polarization current, $\mathbf{J}_{\text{tot}} = \mathbf{J} + j\omega\mathbf{D}$, are given by,

$$\frac{dP_{\text{loss}}}{dV} = \frac{1}{2} \text{Re}[\mathbf{J}_{\text{tot}} \cdot \mathbf{E}^*] = \frac{1}{2} \omega \epsilon_I |\mathbf{E} \cdot \mathbf{E}^*|$$

Integrating over the cross-sectional area of the guide gives the dielectric loss per unit waveguide length (i.e., z-length),

$$P'_{\text{diel}} = \frac{1}{2} \omega \epsilon_I \int_S |\mathbf{E}|^2 dS$$

Applying this to the TE case, we find,

$$\begin{aligned} P'_{\text{diel}} &= \frac{1}{2} \omega \epsilon_I \int_S |\mathbf{E}|^2 dS = \frac{1}{2} \omega \epsilon_I \int_S |\mathbf{E}_T|^2 dS \\ P_T &= \int_S \frac{1}{2} \text{Re}(\mathbf{E}_T \times \mathbf{H}_T^*) \cdot \hat{\mathbf{z}} dS = \frac{1}{2\eta_{TE}} \int_S |\mathbf{E}_T|^2 dS = \frac{\beta}{2\omega\mu} \int_S |\mathbf{E}_T|^2 dS \\ \alpha_d &= \frac{P'_{\text{diel}}}{2P_T} = \frac{\omega^2 \mu \epsilon_I}{2\beta} \end{aligned}$$

The TM case is a bit more involved. Using Eq. (9.13.1) from Problem 9.11, we find, after using the result, $\beta^2 + k_c^2 = \omega^2 \mu \epsilon$,

$$\begin{aligned} P'_{\text{diel}} &= \frac{1}{2} \omega \epsilon_I \int_S |\mathbf{E}|^2 dS = \frac{1}{2} \omega \epsilon_I \int_S [|\mathbf{E}_z|^2 + |\mathbf{E}_T|^2] dS \\ &= \frac{1}{2} \omega \epsilon_I \int_S \left[|\mathbf{E}_z|^2 + \frac{\beta^2}{k_c^4} |\nabla_T \mathbf{E}_z|^2 \right] dS = \frac{1}{2} \omega \epsilon_I \left(1 + \frac{\beta^2}{k_c^2} \right) \int_S |\mathbf{E}_z|^2 dS \\ P_T &= \frac{1}{2\eta_{TM}} \int_S |\mathbf{E}_T|^2 dS = \frac{\omega \epsilon}{2\beta} \int_S \frac{\beta^2}{k_c^4} |\nabla_T \mathbf{E}_z|^2 dS = \frac{\omega \epsilon \beta}{2k_c^2} \int_S |\mathbf{E}_z|^2 dS \\ \alpha_d &= \frac{P'_{\text{diel}}}{2P_T} = \frac{\frac{1}{2} \omega \epsilon_I \left(1 + \frac{\beta^2}{k_c^2} \right)}{\frac{\omega \epsilon \beta}{2k_c^2}} = \frac{\omega^2 \mu \epsilon_I}{\beta} \end{aligned}$$

9.3 TEM, TE, and TM modes

The general solution described by Eqs. (9.1.16) and (9.1.19) is a hybrid solution with non-zero E_z and H_z components. Here, we look at the specialized forms of these equations in the cases of TEM, TE, and TM modes.

One common property of all three types of modes is that the transverse fields $\mathbf{E}_T, \mathbf{H}_T$ are related to each other in the same way as in the case of uniform plane waves propagating in the z-direction, that is, they are perpendicular to each other, their cross-product points in the z-direction, and they satisfy:

$$\mathbf{H}_T = \frac{1}{\eta_T} \hat{\mathbf{z}} \times \mathbf{E}_T \quad (9.3.1)$$

where η_T is the transverse impedance of the particular mode type, that is, $\eta, \eta_{TE}, \eta_{TM}$ in the TEM, TE, and TM cases.

Because of Eq. (9.3.1), the power flow per unit cross-sectional area described by the Poynting vector $\mathcal{P}_z$ of Eq. (9.2.2) takes the simple form in all three cases:

$$\mathcal{P}_z = \frac{1}{2} \text{Re}(\mathbf{E}_T \times \mathbf{H}_T^*) \cdot \hat{\mathbf{z}} = \frac{1}{2\eta_T} |\mathbf{E}_T|^2 = \frac{1}{2} \eta_T |\mathbf{H}_T|^2 \quad (9.3.2)$$

TEM modes

In TEM modes, both E_z and H_z vanish, and the fields are fully transverse. One can set $E_z = H_z = 0$ in Maxwell equations (9.1.5), or equivalently in (9.1.16), or in (9.1.17).

From any point view, one obtains the condition $k_c^2 = 0$, or $\omega = \beta c$. For example, if the right-hand sides of Eq. (9.1.17) vanish, the consistency of the system requires that $\eta_{TE} = \eta_{TM}$, which by virtue of Eq. (9.1.13) implies $\omega = \beta c$. It also implies that η_{TE}, η_{TM} must both be equal to the medium impedance η . Thus, the electric and magnetic fields satisfy:

$$\mathbf{H}_T = \frac{1}{\eta} \hat{\mathbf{z}} \times \mathbf{E}_T \quad (9.3.3)$$

These are the same as in the case of a uniform plane wave, except here the fields are not uniform and may have a non-trivial x, y dependence. The electric field $\mathbf{E}_T$ is determined from the rest of Maxwell's equations (9.1.5), which read:

$$\begin{aligned} \nabla_T \times \mathbf{E}_T &= 0 \\ \nabla_T \cdot \mathbf{E}_T &= 0 \end{aligned} \quad (9.3.4)$$

These are recognized as the field equations of an equivalent two-dimensional *electrostatic* problem. Once this electrostatic solution is found, $\mathbf{E}_T(x, y)$, the magnetic field is constructed from Eq. (9.3.3). The time-varying propagating fields will be given by Eq. (9.1.1), with $\omega = \beta c$. (For backward moving fields, replace β by $-\beta$.)

We explore this electrostatic point of view further in Sec. 11.1 and discuss the cases of the coaxial, two-wire, and strip lines. Because of the relationship between $\mathbf{E}_T$ and $\mathbf{H}_T$, the Poynting vector $\mathcal{P}_z$ of Eq. (9.2.2) will be:

$$\mathcal{P}_z = \frac{1}{2} \text{Re}(\mathbf{E}_T \times \mathbf{H}_T^*) \cdot \hat{\mathbf{z}} = \frac{1}{2\eta} |\mathbf{E}_T|^2 = \frac{1}{2} \eta |\mathbf{H}_T|^2 \quad (9.3.5)$$

TE modes

TE modes are characterized by the conditions $E_z = 0$ and $H_z \neq 0$. It follows from the second of Eqs. (9.1.17) that $\mathbf{E}_T$ is completely determined from $\mathbf{H}_T$, that is, $\mathbf{E}_T = \eta_{TE} \mathbf{H}_T \times \hat{\mathbf{z}}$.

The field $\mathbf{H}_T$ is determined from the second of (9.1.16). Thus, all field components for TE modes are obtained from the equations:

$$\begin{aligned} \nabla_T^2 H_z + k_c^2 H_z &= 0 \\ \mathbf{H}_T &= -\frac{j\beta}{k_c^2} \nabla_T H_z \\ \mathbf{E}_T &= \eta_{TE} \mathbf{H}_T \times \hat{\mathbf{z}} \end{aligned} \quad (\text{TE modes}) \quad (9.3.6)$$

The relationship of $\mathbf{E}_T$ and $\mathbf{H}_T$ is identical to that of uniform plane waves propagating in the z-direction, except the wave impedance is replaced by η_{TE} . The Poynting vector of Eq. (9.2.2) then takes the form:

$$\mathcal{P}_z = \frac{1}{2} \text{Re}(\mathbf{E}_T \times \mathbf{H}_T^*) \cdot \hat{\mathbf{z}} = \frac{1}{2\eta_{TE}} |\mathbf{E}_T|^2 = \frac{1}{2} \eta_{TE} |\mathbf{H}_T|^2 = \frac{1}{2} \eta_{TE} \frac{\beta^2}{k_c^4} |\nabla_T H_z|^2 \quad (9.3.7)$$

The cartesian coordinate version of Eq. (9.3.6) is:

$$\begin{aligned} (\partial_x^2 + \partial_y^2) H_z + k_c^2 H_z &= 0 \\ H_x &= -\frac{j\beta}{k_c^2} \partial_x H_z, \quad H_y = -\frac{j\beta}{k_c^2} \partial_y H_z \\ E_x &= \eta_{TE} H_y, \quad E_y = -\eta_{TE} H_x \end{aligned} \quad (9.3.8)$$

And, the cylindrical coordinate version:

$$\begin{aligned} \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial H_z}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 H_z}{\partial \phi^2} + k_c^2 H_z &= 0 \\ H_\rho &= -\frac{j\beta}{k_c^2} \frac{\partial H_z}{\partial \rho}, \quad H_\phi = -\frac{j\beta}{k_c^2} \frac{1}{\rho} \frac{\partial H_z}{\partial \phi} \\ E_\rho &= \eta_{TE} H_\phi, \quad E_\phi = -\eta_{TE} H_\rho \end{aligned} \quad (9.3.9)$$

where we used $\mathbf{H}_T \times \hat{\mathbf{z}} = (\hat{\boldsymbol{\rho}} H_\rho + \hat{\boldsymbol{\phi}} H_\phi) \times \hat{\mathbf{z}} = -\hat{\boldsymbol{\phi}} H_\rho + \hat{\boldsymbol{\rho}} H_\phi$.

TM modes

TM modes have $H_z = 0$ and $E_z \neq 0$. It follows from the first of Eqs. (9.1.17) that $\mathbf{H}_T$ is completely determined from $\mathbf{E}_T$, that is, $\mathbf{H}_T = \eta_{TM}^{-1} \hat{\mathbf{z}} \times \mathbf{E}_T$. The field $\mathbf{E}_T$ is determined from the first of (9.1.16), so that all field components for TM modes are obtained from the following equations, which are dual to the TE equations (9.3.6):

$$\begin{aligned} \nabla_T^2 E_z + k_c^2 E_z &= 0 \\ \mathbf{E}_T &= -\frac{j\beta}{k_c^2} \nabla_T E_z \\ \mathbf{H}_T &= \frac{1}{\eta_{TM}} \hat{\mathbf{z}} \times \mathbf{E}_T \end{aligned} \quad (\text{TM modes}) \quad (9.3.10)$$

Again, the relationship of $\mathbf{E}_T$ and $\mathbf{H}_T$ is identical to that of uniform plane waves propagating in the z-direction, but the wave impedance is now η_{TM} . The Poynting vector takes the form:

$$\mathcal{P}_z = \frac{1}{2} \text{Re}(\mathbf{E}_T \times \mathbf{H}_T^*) \cdot \hat{\mathbf{z}} = \frac{1}{2\eta_{TM}} |\mathbf{E}_T|^2 = \frac{1}{2\eta_{TM}} \frac{\beta^2}{k_c^4} |\nabla_T E_z|^2 \quad (9.3.11)$$

9.4 Rectangular Waveguides

Next, we discuss in detail the case of a rectangular hollow waveguide with conducting walls, as shown in Fig. 9.4.1. Without loss of generality, we may assume that the lengths a, b of the inner sides satisfy $b \leq a$. The guide is typically filled with air, but any other dielectric material ϵ, μ may be assumed.

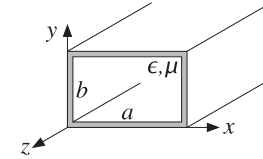


Fig. 9.4.1 Rectangular waveguide.

The simplest and dominant propagation mode is the so-called TE_{10} mode and depends only on the x-coordinate (of the longest side.) Therefore, we begin by looking for solutions of Eq. (9.3.8) that depend only on x. In this case, the Helmholtz equation reduces to:

$$\partial_x^2 H_z(x) + k_c^2 H_z(x) = 0$$

The most general solution is a linear combination of $\cos k_c x$ and $\sin k_c x$. However, only the former will satisfy the boundary conditions. Therefore, the solution is:

$$H_z(x) = H_0 \cos k_c x \quad (9.4.1)$$

where H_0 is a (complex-valued) constant. Because there is no y-dependence, it follows from Eq. (9.3.8) that $\partial_y H_z = 0$, and hence $H_y = 0$ and $E_x = 0$. It also follows that:

$$H_x(x) = -\frac{j\beta}{k_c^2} \partial_x H_z = -\frac{j\beta}{k_c^2} (-k_c) H_0 \sin k_c x = \frac{j\beta}{k_c} H_0 \sin k_c x \equiv H_1 \sin k_c x$$

Then, the corresponding electric field will be:

$$E_y(x) = -\eta_{TE} H_x(x) = -\eta_{TE} \frac{j\beta}{k_c} H_0 \sin k_c x \equiv E_0 \sin k_c x$$

where we defined the constants:

$$H_1 = \frac{j\beta}{k_c} H_0 \quad (9.4.2)$$

$$E_0 = -\eta_{TE} H_1 = -\eta_{TE} \frac{j\beta}{k_c} H_0 = -j\eta \frac{\omega}{\omega_c} H_0$$

where we used $\eta_{TE} = \eta\omega/\beta c$. In summary, the non-zero field components are:

$$\begin{array}{|l} H_z(x) = H_0 \cos k_c x \\ H_x(x) = H_1 \sin k_c x \\ E_y(x) = E_0 \sin k_c x \end{array} \Rightarrow \begin{array}{|l} H_z(x, y, z, t) = H_0 \cos k_c x e^{j\omega t - j\beta z} \\ H_x(x, y, z, t) = H_1 \sin k_c x e^{j\omega t - j\beta z} \\ E_y(x, y, z, t) = E_0 \sin k_c x e^{j\omega t - j\beta z} \end{array} \quad (9.4.3)$$

Assuming perfectly conducting walls, the boundary conditions require that there be no tangential electric field at any of the wall sides. Because the electric field is in the y -direction, it is normal to the top and bottom sides. But, it is parallel to the left and right sides. On the left side, $x = 0$, $E_y(x)$ vanishes because $\sin k_c x$ does. On the right side, $x = a$, the boundary condition requires:

$$E_y(a) = E_0 \sin k_c a = 0 \Rightarrow \sin k_c a = 0$$

This requires that $k_c a$ be an integral multiple of π :

$$k_c a = n\pi \Rightarrow k_c = \frac{n\pi}{a} \quad (9.4.4)$$

These are the so-called TE_{n0} modes. The corresponding cutoff frequency $\omega_c = ck_c$, $f_c = \omega_c/2\pi$, and wavelength $\lambda_c = 2\pi/k_c = c/f_c$ are:

$$\omega_c = \frac{cn\pi}{a}, \quad f_c = \frac{cn}{2a}, \quad \lambda_c = \frac{2a}{n} \quad (TE_{n0} \text{ modes}) \quad (9.4.5)$$

The dominant mode is the one with the lowest cutoff frequency or the longest cutoff wavelength, that is, the mode TE_{10} having $n = 1$. It has:

$$k_c = \frac{\pi}{a}, \quad \omega_c = \frac{c\pi}{a}, \quad f_c = \frac{c}{2a}, \quad \lambda_c = 2a \quad (TE_{10} \text{ mode}) \quad (9.4.6)$$

Fig. 9.4.2 depicts the electric field $E_y(x) = E_0 \sin k_c x = E_0 \sin(\pi x/a)$ of this mode as a function of x .

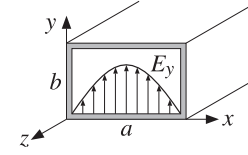


Fig. 9.4.2 Electric field inside a rectangular waveguide.

9.5 Higher TE and TM modes

To construct higher modes, we look for solutions of the Helmholtz equation that are factorable in their x and y dependence:

$$H_z(x, y) = F(x)G(y)$$

Then, Eq. (9.3.8) becomes:

$$F''(x)G(y) + F(x)G''(y) + k_c^2 F(x)G(y) = 0 \Rightarrow \frac{F''(x)}{F(x)} + \frac{G''(y)}{G(y)} + k_c^2 = 0 \quad (9.5.1)$$

Because these must be valid for all x, y (inside the guide), the F - and G -terms must be constants, independent of x and y . Thus, we write:

$$\frac{F''(x)}{F(x)} = -k_x^2, \quad \frac{G''(y)}{G(y)} = -k_y^2 \quad \text{or}$$

$$F''(x) + k_x^2 F(x) = 0, \quad G''(y) + k_y^2 G(y) = 0 \quad (9.5.2)$$

where the constants k_x^2 and k_y^2 are constrained from Eq. (9.5.1) to satisfy:

$$k_c^2 = k_x^2 + k_y^2 \quad (9.5.3)$$

The most general solutions of (9.5.2) that will satisfy the TE boundary conditions are $\cos k_x x$ and $\cos k_y y$. Thus, the longitudinal magnetic field will be:

$$H_z(x, y) = H_0 \cos k_x x \cos k_y y \quad (TE_{nm} \text{ modes}) \quad (9.5.4)$$

It then follows from the rest of the equations (9.3.8) that:

$$\begin{array}{|l} H_x(x, y) = H_1 \sin k_x x \cos k_y y \\ H_y(x, y) = H_2 \cos k_x x \sin k_y y \end{array} \quad \begin{array}{|l} E_x(x, y) = E_1 \cos k_x x \sin k_y y \\ E_y(x, y) = E_2 \sin k_x x \cos k_y y \end{array} \quad (9.5.5)$$

where we defined the constants:

$$H_1 = \frac{j\beta k_x}{k_c^2} H_0, \quad H_2 = \frac{j\beta k_y}{k_c^2} H_0$$

$$E_1 = \eta_{TE} H_2 = j\eta \frac{\omega k_y}{\omega_c k_c} H_0, \quad E_2 = -\eta_{TE} H_1 = -j\eta \frac{\omega k_x}{\omega_c k_c} H_0$$

The boundary conditions are that E_y vanish on the right wall, $x = a$, and that E_x vanish on the top wall, $y = b$, that is,

$$E_y(a, y) = E_{0y} \sin k_x a \cos k_y y = 0, \quad E_x(x, b) = E_{0x} \cos k_x x \sin k_y b = 0$$

The conditions require that $k_x a$ and $k_y b$ be integral multiples of π :

$$k_x a = n\pi, \quad k_y b = m\pi \quad \Rightarrow \quad \boxed{k_x = \frac{n\pi}{a}, \quad k_y = \frac{m\pi}{b}} \quad (9.5.6)$$

These correspond to the TE_{nm} modes. Thus, the cutoff wavenumbers of these modes $k_c = \sqrt{k_x^2 + k_y^2}$ take on the quantized values:

$$\boxed{k_c = \sqrt{\left(\frac{n\pi}{a}\right)^2 + \left(\frac{m\pi}{b}\right)^2}} \quad (TE_{nm} \text{ modes}) \quad (9.5.7)$$

The cutoff frequencies $f_{nm} = \omega_c/2\pi = ck_c/2\pi$ and wavelengths $\lambda_{nm} = c/f_{nm}$ are:

$$f_{nm} = c \sqrt{\left(\frac{n}{2a}\right)^2 + \left(\frac{m}{2b}\right)^2}, \quad \lambda_{nm} = \frac{1}{\sqrt{\left(\frac{n}{2a}\right)^2 + \left(\frac{m}{2b}\right)^2}} \quad (9.5.8)$$

The TE_{0m} modes are similar to the TE_{n0} modes, but with x and a replaced by y and b . The family of TM modes can also be constructed in a similar fashion from Eq. (9.3.10).

Assuming $E_z(x, y) = F(x)G(y)$, we obtain the same equations (9.5.2). Because E_z is parallel to all walls, we must now choose the solutions $\sin k_x$ and $\sin k_y$. Thus, the longitudinal electric fields is:

$$\boxed{E_z(x, y) = E_0 \sin k_x x \sin k_y y} \quad (TM_{nm} \text{ modes}) \quad (9.5.9)$$

The rest of the field components can be worked out from Eq. (9.3.10) and one finds that they are given by the same expressions as (9.5.5), except now the constants are determined in terms of E_0 :

$$E_1 = -\frac{j\beta k_x}{k_c^2} E_0, \quad E_2 = -\frac{j\beta k_y}{k_c^2} E_0$$

$$H_1 = -\frac{1}{\eta_{TM}} E_2 = \frac{j\omega k_y}{\omega_c k_c} \frac{1}{\eta} E_0, \quad H_2 = \frac{1}{\eta_{TM}} E_1 = -\frac{j\omega k_x}{\omega_c k_c} \frac{1}{\eta} H_0$$

where we used $\eta_{TM} = \eta\beta c/\omega$. The boundary conditions on E_x, E_y are the same as before, and in addition, we must require that E_z vanish on all walls.

These conditions imply that k_x, k_y will be given by Eq. (9.5.6), except both n and m must be non-zero (otherwise E_z would vanish identically.) Thus, the cutoff frequencies and wavelengths are the same as in Eq. (9.5.8).

Waveguide modes can be excited by inserting small probes at the beginning of the waveguide. The probes are chosen to generate an electric field that resembles the field of the desired mode.

9.6 Operating Bandwidth

All waveguiding systems are operated in a frequency range that ensures that only the *lowest* mode can propagate. If several modes can propagate simultaneously,[†] one has no control over which modes will actually be carrying the transmitted signal. This may cause undue amounts of dispersion, distortion, and erratic operation.

A mode with cutoff frequency ω_c will propagate only if its frequency is $\omega \geq \omega_c$, or $\lambda < \lambda_c$. If $\omega < \omega_c$, the wave will attenuate exponentially along the guide direction. This follows from the ω, β relationship (9.1.10):

$$\omega^2 = \omega_c^2 + \beta^2 c^2 \quad \Rightarrow \quad \beta^2 = \frac{\omega^2 - \omega_c^2}{c^2}$$

If $\omega \geq \omega_c$, the wavenumber β is real-valued and the wave will propagate. But if $\omega < \omega_c$, β becomes imaginary, say, $\beta = -j\alpha$, and the wave will attenuate in the z -direction, with a penetration depth $\delta = 1/\alpha$:

$$e^{-j\beta z} = e^{-\alpha z}$$

If the frequency ω is greater than the cutoff frequencies of several modes, then all of these modes can propagate. Conversely, if ω is less than all cutoff frequencies, then none of the modes can propagate.

If we arrange the cutoff frequencies in increasing order, $\omega_{c1} < \omega_{c2} < \omega_{c3} < \dots$, then, to ensure single-mode operation, the frequency must be restricted to the interval $\omega_{c1} < \omega < \omega_{c2}$, so that only the lowest mode will propagate. This interval defines the *operating bandwidth* of the guide.

These remarks apply to all waveguiding systems, not just hollow conducting waveguides. For example, in coaxial cables the lowest mode is the TEM mode having no cutoff frequency, $\omega_{c1} = 0$. However, TE and TM modes with non-zero cutoff frequencies do exist and place an upper limit on the usable bandwidth of the TEM mode. Similarly, in optical fibers, the lowest mode has no cutoff, and the single-mode bandwidth is determined by the next cutoff frequency.

In rectangular waveguides, the smallest cutoff frequencies are $f_{10} = c/2a$, $f_{20} = c/a = 2f_{10}$, and $f_{01} = c/2b$. Because we assumed that $b \leq a$, it follows that always $f_{10} \leq f_{01}$. If $b \leq a/2$, then $1/a \leq 1/2b$ and therefore, $f_{20} \leq f_{01}$, so that the two lowest cutoff frequencies are f_{10} and f_{20} .

On the other hand, if $a/2 \leq b \leq a$, then $f_{01} \leq f_{20}$ and the two smallest frequencies are f_{10} and f_{01} (except when $b = a$, in which case $f_{01} = f_{10}$ and the smallest frequencies are f_{10} and f_{20} .) The two cases $b \leq a/2$ and $b \geq a/2$ are depicted in Fig. 9.6.1.

It is evident from this figure that in order to achieve the *widest* possible usable bandwidth for the TE_{10} mode, the guide dimensions must satisfy $b \leq a/2$ so that the bandwidth is the interval $[f_c, 2f_c]$, where $f_c = f_{10} = c/2a$. In terms of the wavelength $\lambda = c/f$, the operating bandwidth becomes: $0.5 \leq a/\lambda \leq 1$, or, $a \leq \lambda \leq 2a$.

We will see later that the total amount of transmitted power in this mode is proportional to the cross-sectional area of the guide, ab . Thus, if in addition to having the

[†]Murphy's law for waveguides states that "if a mode can propagate, it will."

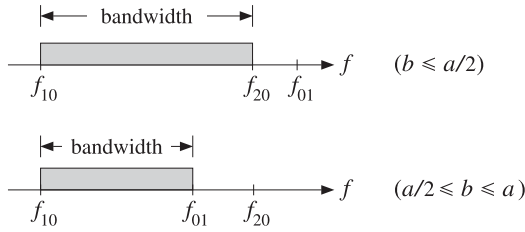


Fig. 9.6.1 Operating bandwidth in rectangular waveguides.

widest bandwidth, we also require to have the *maximum power* transmitted, the dimension b must be chosen to be as large as possible, that is, $b = a/2$. Most practical guides follow these side proportions.

If there is a “canonical” guide, it will have $b = a/2$ and be operated at a frequency that lies in the middle of the operating band $[f_c, 2f_c]$, that is,

$$f = 1.5f_c = 0.75 \frac{c}{a} \quad (9.6.1)$$

Table 9.6.1 lists some standard air-filled rectangular waveguides with their naming designations, inner side dimensions a, b in inches, cutoff frequencies in GHz, minimum and maximum recommended operating frequencies in GHz, power ratings, and attenuations in dB/m (the power ratings and attenuations are representative over each operating band.) We have chosen one example from each microwave band.

name	a	b	f_c	$f_{\min}$	$f_{\max}$	band	P	α
WR-510	5.10	2.55	1.16	1.45	2.20	L	9 MW	0.007
WR-284	2.84	1.34	2.08	2.60	3.95	S	2.7 MW	0.019
WR-159	1.59	0.795	3.71	4.64	7.05	C	0.9 MW	0.043
WR-90	0.90	0.40	6.56	8.20	12.50	X	250 kW	0.110
WR-62	0.622	0.311	9.49	11.90	18.00	Ku	140 kW	0.176
WR-42	0.42	0.17	14.05	17.60	26.70	K	50 kW	0.370
WR-28	0.28	0.14	21.08	26.40	40.00	Ka	27 kW	0.583
WR-15	0.148	0.074	39.87	49.80	75.80	V	7.5 kW	1.52
WR-10	0.10	0.05	59.01	73.80	112.00	W	3.5 kW	2.74

Table 9.6.1 Characteristics of some standard air-filled rectangular waveguides.

9.7 Power Transfer, Energy Density, and Group Velocity

Next, we calculate the time-averaged power transmitted in the TE_{10} mode. We also calculate the energy density of the fields and determine the velocity by which electromagnetic energy flows down the guide and show that it is equal to the group velocity. We recall that the non-zero field components are:

$$H_z(x) = H_0 \cos k_c x, \quad H_x(x) = H_1 \sin k_c x, \quad E_y(x) = E_0 \sin k_c x \quad (9.7.1)$$

where

$$H_1 = \frac{j\beta}{k_c} H_0, \quad E_0 = -\eta_{TE} H_1 = -j\eta \frac{\omega}{\omega_c} H_0 \quad (9.7.2)$$

The Poynting vector is obtained from the general result of Eq. (9.3.7):

$$\mathcal{P}_z = \frac{1}{2\eta_{TE}} |E_T|^2 = \frac{1}{2\eta_{TE}} |E_y(x)|^2 = \frac{1}{2\eta_{TE}} |E_0|^2 \sin^2 k_c x$$

The transmitted power is obtained by integrating $\mathcal{P}_z$ over the cross-sectional area of the guide:

$$P_T = \int_0^a \int_0^b \frac{1}{2\eta_{TE}} |E_0|^2 \sin^2 k_c x dx dy$$

Noting the definite integral,

$$\int_0^a \sin^2 k_c x dx = \int_0^a \sin^2 \left(\frac{\pi x}{a} \right) dx = \frac{a}{2} \quad (9.7.3)$$

and using $\eta_{TE} = \eta \omega / \beta c = \eta / \sqrt{1 - \omega_c^2 / \omega^2}$, we obtain:

$$P_T = \frac{1}{4\eta_{TE}} |E_0|^2 ab = \frac{1}{4\eta} |E_0|^2 ab \sqrt{1 - \frac{\omega_c^2}{\omega^2}} \quad (\text{transmitted power}) \quad (9.7.4)$$

We may also calculate the distribution of electromagnetic energy along the guide, as measured by the time-averaged energy density. The energy densities of the electric and magnetic fields are:

$$w_e = \frac{1}{2} \text{Re} \left(\frac{1}{2} \epsilon \mathbf{E} \cdot \mathbf{E}^* \right) = \frac{1}{4} \epsilon |E_y|^2$$

$$w_m = \frac{1}{2} \text{Re} \left(\frac{1}{2} \mu \mathbf{H} \cdot \mathbf{H}^* \right) = \frac{1}{4} \mu (|H_x|^2 + |H_z|^2)$$

Inserting the expressions for the fields, we find:

$$w_e = \frac{1}{4} \epsilon |E_0|^2 \sin^2 k_c x, \quad w_m = \frac{1}{4} \mu (|H_1|^2 \sin^2 k_c x + |H_0|^2 \cos^2 k_c x)$$

Because these quantities represent the energy per unit volume, if we integrate them over the cross-sectional area of the guide, we will obtain the energy distributions per unit z -length. Using the integral (9.7.3) and an identical one for the cosine case, we find:

$$W'_e = \int_0^a \int_0^b w_e(x, y) dx dy = \int_0^a \int_0^b \frac{1}{4} \epsilon |E_0|^2 \sin^2 k_c x dx dy = \frac{1}{8} \epsilon |E_0|^2 ab$$

$$W'_m = \int_0^a \int_0^b \frac{1}{4} \mu (|H_1|^2 \sin^2 k_c x + |H_0|^2 \cos^2 k_c x) dx dy = \frac{1}{8} \mu (|H_1|^2 + |H_0|^2) ab$$

Although these expressions look different, they are actually equal, $W'_e = W'_m$. Indeed, using the property $\beta^2/k_c^2 + 1 = (\beta^2 + k_c^2)/k_c^2 = k^2/k_c^2 = \omega^2/\omega_c^2$ and the relationships between the constants in (9.7.1), we find:

$$\mu(|H_1|^2 + |H_0|^2) = \mu(|H_0|^2 \frac{\beta^2}{k_c^2} + |H_0|^2) = \mu|H_0|^2 \frac{\omega^2}{\omega_c^2} = \frac{\mu}{\eta^2} |E_0|^2 = \epsilon |E_0|^2$$

The equality of the electric and magnetic energies is a general property of waveguiding systems. We also encountered it in Sec. 2.3 for uniform plane waves. The total energy density per unit length will be:

$$W' = W'_e + W'_m = 2W'_e = \frac{1}{4} \epsilon |E_0|^2 ab \quad (9.7.5)$$

According to the general relationship between flux, density, and transport velocity given in Eq. (1.6.2), the energy transport velocity will be the ratio $v_{en} = P_T/W'$. Using Eqs. (9.7.4) and (9.7.5) and noting that $1/\eta\epsilon = 1/\sqrt{\mu\epsilon} = c$, we find:

$$v_{en} = \frac{P_T}{W'} = c \sqrt{1 - \frac{\omega_c^2}{\omega^2}} \quad (\text{energy transport velocity}) \quad (9.7.6)$$

This is equal to the *group velocity* of the propagating mode. For any dispersion relationship between ω and β , the group and phase velocities are defined by

$$v_{gr} = \frac{d\omega}{d\beta}, \quad v_{ph} = \frac{\omega}{\beta} \quad (\text{group and phase velocities}) \quad (9.7.7)$$

For uniform plane waves and TEM transmission lines, we have $\omega = \beta c$, so that $v_{gr} = v_{ph} = c$. For a rectangular waveguide, we have $\omega^2 = \omega_c^2 + \beta^2 c^2$. Taking differentials of both sides, we find $2\omega d\omega = 2c^2 \beta d\beta$, which gives:

$$v_{gr} = \frac{d\omega}{d\beta} = \frac{\beta c^2}{\omega} = c \sqrt{1 - \frac{\omega_c^2}{\omega^2}} \quad (9.7.8)$$

where we used Eq. (9.1.10). Thus, the energy transport velocity is equal to the group velocity, $v_{en} = v_{gr}$. We note that $v_{gr} = \beta c^2/\omega = c^2/v_{ph}$, or

$$v_{gr} v_{ph} = c^2 \quad (9.7.9)$$

The energy or group velocity satisfies $v_{gr} \leq c$, whereas $v_{ph} \geq c$. Information transmission down the guide is by the group velocity and, consistent with the theory of relativity, it is less than c .

9.8 Power Attenuation

In this section, we calculate the attenuation coefficient due to the ohmic losses of the conducting walls following the procedure outlined in Sec. 9.2. The losses due to the filling dielectric can be determined from Eq. (9.2.5).

The field expressions (9.4.3) were derived assuming the boundary conditions for perfectly conducting wall surfaces. The induced surface currents on the inner walls of the waveguide are given by $\mathbf{J}_s = \hat{\mathbf{n}} \times \mathbf{H}$, where the unit vector $\hat{\mathbf{n}}$ is $\pm \hat{\mathbf{x}}$ and $\pm \hat{\mathbf{y}}$ on the left/right and bottom/top walls, respectively.

The surface currents and tangential magnetic fields are shown in Fig. 9.8.1. In particular, on the bottom and top walls, we have:

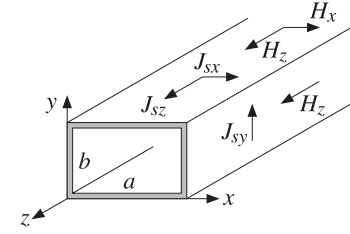


Fig. 9.8.1 Currents on waveguide walls.

$$\mathbf{J}_s = \pm \hat{\mathbf{y}} \times \mathbf{H} = \pm \hat{\mathbf{y}} \times (\hat{\mathbf{x}} H_x + \hat{\mathbf{z}} H_z) = \pm (-\hat{\mathbf{z}} H_x + \hat{\mathbf{x}} H_z) = \pm (-\hat{\mathbf{z}} H_1 \sin k_c x + \hat{\mathbf{x}} H_0 \cos k_c x)$$

Similarly, on the left and right walls:

$$\mathbf{J}_s = \pm \hat{\mathbf{x}} \times \mathbf{H} = \pm \hat{\mathbf{x}} \times (\hat{\mathbf{x}} H_x + \hat{\mathbf{z}} H_z) = \mp \hat{\mathbf{y}} H_z = \mp \hat{\mathbf{y}} H_0 \cos k_c x$$

At $x = 0$ and $x = a$, this gives $\mathbf{J}_s = \mp \hat{\mathbf{y}} (\pm H_0) = \hat{\mathbf{y}} H_0$. Thus, the magnitudes of the surface currents are on the four walls:

$$|\mathbf{J}_s|^2 = \begin{cases} |H_0|^2, & (\text{left and right walls}) \\ |H_0|^2 \cos^2 k_c x + |H_1|^2 \sin^2 k_c x, & (\text{top and bottom walls}) \end{cases}$$

The power loss per unit z -length is obtained from Eq. (9.2.8) by integrating $|\mathbf{J}_s|^2$ around the four walls, that is,

$$\begin{aligned} P'_{\text{loss}} &= 2 \frac{1}{2} R_s \int_0^a |\mathbf{J}_s|^2 dx + 2 \frac{1}{2} R_s \int_0^b |\mathbf{J}_s|^2 dy \\ &= R_s \int_0^a (|H_0|^2 \cos^2 k_c x + |H_1|^2 \sin^2 k_c x) dx + R_s \int_0^b |H_0|^2 dy \\ &= R_s \frac{a}{2} (|H_0|^2 + |H_1|^2) + R_s b |H_0|^2 = \frac{R_s a}{2} (|H_0|^2 + |H_1|^2 + \frac{2b}{a} |H_0|^2) \end{aligned}$$

Using $|H_0|^2 + |H_1|^2 = |E_0|^2/\eta^2$ from Sec. 9.7, and $|H_0|^2 = (|E_0|^2/\eta^2) \omega_c^2/\omega^2$, which follows from Eq. (9.4.2), we obtain:

$$P'_{\text{loss}} = \frac{R_s a |E_0|^2}{2 \eta^2} \left(1 + \frac{2b}{a} \frac{\omega_c^2}{\omega^2} \right)$$

The attenuation constant is computed from Eqs. (9.2.9) and (9.7.4):

$$\alpha_c = \frac{P'_{\text{loss}}}{2P_T} = \frac{\frac{R_s a |E_0|^2}{2\eta^2} \left(1 + \frac{2b}{a} \frac{\omega_c^2}{\omega^2}\right)}{2 \frac{1}{4\eta} |E_0|^2 ab \sqrt{1 - \frac{\omega_c^2}{\omega^2}}}$$

which gives:

$$\alpha_c = \frac{R_s}{\eta b} \frac{\left(1 + \frac{2b}{a} \frac{\omega_c^2}{\omega^2}\right)}{\sqrt{1 - \frac{\omega_c^2}{\omega^2}}} \quad (\text{attenuation of TE}_{10} \text{ mode}) \quad (9.8.1)$$

This is in units of nepers/m. Its value in dB/m is obtained by $\alpha_{\text{dB}} = 8.686\alpha_c$. For a given ratio a/b , α_c increases with decreasing b , thus the smaller the guide dimensions, the larger the attenuation. This trend is noted in Table 9.6.1.

The main tradeoffs in a waveguiding system are that as the operating frequency f increases, the dimensions of the guide must decrease in order to maintain the operating band $f_c \leq f \leq 2f_c$, but then the attenuation increases and the transmitted power decreases as it is proportional to the guide's area.

Example 9.8.1: Design a rectangular air-filled waveguide to be operated at 5 GHz, then, re-design it to be operated at 10 GHz. The operating frequency must lie in the middle of the operating band. Calculate the guide dimensions, the attenuation constant in dB/m, and the maximum transmitted power assuming the maximum electric field is one-half of the dielectric strength of air. Assume copper walls with conductivity $\sigma = 5.8 \times 10^7$ S/m.

Solution: If f is in the middle of the operating band, $f_c \leq f \leq 2f_c$, where $f_c = c/2a$, then $f = 1.5f_c = 0.75c/a$. Solving for a , we find

$$a = \frac{0.75c}{f} = \frac{0.75 \times 30 \text{ GHz cm}}{5} = 4.5 \text{ cm}$$

For maximum power transfer, we require $b = a/2 = 2.25$ cm. Because $\omega = 1.5\omega_c$, we have $\omega_c/\omega = 2/3$. Then, Eq. (9.8.1) gives $\alpha_c = 0.037$ dB/m. The dielectric strength of air is 3 MV/m. Thus, the maximum allowed electric field in the guide is $E_0 = 1.5$ MV/m. Then, Eq. (9.7.4) gives $P_T = 1.12$ MW.

At 10 GHz, because f is doubled, the guide dimensions are halved, $a = 2.25$ and $b = 1.125$ cm. Because R_s depends on f like $f^{1/2}$, it will increase by a factor of $\sqrt{2}$. Then, the factor R_s/b will increase by a factor of $2\sqrt{2}$. Thus, the attenuation will increase to the value $\alpha_c = 0.037 \cdot 2\sqrt{2} = 0.104$ dB/m. Because the area ab is reduced by a factor of four, so will the power, $P_T = 1.12/4 = 0.28$ MW = 280 kW.

The results of these two cases are consistent with the values quoted in Table 9.6.1 for the C-band and X-band waveguides, WR-159 and WR-90. $\square$

Example 9.8.2: WR-159 Waveguide. Consider the C-band WR-159 air-filled waveguide whose characteristics were listed in Table 9.6.1. Its inner dimensions are $a = 1.59$ and $b = a/2 = 0.795$ inches, or, equivalently, $a = 4.0386$ and $b = 2.0193$ cm.

The cutoff frequency of the TE₁₀ mode is $f_c = c/2a = 3.71$ GHz. The maximum operating bandwidth is the interval $[f_c, 2f_c] = [3.71, 7.42]$ GHz, and the recommended interval is $[4.64, 7.05]$ GHz.

Assuming copper walls with conductivity $\sigma = 5.8 \times 10^7$ S/m, the calculated attenuation constant α_c from Eq. (9.8.1) is plotted in dB/m versus frequency in Fig. 9.8.2.

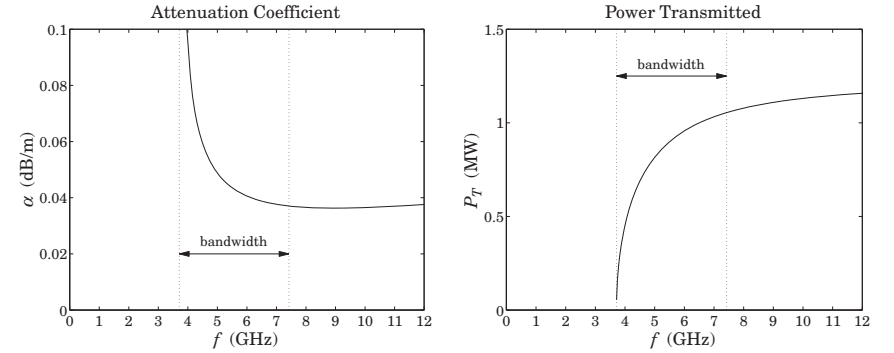


Fig. 9.8.2 Attenuation constant and transmitted power in a WR-159 waveguide.

The power transmitted P_T is calculated from Eq. (9.7.4) assuming a maximum breakdown voltage of $E_0 = 1.5$ MV/m, which gives a safety factor of two over the dielectric breakdown of air of 3 MV/m. The power in megawatt scales is plotted in Fig. 9.8.2.

Because of the factor $\sqrt{1 - \omega_c^2/\omega^2}$ in the denominator of α_c and the numerator of P_T , the attenuation constant becomes very large near the cutoff frequency, while the power is almost zero. A physical explanation of this behavior is given in the next section. $\square$

9.9 Reflection Model of Waveguide Propagation

An intuitive model for the TE₁₀ mode can be derived by considering a TE-polarized uniform plane wave propagating in the z -direction by obliquely bouncing back and forth between the left and right walls of the waveguide, as shown in Fig. 9.9.1.

If θ is the angle of incidence, then the incident and reflected (from the right wall) wavevectors will be:

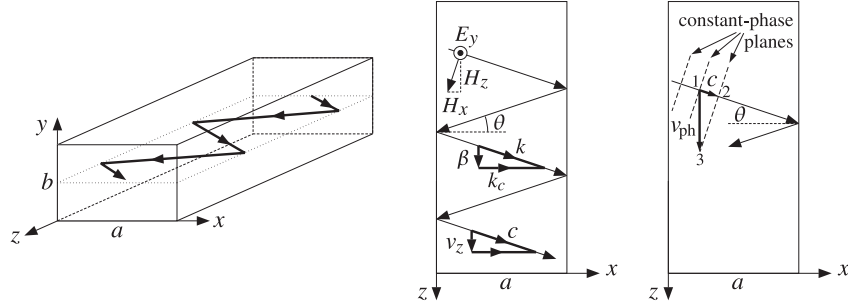
$$\mathbf{k} = \hat{\mathbf{x}} k_x + \hat{\mathbf{z}} k_z = \hat{\mathbf{x}} k \cos \theta + \hat{\mathbf{z}} k \sin \theta$$

$$\mathbf{k}' = -\hat{\mathbf{x}} k_x + \hat{\mathbf{z}} k_z = -\hat{\mathbf{x}} k \cos \theta + \hat{\mathbf{z}} k \sin \theta$$

The electric and magnetic fields will be the sum of an incident and a reflected component of the form:

$$\mathbf{E} = \hat{\mathbf{y}} E_1 e^{-jk \cdot \mathbf{r}} + \hat{\mathbf{y}} E'_1 e^{-jk' \cdot \mathbf{r}} = \hat{\mathbf{y}} E_1 e^{-jk_x x} e^{-jk_z z} + \hat{\mathbf{y}} E'_1 e^{jk_x x} e^{-jk_z z} = \mathbf{E}_1 + \mathbf{E}'_1$$

$$\mathbf{H} = \frac{1}{\eta} \hat{\mathbf{k}} \times \mathbf{E}_1 + \frac{1}{\eta} \hat{\mathbf{k}}' \times \mathbf{E}'_1$$

Fig. 9.9.1 Reflection model of TE₁₀ mode.

where the electric field was taken to be polarized in the y direction. These field expressions become component-wise:

$$\begin{aligned} E_y &= (E_1 e^{-jk_x x} + E'_1 e^{jk_x x}) e^{-jk_z z} \\ H_x &= -\frac{1}{\eta} \sin \theta (E_1 e^{-jk_x x} + E'_1 e^{jk_x x}) e^{-jk_z z} \\ H_z &= \frac{1}{\eta} \cos \theta (E_1 e^{-jk_x x} - E'_1 e^{jk_x x}) e^{-jk_z z} \end{aligned} \quad (9.9.1)$$

The boundary condition on the left wall, $x = 0$, requires that $E_1 + E'_1 = 0$. We may write therefore, $E'_1 = -E_1 = jE_0/2$. Then, the above expressions simplify into:

$$\begin{aligned} E_y &= E_0 \sin k_x x e^{-jk_z z} \\ H_x &= -\frac{1}{\eta} \sin \theta E_0 \sin k_x x e^{-jk_z z} \\ H_z &= \frac{j}{\eta} \cos \theta E_0 \cos k_x x e^{-jk_z z} \end{aligned} \quad (9.9.2)$$

These are identical to Eq. (9.4.3) provided we identify β with k_z and k_c with k_x , as shown in Fig. 9.9.1. It follows from the wavevector triangle in the figure that the angle of incidence θ will be given by $\cos \theta = k_x/k = k_c/k$, or,

$$\cos \theta = \frac{\omega_c}{\omega}, \quad \sin \theta = \sqrt{1 - \frac{\omega_c^2}{\omega^2}} \quad (9.9.3)$$

The ratio of the transverse components, $-E_y/H_x$, is the transverse impedance, which is recognized to be η_{TE} . Indeed, we have:

$$\eta_{TE} = -\frac{E_y}{H_x} = \frac{\eta}{\sin \theta} = \frac{\eta}{\sqrt{1 - \frac{\omega_c^2}{\omega^2}}} \quad (9.9.4)$$

The boundary condition on the right wall requires $\sin k_x a = 0$, which gives rise to the same condition as (9.4.4), that is, $k_c a = n\pi$.

This model clarifies also the meaning of the group velocity. The plane wave is bouncing left and right with the speed of light c . However, the component of this velocity in the z -direction will be $v_z = c \sin \theta$. This is equal to the group velocity. Indeed, it follows from Eq. (9.9.3) that:

$$v_z = c \sin \theta = c \sqrt{1 - \frac{\omega_c^2}{\omega^2}} = v_{gr} \quad (9.9.5)$$

Eq. (9.9.3) implies also that at $\omega = \omega_c$, we have $\sin \theta = 0$, or $\theta = 0$, that is, the wave is bouncing left and right at normal incidence, creating a standing wave, and does not propagate towards the z -direction. Thus, the transmitted power is zero and this also implies, through Eq. (9.2.9), that α_c will be infinite.

On the other hand, for very large frequencies, $\omega \gg \omega_c$, the angle θ will tend to 90° , causing the wave to zoom through guide almost at the speed of light.

The phase velocity can also be understood geometrically. Indeed, we observe in the rightmost illustration of the above figure that the planes of constant phase are moving obliquely with the speed of light c . From the indicated triangle at points 1, 2, 3, we see that the effective speed in the z -direction of the common-phase points will be $v_{ph} = c / \sin \theta$ so that $v_{ph} v_{gr} = c^2$.

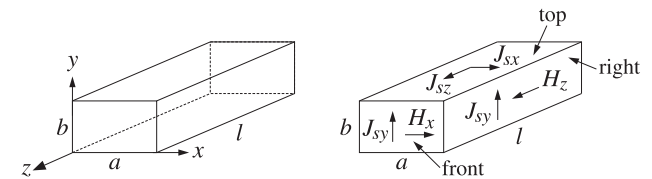
Higher TE and TM modes can also be given similar geometric interpretations in terms of plane waves propagating by bouncing off the waveguide walls [890].

9.10 Resonant Cavities

Cavity resonators are metallic enclosures that can trap electromagnetic fields. The boundary conditions on the cavity walls force the fields to exist only at certain quantized resonant frequencies. For highly conducting walls, the resonances are extremely sharp, having a very high Q of the order of 10,000.

Because of their high Q , cavities can be used not only to efficiently store electromagnetic energy at microwave frequencies, but also to act as precise oscillators and to perform precise frequency measurements.

Fig. 9.10.1 shows a rectangular cavity with z -length equal to l formed by replacing the sending and receiving ends of a waveguide by metallic walls. A forward-moving wave will bounce back and forth from these walls, resulting in a standing-wave pattern along the z -direction.

Fig. 9.10.1 Rectangular cavity resonator (and induced wall currents for the TE_{*n0p*} mode.)

Because the tangential components of the electric field must vanish at the end-walls, these walls must coincide with zero crossings of the standing wave, or put differently, an integral multiple of half-wavelengths must fit along the z -direction, that is, $l = p\lambda_g/2 = p\pi/\beta$, or $\beta = p\pi/l$, where p is a non-zero integer. For the same reason, the standing-wave patterns along the transverse directions require $a = n\lambda_x/2$ and $b = m\lambda_y/2$, or $k_x = n\pi/a$ and $k_y = m\pi/b$. Thus, all three cartesian components of the wave vector are quantized, and therefore, so is the frequency of the wave $\omega = c\sqrt{k_x^2 + k_y^2 + \beta^2}$:

$$\omega_{nmp} = c\sqrt{\left(\frac{n\pi}{a}\right)^2 + \left(\frac{m\pi}{b}\right)^2 + \left(\frac{p\pi}{l}\right)^2} \quad (\text{resonant frequencies}) \quad (9.10.1)$$

Such modes are designated as TE_{nmp} or TM_{nmp} . For simplicity, we consider the case TE_{n0p} . Eqs. (9.3.6) also describe backward-moving waves if one replaces β by $-\beta$, which also changes the sign of $\eta_{TE} = \eta\omega/\beta c$. Starting with a linear combination of forward and backward waves in the TE_{n0} mode, we obtain the field components:

$$\begin{aligned} H_z(x, z) &= H_0 \cos k_c x (Ae^{-j\beta z} + Be^{j\beta z}), \\ H_x(x, z) &= jH_1 \sin k_c x (Ae^{-j\beta z} - Be^{j\beta z}), \quad H_1 = \frac{\beta}{k_c} H_0 \\ E_y(x, z) &= -jE_0 \sin k_c x (Ae^{-j\beta z} + Be^{j\beta z}), \quad E_0 = \frac{\omega}{\omega_c} \eta H_0 \end{aligned} \quad (9.10.2)$$

where $\omega_c = ck_c$. By requiring that $E_y(x, z)$ have z -dependence of the form $\sin \beta z$, the coefficients A, B must be chosen as $A = -B = j/2$. Then, Eq. (9.10.2) specializes into:

$$\begin{aligned} H_z(x, z) &= H_0 \cos k_c x \sin \beta z, \\ H_x(x, z) &= -H_1 \sin k_c x \cos \beta z, \quad H_1 = \frac{\beta}{k_c} H_0 \\ E_y(x, z) &= -jE_0 \sin k_c x \sin \beta z, \quad E_0 = \frac{\omega}{\omega_c} \eta H_0 \end{aligned} \quad (9.10.3)$$

As expected, the vanishing of $E_y(x, z)$ on the front/back walls, $z = 0$ and $z = l$, and on the left/right walls, $x = 0$ and $x = a$, requires the quantization conditions: $\beta = p\pi/l$ and $k_c = n\pi/a$. The Q of the resonator can be calculated from its definition:

$$Q = \omega \frac{W}{P_{\text{loss}}} \quad (9.10.4)$$

where W is the total time-averaged energy stored within the cavity volume and P_{loss} is the total power loss due to the wall ohmic losses (plus other losses, such as dielectric losses, if present.) The ratio $\Delta\omega = P_{\text{loss}}/W$ is usually identified as the 3-dB width of the resonance centered at frequency ω . Therefore, we may write $Q = \omega/\Delta\omega$.

It is easily verified that the electric and magnetic energies are equal, therefore, W may be calculated by integrating the electric energy density over the cavity volume:

$$\begin{aligned} W &= 2W_e = 2 \frac{1}{4} \int_{\text{vol}} \epsilon |E_y(x, z)|^2 dx dy dz = \frac{1}{2} \epsilon |E_0|^2 \int_0^a \int_0^b \int_0^l \sin^2 k_c x \cos^2 \beta z dx dy dz \\ &= \frac{1}{8} \epsilon |E_0|^2 (abl) = \frac{1}{8} \mu |H_0|^2 \frac{\omega^2}{\omega_c^2} (abl) = \frac{1}{8} \mu |H_0|^2 \left[\frac{k_c^2 + \beta^2}{k_c^2} \right] (abl) \end{aligned}$$

where we used the following definite integrals (valid because $k_c = n\pi/a$, $\beta = p\pi/l$):

$$\int_0^a \sin^2 k_c x dx = \int_0^a \cos^2 k_c x dx = \frac{a}{2}, \quad \int_0^l \sin^2 \beta z dz = \int_0^l \cos^2 \beta z dz = \frac{l}{2} \quad (9.10.5)$$

The ohmic losses are calculated from Eq. (9.2.6), integrated over all six cavity sides. The surface currents induced on the walls are related to the tangential magnetic fields by $\mathbf{J}_s = \hat{\mathbf{n}} \times \mathbf{H}_{\text{tan}}$. The directions of these currents are shown in Fig. 9.10.1. Specifically, we find for the currents on the six sides:

$$|\mathbf{J}_s|^2 = \begin{cases} H_0^2 \sin^2 \beta z & (\text{left \& right}) \\ H_0^2 \cos^2 k_c x \sin^2 \beta z + H_1^2 \sin^2 k_c x \cos^2 \beta z & (\text{top \& bottom}) \\ H_1^2 \sin^2 k_c x & (\text{front \& back}) \end{cases}$$

The power loss can be computed by integrating the loss per unit conductor area, Eq. (9.2.6), over the six wall sides, or doubling the answer for the left, top, and front sides. Using the integrals (9.10.5), we find:

$$\begin{aligned} P_{\text{loss}} &= \frac{1}{2} R_s \int_{\text{walls}} |\mathbf{J}_s|^2 dA = R_s \left[H_0^2 \frac{bl}{2} + (H_0^2 + H_1^2) \frac{al}{4} + H_1^2 \frac{ab}{2} \right] \\ &= \frac{1}{4} R_s H_0^2 \left[l(2b + a) + \frac{\beta^2}{k_c^2} a(2b + l) \right] \end{aligned} \quad (9.10.6)$$

where we substituted $H_1^2 = H_0^2 \beta^2 / k_c^2$. It follows that the Q -factor will be:

$$Q = \omega \frac{W}{P_{\text{loss}}} = \frac{\omega \mu}{2R_s} \frac{(k_c^2 + \beta^2)(abl)}{k_c^2 l(2b + a) + \beta^2 a(2b + l)}$$

For the TE_{n0p} mode we have $\beta = p\pi/l$ and $k_c = n\pi/a$. Using Eq. (9.2.7) to replace R_s in terms of the skin depth δ , we find:

$$Q = \frac{1}{\delta} \frac{\frac{n^2}{a^2} + \frac{p^2}{l^2}}{\frac{n^2}{a^2} \left(\frac{2}{a} + \frac{1}{b} \right) + \frac{p^2}{l^2} \left(\frac{2}{l} + \frac{1}{b} \right)} \quad (9.10.7)$$

The lowest resonant frequency corresponds to $n = p = 1$. For a cubic cavity, $a = b = l$, the Q and the lowest resonant frequency are:

$$Q = \frac{a}{3\delta}, \quad \omega_{101} = \frac{c\pi\sqrt{2}}{a}, \quad f_{101} = \frac{\omega}{2\pi} = \frac{c}{a\sqrt{2}} \quad (9.10.8)$$

For an air-filled cubic cavity with $a = 3$ cm, we find $f_{101} = 7.07$ GHz, $\delta = 7.86 \times 10^{-5}$ cm, and $Q = 12724$. As in waveguides, cavities can be excited by inserting small probes that generate fields resembling a particular mode.

9.11 Dielectric Slab Waveguides

A dielectric slab waveguide is a planar dielectric sheet or thin film of some thickness, say $2a$, as shown in Fig. 9.11.1. Wave propagation in the z -direction is by total internal

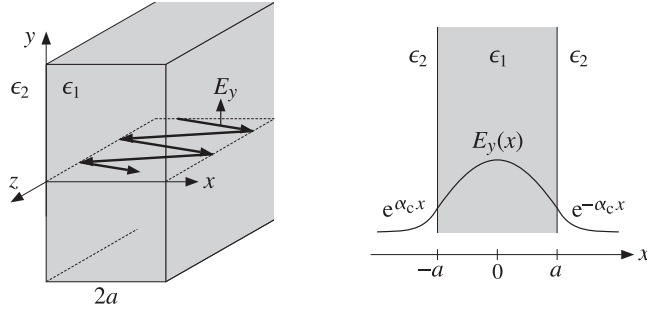


Fig. 9.11.1 Dielectric slab waveguide.

reflection from the left and right walls of the slab. Such waveguides provide simple models for the confining mechanism of waves propagating in optical fibers.

The propagating fields are confined primarily inside the slab, however, they also exist as evanescent waves outside it, decaying exponentially with distance from the slab. Fig. 9.11.1 shows a typical electric field pattern as a function of x .

For simplicity, we assume that the media to the left and right of the slab are the same. To guarantee total internal reflection, the dielectric constants inside and outside the slab must satisfy $\epsilon_1 > \epsilon_2$, and similarly for the refractive indices, $n_1 > n_2$.

We only consider TE modes and look for solutions that depend only on the x coordinate. The cutoff wavenumber k_c appearing in the Helmholtz equation for $H_z(x)$ depends on the dielectric constant of the propagation medium, $k_c^2 = \omega^2 \epsilon \mu - \beta^2$. Therefore, k_c^2 takes different values inside and outside the guide:

$$\begin{aligned} k_{c1}^2 &= \omega^2 \epsilon_1 \mu_0 - \beta^2 = \omega^2 \epsilon_0 \mu_0 n_1^2 - \beta^2 = k_0^2 n_1^2 - \beta^2 & (\text{inside}) \\ k_{c2}^2 &= \omega^2 \epsilon_2 \mu_0 - \beta^2 = \omega^2 \epsilon_0 \mu_0 n_2^2 - \beta^2 = k_0^2 n_2^2 - \beta^2 & (\text{outside}) \end{aligned} \quad (9.11.1)$$

where $k_0 = \omega/c_0$ is the free-space wavenumber. We note that ω, β are the same inside and outside the guide. This follows from matching the tangential fields at all times t and all points z along the slab walls. The corresponding Helmholtz equations in the regions inside and outside the guide are:

$$\begin{aligned} \partial_x^2 H_z(x) + k_{c1}^2 H_z(x) &= 0 & \text{for } |x| \leq a \\ \partial_x^2 H_z(x) + k_{c2}^2 H_z(x) &= 0 & \text{for } |x| \geq a \end{aligned} \quad (9.11.2)$$

Inside the slab, the solutions are $\sin k_{c1}x$ and $\cos k_{c1}x$, and outside, $\sin k_{c2}x$ and $\cos k_{c2}x$, or equivalently, $e^{\pm jk_{c2}x}$. In order for the waves to remain confined in the near vicinity of the slab, the quantity k_{c2} must be *imaginary*, for if it is real, the fields would propagate at large x distances from the slab (they would correspond to the rays refracted from the inside into the outside.)

If we set $k_{c2} = -j\alpha_c$, the solutions outside will be $e^{\pm \alpha_c x}$. If α_c is positive, then only the solution $e^{-\alpha_c x}$ is physically acceptable to the right of the slab, $x \geq a$, and only $e^{\alpha_c x}$ to the left, $x \leq -a$. Thus, the fields attenuate exponentially with the transverse distance

x , and exist effectively within a skin depth distance $1/\alpha_c$ from the slab. Setting $k_{c1} = k_c$ and $k_{c2} = -j\alpha_c$, Eqs. (9.11.1) become in this new notation:

$$\begin{aligned} k_c^2 &= k_0^2 n_1^2 - \beta^2 \\ -\alpha_c^2 &= k_0^2 n_2^2 - \beta^2 \end{aligned} \quad \Rightarrow \quad \begin{aligned} k_c^2 &= k_0^2 n_1^2 - \beta^2 \\ \alpha_c^2 &= \beta^2 - k_0^2 n_2^2 \end{aligned} \quad (9.11.3)$$

Similarly, Eqs. (9.11.2) read:

$$\begin{aligned} \partial_x^2 H_z(x) + k_c^2 H_z(x) &= 0 & \text{for } |x| \leq a \\ \partial_x^2 H_z(x) - \alpha_c^2 H_z(x) &= 0 & \text{for } |x| \geq a \end{aligned} \quad (9.11.4)$$

The two solutions $\sin k_c x$ and $\cos k_c x$ inside the guide give rise to the so-called even and odd TE modes (referring to the evenness or oddness of the resulting electric field.) For the *even modes*, the solutions of Eqs. (9.11.4) have the form:

$$H_z(x) = \begin{cases} H_1 \sin k_c x, & \text{if } -a \leq x \leq a \\ H_2 e^{-\alpha_c x}, & \text{if } x \geq a \\ H_3 e^{\alpha_c x}, & \text{if } x \leq -a \end{cases} \quad (9.11.5)$$

The corresponding x -components H_x are obtained by applying Eq. (9.3.8) using the appropriate value for k_c^2 , that is, $k_{c2}^2 = -\alpha_c^2$ outside and $k_{c1}^2 = k_c^2$ inside:

$$H_x(x) = \begin{cases} -\frac{j\beta}{k_c^2} \partial_x H_z(x) = -\frac{j\beta}{k_c} H_1 \cos k_c x, & \text{if } -a \leq x \leq a \\ -\frac{j\beta}{-\alpha_c^2} \partial_x H_z(x) = -\frac{j\beta}{\alpha_c} H_2 e^{-\alpha_c x}, & \text{if } x \geq a \\ -\frac{j\beta}{-\alpha_c^2} \partial_x H_z(x) = \frac{j\beta}{\alpha_c} H_3 e^{\alpha_c x}, & \text{if } x \leq -a \end{cases} \quad (9.11.6)$$

The electric fields are $E_y(x) = -\eta_{TE} H_x(x)$, where $\eta_{TE} = \omega \mu_0 / \beta$ is the *same* inside and outside the slab. Thus, the electric field has the form:

$$E_y(x) = \begin{cases} E_1 \cos k_c x, & \text{if } -a \leq x \leq a \\ E_2 e^{-\alpha_c x}, & \text{if } x \geq a \\ E_3 e^{\alpha_c x}, & \text{if } x \leq -a \end{cases} \quad (\text{even TE modes}) \quad (9.11.7)$$

where we defined the constants:

$$E_1 = \frac{j\beta}{k_c} \eta_{TE} H_1, \quad E_2 = \frac{j\beta}{\alpha_c} \eta_{TE} H_2, \quad E_3 = -\frac{j\beta}{\alpha_c} \eta_{TE} H_3 \quad (9.11.8)$$

The boundary conditions state that the tangential components of the magnetic and electric fields, that is, H_z, E_y , are continuous across the dielectric interfaces at $x = -a$ and $x = a$. Similarly, the normal components of the magnetic field $B_x = \mu_0 H_x$ and therefore also H_x must be continuous. Because $E_y = -\eta_{TE} H_x$ and η_{TE} is the same in both media, the continuity of E_y follows from the continuity of H_x . The continuity of H_z at $x = a$ and $x = -a$ implies that:

$$H_1 \sin k_c a = H_2 e^{-\alpha_c a} \quad \text{and} \quad -H_1 \sin k_c a = H_3 e^{-\alpha_c a} \quad (9.11.9)$$

Similarly, the continuity of H_x implies (after canceling a factor of $-j\beta$):

$$\frac{1}{k_c} H_1 \cos k_c a = \frac{1}{\alpha_c} H_2 e^{-\alpha_c a} \quad \text{and} \quad \frac{1}{k_c} H_1 \cos k_c a = -\frac{1}{\alpha_c} H_3 e^{-\alpha_c a} \quad (9.11.10)$$

Eqs. (9.11.9) and (9.11.10) imply:

$$H_2 = -H_3 = H_1 e^{\alpha_c a} \sin k_c a = H_1 e^{\alpha_c a} \frac{\alpha_c}{k_c} \cos k_c a \quad (9.11.11)$$

Similarly, we find for the electric field constants:

$$E_2 = E_3 = E_1 e^{\alpha_c a} \cos k_c a = E_1 e^{\alpha_c a} \frac{k_c}{\alpha_c} \sin k_c a \quad (9.11.12)$$

The consistency of the last equations in (9.11.11) or (9.11.12) requires that:

$$\cos k_c a = \frac{k_c}{\alpha_c} \sin k_c a \quad \Rightarrow \quad \alpha_c = k_c \tan k_c a \quad (9.11.13)$$

For the *odd* TE modes, we have for the solutions of Eq. (9.11.4):

$$H_z(x) = \begin{cases} H_1 \cos k_c x, & \text{if } -a \leq x \leq a \\ H_2 e^{-\alpha_c x}, & \text{if } x \geq a \\ H_3 e^{\alpha_c x}, & \text{if } x \leq -a \end{cases} \quad (9.11.14)$$

The resulting electric field is:

$$E_y(x) = \begin{cases} E_1 \sin k_c x, & \text{if } -a \leq x \leq a \\ E_2 e^{-\alpha_c x}, & \text{if } x \geq a \\ E_3 e^{\alpha_c x}, & \text{if } x \leq -a \end{cases} \quad (\text{odd TE modes}) \quad (9.11.15)$$

The boundary conditions imply in this case:

$$H_2 = H_3 = H_1 e^{\alpha_c a} \cos k_c a = -H_1 e^{\alpha_c a} \frac{\alpha_c}{k_c} \sin k_c a \quad (9.11.16)$$

and, for the electric field constants:

$$E_2 = -E_3 = E_1 e^{\alpha_c a} \sin k_c a = -E_1 e^{\alpha_c a} \frac{k_c}{\alpha_c} \cos k_c a \quad (9.11.17)$$

The consistency of the last equation requires:

$$\alpha_c = -k_c \cot k_c a \quad (9.11.18)$$

We note that the electric fields $E_y(x)$ given by Eqs. (9.11.7) and (9.11.15) are even or odd functions of x for the two families of modes. Expressing E_2 and E_3 in terms of E_1 , we summarize the forms of the electric fields in the two cases:

$$E_y(x) = \begin{cases} E_1 \cos k_c x, & \text{if } -a \leq x \leq a \\ E_1 \cos k_c a e^{-\alpha_c(x-a)}, & \text{if } x \geq a \\ E_1 \cos k_c a e^{\alpha_c(x+a)}, & \text{if } x \leq -a \end{cases} \quad (\text{even TE modes}) \quad (9.11.19)$$

$$E_y(x) = \begin{cases} E_1 \sin k_c x, & \text{if } -a \leq x \leq a \\ E_1 \sin k_c a e^{-\alpha_c(x-a)}, & \text{if } x \geq a \\ -E_1 \sin k_c a e^{\alpha_c(x+a)}, & \text{if } x \leq -a \end{cases} \quad (\text{odd TE modes}) \quad (9.11.20)$$

Given the operating frequency ω , Eqs. (9.11.3) and (9.11.13) or (9.11.18) provide three equations in the three unknowns k_c, α_c, β . To solve them, we add the two equations (9.11.3) to eliminate β :

$$\alpha_c^2 + k_c^2 = k_0^2 (n_1^2 - n_2^2) = \frac{\omega^2}{c_0^2} (n_1^2 - n_2^2) \quad (9.11.21)$$

Next, we discuss the numerical solutions of these equations. Defining the dimensionless quantities $u = k_c a$ and $v = \alpha_c a$, we may rewrite Eqs. (9.11.13), (9.11.18), and (9.11.21) in the equivalent forms:

$$\begin{aligned} v &= u \tan u & v &= -u \cot u \\ v^2 + u^2 &= R^2 & v^2 + u^2 &= R^2 \end{aligned} \quad (\text{even modes}), \quad (\text{odd modes}) \quad (9.11.22)$$

where R is the normalized frequency variable:

$$R = k_0 a N_A = \frac{\omega a}{c_0} N_A = \frac{2\pi f a}{c_0} N_A = \frac{2\pi a}{\lambda} N_A \quad (9.11.23)$$

where $N_A = \sqrt{n_1^2 - n_2^2}$ is the *numerical aperture* of the slab and $\lambda = c_0/f$, the free-space wavelength.

Because the functions $\tan u$ and $\cot u$ have many branches, there may be several possible solution pairs u, v for each value of R . These solutions are obtained at the intersections of the curves $v = u \tan u$ and $v = -u \cot u$ with the circle of radius R , that is, $v^2 + u^2 = R^2$. Fig. 9.11.2 shows the solutions for various values of the radius R corresponding to various values of ω .

It is evident from the figure that for small enough R , that is, $0 \leq R < \pi/2$, there is only one solution and it is even.[†] For $\pi/2 \leq R < \pi$, there are two solutions, one even and one odd. For $\pi \leq R < 3\pi/2$, there are three solutions, two even and one odd, and

[†]for an optical fiber, the single-mode condition reads $2\pi a N_A / \lambda < 2.405$, where a is the core radius.

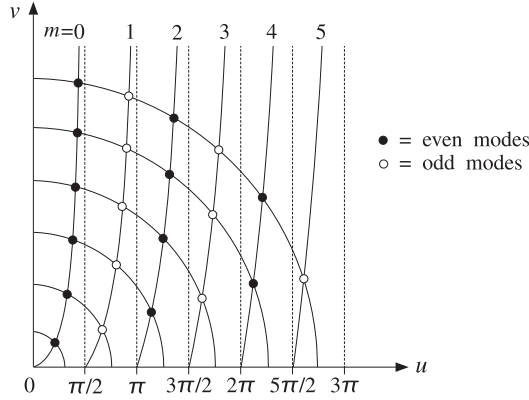


Fig. 9.11.2 Even and odd TE modes at different frequencies.

so on. In general, there will be $M + 1$ solutions, alternating between even and odd, if R falls in the interval:

$$\frac{M\pi}{2} \leq R < \frac{(M+1)\pi}{2} \quad (9.11.24)$$

Given a value of R , we determine M as that integer satisfying Eq. (9.11.24), or, $M \leq 2R/\pi < M + 1$, that is, the largest integer less than $2R/\pi$:

$$M = \text{floor}\left(\frac{2R}{\pi}\right) \quad (\text{maximum mode number}) \quad (9.11.25)$$

Then, there will be $M + 1$ solutions indexed by $m = 0, 1, \dots, M$, which will correspond to even modes if m is even and to odd modes if m is odd. The $M + 1$ branches of $\tan u$ and $\cot u$ being intersected by the R -circle are those contained in the u -ranges:

$$R_m \leq u < R_{m+1}, \quad m = 0, 1, \dots, M \quad (9.11.26)$$

where

$$R_m = \frac{m\pi}{2}, \quad m = 0, 1, \dots, M \quad (9.11.27)$$

If m is even, the u -range (9.11.26) defines a branch of $\tan u$, and if m is odd, a branch of $\cot u$. We can combine the even and odd cases of Eq. (9.11.22) into a single case by noting the identity:

$$\tan(u - R_m) = \begin{cases} \tan u, & \text{if } m \text{ is even} \\ -\cot u, & \text{if } m \text{ is odd} \end{cases} \quad (9.11.28)$$

This follows from the trigonometric identity:

$$\tan(u - m\pi/2) = \frac{\sin u \cos(m\pi/2) - \cos u \sin(m\pi/2)}{\cos u \cos(m\pi/2) + \sin u \sin(m\pi/2)}$$

Therefore, to find the m th mode, whether even or odd, we must find the unique solution of the following system in the u -range $R_m \leq u < R_{m+1}$:

$$\begin{cases} v = u \tan(u - R_m) \\ v^2 + u^2 = R^2 \end{cases} \quad (m\text{th mode}) \quad (9.11.29)$$

If one had an approximate solution u, v for the m th mode, one could refine it by using Newton's method, which converges very fast provided it is close to the true solution. Just such an approximate solution, accurate to within one percent of the true solution, was given by Lotspeich [930]. Without going into the detailed justification of this method, the approximation is as follows:

$$u = R_m + w_1(m)u_1(m) + w_2(m)u_2(m), \quad m = 0, 1, \dots, M \quad (9.11.30)$$

where $u_1(m)$, $u_2(m)$ are approximate solutions near and far from the cutoff R_m , and $w_1(m)$, $w_2(m)$ are weighting factors:

$$\begin{aligned} u_1(m) &= \frac{\sqrt{1 + 2R(R - R_m)} - 1}{R}, & u_2(m) &= \frac{\pi}{2} \frac{R - m}{R + 1} \\ w_1(m) &= \exp(-(R - R_m)^2/V_m^2), & w_2(m) &= 1 - w_1(m) \end{aligned} \quad (9.11.31)$$

$$V_m = \frac{1}{\sqrt{\ln 1.25}} \left(\frac{\pi/4 + R_m}{\cos(\pi/4)} - R_m \right)$$

This solution serves as the starting point to Newton's iteration for solving the equation $F(u) = 0$, where $F(u)$ is defined by

$$F(u) = u \tan(u - R_m) - v = u \tan(u - R_m) - \sqrt{R^2 - u^2} \quad (9.11.32)$$

Newton's iteration is:

$$\begin{aligned} &\text{for } i = 1, 2, \dots, N_{\text{it}} \text{ do:} \\ &\quad u = u - \frac{F(u)}{G(u)} \end{aligned} \quad (9.11.33)$$

where $G(u)$ is the derivative $F'(u)$, correct to order $O(F)$:

$$G(u) = \frac{v}{u} + \frac{u}{v} + \frac{R^2}{u^2} \quad (9.11.34)$$

The solution steps defined in Eqs. (9.11.29)–(9.11.34) have been implemented in the MATLAB function `ds1ab.m`, with usage:

```
[u,v,err] = ds1ab(R,Nit); % TE-mode cutoff wavenumbers in a dielectric slab
```

where N_{it} is the desired number of Newton iterations (9.11.33), err is the value of $F(u)$ at the end of the iterations, and u, v are the $(M + 1)$ -dimensional vectors of solutions. The number of iterations is typically very small, $N_{it} = 2-3$.

The related MATLAB function `dguide.m` uses `ds1ab` to calculate the solution parameters β, k_c, α_c , given the frequency f , the half-length a , and the refractive indices n_1, n_2 of the slab. It has usage:

```
[be, kc, ac, fc, err] = dguide(f, a, n1, n2, Nit); % dielectric slab guide
```

where f is in GHz, a in cm, and β, k_c, α_c in cm^{-1} . The quantity f_c is the vector of the $M + 1$ cutoff frequencies defined by the branch edges $R_m = m\pi/2$, that is, $R_m = \omega_m a N_A / c_0 = 2\pi f_m a N_A / c_0 = m\pi/2$, or,

$$f_m = \frac{mc_0}{4aN_A}, \quad m = 0, 1, \dots, M \quad (9.11.35)$$

The meaning of f_m is that there are $m + 1$ propagating modes for each f in the interval $f_m \leq f < f_{m+1}$.

Example 9.11.1: Dielectric Slab Waveguide. Determine the propagating TE modes of a dielectric slab of half-length $a = 0.5$ cm at frequency $f = 30$ GHz. The refractive indices of the slab and the surrounding dielectric are $n_1 = 2$ and $n_2 = 1$.

Solution: The solution is obtained by the MATLAB call:

```
f = 30; a = 0.5; n1 = 2; n2 = 1; Nit = 3;
[be, kc, ac, fc, err] = dguide(f, a, n1, n2, Nit)
```

The frequency radius is $R = 5.4414$, which gives $2R/\pi = 3.4641$, and therefore, $M = 3$. The resulting solutions, depicted in Fig. 9.11.3, are as follows:

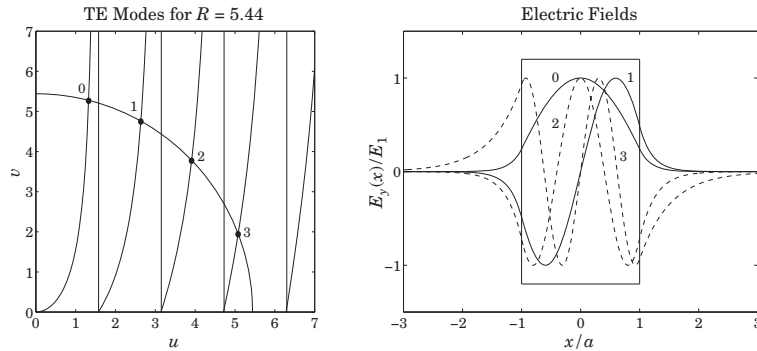


Fig. 9.11.3 TE modes and corresponding E -field patterns.

m	u	v	β	k_c	α_c	f_m
0	1.3248	5.2777	12.2838	2.6497	10.5553	0.0000
1	2.6359	4.7603	11.4071	5.2718	9.5207	8.6603
2	3.9105	3.7837	9.8359	7.8210	7.5675	17.3205
3	5.0793	1.9519	7.3971	10.1585	3.9037	25.9808

The cutoff frequencies f_m are in GHz. We note that as the mode number m increases, the quantity α_c decreases and the effective skin depth $1/\alpha_c$ increases, causing the fields outside the slab to be less confined. The electric field patterns are also shown in the figure as functions of x .

The approximation error, err , is found to be 4.885×10^{-15} using only three Newton iterations. Using two, one, and no (the Lotspeich approximation) iterations would result in the errors 2.381×10^{-8} , 4.029×10^{-4} , and 0.058.

The lowest non-zero cutoff frequency is $f_1 = 8.6603$ GHz, implying that there will be a single solution if f is in the interval $0 \leq f < f_1$. For example, if $f = 5$ GHz, the solution is $\beta = 1.5649$ rad/cm, $k_c = 1.3920$ rad/cm, and $\alpha_c = 1.1629$ nepers/cm.

The frequency range over which there are only four solutions is $[25.9808, 34.6410]$ GHz, where the upper limit is $4f_1$.

We note that the function `dguide` assumes internally that $c_0 = 30$ GHz cm, and therefore, the calculated values for k_c, α_c would be slightly different if a more precise value of c_0 is used, such as 29.9792458 of Appendix A. Problem 9.13 studies the sensitivity of the solutions to small changes of the parameters f, a, c_0, n_1, n_2 . $\square$

In terms of the ray picture of the propagating wave, the angles of total internal reflection are quantized according to the values of the propagation wavenumber β for the various modes.

If we denote by $k_1 = k_0 n_1$ the wavenumber within the slab, then the wavenumbers β, k_c are the z - and x -components k_z, k_x of k_1 with an angle of incidence θ . (The vectorial relationships are the same as those in Fig. 9.9.1.) Thus, we have:

$$\begin{aligned} \beta &= k_1 \sin \theta = k_0 n_1 \sin \theta \\ k_c &= k_1 \cos \theta = k_0 n_1 \cos \theta \end{aligned} \quad (9.11.36)$$

The value of β for each mode will generate a corresponding value for θ . The attenuation wavenumber α_c outside the slab can also be expressed in terms of the total internal reflection angles:

$$\alpha_c = \sqrt{\beta^2 - k_0^2 n_2^2} = k_0 \sqrt{n_1^2 \sin^2 \theta - n_2^2}$$

Since the critical angle is $\sin \theta_c = n_2/n_1$, we may also express α_c as:

$$\alpha_c = k_0 n_1 \sqrt{\sin^2 \theta - \sin^2 \theta_c} \quad (9.11.37)$$

Example 9.11.2: For the Example 9.11.1, we calculate $k_0 = 6.2832$ and $k_1 = 12.5664$ rad/cm. The critical and total internal reflection angles of the four modes are found to be:

$$\theta_c = \text{asin}\left(\frac{n_2}{n_1}\right) = 30^\circ$$

$$\theta = \text{asin}\left(\frac{\beta}{k_1}\right) = \{77.8275^\circ, 65.1960^\circ, 51.5100^\circ, 36.0609^\circ\}$$

As required, all θ s are greater than θ_c . $\square$

9.12 Asymmetric Dielectric Slab

The three-layer asymmetric dielectric slab waveguide shown in Fig. 9.12.1 is a typical component in integrated optics applications [912–933].

A thin dielectric film n_f of thickness $2a$ is deposited on a dielectric substrate n_s . Above the film is a dielectric cover or cladding n_c , such as air. To achieve propagation by total internal reflection within the film, we assume that the refractive indices satisfy $n_f > n_s \geq n_c$. The case of the symmetric dielectric slab of the previous section is obtained when $n_c = n_s$.

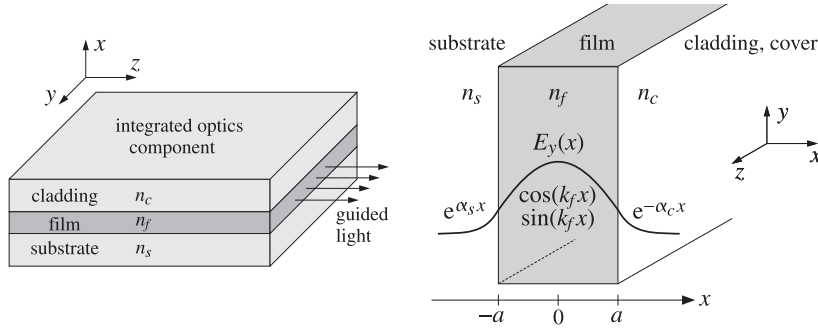


Fig. 9.12.1 Three-layer asymmetric dielectric slab waveguide.

In this section, we briefly discuss the properties of the TE and TM propagation modes. Let $k_0 = \omega \sqrt{\mu_0 \epsilon_0} = \omega/c_0 = 2\pi f/c_0 = 2\pi/\lambda_0$ be the free-space wavenumber at the operating frequency ω or f in Hz. The t, z dependence of the fields is assumed to be the usual $e^{j\omega t - j\beta t}$. If we orient the coordinate axes as shown in the above figure, then the decay constants α_s and α_c within the substrate and cladding must be positive so that the fields attenuate exponentially with x within both the substrate and cladding, hence, the corresponding transverse wavenumbers will be $j\alpha_s$ and $-j\alpha_c$. On the other hand, the transverse wavenumber k_f within the film will be real-valued. These quantities satisfy the relations (we assume $\mu = \mu_0$ in all three media):

$$\begin{aligned} k_f^2 &= k_0^2 n_f^2 - \beta^2 & k_f^2 + \alpha_s^2 &= k_0^2 (n_f^2 - n_s^2) \\ \alpha_s^2 &= \beta^2 - k_0^2 n_s^2 & k_f^2 + \alpha_c^2 &= k_0^2 (n_f^2 - n_c^2) (1 + \delta) = k_0^2 (n_f^2 - n_c^2) \\ \alpha_c^2 &= \beta^2 - k_0^2 n_c^2 & \alpha_c^2 - \alpha_s^2 &= k_0^2 (n_f^2 - n_s^2) \delta = k_0^2 (n_s^2 - n_c^2) \end{aligned} \quad (9.12.1)$$

where we defined the *asymmetry parameter* δ :

$$\delta = \frac{n_s^2 - n_c^2}{n_f^2 - n_s^2} \quad (9.12.2)$$

Note that $\delta \geq 0$ since we assumed $n_f > n_s \geq n_c$. Because k_f, α_s, α_c are assumed to be real, it follows that β must satisfy the inequalities, $\beta \leq k_0 n_f$, $\beta \geq k_0 n_s$, and $\beta \geq k_0 n_c$,

which combine to define the allowed range of β for the guided modes:

$$n_c \leq n_s \leq \frac{\beta}{k_0} \leq n_f \quad (9.12.3)$$

where the lower limit $\beta = k_0 n_s$ defines the cutoff frequencies, see Eq. (9.12.13).

TE modes

We consider the TE modes first. Assuming only x -dependence for the H_z component, it must satisfy the Helmholtz equations in the three regions:

$$\begin{aligned} (\partial_x^2 + k_f^2) H_z(x) &= 0, & |x| &\leq a \\ (\partial_x^2 - \alpha_c^2) H_z(x) &= 0, & x &\geq a \\ (\partial_x^2 - \alpha_s^2) H_z(x) &= 0, & x &\leq -a \end{aligned}$$

The solutions, decaying exponentially in the substrate and cover, can be written in the following form, which automatically satisfies the continuity conditions at the two boundaries $x = \pm a$:

$$H_z(x) = \begin{cases} H_1 \sin(k_f x + \phi), & |x| \leq a \\ H_1 \sin(k_f a + \phi) e^{-\alpha_c(x-a)}, & x \geq a \\ -H_1 \sin(k_f a - \phi) e^{\alpha_s(x+a)}, & x \leq -a \end{cases} \quad (9.12.4)$$

where ϕ is a parameter to be determined. The remaining two components, H_x and E_y , are obtained by applying Eq. (9.3.8), that is,

$$H_x = -\frac{j\beta}{k_f^2} \partial_x H_z, \quad E_y = -\eta_{TE} H_x \quad \eta_{TE} = \frac{\omega\mu}{\beta}$$

This gives in the three regions:

$$H_x(x) = \begin{cases} -j\frac{\beta}{k_f} H_1 \cos(k_f x + \phi), & |x| \leq a \\ -j\frac{\beta}{\alpha_c} H_1 \sin(k_f a + \phi) e^{-\alpha_c(x-a)}, & x \geq a \\ -j\frac{\beta}{\alpha_s} H_1 \sin(k_f a - \phi) e^{\alpha_s(x+a)}, & x \leq -a \end{cases} \quad (9.12.5)$$

Since we assumed that $\mu = \mu_0$ in all three regions, the continuity of E_y across the boundaries $x = \pm a$ implies the same for the H_x components, resulting in the two conditions:

$$\begin{aligned} \frac{1}{k_f} \cos(k_f a + \phi) &= \frac{1}{\alpha_c} \sin(k_f a + \phi) & \tan(k_f a + \phi) &= \frac{\alpha_c}{k_f} \\ \frac{1}{k_f} \cos(k_f a - \phi) &= \frac{1}{\alpha_s} \sin(k_f a - \phi) & \tan(k_f a - \phi) &= \frac{\alpha_s}{k_f} \end{aligned} \quad (9.12.6)$$

Since the argument of the tangent is unique up to an integral multiple of π , we may invert the two tangents as follows without loss of generality:

$$k_f a + \phi = \arctan\left(\frac{\alpha_c}{k_f}\right) + m\pi$$

$$k_f a - \phi = \arctan\left(\frac{\alpha_s}{k_f}\right)$$

which result in the characteristic equation of the slab and the solution for ϕ :

$$k_f a = \frac{1}{2}m\pi + \frac{1}{2}\arctan\left(\frac{\alpha_s}{k_f}\right) + \frac{1}{2}\arctan\left(\frac{\alpha_c}{k_f}\right) \quad (9.12.7)$$

$$\phi = \frac{1}{2}m\pi + \frac{1}{2}\arctan\left(\frac{\alpha_c}{k_f}\right) - \frac{1}{2}\arctan\left(\frac{\alpha_s}{k_f}\right) \quad (9.12.8)$$

where the integer $m = 0, 1, 2, \dots$, corresponds to the m th mode. Eq. (9.12.7) and the three equations (9.12.1) provide four equations in the four unknowns $\{\beta, k_f, \alpha_s, \alpha_c\}$. Using the trig identities $\tan(\theta_1 + \theta_2) = (\tan \theta_1 + \tan \theta_2) / (1 - \tan \theta_1 \tan \theta_2)$ and $\tan(\theta) = \tan(\theta + m\pi)$, Eqs. (9.12.7) and (9.12.8) may also be written in the following forms:

$$\tan(2k_f a) = \frac{k_f(\alpha_c + \alpha_s)}{k_f^2 - \alpha_c \alpha_s}, \quad \tan(2\phi) = \frac{k_f(\alpha_c - \alpha_s)}{k_f^2 + \alpha_c \alpha_s} \quad (9.12.9)$$

The form of Eq. (9.12.7) is preferred for the numerical solution. To this end, we introduce the dimensionless variables:

$$R = k_0 a \sqrt{n_f^2 - n_s^2} = \frac{2\pi f a}{c_0} \sqrt{n_f^2 - n_s^2} = 2\pi \frac{a}{\lambda_0} \sqrt{n_f^2 - n_s^2} \quad (9.12.10)$$

$$u = k_f a, \quad v = \alpha_s a, \quad w = \alpha_c a$$

Then, Eqs. (9.12.7) and (9.12.1) can be written in the normalized forms:

$$\begin{aligned} u &= \frac{1}{2}m\pi + \frac{1}{2}\arctan\left(\frac{v}{u}\right) + \frac{1}{2}\arctan\left(\frac{w}{u}\right) \\ u^2 + v^2 &= R^2 \\ w^2 - v^2 &= R^2 \delta \end{aligned} \quad (9.12.11)$$

Once these are solved for the three unknowns u, v, w , or k_f, α_s, α_c , the propagation constant β , or equivalently, the *effective index* $n_\beta = \beta/k_0$ can be obtained from:

$$\beta = \sqrt{k_0^2 n_f^2 - k_f^2} \Rightarrow n_\beta = \frac{\beta}{k_0} = \sqrt{n_f^2 - \frac{k_f^2}{k_0^2}} = \sqrt{n_f^2 - \frac{u^2}{k_0^2 a^2}} \quad (9.12.12)$$

To determine the number of propagating modes and the range of the mode index m , we set $v = 0$ in the characteristic equation (9.12.11) to find the radius R_m of the m th mode. Then, $u = R_m$ and $w = R_m \sqrt{\delta}$, and we obtain:

$$R_m = \frac{1}{2}m\pi + \frac{1}{2}\arctan(\sqrt{\delta}), \quad m = 0, 1, 2, \dots \quad (9.12.13)$$

For a given operating frequency f , the value of R is fixed. All allowed propagating modes must satisfy $R_m \leq R$, or,

$$\frac{1}{2}m\pi + \frac{1}{2}\arctan(\sqrt{\delta}) \leq R \Rightarrow m \leq \frac{2R - \arctan(\sqrt{\delta})}{\pi}$$

This fixes the maximum mode index M to be:

$$M = \text{floor}\left(\frac{2R - \arctan(\sqrt{\delta})}{\pi}\right) \quad (\text{maximum TE mode index}) \quad (9.12.14)$$

Thus, there are $(M + 1)$ modes labeled by $m = 0, 1, \dots, M$. In the symmetric case, $\delta = 0$, and (9.12.14) reduces to Eq. (9.11.25) of the previous section. The corresponding cutoff frequencies are obtained by setting:

$$R_m = \frac{2\pi f_m a}{c_0} \sqrt{n_f^2 - n_s^2} \Rightarrow f_m = \frac{\frac{1}{2}m\pi + \frac{1}{2}\arctan(\sqrt{\delta})}{\frac{2\pi a}{c_0} \sqrt{n_f^2 - n_s^2}} \quad (9.12.15)$$

which can be written more simply as $f_m = f R_m / R$, $m = 0, 1, \dots, M$, where $f = c_0 / \lambda_0$.

For each of the $M + 1$ propagating modes one can calculate the corresponding angle of total internal reflection of the equivalent ray model by identifying k_f with the transverse propagation wavenumber, that is, $k_f = k_0 n_f \cos \theta$, as shown in Fig. 9.12.2.

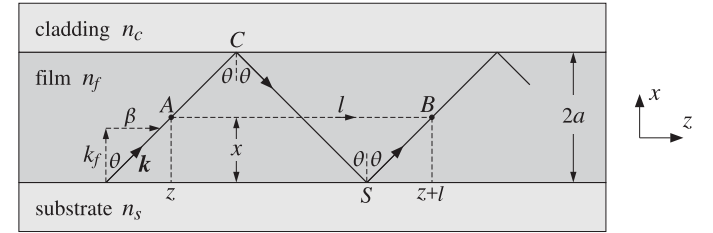


Fig. 9.12.2 Ray propagation model.

The characteristic equation (9.12.7) can be given a nice interpretation in terms of the ray model [925]. The field of the upgoing ray at a point A at (x, z) is proportional, up to a constant amplitude, to

$$e^{-jk_f x} e^{-j\beta z}$$

Similarly, the field of the upgoing ray at the point B at $(x, z + l)$ should be

$$e^{-jk_f x} e^{-j\beta(z+l)} \quad (9.12.16)$$

But if we follow the ray starting at A along the zig-zag path $AC \rightarrow CS \rightarrow SB$, the ray will have traveled a total vertical roundtrip distance of $4a$ and will have suffered two total internal reflection phase shifts at points C and S , denoted by $2\psi_c$ and $2\psi_s$. We

recall that the reflection coefficients have the form $\rho = e^{2j\psi}$ for total internal reflection, as given for example by Eq. (7.8.3). Thus, the field at point B would be

$$e^{-jk_f(x+4a)} e^{2j\psi_s} e^{2j\psi_c} e^{-j\beta(z+l)}$$

This must match (9.12.16) and therefore the extra accumulated phase $4k_f a - 2\psi_s - 2\psi_c$ must be equal to a multiple of 2π , that is,

$$4k_f a - 2\psi_s - 2\psi_c = 2m\pi \Rightarrow k_f a = \frac{1}{2}m\pi + \frac{1}{2}\psi_s + \frac{1}{2}\psi_c$$

As seen from Eq. (7.8.3), the phase terms are exactly those appearing in Eq. (9.12.7):

$$\tan \psi_c = \frac{\alpha_c}{k_f}, \quad \tan \psi_s = \frac{\alpha_s}{k_f} \Rightarrow \psi_c = \arctan\left(\frac{\alpha_c}{k_f}\right), \quad \psi_s = \arctan\left(\frac{\alpha_s}{k_f}\right)$$

A similar interpretation can be given for the TM modes.

It is common in the literature to represent the characteristic equation (9.12.11) by means of a *universal mode curve* [927] defined in terms of the following scaled variable:

$$b = \frac{v^2}{R^2} = \frac{\beta^2 - k_0^2 n_s^2}{k_0^2 (n_f^2 - n_s^2)} \quad (9.12.17)$$

which ranges over the standardized interval $0 \leq b \leq 1$, so that

$$u = R\sqrt{1-b}, \quad v = R\sqrt{b}, \quad w = R\sqrt{b+\delta} \quad (9.12.18)$$

Then, Eq. (9.12.11) takes the universal form in terms of the variables b, R :[†]

$$2R\sqrt{1-b} = m\pi + \arctan\left(\sqrt{\frac{b}{1-b}}\right) + \arctan\left(\sqrt{\frac{b+\delta}{1-b}}\right) \quad (9.12.19)$$

It is depicted in Fig. 9.12.3 with one branch for each value of $m = 0, 1, 2, \dots$, and for the three asymmetry parameter values $\delta = 0, 1, 10$.

A vertical line drawn at each value of R determines the values of b for the propagating modes. Similar curves can be developed for TM modes. See Example 9.12.1 for a concrete example that includes both TE and TM modes.

TM modes

The TM modes are obtained by solving Eqs. (9.3.10) in each region and applying the boundary conditions. Assuming x -dependence only, we must solve in each region:

$$(\partial_x^2 + k_f^2)E_z = 0, \quad E_x = -\frac{j\beta}{k_f^2} \partial_x E_z, \quad H_y = \frac{1}{\eta_{TM}} E_x, \quad \eta_{TM} = \frac{\beta}{\omega\epsilon}$$

The solution for $E_z(x)$ is given by a similar expression as Eq. (9.12.4):

$$E_z(x) = \begin{cases} E_1 \sin(k_f x + \phi), & |x| \leq a \\ E_1 \sin(k_f a + \phi) e^{-\alpha_c(x-a)}, & x \geq a \\ -E_1 \sin(k_f a - \phi) e^{\alpha_s(x+a)}, & x \leq -a \end{cases} \quad (9.12.20)$$

[†] R is usually denoted by the variable V .

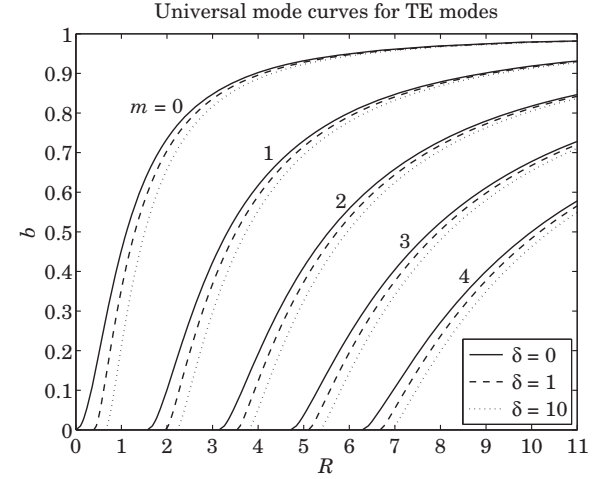


Fig. 9.12.3 Universal mode curves.

where ϕ is a parameter to be determined. Then, the E_x component is:

$$E_x(x) = \begin{cases} -j\frac{\beta}{k_f} E_1 \cos(k_f x + \phi), & |x| \leq a \\ -j\frac{\beta}{\alpha_c} E_1 \sin(k_f a + \phi) e^{-\alpha_c(x-a)}, & x \geq a \\ -j\frac{\beta}{\alpha_s} E_1 \sin(k_f a - \phi) e^{\alpha_s(x+a)}, & x \leq -a \end{cases} \quad (9.12.21)$$

The boundary conditions require the continuity of the normal component of displacement field $D_x = \epsilon E_x$ across the interfaces at $x = \pm a$, which is equivalent to the continuity of the tangential field H_y because $H_y = E_x / \eta_{TM} = \epsilon E_x \omega / \beta = D_x \omega / \beta$. Thus, the boundary conditions at $x = \pm a$ require:

$$\begin{aligned} \frac{\epsilon_f}{k_f} \cos(k_f a + \phi) &= \frac{\epsilon_c}{\alpha_c} \sin(k_f a + \phi) & \tan(k_f a + \phi) &= p_c \frac{\alpha_c}{k_f} \\ \frac{\epsilon_f}{k_f} \cos(k_f a - \phi) &= \frac{\epsilon_s}{\alpha_s} \sin(k_f a - \phi) & \tan(k_f a - \phi) &= p_s \frac{\alpha_s}{k_f} \end{aligned} \Rightarrow \quad (9.12.22)$$

where we defined the ratios:

$$p_c = \frac{\epsilon_f}{\epsilon_c} = \frac{n_f^2}{n_c^2}, \quad p_s = \frac{\epsilon_f}{\epsilon_s} = \frac{n_f^2}{n_s^2} \quad (9.12.23)$$

Inverting the tangents we obtain:

$$k_f a + \phi = \arctan\left(p_c \frac{\alpha_c}{k_f}\right) + m\pi$$

$$k_f a - \phi = \arctan\left(p_s \frac{\alpha_s}{k_f}\right)$$

These give the characteristic equation of the slab and ϕ :

$$k_f a = \frac{1}{2}m\pi + \frac{1}{2}\arctan\left(p_s \frac{\alpha_s}{k_f}\right) + \frac{1}{2}\arctan\left(p_c \frac{\alpha_c}{k_f}\right) \quad (9.12.24)$$

$$\phi = \frac{1}{2}m\pi + \frac{1}{2}\arctan\left(p_c \frac{\alpha_c}{k_f}\right) - \frac{1}{2}\arctan\left(p_s \frac{\alpha_s}{k_f}\right) \quad (9.12.25)$$

and, as in Eq. (9.12.9), we can write:

$$\tan(2k_f a) = \frac{k_f(p_s \alpha_s + p_c \alpha_c)}{k_f^2 - p_s p_c \alpha_s \alpha_c}, \quad \tan(2\phi) = \frac{k_f(p_c \alpha_c - p_s \alpha_s)}{k_f^2 + p_s p_c \alpha_s \alpha_c} \quad (9.12.26)$$

In terms of the normalized variables u, v, w, R , we have:

$$\begin{aligned} u &= \frac{1}{2}m\pi + \frac{1}{2}\arctan\left(p_s \frac{v}{u}\right) + \frac{1}{2}\arctan\left(p_c \frac{w}{u}\right) \\ u^2 + v^2 &= R^2 \\ w^2 - v^2 &= R^2 \delta \end{aligned} \quad (9.12.27)$$

The number of propagating modes and the range of the mode index m , are again determined by setting $v = 0$, $u = R_m$, and $w = R_m \sqrt{\delta}$:

$$R_m = \frac{1}{2}m\pi + \frac{1}{2}\arctan(p_c \sqrt{\delta}), \quad m = 0, 1, 2, \dots$$

The allowed propagating modes must satisfy $R_m \leq R$, or,

$$\frac{1}{2}m\pi + \frac{1}{2}\arctan(p_c \sqrt{\delta}) \leq R \Rightarrow m \leq \frac{2R - \arctan(p_c \sqrt{\delta})}{\pi}$$

which fixes the maximum mode index M to be:

$$M = \text{floor}\left(\frac{2R - \arctan(p_c \sqrt{\delta})}{\pi}\right) \quad (\text{maximum TM mode index}) \quad (9.12.28)$$

The $(M + 1)$ modes are again labeled by $m = 0, 1, \dots, M$. The corresponding cutoff frequencies are obtained by setting:

$$R_m = \frac{2\pi f_m a}{c_0} \sqrt{n_f^2 - n_s^2} \Rightarrow f_m = \frac{\frac{1}{2}m\pi + \frac{1}{2}\arctan(p_c \sqrt{\delta})}{\frac{2\pi a}{c_0} \sqrt{n_f^2 - n_s^2}} \quad (9.12.29)$$

which can be written more simply as $f_m = f R_m / R$, $m = 0, 1, \dots, M$, where $f = c_0 / \lambda_0$. The corresponding angles of total internal reflection in the equivalent ray model are obtained by solving $k_f = k_0 n_f \cos \theta$.

Because $p_c > 1$, we observe that the maximum mode index M and the cutoff frequencies f_m will satisfy the following inequalities for the TE and TM cases:

$$M_{TM} \leq M_{TE}, \quad f_{m,TE} \leq f_{m,TM} \quad (9.12.30)$$

Numerical Solutions

Next we look at the numerical solutions of Eqs. (9.12.27). The TE case is also included by setting $p_s = p_c = 1$. A simple and effective iterative method for solving such characteristic equations was given in Ref. [963]. By replacing v, w in terms of u , let $F(u)$ denote the right-hand side of Eq. (9.12.27):

$$F(u) = \frac{1}{2}m\pi + \frac{1}{2}\arctan\left(p_s \frac{v}{u}\right) + \frac{1}{2}\arctan\left(p_c \frac{w}{u}\right)$$

The problem then becomes that of finding the fixed-point solutions $u = F(u)$. The method suggested in Ref. [963] is to use the iteration:

$$u_{n+1} = F(u_n), \quad n = 0, 1, 2, \dots$$

initialized at $u_0 = R$. This simple iteration does converge in many cases, but fails in others. We have found that a simple modification that involves the introduction of a “relaxation” parameter r such that $0 \leq r \leq 1$, enables the convergence of even the most difficult cases. The modified iteration has the form:

$$u_{n+1} = rF(u_n) + (1 - r)u_n$$

Explicitly, the iteration starts with the initial values:

$$u_0 = R, \quad v_0 = 0, \quad w_0 = R\sqrt{\delta} \quad (9.12.31)$$

and proceeds iteratively, for $n = 0, 1, 2, \dots$, until two successive u_n values become closer to each other than some specified error tolerance, say tol , such as $\text{tol} = 10^{-10}$:

$$u_{n+1} = r \left[\frac{1}{2}m\pi + \frac{1}{2}\arctan\left(p_s \frac{v_n}{u_n}\right) + \frac{1}{2}\arctan\left(p_c \frac{w_n}{u_n}\right) \right] + (1 - r)u_n$$

if $|u_{n+1} - u_n| < \text{tol}$, then exit, else continue (9.12.32)

$$v_{n+1} = \sqrt{R^2 - u_{n+1}^2}$$

$$w_{n+1} = \sqrt{R^2 \delta - v_{n+1}^2}$$

The MATLAB function `dguide3.m` implements the method and has usage:

$$[\text{be}, k_f, \text{as}, \text{ac}, \text{fm}, \text{Nit}] = \text{dguide3}(\text{a}, \text{ns}, \text{nf}, \text{nc}, \text{mode}, \text{r}, \text{tol});$$

where the inputs and outputs have the following meanings:

a	=	half-width of slab in units of the free-space wavelength λ_0
n_s, n_f, n_c	=	refractive indices of substrate, film, and cladding ($n_f > n_s > n_c$)
mode	=	'TE' or 'TM'
r	=	relaxation parameter (default $r = 0.5$)
tol	=	error tolerance (default tol = 10^{-10})
β	=	propagation wavenumbers in units of $k_0 = 2\pi/\lambda_0$
k_f	=	transverse wavenumbers inside slab in units of k_0
α_s, α_c	=	decay wavenumbers in substrate and cladding in units of k_0
f_m	=	cutoff frequencies in units of $f = c_0/\lambda_0$
N_{it}	=	number of iterations it takes to converge to within tol

Internally, the function determines M from Eq. (9.12.14) or (9.12.28) and calculates $\beta, k_f, \alpha_s, \alpha_c, f_m$ as $(M + 1)$ -dimensional column vectors representing the $M + 1$ modes. To clarify the computations, the essential part of the code is listed below:

```

k0 = 2*pi; % la0 = 2*pi/k0 = 1 in the assumed units
R = k0*a * sqrt(nf^2-ns^2); % (u,v) circle radius, note k0*a = 2*pi*(a/la0)
d = (ns^2-nc^2)/(nf^2-ns^2); % asymmetry parameter

if strcmpi(mode,'TE') % mode can also be entered in lower case
    ps = 1; pc = 1;
else
    ps = nf^2/ns^2; pc = nf^2/nc^2;
end

M = floor((2*R - atan(pc*sqrt(d)))/pi); % highest mode index
m = (0:M)'; % vector of mode indices

u = R*ones(M+1,1); % initialize iteration variables u,v,w
v = zeros(M+1,1); % u,v,w are (M+1)x1 vectors
w = R*sqrt(d)*ones(M+1,1);

Nit = 1; % number of iterations
% while loop repeats till convergence
while 1
    unew = r*(m*pi/2 + atan(ps*v./u)/2 + atan(pc*w./u)/2) + (1-r)*u;
    if norm(unew-u) <= tol, break; end
    Nit = Nit + 1;
    u = unew;
    v = sqrt(R^2 - u.^2);
    w = sqrt(R^2*d + v.^2);
    if Nit>1000, break; end % safeguard against possible divergence
end

kf = u/(k0*a); % kf in units of k0, i.e., kf/k0 = u/(k0*a)
as = v/(k0*a);
ac = w/(k0*a);
be = sqrt(nf^2 - kf.*kf); % beta in units of k0, i.e., beta/k0

Rm = m*pi/2 + atan(pc*sqrt(d))/2; % cutoff radius for m-th mode
fm = Rm/R; % cutoff frequencies in units of f = c0/la0
thm = acos(kf/nf); % angles of total internal reflection

```

Example 9.12.1: For comparison purposes, we consider the same benchmark example discussed in [963] consisting of a silicon film of thickness of $1 \mu\text{m}$ with $n_f = 3.5$, an oxide substrate with $n_s = 1.45$, and air cover, with operating wavelength $\lambda_0 = 1.55 \mu\text{m}$. The following MATLAB code generates both the TE and TM modes, with the numerical outputs listed in the tables below.

```

nf=3.5; ns=1.45; nc=1; % oxide substrate | silicon film | air cover
la0 = 1.55; a = 0.5; % units of microns
a = a/la0; % half-thickness in units of la0

r=0.3; % default value r=0.5 fails to converge for the TE modes
tol=1e-10;

[be,kf,as,ac,fm,Nit] = dguide3(a,ns,nf,nc,'te',r,tol); % TE modes
[be,kf,as,ac,fm,Nit] = dguide3(a,ns,nf,nc,'tm',r,tol); % TM modes

```

m	β/k_0	k_f/k_0	α_s/k_0	α_c/k_0	f_m/f	θ_m°
0	3.434746	0.6727	3.1137	3.2860	0.0247	78.92
1	3.232789	1.3413	2.8894	3.0742	0.2679	67.47
2	2.872310	2.0000	2.4794	2.6926	0.5112	55.15
3	2.302025	2.6364	1.7880	2.0735	0.7545	41.13
4	1.451972	3.1846	0.0756	1.0527	0.9978	24.51

(TE modes)

m	β/k_0	k_f/k_0	α_s/k_0	α_c/k_0	f_m/f	θ_m°
0	3.416507	0.7599	3.0935	3.2669	0.1028	77.46
1	3.154191	1.5169	2.8011	2.9915	0.3461	64.32
2	2.668932	2.2642	2.2407	2.4745	0.5894	49.69
3	1.865244	2.9616	1.1733	1.5745	0.8327	32.20
4	-	-	-	-	1.0760	-

(TM modes)

The β/k_0 column is the effective phase index of the modes. The default value of the relaxation parameter $r = 0.5$ did not work in this case and caused the TE iteration to diverge and the smaller value $r = 0.3$ was chosen. The number of iterations were $N_{it} = 57$ for TE and $N_{it} = 66$ for TM. The TIR angles were computed by the following command:

```
thm = acos(kf/nf)*180/pi; % degrees
```

We note that all TIR angles are greater than the critical angles computed by:

$$\theta_s = \arcsin\left(\frac{n_s}{n_f}\right) = 24.47^\circ, \quad \theta_c = \arcsin\left(\frac{n_c}{n_f}\right) = 16.60^\circ$$

There are five TE modes and four TM ones. The fifth TE mode is very weakly bound to the substrate side because its decay parameter α_s is very small, its cutoff frequency is very near the operating frequency $f = c_0/\lambda_0$, and its TIR angle, very close to the critical angle.

With reference to the inequality (9.12.30), it so happened that in this example f falls in the range $f_{4,TE} < f < f_{4,TM}$, and therefore, the fifth TM mode $f_{4,TM}$ is not excited, but $f_{4,TE}$ is.

The convergence can be verified for all modes at once by computing the vector error norm of the characteristic equations, that is,

```
M = length(be)-1; m = (0:M)';
Err = norm(kf*2*pi*a - m*pi/2 - atan(ps*as./kf)/2 - atan(pc*ac./kf)/2);
```

This error is of the order of the assumed tolerance, indeed, we have $\text{Err} = 2.94 \cdot 10^{-10}$ for TE, and $\text{Err} = 2.53 \cdot 10^{-10}$ for TM. We note that the quantity $k_f \cdot 2\pi \cdot a$ represents the variable u in our units, indeed, $u = k_f a = (k_f/k_0) k_0 a = (k_f/k_0) 2\pi (a/\lambda_0)$.

Finally, Fig. 9.12.4 displays the TE and TM solutions on the universal mode curves, see e.g. Eq. (9.12.19). $\square$

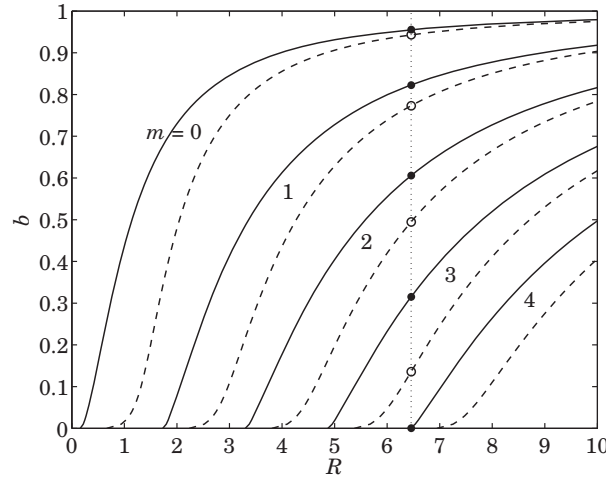


Fig. 9.12.4 Universal mode curves. TE (solid lines/filled circles), TM (dashed lines/open circles).

Example 9.12.2: A second, more difficult, example from [963] has the parameters $\lambda_0 = 1.55 \mu\text{m}$, $a = 0.5 \mu\text{m}$, $n_f = 3.3$, $n_s = 3.256$, $n_c = 1$.

The same MATLAB code applies here, but we used the default value $r = 0.5$, which converges in 8 and 10 iterations respectively for the TE and TM modes. Only one ($M = 0$) TE and one TM mode are supported with parameters given in the table below. The critical TIR angles are in this example:

$$\theta_s = \arcsin\left(\frac{n_s}{n_f}\right) = 80.63^\circ, \quad \theta_c = \arcsin\left(\frac{n_c}{n_f}\right) = 17.64^\circ$$

mode	β/k_0	k_f/k_0	α_s/k_0	α_c/k_0	f_m/f	θ_m°
TE	3.265996	0.4725	0.2553	3.1091	0.6427	81.77
TM	3.263384	0.4902	0.2194	3.1064	0.7142	81.46

The computational errors in the characteristic equations were $\text{Err} = 1.63 \cdot 10^{-11}$ for TE, and $\text{Err} = 1.52 \cdot 10^{-11}$ for TM. $\square$

9.13 Problems

- 9.1 An air-filled $1.5 \text{ cm} \times 3 \text{ cm}$ waveguide is operated at a frequency that lies in the middle of its TE_{10} mode band. Determine this operating frequency in GHz and calculate the maximum power in Watts that can be transmitted without causing dielectric breakdown of air. The dielectric strength of air is 3 MV/m.
- 9.2 It is desired to design an air-filled rectangular waveguide such that (a) it operates only in the TE_{10} mode with the widest possible bandwidth, (b) it can transmit the maximum possible power, and (c) the operating frequency is 12 GHz and it lies in the middle of the operating band. What are the dimensions of the guide in cm?
- 9.3 An air-filled rectangular waveguide is used to transfer power to a radar antenna. The guide must meet the following specifications: The two lowest modes are TE_{10} and TE_{20} . The operating frequency is 3 GHz and must lie exactly halfway between the cutoff frequencies of these two modes. The maximum electric field within the guide may not exceed, by a safety margin of 3, the breakdown field of air 3 MV/m.
 - a. Determine the smallest dimensions a, b for such a waveguide, if the transmitted power is required to be 1 MW.
 - b. What are the dimensions a, b if the transmitted power is required to be maximum? What is that maximum power in MW?
- 9.4 It is desired to design an air-filled rectangular waveguide operating at 5 GHz, whose group velocity is $0.8c$. What are the dimensions a, b of the guide (in cm) if it is also required to carry maximum power and have the widest possible bandwidth? What is the cutoff frequency of the guide in GHz and the operating bandwidth?
- 9.5 Show the following relationship between guide wavelength and group velocity in an arbitrary air-filled waveguide: $v_g \lambda_g = c\lambda$, where $\lambda_g = 2\pi/\beta$ and λ is the free-space wavelength. Moreover, show that the λ and λ_g are related to the cutoff wavelength λ_c by:

$$\frac{1}{\lambda^2} = \frac{1}{\lambda_g^2} + \frac{1}{\lambda_c^2}$$

- 9.6 Determine the four lowest modes that can propagate in a WR-159 and a WR-90 waveguide. Calculate the cutoff frequencies (in GHz) and cutoff wavelengths (in cm) of these modes.
- 9.7 An air-filled WR-90 waveguide is operated at 9 GHz. Calculate the maximum power that can be transmitted without causing dielectric breakdown of air. Calculate the attenuation constant in dB/m due to wall ohmic losses. Assume copper walls.
- 9.8 A rectangular waveguide has sides a, b such that $b \leq a/2$. Determine the cutoff wavelength λ_c of this guide. Show that the operating wavelength band of the lowest mode is $0.5\lambda_c \leq \lambda \leq \lambda_c$. Moreover, show that the allowed range of the guide wavelength is $\lambda_g \geq \lambda_c/\sqrt{3}$.
- 9.9 The TE_{10} mode operating bandwidth of an air-filled waveguide is required to be 4-7 GHz. What are the dimensions of the guide?
- 9.10 *Computer Experiment: WR-159 Waveguide.* Reproduce the two graphs of Fig. 9.8.2.
- 9.11 A TM mode is propagated along a hollow metallic waveguide of arbitrary but uniform cross section. Assume perfectly conducting walls.

- a. Show that the $E_z(x, y)$ component satisfies:

$$\int_S |\nabla_T E_z|^2 dS = k_c^2 \int_S |E_z|^2 dS, \quad (k_c = \text{cutoff wavenumber}) \quad (9.13.1)$$

b. Using the above result, show that the energy velocity is equal to the group velocity.

Hint: Use the identity: $\nabla_T \cdot (A \nabla_T B) = \nabla_T A \cdot \nabla_T B + A \nabla_T^2 B$, for scalar A, B .

9.12 *Computer Experiment: Dielectric Slab Waveguide.* Using the MATLAB functions `ds1ab` and `dguide`, write a program that reproduces all the results and graphs of Examples 9.11.1 and 9.11.2.

9.13 Show that if the speed of light c_0 is slightly changed to $c = c_0 + \Delta c_0$ (e.g. representing a more exact value), then the solutions of Eq. (9.11.29) for k_c, α_c change into:

$$k_c + \Delta k_c = k_c - \left(\frac{k_c}{1 + \alpha_c a} \right) \left(\frac{\Delta c_0}{c_0} \right)$$

$$\alpha_c + \Delta \alpha_c = \alpha_c - \left(\alpha_c + \frac{k_c^2 a}{1 + \alpha_c a} \right) \left(\frac{\Delta c_0}{c_0} \right)$$

For Example 9.11.1, calculate the corrected values when $c_0 = 30$ and $c = 29.9792458$ GHz cm. Compare with the values obtained if c_0 is replaced by c inside the function `dguide`. More generally, consider the sensitivity of the solutions of Eq. (9.11.29) to any of the parameters $\omega_0, a, c_0, n_1, n_2$, which affect the solution through the value of $R = a \omega_0 c_0^{-1} \sqrt{n_1^2 - n_2^2}$. A small change in one or all of the parameters will induce a small change $R \rightarrow R + \Delta R$. Show that the solutions are changed to

$$u + \Delta u = u + \left(\frac{u}{1 + v} \right) \left(\frac{\Delta R}{R} \right)$$

$$v + \Delta v = v + \left(v + \frac{u^2}{1 + v} \right) \left(\frac{\Delta R}{R} \right)$$

In particular, for simultaneous changes in all of the parameters, show that

$$\frac{\Delta R}{R} = \frac{\Delta a}{a} + \frac{\Delta \omega_0}{\omega_0} - \frac{\Delta c_0}{c_0} + \frac{2n_1 \Delta n_1 - 2n_2 \Delta n_2}{n_1^2 - n_2^2}$$

From these results, show that the changes due to a change $a \rightarrow a + \Delta a$ of the slab thickness are given by,

$$k_c + \Delta k_c = k_c - \frac{k_c \alpha_c}{1 + \alpha_c a} \Delta a$$

$$\alpha_c + \Delta \alpha_c = \alpha_c + \frac{k_c^2}{1 + \alpha_c a} \Delta a$$

9.14 For the dielectric slab waveguide shown in Fig. 9.11.1, demonstrate that the energy transport velocity is equal to the group velocity. Specifically, consider the case of even TE modes defined by Eqs. (9.11.3)–(9.11.13), and show that $v_{\text{en}} = v_{\text{gr}}$, where

$$v_{\text{en}} = \frac{P_T}{W'}, \quad v_{\text{gr}} = \frac{d\omega}{d\beta} \quad (9.13.2)$$

where P_T is the time-averaged power transmitted in the z -direction through the cross-sectional area defined by $0 \leq y \leq 1$ and $-\infty < x < \infty$, and W' is the energy contained in the volume defined by the above area and unit- z -length, i.e.,

$$P_T = \frac{1}{2\eta_{TE}} \int_{-\infty}^{\infty} |E_y(x)|^2 dx, \quad W' = \frac{1}{4} \int_{-\infty}^{\infty} [\epsilon |E_y(x)|^2 + |H_x(x)|^2 + \mu |H_z(x)|^2] dx$$

Because of the substantial amount of algebra involved, break the calculation as follows:

a. Show that P_T is given as the sum of the following two terms, where the first one represents the power flowing within the slab, and the second, the power flowing outside the slab:

$$P_T = |H_1|^2 \frac{\beta \omega \mu_0 [ak_c + \sin(ak_c) \cos(ak_c)]}{2k_c^3} + |H_1|^2 \frac{\beta \omega \mu_0 \sin^2(ak_c)}{2\alpha_c^3}$$

where H_1 is the amplitude defined in Eq. (9.11.5). Without loss of generality, from now on set, $H_1 = 1$.

b. Show that the electric and magnetic energy densities are given as follows, where again, the first terms represent the energy contained within the slab, and the second, the energy outside the slab:

$$W'_e = \frac{\mu_0 (\beta^2 + k_c^2) [ak_c + \sin(ak_c) \cos(ak_c)]}{4k_c^3} + \frac{\mu_0 (\beta^2 - \alpha_c^2) \sin^2(ak_c)}{4\alpha_c^3}$$

$$W'_m = \frac{\mu_0 (\beta^2 + k_c^2) ak_c + \mu_0 (\beta^2 - k_c^2) \sin(ak_c) \cos(ak_c)}{4k_c^3} + \frac{\mu_0 (\beta^2 + \alpha_c^2) \sin^2(ak_c)}{4\alpha_c^3}$$

c. Using the above expressions and Eq. (9.11.13), show the equality

$$W'_e = W'_m$$

Thus, the total energy density is $W' = W'_e + W'_m = 2W'_e$.

d. From parts (a,b,c), show that

$$P_T = \frac{\omega \beta \mu_0 (1 + a \alpha_c)}{2 \alpha_c k_c^2}$$

$$W' = \frac{\mu_0 [(\beta^2 + k_c^2) a \alpha_c + \beta^2]}{2 \alpha_c k_c^2}$$

e. By differentiating Eqs. (9.11.3) and (9.11.13) with respect to ω , show that

$$\omega \beta \frac{d\beta}{d\omega} = \frac{(\beta^2 + k_c^2) a \alpha_c + \beta^2}{1 + a \alpha_c}$$

f. Combining the results of parts (e,f), show finally that

$$v_{\text{en}} = v_{\text{gr}} = \frac{\omega \beta (1 + a \alpha_c)}{(\beta^2 + k_c^2) a \alpha_c + \beta^2} = \frac{\omega \beta (1 + a \alpha_c)}{\omega^2 \epsilon_0 \mu_0 n_1^2 a \alpha_c + \beta^2}$$

9.15 *Computer Experiment. Asymmetric Slab Waveguide.* Reproduce all the results and Fig. 9.12.4 of Example 9.12.1. Moreover, make a separate graph of Fig. 9.12.4 that zooms into the neighborhood of the fifth TE mode to make sure that it is indeed below cutoff.