CHAPTER 14

WAVEGUIDE AND ANTENNA FUNDAMENTALS

As a conclusion to our study of electromagnetics, we investigate the basic principles of two important classes of devices: waveguides and antennas. In broad definitions, a waveguide is a structure through which electromagnetic waves can be transmitted from point to point, and within which the fields are confined to a certain extent. An antenna is any device that radiates electromagnetic fields into space, where the fields originate from a source that feeds the antenna through a transmission line or waveguide. The antenna thus serves as an interface between the confining line and space when used as a transmitter—or between space and the line when used as a receiver.

In our study of waveguides, we will first take a broad view of waveguide devices, to obtain a physical understanding of how they work and the conditions under which they are used. We will next explore the simple parallel-plate waveguide and study the concept of waveguide modes and the conditions under which these will occur. We will study the electric and magnetic field configurations of the guided modes using simple plane wave models and through use of the wave equation. We will then study more complicated structures, including the rectangular waveguide and the dielectric slab guide.

Our study of antennas will include the derivation of the radiated fields from an elemental dipole, beginning with the retarded vector potentials that we studied in Chap. 10. We will address issues that include the efficiency of power radiation from an antenna, and the parameters that govern this.

14.1 BASIC WAVEGUIDE OPERATION

Waveguides assume many different forms that depend on the purpose of the guide, and on the frequency of the waves to be transmitted. The simplest form (in terms of analysis) is the parallel-plate guide shown in Fig. 14.1. Other forms are the hollow-pipe guides, including the rectangular waveguide of Fig. 14.2, and the cylindrical guide, shown in Fig. 14.3. Dielectric waveguides, used primarily at optical frequencies, include the slab waveguide of Fig. 14.4 and the optical fiber, shown in Fig. 14.5. Each of these structures possesses certain advantages over the others, depending on the application and the frequency of the waves to be transmitted. All guides, however, exhibit the same basic operating principles, which we will explore in this section.

To develop an understanding of waveguide behavior, we consider the parallel-plate waveguide of Fig. 14.1. At first, we recognize this as one of the transmission line structures that we investigated in Chap. 13. So the first question that arises is: How does a waveguide differ from a transmission line to begin with? The difference lies in the form of the electric and magnetic fields within the line. To see this, consider Fig. 14.6*a*, which shows the fields when the line operates as a transmission line. A sinusoidal voltage wave, with voltage applied between conductors, leads to an electric field that is directed vertically between the conductors as shown. Since current flows only in the z direction, magnetic field will be oriented in and out of the page (in the y direction). The interior fields comprise a plane electromagnetic wave which propagates in the z direction (as the Poynting vector will show), since both fields lie in the transverse plane. We refer to this as a transmission line wave, which, as discussed in Chap. 13, is a transverse electromagnetic (TEM) wave. The wavevector **k**, shown in the figure, indicates the direction of wave travel, as well as the direction of power flow. With perfectly conducting plates, the electric field between plates is found by solving Eq. (29), Chap. 11, leading to Eq. (31) in that chapter.

As the frequency is increased, a remarkable change occurs in the way the fields progagate down the line. Although the original field configuration of Fig. 14.6*a* may still be present, another possibility emerges which is shown in Fig.

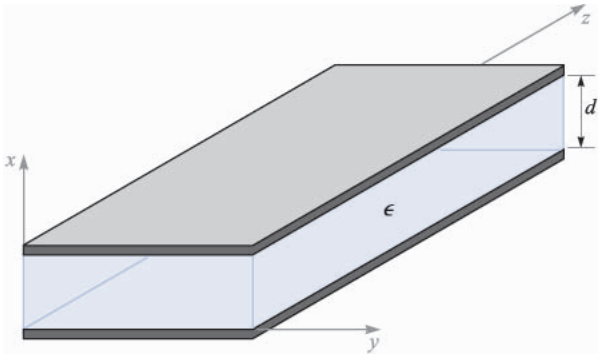


FIGURE 14.1
Parallel-plate waveguide, with metal plates at $x = 0, d$. Between the plates is a dielectric of permittivity ϵ.

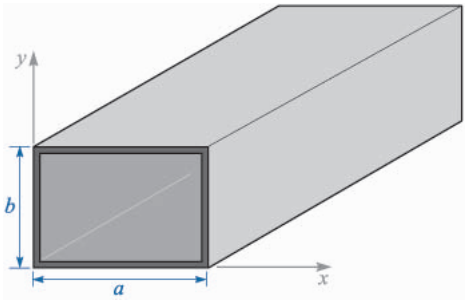


FIGURE 14.2
Rectangular waveguide.

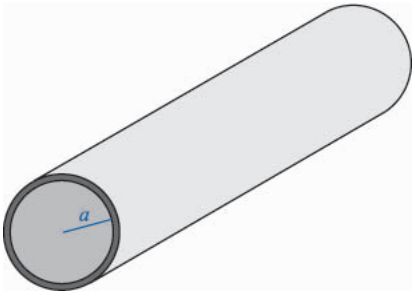


FIGURE 14.3
Cylindrical waveguide.

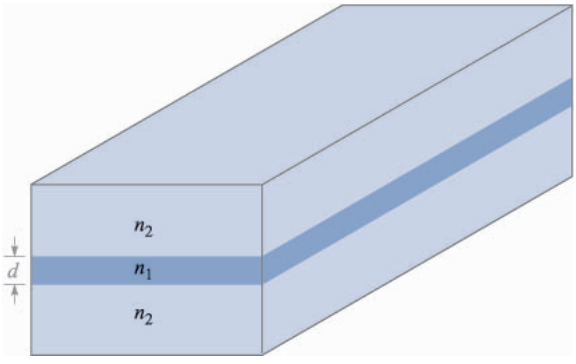


FIGURE 14.4
Symmetric dielectric slab waveguide, with slab region (refractive index n_1) surrounded by two dielectrics of index $n_2 < n_1$.

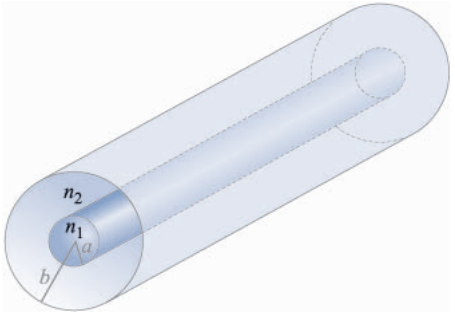


FIGURE 14.5
Optical fiber waveguide, with the core dielectric ($r < a$) of refractive index n_1. The cladding dielectric ($a < r < b$) is of index $n_2 < n_1$.

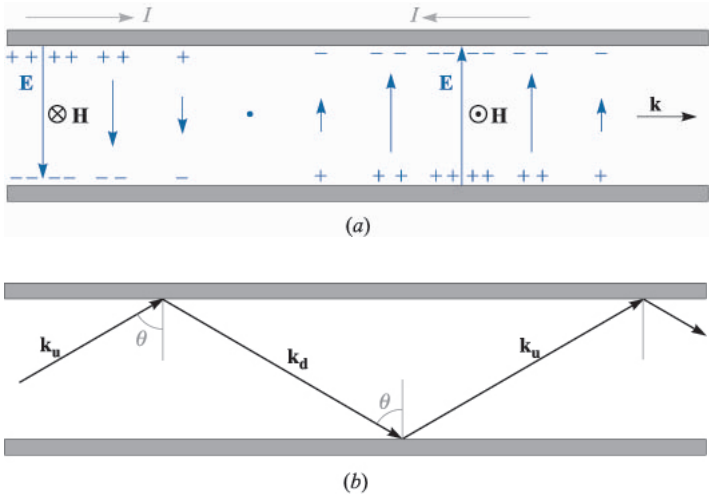


FIGURE 14.6
(*a*) Electric and magnetic fields of a TEM (transmission line) mode in a parallel-plate waveguide, forming a plane wave that propagates down the guide axis. (*b*) Plane waves that reflect from the conducting walls can produce a waveguide mode that is no longer TEM.

14.6*b*. Again, a plane wave is guided in the z direction, but does so by means of a progression of zig-zag reflections at the upper and lower plates. Wavevectors $\mathbf{k_u}$ and $\mathbf{k_d}$ are associated with the upward and downward-propagating waves, respectively, and these have identical magnitudes,

$$|\mathbf{k_u}| = |\mathbf{k_d}| = k = \omega\sqrt{\mu\epsilon}$$

For such a wave to propagate, all upward-propagating waves must be *in phase* (as must be true of all downward-propagating waves). This condition can only be satisfied at certain discrete angles of incidence, shown as θ in the figure. An allowed value of θ, along with the resulting field configuration, comprise a *waveguide mode* of the structure. Associated with each guided mode is a *cutoff frequency*. If the operating frequency is below the cutoff frequency, the mode will not propagate. If above cutoff, the mode propagates. The TEM mode, however, has no cutoff; it will be supported at any frequency. At a given frequency, the guide may support several modes, the quantity of which depends on the plate separation and on the dielectric constant of the interior medium, as will be shown. The number of modes increases as the frequency is raised.

So to answer our initial question on the distinction between transmission lines and waveguides, we can state the following: Transmission lines consist of two or more conductors and as a rule will support TEM waves (or something which could approximate such a wave). A waveguide may consist of one or more conductors, or no conductors at all, and will support waveguide modes, of forms similar to those described above. Waveguides may or may not support TEM waves, depending on the design.

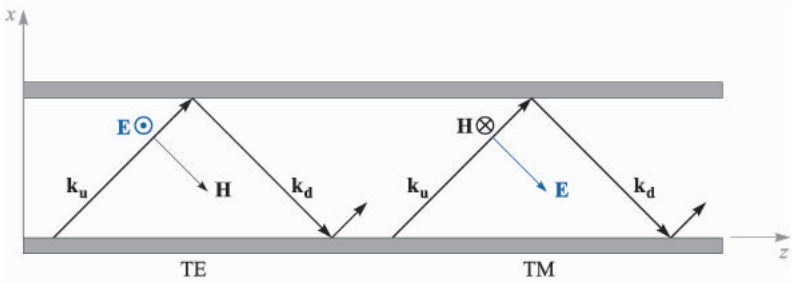


FIGURE 14.7
Plane wave representation of TE and TM modes in a parallel-plate guide.

In the parallel-plate guide, two types of waveguide modes can be supported. These are shown in Fig. 14.7 as arising from the *s* and *p* orientations of the plane wave polarizations. In a manner consistent with our previous discussions on oblique reflection (Sec. 12.5), we identify a *transverse electric* (*TE*) mode when **E** is perpendicular to the plane of incidence (*s* polarized); this positions **E** parallel to the transverse plane of the waveguide, as well as to the boundaries. Similarly, a *transverse magnetic* (*TM*) mode results with a *p* polarized wave; the entire **H** field is in the *y* direction and is thus within the transverse plane of the guide. Both possibilities are illustrated in the figure. Note, for example, that with **E** in the *y* direction (TE mode), **H** will have *x* and *z* components. Likewise, a TM mode will have *x* and *z* components of **E**.[1] In any event, the reader can verify from the geometry of Fig. 14.7 that it is not possible to achieve a purely TEM mode for values of θ other than 90°. Other wave polarizations are possible that lie between the TE and TM cases, but these can always be expressed as superpositions of TE and TM modes.

14.2 PLANE WAVE ANALYSIS OF THE PARALLEL-PLATE WAVEGUIDE

Let us now investigate the conditions under which waveguide modes will occur, using our plane wave model for the mode fields. In Fig. 14.8*a*, a zig-zag path is again shown, but this time phase fronts are drawn that are associated with two of the upward-propagating waves. The first wave has reflected twice (at the top and bottom surfaces) to form the second wave (the downward-propagating phase fronts are not shown). Note that the phase fronts of the second wave do not coincide with those of the first wave, and so the two waves are out of phase. In Fig. 14.8*b*, the wave angle has been adjusted so that the two waves are now in phase. Having satisfied this condition for the two waves, we will find that *all*

[1] Other types of modes can exist in other structures (not the parallel-plate guide) in which *both* **E** and **H** have *z* components. These are known as *hybrid* modes, and typically occur in guides with cylindrical cross sections, such as the optical fiber.

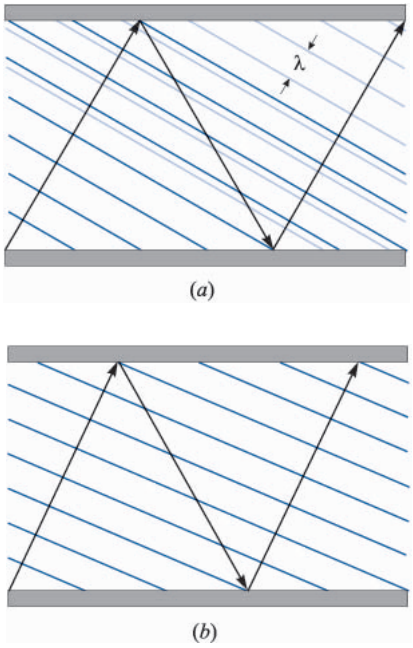


FIGURE 14.8
(*a*) Plane wave propagation in a parallel-plate guide in which the wave angle is such that the upward-propagating waves are not in phase. (*b*) The wave angle has been adjusted so that upward waves are in phase, resulting in a guided mode.

upward-propagating waves will have coincident phase fronts. The same condition will automatically occur for all downward-propagating waves. This is the requirement to establish a guided mode.

In Fig. 14.9 we show the wavevector, $\mathbf{k_u}$, and its components, along with a series of phase fronts. A drawing of this kind for $\mathbf{k_d}$ would be the same, except the x component, κ_m, would be reversed. In Sec. 12.4, we measured the phase shift per unit distance along the x and z directions by the components, k_x and k_z, which varied continuously as the direction of $\mathbf{k}$ changed. In our discussion of waveguides, we introduce a different notation, where κ_m and β_m are used for k_x and k_z. The subscript m is an integer, indicating the *mode number*. This provides a subtle hint that β_m and κ_m will assume only certain discrete values that correspond to certain allowed directions of $\mathbf{k_u}$ and $\mathbf{k_d}$, such that our coincident phase front requirement is satisfied.[2] From the geometry we see that for any value of m,

$$\beta_m = \sqrt{k^2 - \kappa_m^2} \tag{1}$$

Use of the symbol β_m for the z components of $\mathbf{k_u}$ and $\mathbf{k_d}$ is appropriate because β_m will ultimately be the phase constant for the mth waveguide mode, measuring

[2] Subscripts (m) are not shown on $\mathbf{k_u}$ and $\mathbf{k_d}$, but are understood. Changing m does not affect the magnitudes of these vectors—only their directions.

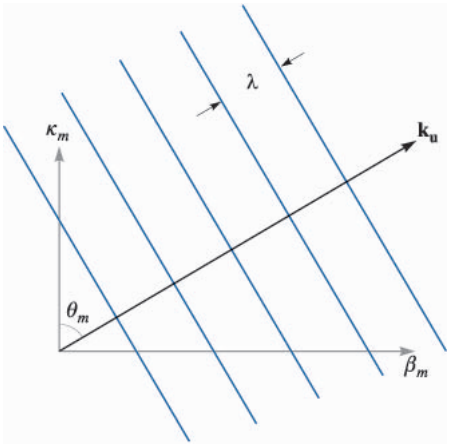


FIGURE 14.9
The components of the upward wavevector are κ_m and β_m, the transverse and axial phase constants. To form the downward wavevector, $\mathbf{k_d}$, the direction of κ_m is reversed.

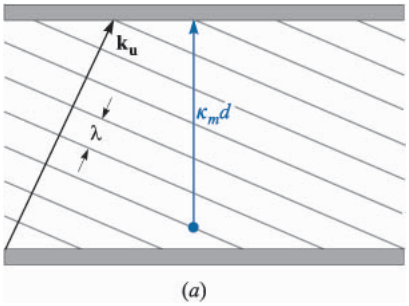
phase shift per distance down the guide; it is also used to determine the phase velocity of the mode, ω/β_m, and the group velocity, $d\omega/d\beta_m$.

Throughout our discussion, we will assume that the medium within the guide is lossless and nonmagnetic, so that

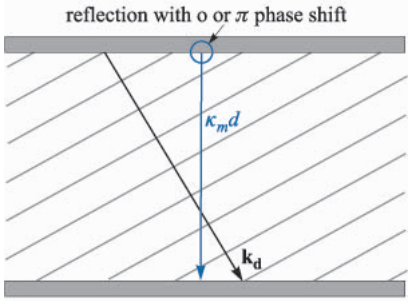
$$k = \omega\sqrt{\mu_0\epsilon'} = \frac{\omega\sqrt{\epsilon'_R}}{c} = \frac{\omega n}{c} \tag{2}$$

which we express either in terms of the dielectric constant, ϵ'_R, or the refractive index, n, of the medium.

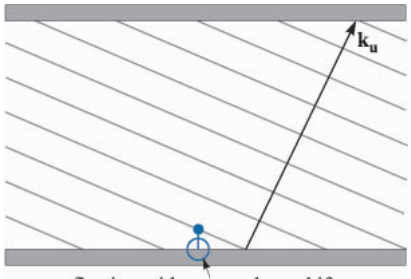
It is κ_m, the x component of $\mathbf{k_u}$ and $\mathbf{k_d}$, that will be useful to us in quantifying our requirement on coincident phase fronts through a condition known as *transverse resonance*. This condition states that the net phase shift measured during a round-trip over the full transverse dimension of the guide must be an integer multiple of 2π radians. This is another way of stating that all upward (or downward) propagating plane waves must have coincident phases. The various segments of this round-trip are illustrated in Fig. 14.10. We assume for this exercise that the waves are frozen in time, and that an observer moves vertically over the round-trip, measuring phase shift along the way. In the first segment (Fig. 14.10*a*) the observer starts at a position just above the lower conductor and moves vertically to the top conductor through distance d. The measured phase shift over this distance is $\kappa_m d$ rad. On reaching the top surface, the observer will note a possible phase shift on reflection (Fig. 14.10*b*). This will be π if the wave is TE polarized, and will be zero if the wave is TM polarized (see Fig. 14.11 for a demonstration of this). Next, the observer moves along the reflected wave phases down to the lower conductor and again measures a phase shift of $\kappa_m d$ (Fig. 14.10*c*). Finally, after including the phase shift on reflection at the bottom conductor, the observer is back at the original starting point, and is noting the phase of the next upward-propagating wave.



(*a*)



(*b*)



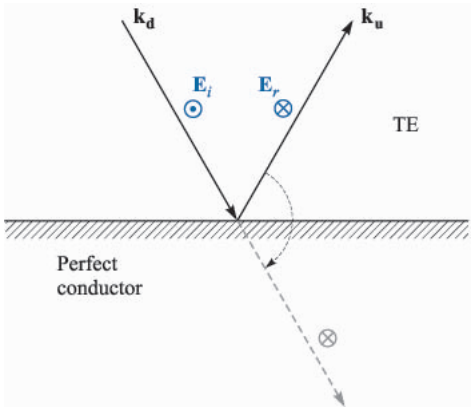
(*c*)

FIGURE 14.10
The net phase shift over a round-trip in the parallel-plate guide is found by first measuring the transverse phase shift between plates of the initial upward wave (*a*); next, the transverse phase shift in the reflected (downward) wave is measured, while accounting for the reflective phase shift at the top plate (*b*); finally, the phase shift on reflection at the bottom plate is added, thus returning to the starting position, but with a new upward wave (*c*). Transverse resonance occurs if the phase at the final point is the same as that at the starting point (the two upward waves are in phase)

The total phase shift over the round-trip is required to be an integer multiple of 2π:

$$\kappa_m d + \phi + \kappa_m d + \phi = 2m\pi \tag{3}$$

where ϕ is the phase shift on reflection at each boundary. Note that with $\phi = \pi$ (TE waves) or 0 (TM waves) the net reflective phase shift over a round-trip is 2π



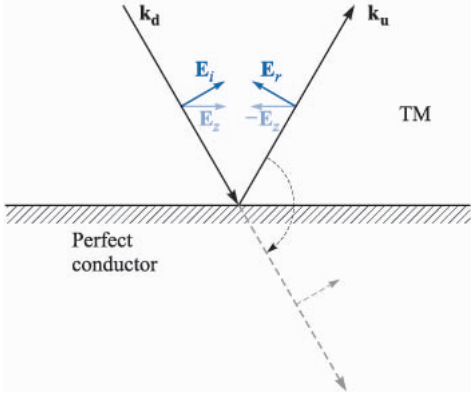


FIGURE 14.11
The phase shift of a wave on reflection from a perfectly conducting surface depends on whether the incident wave is TE (*s*-polarized) or TM (*p*-polarized). In both drawings, electric fields are shown as they would appear immediately adjacent to the conducting boundary. In (*a*) the field of a TE wave reverses direction upon reflection to establish a zero net field at the boundary. This constitutes a π phase shift, as is evident by considering a fictitious transmitted wave (*dashed line*), formed by a simple rotation of the reflected wave into alignment with the incident wave. In (*b*) an incident TM wave experiences a reversal of the *z component* of its electric field. The resultant field of the reflected wave, however, has not been phase-shifted; rotating the reflected wave into alignment with the incident wave (*dashed line*) shows this.

or 0, regardless of the angle of incidence. Thus the reflective phase shift has no bearing on the current problem, and we may simplify (3) to read:

$$\boxed{\kappa_m = \frac{m\pi}{d}} \tag{4}$$

which is valid for *both* TE and TM modes. Note from Fig. 14.9 that $\kappa_m = k\cos\theta_m$. Thus the wave angles for the allowed modes are readily found from (4) with (2):

$$\theta_m = \cos^{-1}\left(\frac{m\pi}{kd}\right) = \cos^{-1}\left(\frac{m\pi c}{\omega nd}\right) = \cos^{-1}\left(\frac{m\lambda}{2nd}\right) \tag{5}$$

where λ is the wavelength in free space.

We can next solve for the phase constant for each mode, using (1) with (4):

$$\beta_m = \sqrt{k^2 - \kappa_m^2} = k\sqrt{1 - \left(\frac{m\pi}{kd}\right)^2} = k\sqrt{1 - \left(\frac{m\pi c}{\omega nd}\right)^2} \tag{6}$$

We define the radian *cutoff frequency* for mode m as

$$\boxed{\omega_{cm} = \frac{m\pi c}{nd}} \tag{7}$$

so that (6) becomes

$$\boxed{\beta_m = \frac{n\omega}{c}\sqrt{1 - \left(\frac{\omega_{cm}}{\omega}\right)^2}} \tag{8}$$

The significance of the cutoff frequency is readily seen from (8): If the operating frequency, ω, is greater than the cutoff frequency for mode m, then that mode will have phase constant β_m that is real-valued, and so the mode will propagate. For $\omega < \omega_{cm}$, β_m is imaginary, and the mode does not propagate.

Associated with the cutoff frequency is the *cutoff wavelength*, λ_{cm}, defined as the free-space wavelength at which cutoff for mode m occurs. This will be

$$\boxed{\lambda_{cm} = \frac{2\pi c}{\omega_{cm}} = \frac{2nd}{m}} \tag{9}$$

Note, for example, that in an air-filled guide (n = 1) the wavelength at which the lowest-order mode first starts to propagate is $\lambda_{c1} = 2d$, or the plate separation is one-half wavelength. Mode m will propagate whenever $\omega > \omega_{cm}$, or equivalently whenever $\lambda < \lambda_{cm}$. Use of the cutoff wavelength enables us to construct a second useful form of Eq. (8):

$$\boxed{\beta_m = \frac{2\pi n}{\lambda}\sqrt{1 - \left(\frac{\lambda}{\lambda_{cm}}\right)^2}} \tag{10}$$

Example 14.1

A parallel-plate transmission line has plate separation $d = 1$ cm, and is filled with teflon, having dielectric constant $\epsilon'_R = 2.1$. Determine the maximum operating frequency such that only the TEM mode will propagate. Also find the range of frequencies over which the TE_1 and $TM_1 (m = 1)$ modes will propagate, and no higher order modes.

Solution. Using (7), the cutoff frequency for the first waveguide mode ($m = 1$), will be

$$f_{c1} = \frac{\omega_{c1}}{2\pi} = \frac{2.99 \times 10^{10}}{2\sqrt{2.1}} = 1.03 \times 10^{10} \text{ Hz} = 10.3 \text{ GHz}$$

To propagate only TEM waves, we must have $f < 10.3$ GHz. To allow TE_1 and TM_1 (along with TEM) only, the frequency range must be $\omega_{c1} < \omega < \omega_{c2}$, where $\omega_{c2} = 2\omega_{c1}$, from (7). Thus, the frequencies at which we will have the $m = 1$ modes and TEM will be $10.3\,\text{GHz} < f < 20.6\,\text{GHz}$.

Example 14.2

In the parallel-plate guide of Example 14.1, the operating wavelength is $\lambda = 2$ mm. How many waveguide modes will propagate?

Solution. For mode m to propagate, the requirement is $\lambda < \lambda_{cm}$. For the given waveguide and wavelength, the inequality becomes, using (9),

$$2 \text{ mm} < \frac{2\sqrt{2.1}\ (10\,\text{mm})}{m}$$

from which

$$m < \frac{2\sqrt{2.1}\ (10\,\text{mm})}{2\,\text{mm}} = 14.5$$

Thus the guide will support modes at the given wavelength up to order $m = 14$. Since there will be a TE and a TM mode for each value of m, this gives, in addition to the TEM mode, a total of 28 guided modes that are above cutoff.

The field configuration for a given mode can be found through the superposition of the fields of all the reflected waves. We can do this for the TE waves, for example, by writing the electric field phasor in the guide in terms of incident and reflected fields through

$$E_{ys} = E_0 e^{-j\mathbf{k_u}\cdot\mathbf{r}} - E_0 e^{-j\mathbf{k_d}\cdot\mathbf{r}} \tag{11}$$

where the wavevectors, $\mathbf{k_u}$ and $\mathbf{k_d}$, are indicated in Fig. 14.7. The minus sign in front of the second term arises from the π phase shift on reflection. From the geometry depicted in Fig. 14.9, we write

$$\mathbf{k_u} = \kappa_m \mathbf{a}_x + \beta_m \mathbf{a}_z \tag{12}$$

and

$$\mathbf{k_d} = -\kappa_m \mathbf{a}_x + \beta_m \mathbf{a}_z \tag{13}$$

Then, using

$$\mathbf{r} = x\mathbf{a}_x + z\mathbf{a}_z$$

Eq. (11) becomes

$$E_{ys} = E_0\left(e^{-j\kappa_m x} - e^{j\kappa_m x}\right)e^{-j\beta_m z} = 2jE_0 \sin(\kappa_m x)e^{-j\beta_m z} = E_0' \sin(\kappa_m x)e^{-j\beta_m z} \tag{14}$$

where the plane wave amplitude, E_0, and the overall phase are absorbed into E_0'. In real instantaneous form, (14) becomes

$$E_y(z, t) = \mathrm{Re}\left(E_{ys}e^{j\omega t}\right) = E_0' \sin(\kappa_m x)\cos(\omega t - \beta_m z) \qquad \text{(TE mode above cutoff)} \tag{15}$$

We interpret this as a wave that propagates in the positive z direction (down the guide) while having a field profile that varies with x.[3] The TE mode field is the *interference pattern* resulting from the superposition of the upward and downward plane waves. Note that if $\omega < \omega_{cm}$, then (8) yields an imaginary value for β_m, which we may write as $-j|\beta_m| = -j\alpha_m$. Eqs. (14) and (15) then become

$$E_{ys} = E_0' \sin(\kappa_m x)e^{-\alpha_m z} \tag{16}$$

$$E(z, t) = E_0' \sin(\kappa_m x)e^{-\alpha_m z}\cos(\omega t) \qquad \text{(TE mode below cutoff)} \tag{17}$$

This mode does not propagate, but simply oscillates at frequency ω, while exhibiting a field pattern that decreases in strength with increasing z. The attenuation coefficient, α_m, is found from (8) with $\omega < \omega_{cm}$:

$$\alpha_m = \frac{n\omega_{cm}}{c}\sqrt{1 - \left(\frac{\omega}{\omega_{cm}}\right)^2} = \frac{2\pi n}{\lambda_{cm}}\sqrt{1 - \left(\frac{\lambda_{cm}}{\lambda}\right)^2} \tag{18}$$

We note from (5) and (7) that the plane wave angle is related to the cutoff frequency and cutoff wavelength through

$$\boxed{\cos\theta_m = \frac{\omega_{cm}}{\omega} = \frac{\lambda}{\lambda_{cm}}} \tag{19}$$

So we see that at cutoff ($\omega = \omega_{cm}$), $\theta_m = 0$, and the plane waves are just reflecting back and forth over the cross section; they are making no forward progress down the guide. As ω is increased beyond cutoff (or λ is decreased), the wave angle increases, approaching 90° as ω approaches infinity (or as λ approaches zero). From Fig. 14.9, we have

$$\boxed{\beta_m = k\sin\theta_m = \frac{n\omega}{c}\sin\theta_m} \tag{20}$$

[3] We can also interpret this field as that of a standing wave in x, while it is a traveling wave in z.

and so the phase velocity of mode m will be

$$v_{pm} = \frac{\omega}{\beta_m} = \frac{c}{n \sin \theta_m} \tag{21}$$

The velocity minimizes at c/n for all modes, approaching this value at frequencies far above cutoff; v_{pm} approaches infinity as the frequency is reduced to approach the cutoff frequency. Again, phase velocity is the speed of the phases in the z direction, and the fact that this velocity may exceed the speed of light in the medium is not a violation of relativistic principles, as discussed in Sec. 12.5.

The energy will propagate at the group velocity, $v_g = d\omega/d\beta$. Using (8), we have

$$v_{gm}^{-1} = \frac{d\beta_m}{d\omega} = \frac{d}{d\omega}\left[\frac{n\omega}{c}\sqrt{1 - \left(\frac{\omega_{cm}}{\omega}\right)^2}\right] \tag{22}$$

The derivative is straightforward. Carrying it out and taking the reciprocal of the result yields:

$$v_{gm} = \frac{c}{n}\sqrt{1 - \left(\frac{\omega_{cm}}{\omega}\right)^2} = \frac{c}{n}\sin\theta_m \tag{23}$$

Group velocity is thus identified as the projection of the velocity associated with $\mathbf{k_u}$ or $\mathbf{k_d}$ into the z direction. This will be less than or equal to the velocity of light in the medium, c/n, as expected.

Example 14.3

In the guide of Example 14.1, the operating frequency is 25 GHz. Consequently, modes for which $m = 1$ and $m = 2$ will be above cutoff. Determine the *group delay difference* between these two modes over a distance of 1 cm. This is the difference in propagation times between the two modes when energy in each propagates over the 1 cm distance.

Solution. The group delay difference is expressed as

$$\Delta t = \left(\frac{1}{v_{g2}} - \frac{1}{v_{g1}}\right) \text{(s/cm)}$$

From (23), along with the results of Example 14.1, we have

$$v_{g1} = \frac{c}{\sqrt{2.1}}\sqrt{1 - \left(\frac{10.3}{25}\right)^2} = 0.63c$$

$$v_{g2} = \frac{c}{\sqrt{2.1}}\sqrt{1-\left(\frac{20.6}{25}\right)^2} = 0.39c$$

Then

$$\Delta t = \frac{1}{c}\left[\frac{1}{.39} - \frac{1}{.63}\right] = 3.3 \times 10^{-11} \text{ s/cm} = 33 \text{ ps/cm}$$

This computation gives a rough measure of the *modal dispersion* in the guide, applying to the case of having only two modes propagating. A pulse, for example, whose center frequency is 25 GHz would have its energy divided between the two modes. The pulse would broaden by approximately 33 ps/cm of propagation distance as the energy in the modes separates. If, however, we include the TEM mode (as we really must), then the broadening will be even greater. The group velocity for TEM will be $c/\sqrt{2.1}$. The group delay difference of interest will then be between the TEM mode and the $m = 2$ mode (TE or TM). We would thus have

$$\Delta t_{net} = \frac{1}{c}\left[\frac{1}{.39} - 1\right] = 52 \text{ ps/cm}$$

✔ **D14.1.** Determine the wave angles, θ_m, for the first four modes ($m = 1,2,3,4$) in a parallel-plate guide with $d = 2$ cm, $\epsilon'_R = 1$, and $f = 30$ GHz.

Ans. 76°; 60°; 41°; 0°.

✔ **D14.2.** A parallel-plate guide has plate spacing $d = 5$ mm, and is filled with glass ($n = 1.45$). What is the maximum frequency at which the guide will operate in the TEM mode only?

Ans. 20.7 GHz.

✔ **D14.3.** A parallel-plate guide having $d = 1$ cm is filled with air. Find the cutoff wavelength for the $m = 2$ mode (TE or TM).

Ans. 1 cm.

14.3 PARALLEL-PLATE GUIDE ANALYSIS USING THE WAVE EQUATION

The most direct approach in the analysis of any waveguide is through the wave equation, which we solve subject to the boundary conditions at the conducting walls. The form of the equation that we will use is that of Eq. (12) in Sec. 11.1, which was written for the case of free-space propagation. We account for the dielectric properties in the waveguide by replacing k_0 in that equation with k to obtain:

$$\boxed{\nabla^2 \mathbf{E_s} = -k^2 \mathbf{E_s}} \tag{24}$$

where $k = n\omega/c$ as before.

We can use the results of the last section to help visualize the process of solving the wave equation. For example, we may consider TE modes first, in which there will be only a y component of $\mathbf{E}$. The wave equation becomes:

$$\frac{\partial^2 E_{ys}}{\partial x^2} + \frac{\partial^2 E_{ys}}{\partial y^2} + \frac{\partial^2 E_{ys}}{\partial z^2} + k^2 E_{ys} = 0 \tag{25}$$

We assume that the width of the guide (in the y direction) is very large compared to the plate separation, d. Therefore we can assume no y variation in the fields (fringing fields are ignored), and so $\partial^2 E_{ys}/\partial y^2 = 0$. We also know that the z variation will be of the form $e^{-j\beta_m z}$. The form of the field solution will thus be

$$\boxed{E_{ys} = E_0 f_m(x) e^{-j\beta_m z}} \tag{26}$$

where E_0 is a constant, and where $f_m(x)$ is a normalized function to be determined (whose maximum value is unity). We have included subscript m on β, κ, and $f(x)$, since we anticipate several solutions that correspond to discrete modes, to which we associate mode number, m. We now substitute (26) into (25) to obtain:

$$\boxed{\frac{d^2 f_m(x)}{dx^2} + (k^2 - \beta_m^2) f_m(x) = 0} \tag{27}$$

where E_0 and $e^{-j\beta_m z}$ have divided out, and where we have used the fact that

$$\frac{d^2}{dz^2} e^{-j\beta_m z} = -\beta_m^2 e^{-j\beta_m z}$$

Note also that we have written (27) using the total derivative, d^2/dx^2 since f_m is a function only of x. We next make use of the geometry of Fig. 14.9, and note that $k^2 - \beta_m^2 = \kappa_m^2$. Using this in (27) we obtain

$$\boxed{\frac{d^2 f_m(x)}{dx^2} + \kappa_m^2 f_m(x) = 0} \tag{28}$$

The general solution of (28) will be

$$\boxed{f_m(x) = \cos(\kappa_m x) + \sin(\kappa_m x)} \tag{29}$$

We next apply the appropriate boundary conditions in our problem to evaluate κ_m. From Fig. 14.1, conducting boundaries appear at $x = 0$ and $x = d$, at which the tangential electric field (E_y) must be zero. In Eq. (29), only the $\sin(\kappa_m x)$ term will allow the boundary conditions to be satisfied, so

we retain it and drop the cosine term. The $x = 0$ condition is automatically satisfied by the sine function. The $x = d$ condition is met when we choose the value of κ_m such that

$$\kappa_m = \frac{m\pi}{d} \tag{30}$$

We recognize Eq. (30) as the same result that we obtained using the transverse resonance condition of the last section. The final form of E_{ys} is obtained by substituting $f_m(x)$ as expressed through (29) and (30) into (26), yielding a result that is consistent with the one expressed in Eq. (14):

$$\boxed{E_{ys} = E_0 \sin\left(\frac{m\pi x}{d}\right) e^{-j\beta_m z}} \tag{31}$$

An additional significance of the mode number m is seen when considering the form of the electric field of (31). Specifically, m is the number of spatial half-cycles of electric field that occur over the distance d in the transverse plane. This can be understood physically by considering the behavior of the guide at cutoff. As we learned in the last section, the plane wave angle of incidence in the guide at cutoff is zero, meaning that the wave simply bounces up and down between the conducting walls. The wave must be resonant in the structure, which means that the net round-trip phase shift is $2m\pi$. With the plane waves oriented vertically, $\beta_m = 0$, and so $\kappa_m = k = 2n\pi/\lambda_{cm}$. So at cutoff,

$$\frac{m\pi}{d} = \frac{2n\pi}{\lambda_{cm}} \tag{32}$$

which leads to

$$d = \frac{m\lambda_{cm}}{2n} \qquad \text{at cutoff} \tag{33}$$

Eq. (31) at cutoff then becomes

$$E_{ys} = E_0 \sin\left(\frac{m\pi x}{d}\right) = E_0 \sin\left(\frac{2n\pi x}{\lambda_{cm}}\right) \tag{34}$$

The waveguide is simply a one-dimensional *resonant cavity*, in which a wave can oscillate in the x direction if its wavelength as measured in the medium is an integer multiple of $2d$, where the integer is m.

Now, as the frequency increases, wavelength will decrease, and so the above requirement of wavelength equaling an integer multiple of $2d$ is no longer met. The response of the mode is to establish *z components* of $\mathbf{k_u}$ and $\mathbf{k_d}$ which results in the decreased wavelength being compensated by an increase in wavelength *as measured in the x direction*. Figure 14.12*a* and *b* show this effect for the $m = 4$ mode, in which the wave angle, θ_4, steadily increases with increasing frequency. Thus, the mode retains precisely the functional form of its field in the x direction, but establishes an increasing value of β_m as the frequency is raised. This invar-

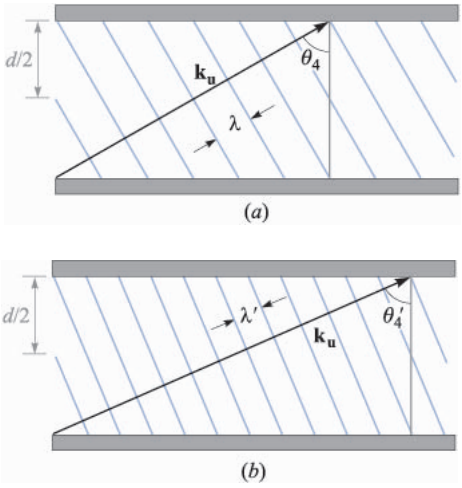


FIGURE 14.12
(*a*) A plane wave associated with an $m = 4$ mode, showing a net phase shift of 4π (two wavelengths measured in x) occurring over distance d in the transverse plane. (*b*) As frequency increases, an increase in wave angle is required to maintain the 4π transverse phase shift.

iance in the transverse spatial pattern means that the mode will retain its identity at all frequencies. Group velocity, expressed in (23), is changing as well, meaning that the changing wave angle with frequency is a mechanism for group velocity dispersion, known simply as *waveguide dispersion*. Pulses, for example, that propagate in a single waveguide mode will thus experience broadening in the manner considered in Sec. 12.6.

Having found the electric field, we can find the magnetic field using Maxwell's equations. We note from our plane wave model that we expect to obtain x and z components of $\mathbf{H_s}$ for a TE mode. We use the Maxwell equation

$$\nabla \times \mathbf{E_s} = -j\omega\mu\mathbf{H_s} \tag{35}$$

where, in the present case of having only a y component of $\mathbf{E_s}$, we have

$$\nabla \times \mathbf{E_s} = \frac{\partial E_{ys}}{\partial x}\mathbf{a}_z - \frac{\partial E_{ys}}{\partial z}\mathbf{a}_x = \kappa_m E_0 \cos(\kappa_m x)e^{-j\beta_m z}\mathbf{a}_z + j\beta_m E_0 \sin(\kappa_m x)e^{-j\beta_m z}\mathbf{a}_x \tag{36}$$

We solve for $\mathbf{H_s}$ by dividing both sides of (35) by $-j\omega\mu$. Performing this operation on (36), we obtain the two magnetic field components:

$$H_{xs} = -\frac{\beta_m}{\omega\mu}E_0 \sin(\kappa_m x)e^{-j\beta_m z} \tag{37}$$

$$H_{zs} = j\frac{\kappa_m}{\omega\mu}E_0 \cos(\kappa_m x)e^{-j\beta_m z} \tag{38}$$

Together, these two components form closed loop patterns for $\mathbf{H_s}$ in the x, z plane, as can be verified using the streamline plotting methods developed in Sec. 2.6.

It is interesting to consider the magnitude of $\mathbf{H_s}$, which is found through

$$|\mathbf{H_s}| = \sqrt{\mathbf{H_s} \cdot \mathbf{H_s^*}} = \sqrt{H_{xs}H_{xs}^* + H_{zs}H_{zs}^*} \tag{39}$$

Carrying this out using (37) and (38) results in

$$|\mathbf{H_s}| = \frac{E_o}{\omega\mu}\left(\kappa_m^2 + \beta_m^2\right)^{1/2}\left(\sin^2(\kappa_m x) + \cos^2(\kappa_m x)\right)^{1/2} \tag{40}$$

Using the fact that $\kappa_m^2 + \beta_m^2 = k^2$ and using the identity $\sin^2(\kappa_m x) + \cos^2(\kappa_m x) = 1$, (40) becomes

$$|\mathbf{H_s}| = \frac{k}{\omega\mu}E_0 = \frac{\omega\sqrt{\mu\epsilon}}{\omega\mu} = \frac{E_0}{\eta} \tag{41}$$

where $\eta = \sqrt{\mu/\epsilon}$. This result is consistent with our understanding of waveguide modes based on the superposition of plane waves, in which the relation between $\mathbf{E_s}$ and $\mathbf{H_s}$ is through the medium intrinsic impedance, η.

✔ **D14.4.** Determine the group velocity of the $m = 1$ (TE or TM) mode in an air-filled parallel-plate guide with $d = 0.5$ cm at $f =$ (a) 30 GHz, (b) 60 GHz, and (c) 100 GHz.

Ans. 0; 2.6×10^8 m/s; 2.9×10^8 m/s.

✔ **D14.5.** A TE mode in a parallel-plate guide is observed to have three maxima in its electric field pattern between $x = 0$ and $x = d$. What is the value of m?

Ans. 3.

14.4 RECTANGULAR WAVEGUIDES

In this section we consider the rectangular waveguide, a widely used structure that is usually used in the microwave region of the electromagnetic spectrum. A brief analysis of the structure will be presented here, with the goal of understanding the key operational features and special attributes of the guide. The reader is referred to Ref. 1 for further study.

The rectangular guide is shown in Fig. 14.2. We can relate this structure to that of the parallel-plate guide of the previous sections by thinking of it as two parallel-plate guides of orthogonal orientation that are assembled to form one unit. We thus have a pair of horizontal conducting walls (along the x direction) and a pair of vertical walls (along y), all of which now form one continuous boundary. The wave equation in its full three-dimensional form (Eq. (25)) must now be solved, since in general we may have field variations in all three coordinate directions. Assuming that the variation with z will be just $e^{-j\beta z}$ as before,

and assuming, for example, the existence of a y component of $\mathbf{E_s}$, Eq. (25) will take the form:

$$\frac{\partial^2 E_{ys}}{\partial x^2} + \frac{\partial^2 E_{ys}}{\partial y^2} + (k^2 - \beta_{mp}^2)E_{ys} = 0 \tag{42}$$

This equation brings to mind the two-dimensional Laplace equation problem of Sec. 7.5, in which the product solution method was used. The same basic method is used to solve (42) as well, and the resulting solution takes the general form:

$$E_{ys} = E_0 f_m(x) f_p(y) e^{-j\beta_{mp}z} \tag{43}$$

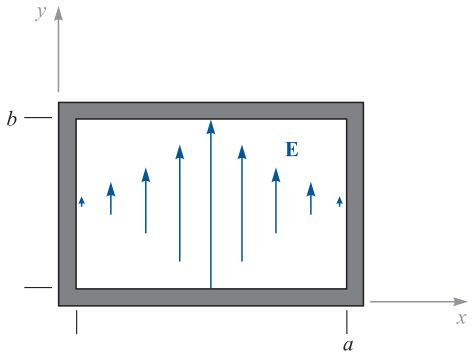
where f_m and f_p are sines or cosines. Two integers, m and p, are now needed to describe field variations in the x and y directions. Again, we are concerned with TE and TM modes, and the wave equation is solved separately for each type.

From here, the problem gets complicated and goes beyond the scope (and purpose) of our present treatment. Instead, much can be learned about this guide through our intuition, and through our knowledge of the parallel-plate guide. It turns out that the most important modes in the rectangular waveguide are of the same form as those of the parallel-plate structure. Consider, for example, the appearence of a TE mode in the rectangular guide. The electric field of such a mode could appear as shown in Fig. 14.13*a*, in which the field is vertically polarized, and terminates on the top and bottom plates. The field also becomes zero at the two vertical walls, as is required from our boundary condition on tangential electric field at a conducting surface. Let us consider the case in which the field exhibits no variation with y, but does vary with x and z (the latter according to $e^{-j\beta z}$). Consequently the $\partial^2/\partial y^2$ term in the wave equation (42) drops out and the equation becomes identical in form to (27), which was used for the parallel-plate guide. The field solutions are thus identical in form to (31), (37) and (38), but with a few minor notation differences:

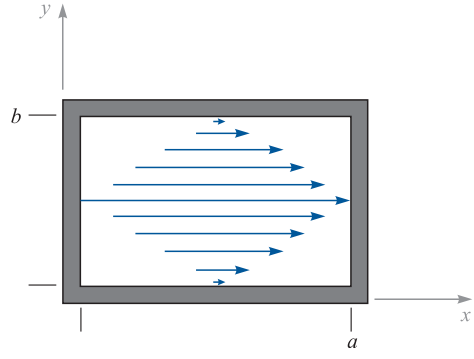
$$E_{ys} = E_0 \sin(\kappa_{m0}x) e^{-j\beta_{m0}z} \tag{44}$$

$$H_{xs} = -\frac{\beta_{m0}}{\omega\mu} E_0 \sin(\kappa_{m0}x) e^{-j\beta_{m0}z} \tag{45}$$

$$H_{zs} = j\frac{\kappa_{m0}}{\omega\mu} E_0 \cos(\kappa_{m0}x) e^{-j\beta_{m0}z} \tag{46}$$



(*a*)



(*b*)

FIGURE 14.13
(*a*) TE_{10} and (*b*) TE_{01} mode electric field configurations in a rectangular waveguide.

where

$$\kappa_{m0} = \frac{m\pi}{a} \tag{47}$$

The fields of Eqs. (44) through (46) are those of a general mode of designation TE_{m0}, where the subscripts indicate that there are m half cycles of the electric field over the x dimension and zero variation in y. The phase constant is subscripted $m0$, and it is still true that

$$\kappa_{m0}^2 + \beta_{m0}^2 = k^2$$

The cutoff frequency for the TE_{m0} mode is given by (7), appropriately modified:

$$\boxed{\omega_c(m0) = \frac{m\pi c}{na}} \tag{48}$$

The phase constant, β_{m0}, is given by (8); all of the implications on mode behavior above and below cutoff are exactly the same as we found for the parallel-plate guide. The plane wave analysis is also carried out in the same manner. TE_{m0} modes can be modeled as plane waves that propagate down the guide by reflecting between the vertical side walls.

Another possibility is the TE_{0p} field configuration shown in Fig. 14.13*b*, which shows a horizontally polarized electric field. The wave equation (42) would now contain the $\partial^2/\partial y^2$ term and the $\partial^2/\partial x^2$ term would be dropped. The resulting fields would be those of (44) through (46) after a rotation through 90°, along with a few notational changes:

$$E_{xs} = E_0 \sin(\kappa_{0p} y) e^{-j\beta_{0p} z} \tag{49}$$

$$H_{ys} = \frac{\beta_{0p}}{\omega\mu} E_0 \sin(\kappa_{0p} y) e^{-j\beta_{0p} z} \tag{50}$$

$$H_{zs} = -j\frac{\kappa_{0p}}{\omega\mu} E_0 \cos(\kappa_{0p} y) e^{-j\beta_{0p} z} \tag{51}$$

where

$$\kappa_{0p} = \frac{p\pi}{b} \tag{52}$$

and where the cutoff frequency will be

$$\boxed{\omega_c(0p) = \frac{p\pi c}{nb}} \tag{53}$$

Other modes are possible which exhibit variations in both x and y. In general, the cutoff frequency for these modes is given by

$$\boxed{\omega_c(mp) = \sqrt{\left(\frac{m\pi c}{na}\right)^2 + \left(\frac{p\pi c}{nb}\right)^2}} \tag{54}$$

Modes having variations in both transverse directions include TE and TM, but only TE modes can have zero variation in x or y.

Of considerable practical interest is the mode that has the lowest cutoff frequency. If the guide dimensions are such that $b < a$, then inspection of (48) and (54) indicates that the lowest cutoff will occur for the TE_{10} mode. As such, this is the dominant (and most important) mode in the rectangular waveguide, since it can propagate alone if the operating frequency is appropriately chosen.

Example 14.4

An air-filled rectangular waveguide has dimensions $a = 2$ cm and $b = 1$ cm. Determine the range of frequencies over which the guide will operate single mode (TE_{10}).

Solution. Since the guide is air-filled, $n = 1$, and (48) gives, for $m = 1$:

$$f_c(10) = \frac{\omega_c}{2\pi} = \frac{c}{2a} = \frac{3 \times 10^{10}}{2(2)} = 7.5 \text{ GHz}$$

The next higher-order mode will be either TE_{20} or TE_{01}, which, from (48) and (54), will have the same cutoff frequency, since $a = 2b$. This frequency will be twice that found for TE_{10}, or 15 GHz. Thus the operating frequency range over which the guide will be single mode is $7.5 \text{ GHz} < f < 15 \text{ GHz}$.

Having seen how rectangular waveguides work, the questions arise: Why are they used and when are they useful? Let us consider for a moment the operation of a transmission line at frequencies high enough such that waveguide modes can occur. The onset of guided modes in a transmission line, known as *moding*, is in fact a problem that needs to be avoided, because signal distortion may result. A signal that is input to such a line will find its power divided in some proportions among the various modes. The signal power in each mode propagates at a group velocity unique to that mode. With the power thus distributed, distortion will occur over sufficient distances, as the signal components among the modes lose synchronization with each other, owing to the different delay times (group delays) associated with the different modes. We encountered this concept in Example 14.3.

The above problem of *modal dispersion* in transmission lines is avoided by ensuring that only the TEM mode propagates, and that all waveguide modes are below cutoff. This is accomplished either by using line dimensions that are smaller than one-half the signal wavelength, or by ensuring an upper limit to the operating frequency in a given line. But it is more complicated than this.

In Chap. 13, we saw that increasing the frequency increases the line loss as a result of the skin effect. This is manifested through the increase in the series resistance per unit length, R. One can compensate by increasing one or more dimensions in the line cross section, as Eqs. (17) and (33) in Chap. 13 show, but only to the point at which moding may occur. Typically, the increasing loss with increasing frequency will render the transmission line useless before the onset of moding, but one still cannot increase the line dimensions to reduce losses without considering the possibility of moding. This limitation on dimensions also limits the power-handling capability of the line, since the voltage at which dielectric breakdown occurs decreases with decreasing conductor separation. Consequently, the use of transmission lines as frequencies increase beyond a certain point becomes undesirable, since losses will become excessive, and since the limitation on dimensions will limit the power-handling capability. Instead, we look to other guiding structures, among which is the rectangular guide.

The important fundamental difference between the rectangular waveguide (or any hollow pipe guide) and the transmission line is that the rectangular guide *will not support a TEM mode*. We have already demonstrated this in our study of the TE wave. The fact that the guide is formed from a completely enclosed metal structure means that any electric field distribution in the transverse plane must exhibit variations in the plane; this is because all electric field components that are tangent to the conductors must be zero at the conducting boundaries. Since **E** varies in the transverse plane, the computation of **H** through $\nabla \times \mathbf{E} = -j\omega\mu\mathbf{H}$ *must* lead to a z component of **H**, and so we cannot have a TEM mode. We cannot find any other orientation of a completely transverse **E** in the guide that will allow a completely transverse **H**.

Since the rectangular guide will not support a TEM mode, it will not operate until the frequency exceeds the cutoff frequency of the lowest-order guided mode of the structure. Thus, it must be constructed of large enough size to accomplish this for a given frequency; the required transverse dimensions will consequently be larger than those of a transmission line that is designed to support only the TEM mode. The increased size, coupled with the fact that there is more conductor surface area than in a transmission line of equal volume means that losses will be substantially lower in the rectangular waveguide structure. Additionally, the guides will support more power at a given electric field strength than a transmission line, since the rectangular guide will have a higher cross-sectional area.

Still, hollow pipe guides must operate in a single mode in order to avoid the signal distortion problems arising from multimode transmission. This means that the guides must be of dimension such that they operate above the cutoff frequency of the lowest-order mode, but below the cutoff frequency of the next higher-order mode, as demonstrated in Example 14.4. Increasing the operating frequency again means that the guide transverse dimensions must be decreased to maintain single-mode operation. This can be accomplished to a point at which skin effect losses again become problematic (remember that the skin depth is decreasing with increasing frequency, in addition to the decrease in metal surface area with diminishing guide size). In addition, the guides become too difficult to fabricate, with machining tolerances becoming more stringent. So again, as frequencies are further increased, we look for another type of structure.

✔ **D14.6.** Specify the minimum width, a, and the maximum height, b, of an air-filled rectangular guide so that it will operate single mode over the frequency range $15\,\text{GHz} < f < 20\,\text{GHz}$.

Ans. 1 cm; 0.75 cm.

14.5 DIELECTRIC WAVEGUIDES

When skin effect losses become excessive, a good way to remove them is to remove the metal in the structure entirely, and use interfaces between dielectrics

for the confining surfaces. We thus obtain a *dielectric waveguide*; a basic form, the *symmetric slab waveguide*, is shown in Fig. 14.14. The structure is so-named because of its vertical symmetry about the z axis. The guide is assumed to have width in y much greater than the slab thickness, d, so the problem becomes two-dimensional, with fields presumed to vary with x and z while being independent of y. The slab guide works in very much the same way as the parallel plate waveguide, except wave reflections occur at the interfaces between dielectrics, having different refractive indices, n_1 for the slab, and n_2 for the surrounding regions above and below. In the dielectric guide, total reflection is needed, so the incident angle must exceed the critical angle. Consequently, as discussed in Sec. 12.5, the slab index, n_1, must be greater than that of the surrounding materials, n_2. Dielectric guides differ from conducting guides in that power is not completely confined to the slab, but resides partially above and below.

Dielectric guides are used primarily at optical frequencies (on the order of 10^{14} Hz). Again, guide transverse dimensions must be kept on the order of a wavelength to achieve operation in a single mode. A number of fabrication methods can be used to accomplish this. For example, a glass plate can be doped with materials that will raise the refractive index. The doping process allows materials to be introduced only within a thin layer adjacent to the surface that is a few micrometers thick.

To understand how the guide operates, consider Fig. 14.15, which shows a wave propagating through the slab by multiple reflections, but where *partial transmission* into the upper and lower regions occurs at each bounce. Wavevectors are shown in the middle and upper regions, along with their components in the x and z directions. As we found in Chap. 12, the z components (β) of all wavevectors are equal, as must be true if the field boundary conditions at the interfaces are to be satisfied for all positions and times. Partial transmission at the boundaries is, of course, an undesirable situation, since power in the slab will eventually leak away. We thus have a *leaky wave* propagating in the struc-

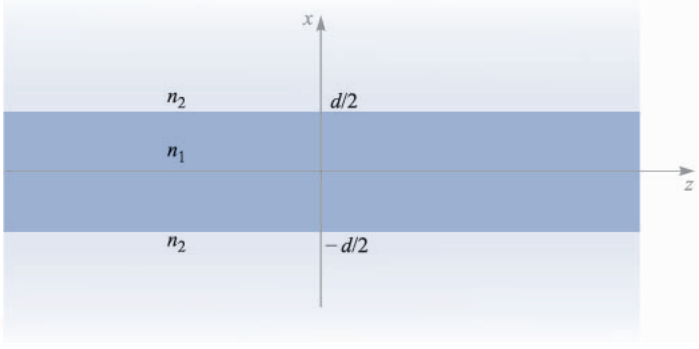


FIGURE 14.14
Symmetric dielectric slab waveguide structure, where waves propagate along z. The guide is assumed infinite in the y direction, thus making the problem two-dimensional.

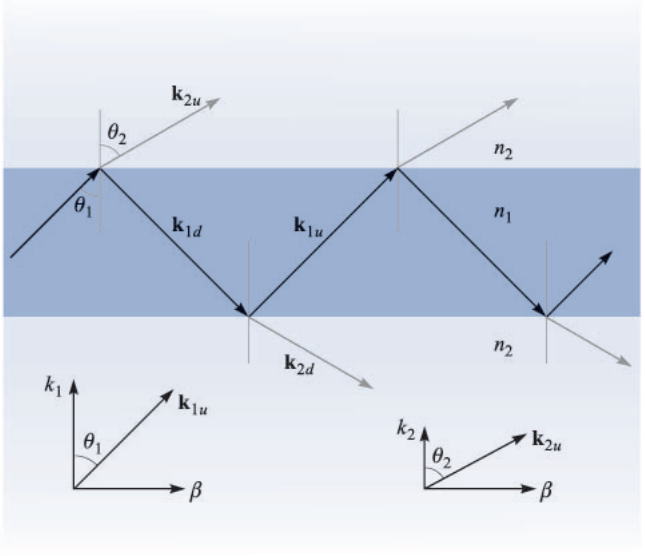


FIGURE 14.15
Plane wave geometry of a leaky wave in a symmetric slab waveguide. For a guided mode, total reflection occurs in the interior, and the x components of $\mathbf{k}_{2u}$ and $\mathbf{k}_{2d}$ are imaginary.

ture, whereas we need to have a guided mode. Note that in either case, we still have the two possibilities on wave polarization, and the resulting mode designation—either TE or TM.

Total power reflection at the boundaries for TE or TM waves implies, respectively, that $|\Gamma_s|^2$ or $|\Gamma_p|^2$ is unity, where the reflection coefficients are given in Eqs. (71) and (69) in Chap. 12:

$$\Gamma_s = \frac{\eta_{2s} - \eta_{1s}}{\eta_{2s} + \eta_{1s}} \tag{55}$$

and

$$\Gamma_p = \frac{\eta_{2p} - \eta_{1p}}{\eta_{2p} + \eta_{1p}} \tag{56}$$

As discussed in Sec. 12.5, we require that the effective impedances, η_{2s} or η_{2p} be purely imaginary, zero, or infinite if (55) or (56) are to have unity magnitudes. Knowing that

$$\eta_{2s} = \frac{\eta_2}{\cos\theta_2} \tag{57}$$

and

$$\eta_{2p} = \eta_2 \cos\theta_2 \tag{58}$$

the requirement is that $\cos\theta_2$ be zero or imaginary, where, from Eq. (75), Sec. 12.5,

$$\cos\theta_2 = \left[1 - \sin^2\theta_2\right]^{1/2} = \left[1 - \left(\frac{n_1}{n_2}\right)^2 \sin^2\theta_1\right]^{1/2} \tag{59}$$

As a result, we require that

$$\theta_1 \geq \theta_c \tag{60}$$

where the critical angle is defined through

$$\sin\theta_c = \frac{n_2}{n_1} \tag{61}$$

Now, from the geometry of Fig. 14.15, we can construct the field distribution of a TE wave in the guide using plane wave superposition. In the slab region $(-d/2 < x < d/2)$, we have

$$E_{y1s} = E_0 e^{-j\mathbf{k}_{1u}\cdot\mathbf{r}} \pm E_0 e^{-j\mathbf{k}_{1d}\cdot\mathbf{r}} \qquad \left(-\frac{d}{2} < x < \frac{d}{2}\right) \tag{62}$$

where

$$\mathbf{k}_{1u} = \kappa_1 \mathbf{a_x} + \beta \mathbf{a_z} \tag{63}$$

and

$$\mathbf{k}_{1d} = -\kappa_1 \mathbf{a_x} + \beta \mathbf{a_z} \tag{64}$$

The second term in (62) may either add to or subtract from the first term, since either case would result in a symmetric intensity distribution in the x direction. We expect this, since the guide is symmetric. Now, using $\mathbf{r} = x\mathbf{a_x} + z\mathbf{a_z}$, (62) becomes

$$\boxed{E_{y1s} = E_0[e^{j\kappa_1 x} + e^{-j\kappa_1 x}]e^{-j\beta z} = 2E_0\cos(\kappa_1 x)e^{-j\beta z}} \tag{65}$$

for the choice of the plus sign in (62), and

$$\boxed{E_{y1s} = E_0[e^{j\kappa_1 x} - e^{-j\kappa_1 x}]e^{-j\beta z} = 2jE_0\sin(\kappa_1 x)e^{-j\beta z}} \tag{66}$$

if the minus sign is chosen. Since $\kappa_1 = n_1 k_0 \cos\theta_1$, we see that larger values of κ_1 imply smaller values of θ_1 at a given frequency. In addition, larger κ_1 values result in a greater number of spatial oscillations of the electric field over the transverse dimension, as (65) and (66) show. We found similar behavior in the parallel-plate guide. In the slab waveguide, as with the parallel-plate guide, we associate higher-order modes with increasing values of κ_1.[4]

[4] It would be appropriate to add the mode number subscript, m, to κ_1, κ_2, β, and θ_1, since, as was true with the metal guides, we will obtain discrete values of these quantities. To keep notation simple, the m subscript is suppressed, and we will assume it to be understood. Again, the subscripts 1 and 2 in this section indicate respectively the slab and surrounding *regions*, and have nothing to do with mode number.

In the regions above and below the slab, waves propagate according to wavevectors $\mathbf{k}_{2u}$ and $\mathbf{k}_{2d}$ as shown in Fig. 14.15. For example, above the slab ($x > d/2$), the TE electric field will be of the form:

$$E_{y2s} = E_{02}e^{-j\mathbf{k}_2\cdot\mathbf{r}} = E_{02}e^{-j\kappa_2 x}e^{-j\beta z} \tag{67}$$

However, $\kappa_2 = n_2 k_0 \cos\theta_2$, where $\cos\theta_2$, given in (59), is imaginary. We may therefore write

$$\kappa_2 = -j\gamma_2 \tag{68}$$

where γ_2 is real, and is given by (using 59)

$$\gamma_2 = j\kappa_2 = jn_2k_0\cos\theta_2 = jn_2k_0(-j)\left[\left(\frac{n_1}{n_2}\right)^2 \sin^2\theta_1 - 1\right]^{1/2} \tag{69}$$

Eq. (67) now becomes

$$\boxed{E_{y2s} = E_{02}e^{-\gamma_2(x-d/2)}e^{-j\beta z} \qquad \left(x > \frac{d}{2}\right)} \tag{70}$$

where the x variable in (67) has been replaced by $x - (d/2)$ to position the field magnitude, E_{02}, at the boundary. Using similar reasoning, the field in the region below the lower interface, where x is negative, and where $\mathbf{k}_{2d}$ is involved, will be

$$\boxed{E_{y2s} = E_{02}e^{\gamma_2(x+d/2)}e^{-j\beta z} \qquad \left(x < -\frac{d}{2}\right)} \tag{71}$$

The fields expressed in (70) and (71) are those of *surface* waves. Note that they propagate in the z direction only, according to $e^{-j\beta z}$, but simply reduce in amplitude with increasing $|x|$, according to the $e^{-\gamma_2(x-d/2)}$ term in (70) and the $e^{\gamma_2(x+d/2)}$ term in (71). These waves represent a certain fraction of the total power in the mode, and so we see an important fundamental difference between dielectric waveguides and metal waveguides: in the dielectric guide, the fields (and guided power) exist over a cross section that extends beyond the confining boundaries, and in principle exist over an infinite cross section. In practical situations, the exponential decay of the fields above and below the boundaries is typically sufficient to render the fields negligible within a few slab thicknesses from each boundary.

The total electric field distribution is composed of the field in all three regions, and is sketched in Fig. 14.16 for the first few modes. Within the slab, the field is oscillatory and is of a similar form to that of the parallel-plate waveguide. The difference is that the fields in the slab guide do not reach zero at the boundaries, but connect to the evanescent fields above and below the slab. The restriction is that the TE fields on either side of a boundary (being tangent to the interface) must match at the boundary. Specifically,

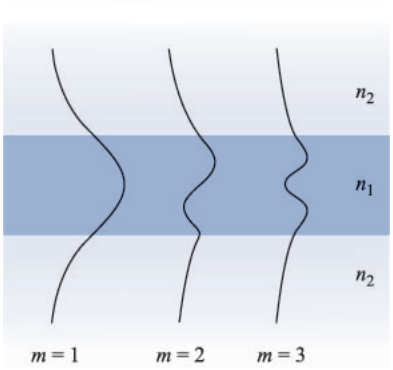


FIGURE 14.16
Electric field amplitude distributions over the transverse plane for the first three modes in a symmetric slab guide.

$$E_{y1s}|_{x=\pm d/2} = E_{y2s}|_{x=\pm d/2} \tag{72}$$

Applying the above condition to (65), (66), (70), and (71) results in the final expressions for the TE electric field in the symmetric slab waveguide, for the cases of even and odd symmetry:

$$E_{se}(\text{even TE}) = \begin{cases} E_{0e}\cos(\kappa_1 x)e^{-j\beta z} & \left(-\frac{d}{2} < x < \frac{d}{2}\right) \\ E_{0e}\cos\left(\kappa_1 \frac{d}{2}\right)e^{-\gamma_2(x-d/2)}e^{-j\beta z} & \left(x > \frac{d}{2}\right) \\ E_{0e}\cos\left(\kappa_1 \frac{d}{2}\right)e^{\gamma_2(x+d/2)}e^{-j\beta z} & \left(x < -\frac{d}{2}\right) \end{cases} \tag{73}$$

$$E_{so}(\text{odd TE}) = \begin{cases} E_{0o}\sin(\kappa_1 x)e^{-j\beta z} & \left(-\frac{d}{2} < x < \frac{d}{2}\right) \\ E_{0o}\sin\left(\kappa_1 \frac{d}{2}\right)e^{-\gamma_2(x-d/2)}e^{-j\beta z} & \left(x > \frac{d}{2}\right) \\ -E_{0o}\sin\left(\kappa_1 \frac{d}{2}\right)e^{\gamma_2(x+d/2)}e^{-j\beta z} & \left(x < -\frac{d}{2}\right) \end{cases} \tag{74}$$

Solution of the wave equation yields (as it must) results identical to these. The reader is referred to Refs. 2 and 3 for the details. The magnetic field for the TE modes will consist of x and z components, as was true for the parallel-plate guide. Finally, the TM mode fields will be nearly the same in form as those of TE modes, but with a simple rotation in polarization of the plane wave components by 90°. Thus, in TM modes, H_y will result, and will have the same form as E_y for TE, as presented in (73) and (74).

Apart from the differences in the field structures, the dielectric slab waveguide operates in a manner that is qualitatively similar to the parallel-plate guide. Again, a finite number of discrete modes will be allowed at a given frequency, and this number increases as frequency increases. Higher order modes are characterized by successively smaller values of θ_1.

An important difference in the slab guide occurs at cutoff for any mode. We know that $\theta = 0$ at cutoff in the metal guides. In the dielectric guide at cutoff, the wave angle, θ_1, is equal to the *critical angle*, θ_c. Then, as the frequency of a given mode is raised, its θ_1 value increases beyond θ_c in order to maintain transverse resonance, while maintaining the same number of field oscillations in the transverse plane.

As wave angle increases, however, the character of the evanescent fields changes significantly. This can be understood by considering the wave angle dependence on evanescent decay coefficient, γ_2, as given by (69). Note in that equation, that as θ_1 increases (as frequency goes up), γ_2 also increases, leading to a more rapid fall-off of the fields with increasing distance above and below the slab. Hence, the mode becomes more tightly confined to the slab as frequency is raised. Also, at a given frequency, higher-order modes, having smaller wave angles, will have lower values of γ_2 as (69) indicates. Consequently, when considering several modes propagating together at a single frequency, the higher order modes will carry a greater percentage of their power in the upper and lower regions surrounding the slab, than modes of lower order.

One can determine the conditions under which modes will propagate by using the transverse resonance condition, as we did with the parallel-plate guide. We perform the transverse round-trip analysis in the slab region in the same manner as was done in Sec. 14.2, and obtain an equation similar to (3):

$$\kappa_1 d + \phi_{TE} + \kappa_1 d + \phi_{TE} = 2m\pi \tag{75}$$

for TE waves and

$$\kappa_1 d + \phi_{TM} + \kappa_1 d + \phi_{TM} = 2m\pi \tag{76}$$

for the TM case. The phase shifts on reflection, ϕ_{TE} and ϕ_{TM} are the phases of the reflection coefficients, Γ_s and Γ_p, given in (55) and (56). These are readily found, but turn out to be functions of θ_1. As we know, κ_1 also depends on θ_1, but in a different way than ϕ_{TE} and ϕ_{TM}. Consequently, (75) and (76) are *transcendental* in θ_1, and cannot be solved in closed form. Instead, numerical or graphical methods must be used (see Refs. 2 or 3). Emerging from this solution process, however, is a fairly simple cutoff condition for any TE or TM mode:

$$\boxed{k_0 d\sqrt{n_1^2 - n_2^2} \geq (m-1)\pi \qquad (m = 1, 2, 3, \ldots)} \tag{77}$$

For mode m to propagate, (77) must hold. The physical interpretation of the mode number m is again the number of half-cycles of the electric field (for TE modes) or magnetic field (for TM modes) that occur over the transverse dimension. The lowest order mode ($m = 1$) is seen to have no cutoff—it will propagate from zero frequency on up. We will thus achieve single mode operation (actually a single pair of TE and TM modes) if we can ensure that the $m = 2$ modes are below cutoff. Using (77), our single mode condition will thus be:

$$\boxed{k_0 d\sqrt{n_1^2 - n_2^2} < \pi} \tag{78}$$

Using $k_0 = 2\pi/\lambda$, the wavelength range over which single mode operation occurs is

$$\lambda > 2d\sqrt{n_1^2 - n_2^2} \tag{79}$$

Example 14.5

A symmetric dielectric slab waveguide is to guide light at wavelength $\lambda = 1.30\ \mu$m. The slab thickness is to be $d = 5.00\ \mu$m, and the refractive index of the surrounding material is $n_2 = 1.450$. Determine the maximum allowable refractive index of the slab material that will allow single TE and TM mode operation.

Solution. Eq. (79) can be rewritten in the form:

$$n_1 < \sqrt{\left(\frac{\lambda}{2d}\right)^2 + n_2^2}$$

So

$$n_1 < \sqrt{\left(\frac{1.30}{2(5.00)}\right)^2 + (1.450)^2} = 1.456$$

Clearly, fabrication tolerances are very exacting when constructing dielectric guides for single mode operation!

Optical fiber waveguides work on the same principle as the basic dielectric guide, except, of course for the round cross section. A *step index* fiber is shown in Fig. 14.5 in which a high index *core* of radius a is surrounded by a lower-index *cladding* of radius b. Light is confined to the core through the mechanism of total reflection, but again some fraction of the power resides in the cladding as well. As we found in the slab guide, the cladding power again moves in toward the core as frequency is raised. Additionally, as is true in the slab waveguide, the fiber supports a mode that has no cutoff. Again, without proof (but see Refs. 2 or 3), the condition for single-mode operation in a step index fiber is found to be similar to that in the slab:

$$\lambda > \frac{2\pi a}{2.405}\sqrt{n_1^2 - n_2^2} \tag{80}$$

Typical dimensions of a single mode fiber include core diameters between 5 and 10 μm, with the cladding diameter usually 125 μm. As we found in the above example, refractive index differences between core and cladding are quite small, and are typically a small fraction of a percent.

✔ **D14.7.** A 0.5 mm thick slab of glass ($n_1 = 1.45$) is surrounded by air ($n_2 = 1$). The slab guides infrared light at wavelength $\lambda = 1.0$ μm. How many TE and TM modes will propagate?

Ans. 2102.

14.6 BASIC ANTENNA PRINCIPLES

In this final section we explore some concepts on radiation of electromagnetic energy from a simple dipole antenna. A complete discussion of antennas and their applications would require several chapters or entire books. Our purpose is to produce a fundamental understanding of how electromagnetic fields radiate from *current distributions*. So for the first time, we shall have the specific field which results from a specific time-varying *source*. In the discussion of waves and fields in bulk media and in waveguides, only the wave motion in the medium was investigated, and the sources of the fields were not considered. The current distribution in a conductor was a similar problem, although we did at least relate the current to an assumed electric field intensity at the conductor surface. This might be considered as a source, but it is not a very practical one for it is infinite in extent.

We now assume a current filament (of infinitesimally-small cross section) as the source, positioned within an infinite lossless medium. The filament is taken as a differential length, but we shall be able to extend the results easily to a filament which is short compared to a wavelength, specifically less than about one-quarter of a wavelength overall. The differential filament is shown at the origin and is oriented along the z axis in Fig. 14.17. The positive sense of the current is taken in the $\mathbf{a}_z$ direction. We assume a uniform current $I_0 \cos \omega t$ in this short length d and do not concern ourselves at present with the apparent discontinuity at each end.

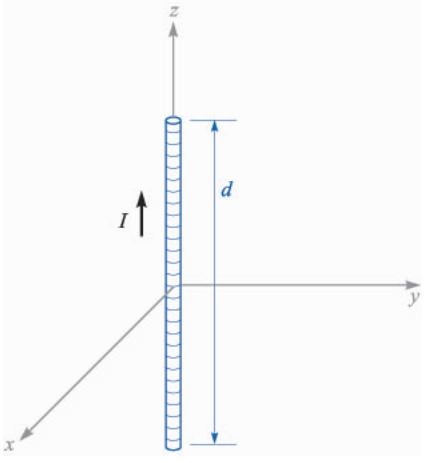


FIGURE 14.17
A differential current filament of length d carries a current $I = I_0 \cos \omega t$.

We shall not attempt at this time to discover the "source of the source," but shall merely assume that the current distribution cannot be changed by any field which it produces.

The first step is the application of the retarded vector magnetic potential expression, as presented in Sec. 10.5,

$$\mathbf{A} = \int \frac{\mu[I]d\mathbf{L}}{4\pi R}$$

where $[I]$ is a function of the retarded time $t - R/v$. When a single frequency is used to drive the antenna, v is the phase velocity at that frequency. Since no integration is required for the very short filament assumed, we have

$$\mathbf{A} = \frac{\mu[I]d}{4\pi R}\mathbf{a}_z$$

Only the z component of $\mathbf{A}$ is present, for the current is only in the $\mathbf{a}_z$ direction. At any point P, distant R from the origin, the vector potential is retarded by R/v and

$$I = I_0 \cos \omega t$$

becomes

$$[I] = I_0 \cos\left[\omega\left(t - \frac{R}{v}\right)\right]$$

$$[I_s] = I_0 e^{-j\omega R/v}$$

Thus

$$A_{zs} = \frac{\mu I_0 d}{4\pi R} e^{-j\omega R/v}$$

Using a mixed coordinate system for the moment, let us replace R by the small r of the spherical coordinate system and then determine which spherical components are represented by A_{zs}. Figure 14.18 helps to determine that

$$A_{rs} = A_{zs} \cos\theta$$

$$A_{\theta s} = -A_{zs} \sin\theta$$

and therefore

$$A_{rs} = \frac{\mu I_0 d}{4\pi r} \cos\theta\, e^{-j\omega r/v}$$

$$A_{\theta s} = -\frac{\mu I_0 d}{4\pi r} \sin\theta\, e^{-j\omega r/v}$$

From these two components of the vector magnetic potential at P we may find $\mathbf{B_s}$ or $\mathbf{H_s}$ from the definition of $\mathbf{A_s}$,

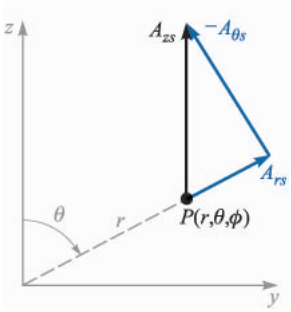


FIGURE 14.18
The resolution of A_{zs} at $P(r, \theta, \phi)$ into the two spherical components A_{rs} and $A_{\theta s}$. The sketch is arbitrarily drawn in the $\phi = 90^\circ$ plane.

$$\mathbf{B_s} = \mu \mathbf{H_s} = \nabla \times \mathbf{A_s}$$

by merely taking the indicated partial derivatives. Thus

$$H_{\phi s} = \frac{1}{\mu r}\frac{\partial}{\partial r}(rA_{\theta s}) - \frac{1}{\mu r}\frac{\partial A_{rs}}{\partial \theta}$$

$$H_{rs} = H_{\theta s} = 0$$

and

$$H_{\phi s} = \frac{I_0 d}{4\pi}\sin\theta\, e^{-j\omega r/v}\left(j\frac{\omega}{vr} + \frac{1}{r^2}\right)$$

The components of the electric field which must be associated with this magnetic field are found from the point form of Ampere's circuital law as it applies to a region in which conduction and convection current are absent,

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t}$$

or in complex notation,

$$\nabla \times \mathbf{H_s} = j\omega\epsilon \mathbf{E_s}$$

Expansion of the curl in spherical coordinates leads to

$$E_{rs} = \frac{1}{j\omega\epsilon}\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(H_{\phi s}\sin\theta)$$

$$E_{\theta s} = \frac{1}{j\omega\epsilon}\left(-\frac{1}{r}\right)\frac{\partial}{\partial r}(rH_{\phi s})$$

or

$$E_{rs} = \frac{I_0 d}{2\pi}\cos\theta\, e^{-j\omega r/v}\left(\frac{1}{\epsilon v r^2} + \frac{1}{j\omega\epsilon r^3}\right)$$

$$E_{\theta s} = \frac{I_0 d}{4\pi}\sin\theta\, e^{-j\omega r/v}\left(\frac{j\omega}{\epsilon v^2 r} + \frac{1}{\epsilon v r^2} + \frac{1}{j\omega\epsilon r^3}\right)$$

In order to simplify the interpretation of the terms enclosed in parentheses above, we make the substitutions $\omega = 2\pi f$, $f\lambda = v$, $v = 1/\sqrt{\mu\epsilon}$, and $\eta = \sqrt{\mu/\epsilon}$, producing

$$H_{\phi s} = \frac{I_0 d}{4\pi} \sin\theta \, e^{-j2\pi r/\lambda} \left(j\frac{2\pi}{\lambda r} + \frac{1}{r^2} \right) \tag{81}$$

$$E_{rs} = \frac{I_0 d\eta}{2\pi} \cos\theta \, e^{-j2\pi r/\lambda} \left(\frac{1}{r^2} + \frac{\lambda}{j2\pi r^3} \right) \tag{82}$$

$$E_{\theta s} = \frac{I_0 d\eta}{4\pi} \sin\theta \, e^{-j2\pi r/\lambda} \left(j\frac{2\pi}{\lambda r} + \frac{1}{r^2} + \frac{\lambda}{j2\pi r^3} \right) \tag{83}$$

These three equations are indicative of the reason that so many problems involving antennas are solved by experimental rather than theoretical methods. They have resulted from three general steps: an integration (atypically trivial) and two differentiations. These steps are sufficient to cause the simple current element and its simple current expression to "blow up" into the complicated field described by (81) to (83). In spite of this complexity, several interesting observations are possible.

We might notice first the $e^{-j2\pi r/\lambda}$ factor appearing with each component. This indicates propagation outward from the origin in the positive r direction with a phase factor $\beta = 2\pi/\lambda$; thus the wavelength is λ and the velocity $v = 1/\sqrt{\mu\epsilon}$. We use the term "wavelength" now in a somewhat broader sense than the original definition, which identified the wavelength of a uniform plane wave with the distance between two points, measured in the direction of propagation, at which the wave has identical instantaneous values. Here there are additional complications caused by the terms enclosed in parentheses, which are complex functions of r. These variations must now be neglected in determining the wavelength. This is equivalent to a determination of the wavelength at a large distance from the origin, and we may demonstrate this by sketching the H_ϕ component as a function of r under the following conditions:

$$I_0 d = 4\pi \qquad \theta = 90^\circ \qquad t = 0 \qquad f = 300 \text{ MHz}$$

$$v = 3 \times 10^8 \text{ m/s (free space)} \qquad \lambda = 1 \text{ m}$$

Therefore

$$H_{\phi s} = \left(j\frac{2\pi}{r} + \frac{1}{r^2}\right)e^{-j2\pi r}$$

and the real part may be determined at $t = 0$,

$$H_\phi = \sqrt{\left(\frac{2\pi}{r}\right)^2 + \frac{1}{r^4}}\cos[(\tan^{-1} 2\pi r) - 2\pi r]$$

Knowing that $\cos(a - b) = \cos a \cos b + \sin a \sin b$ and that $\cos(\tan^{-1} x) = 1/\sqrt{1 + x^2}$, we may simplify this result to

$$H_\phi = \frac{1}{r^2}(\cos 2\pi r + 2\pi r \sin 2\pi r)$$

Values obtained from this last equation are plotted against r in the range $1 \leq r \leq 2$ in Fig. 14.19*a*; the curve is noticeably nonsinusoidal. At $r = 1$, $H_\phi = 1$, while at $r = 2$, one wavelength greater, $H_\phi = 0.25$. Moreover, the curve crosses the axis (with positive slope) at $r = 1 - 0.0258$ and $r = 2 - 0.0127$, again a distance not equal to a wavelength. If a similar sketch is made in the range $101 \leq r \leq 102$, shown in Fig. 14.19*b* on a different amplitude scale, an essentially sinusoidal wave is obtained and the instantaneous values of H_ϕ at $r = 101$ and $r = 102$ are 0.000 099 8 and 0.000 099 6. The maximum amplitudes of the positive and negative portions of the waveform differ by less than 1 percent, and we may say that for all practical purposes the wave in this region is a uniform plane wave having a sinusoidal variation with distance (and time, of course) and a well-defined wavelength. This wave evidently carries energy away from the differential antenna, and we shall calculate this power shortly.

Continuing the investigation of (81) to (83), let us now take a more careful look at the expressions containing terms varying as $1/r^3$, $1/r^2$, and $1/r$. At points very close to the current element the $1/r^3$ term must be dominant. In the numerical example we have used, the relative values of the terms in $1/r^3$, $1/r^2$, and $1/r$

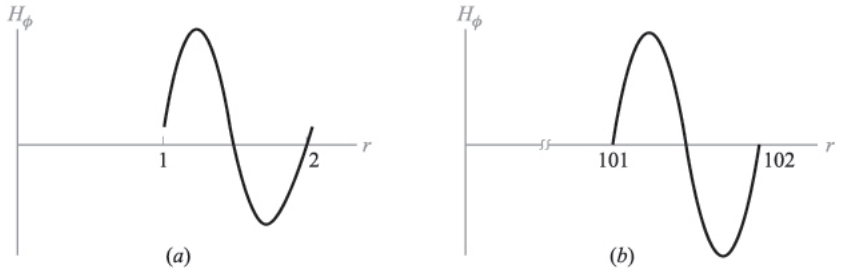


FIGURE 14.19
The instantaneous amplitude of H_ϕ for the special case of a current element having $I_0 d = 4\pi$ and $\lambda = 1$ is plotted at $\theta = 90°$ and $t = 0$ (*a*) in the region $1 \leq r \leq 2$ near the antenna, and (*b*) in the region $101 \leq r \leq 102$ distant from the antenna. The left curve is noticeably nonsinusoidal, since if it were a sinusoid, its end points should reach the r axis precisely at 1 and 2.

in the $E_{\theta s}$ expression are about 250, 16, and 1, respectively, when r is 1 cm. The variation of an electric field as $1/r^3$ should remind us of the *electrostatic* field of the dipole (Chap. 4). This term represents energy stored in a reactive (capacitive) field, and it does not contribute to the radiated power. The inverse-square term in the $H_{\phi s}$ expression is similarly important only in the region very near to the current element and corresponds to the *induction* field of the dc element given by the Biot-Savart law. At distances corresponding to 10 or more wavelengths from the oscillating current element, all terms except the inverse-distance $(1/r)$ term may be neglected and the *distant* or *radiation* fields become

$$E_{rs} = 0$$

$$E_{\theta s} = j\frac{I_0 d\eta}{2\lambda r}\sin\theta\, e^{-j2\pi r/\lambda} \tag{84}$$

$$H_{\phi s} = j\frac{I_0 d}{2\lambda r}\sin\theta\, e^{-j2\pi r/\lambda} \tag{85}$$

or

$$E_{\theta s} = \eta H_{\phi s}$$

The relationship between $E_{\theta s}$ and $H_{\phi s}$ is thus seen to be that between the electric and magnetic fields of the uniform plane wave, thus substantiating the conclusion we reached when investigating the wavelength.

The variation of both radiation fields with the polar angle θ is the same; the fields are maximum in the equatorial plane (xy plane) of the current element and vanish off the ends of the element. The variation with angle may be shown by plotting a *vertical pattern* (assuming a vertical orientation of the current element) in which the relative magnitude of $E_{\theta s}$ is plotted against θ for a constant r. The pattern is usually shown on polar coordinates, as in Fig. 14.20. A *horizontal pattern* may also be plotted for more complicated antenna systems and shows the variation of field intensity with ϕ. The horizontal pattern of the current element is a circle centered at the origin since the field is not a function of the azimuth angle.

In order to obtain a quantitative expression for the power radiated, we need to apply the Poynting vector $\mathcal{P} = \mathbf{E} \times \mathbf{H}$ developed in Sec. 11.3. The instantaneous expressions for the radiation components of the electric and magnetic field intensities are

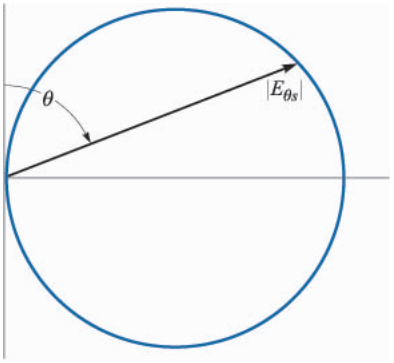


FIGURE 14.20
The polar plot of the vertical pattern of a vertical current element. The crest amplitude of $E_{\theta s}$ is plotted as a function of the polar angle θ at a constant distance r. The locus is a circle.

$$E_\theta = \eta H_\phi$$

$$H_\phi = -\frac{I_0 d}{2\lambda r}\sin\theta \sin\left(\omega t - \frac{2\pi r}{\lambda}\right)$$

and thus

$$\mathcal{P}_r = E_\theta H_\phi = \left(\frac{I_0 d}{2\lambda r}\right)^2 \eta \sin^2\theta \sin^2\left(\omega t - \frac{2\pi r}{\lambda}\right)$$

The total (in space) instantaneous (in time) power crossing the surface of a sphere of radius r_0 is then

$$\begin{aligned} P &= \int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi} \mathcal{P}_r r_0^2 \sin\theta d\theta d\phi \\ &= \left(\frac{I_0 d}{\lambda}\right)^2 \eta \frac{2\pi}{3}\sin^2\left(\omega t - \frac{2\pi r_0}{\lambda}\right) \end{aligned}$$

and the time-average power is given by one-half the maximum amplitude,

$$P_{av} = \left(\frac{I_0 d}{\lambda}\right)^2 \eta \frac{\pi}{3} = 40\pi^2\left(\frac{I_0 d}{\lambda}\right)^2$$

where $\eta = 120\pi\ \Omega$ in free space.

This is the same power as that which would be dissipated in a resistance R_{rad} by the current I_0 in the absence of any radiation, where

$$\boxed{P_{av} = \frac{1}{2} I_0^2 R_{rad}}$$

$$R_{rad} = \frac{2P_{av}}{I_0^2} = 80\pi^2\left(\frac{d}{\lambda}\right)^2 \tag{86}$$

If we assume the differential length is 0.01λ, then R_{rad} is about $0.08\,\Omega$. This small resistance is probably comparable to the *ohmic* resistance of a practical antenna, and thus the efficiency of the antenna might be unsatisfactorily low. Effective matching to the source also becomes very difficult to achieve, for the input reactance of a short antenna is much greater in magnitude than the input resistance R_{rad}. This is the basis for the statement that an effective antenna should be an appreciable fraction of a wavelength long.

The actual current distribution on a thin linear antenna is very nearly sinusoidal, even for antennas that may be several wavelengths long. Note that if the conductors of an open-circuited two-wire transmission line are folded back 90°, the standing-wave distribution on the line is sinusoidal. The current is zero at each end and maximizes one-quarter wavelength from each end, and the current continues to vary in this manner toward the center. The current at the center, therefore, will be very small for an antenna whose length is an integral number of wavelengths, but it will be equal to the maximum found at any point on the antenna if the antenna length is $\lambda/2$, $3\lambda/2$, $5\lambda/2$, and so forth.

On a short antenna, then, we see only the first portion of the sine wave; the amplitude of the current is zero at each end and increases approximately in a linear manner to a maximum value of I_0 at the center. This is suggested by the sketch of Fig. 14.21. Note that this antenna has identical currents in the two halves and it may be fed conveniently by a two-wire line, where the currents in the two conductors are equal in amplitude but opposite in direction. The gap at

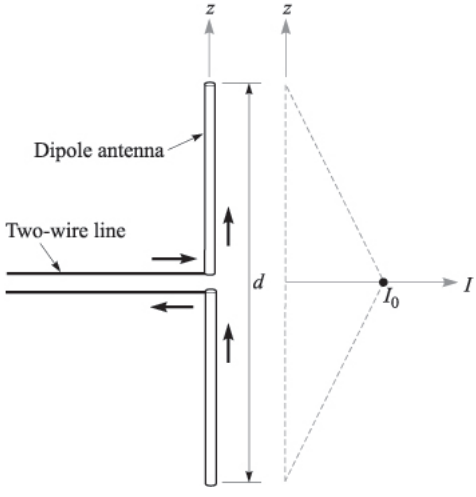


FIGURE 14.21
A short antenna ($d < \lambda/4$) has a linear current distribution and may be driven by a two-wire line.

the feed point is small and has negligible effects. A symmetrical antenna of this type is called a *dipole*. The linear current variation with distance is a reasonable assumption for antennas having an overall length less than about one-quarter wavelength.

It is possible to extend the analysis of the differential current element to the short dipole if we assume that the length is short enough that retardation effects may be neglected. That is, we consider that signals arriving at any field point P from the two ends of the antenna are in phase. The average current along the antenna is $I_0/2$, where I_0 is the input current at the center terminals. Thus, the electric and magnetic field intensities will be one-half the values given in (84) and (85), and there are no changes in the vertical and horizontal patterns. The power will be one-quarter of its previous value, and thus the radiation resistance will also be one-quarter of the value given by (86).

If we try to improve our results by assuming a sinusoidal variation of current amplitude with distance along the antenna, and if the effects of retardation are included in the analysis, then the integrations required to find $\mathbf{A}$ and the power radiated become enormously more difficult. Since we want to hurry along toward the last page, let us merely note that for the world's most popular antenna, the half-wave dipole ($d = \lambda/2$), the following results are eventually obtained:

$$E_{\theta s} = \frac{I_0 \eta}{2\pi r} \frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta} \tag{87}$$

$$H_{\phi s} = \frac{E_{\theta s}}{\eta} \tag{88}$$

$$R_{rad} = 30\left[\frac{(2\pi)^2}{2 \cdot 2!} - \frac{(2\pi)^4}{4 \cdot 4!} + \frac{(2\pi)^6}{6 \cdot 6!} - \frac{(2\pi)^8}{8 \cdot 8!} + \ldots\right] = 73.1\ \Omega \tag{89}$$

Let us compare this accurate value with results obtained by more approximate means. Suppose we first attempt to find the radiation resistance by assuming a uniform current distribution and neglecting the effects of retardation. The result is obtained from (86) with $d/\lambda = 1/2$; $R_{rad} = 20\pi^2 = 197.4\,\Omega$. This is much greater than $73.1\,\Omega$, but we have also assumed a much greater current on the antenna than is actually present.

The result may be improved by considering a linear current distribution while still ignoring retardation. The average current is half the maximum value, the power is one-quarter, and the radiation resistance drops to $5\pi^2$ or $49.3\,\Omega$. Now the result is too small, primarily because the average value of a triangular wave is less than the average value of a sine wave.

Finally, if we assume a sinusoidal current distribution, we have an average value of $2/\pi$ times the maximum, and the radiation resistance comes to $(2/\pi)^2(20\pi^2)$, or $80\,\Omega$. This is reasonably close to the true value, and the discrepancy lies in neglecting retardation. In a linear antenna, the effect of retardation

is always one of cancellation, and therefore its consideration must always lead to smaller values of radiation resistance. This decrease is of relatively small magnitude here (from 80 to 73.1 Ω) because the current elements tending to cancel each other are those at the ends of the dipole, and these are of small amplitude; moreover, the cancellation is greatest in a direction along the antenna axis where all radiation fields are zero for a linear antenna.

Familiar antennas that fall into the dipole classification are the elements used in the common TV and FM receiving antennas.

As a final example of a practical antenna, let us collect a few facts about the *monopole* antenna. This is one-half a dipole plus a perfectly conducting plane, as shown in Fig. 14.22*a*. The image principle discussed in Sec. 5.5 provides the image shown in Fig. 14.22*b* and assures us that the fields above the plane are the same for the monopole and the dipole. Hence, the expressions of (84) and (85) are equally valid for the monopole. The Poynting vector is therefore also the same above the plane, but the integration to find the total power radiated is extended through but one-half the volume. Thus the power radiated and the radiation resistance for the monopole are half the corresponding values for the dipole. As an example, a monopole with an assumed uniform current distribution has $R_{rad} = 40\pi^2(d/\lambda)^2$; a triangular current leads to $R_{rad} = 10\pi^2(d/\lambda)^2$; and the sinusoidal current distribution of a $\lambda/4$ monopole leads to $R_{rad} = 36.5\,\Omega$.

Monopole antennas may be driven by a coaxial cable below the plane, having its center conductor connected to the antenna through a small hole, and having its outer conductor connected to the plane. If the region below the plane is inaccessible or inconvenient, the coax may be laid on top of the plane and its outer conductor connected to it. Examples of this type of antenna include AM broadcasting towers and CB antennas.

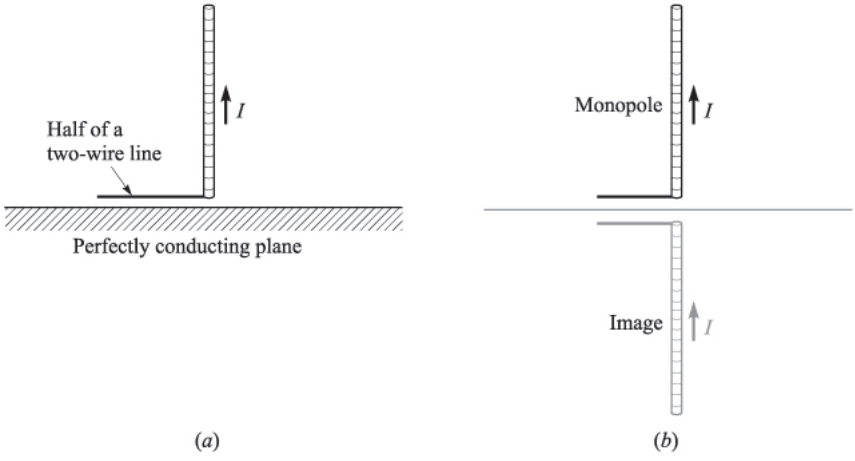


FIGURE 14.22
(*a*) An ideal monopole is always associated with a perfectly conducting plane. (*b*) The monopole plus its image form a dipole.

✔ **D14.8.** Calculate values for the curve shown in Fig. 14.19*a* at $r = 1$, 1.2, 1.4, 1.6, 1.8, and 2.

Ans. 1.00; 5.19; 2.23; −2.62; −3.22; 0.25.

✔ **D14.9.** A short antenna with a uniform current distribution in air has $I_0 d = 3 \times 10^{-4}$ A · m and $\lambda = 10$ cm. Find $|E_{\theta s}|$ at $\theta = 90°$, $\phi = 0°$, and $r =$: (a) 2 cm; (b) 20 cm; (c) 200 cm.

Ans. 24.8 V/m; 2.82 V/m; 0.283 V/m.

✔ **D14.10.** The monopole antenna of Fig. 14.22*a* has a length $d/2 = 0.080$ m and may be assumed to carry a triangular current distribution for which the feed current I_0 is 16.0 A at a frequency of 375 MHz in free space. At point P ($r = 400$ m, $\theta = 60°$, $\phi = 45°$) find: (a) $H_{\phi s}$; (b) $E_{\theta s}$; (c) the amplitude of $\mathcal{P}_r$.

Ans. j1.73 mA/m; j0.653 V/m; 1.13 mW/m^2

SUGGESTED REFERENCES

1. Ramo, S., J. R. Whinnery, and T. Van Duzer: "Fields and Waves in Communication Electronics," 3d ed., John Wiley and Sons, New York, 1990. In-depth treatment of parallel-plate and rectangular waveguides is presented in Chapter 8.
2. Marcuse, D.: "Theory of Dielectric Optical Waveguides," 2d ed., Academic Press, New York, 1990. This book provides a very general and complete discussion of dielectric slab waveguides, plus other types.
3. Buck, J. A.: "Fundamentals of Optical Fibers," Wiley-Interscience, New York, 1995. Symmetric slab dielectric guides are emphasized in this book by one of the authors.
4. Weeks, W. L.: "Antenna Engineering," McGraw-Hill Book Company, New York, 1968. This excellent text probably contains more information about antennas than you want to know.
5. Jordan, E. C. and K. G. Balmain: "Electromagnetic Waves and Radiating Systems," 2d ed., Prentice-Hall, Inc., Englewood Cliffs, N.J., 1968. This classic text provides an excellent treatment of waveguides and antennas.

PROBLEMS

14.1 A parallel-plate waveguide is known to have a cutoff wavelength for the $m = 1$ TE and TM modes of $\lambda_{c1} = 0.4$ cm. The guide is operated at wavelength $\lambda = 1$ mm. How many modes propagate?

14.2 A parallel-plate guide is to be constructed for operation in the TEM mode only over the frequency range $0 < f < 3$ GHz. The dielectric between plates is to be teflon ($\epsilon'_R = 2.1$). Determine the maximum allowable plate separation, d.

14.3 A lossless parallel-plate waveguide is known to propagate the $m = 2$ TE and TM modes at frequencies as low as 10 GHz. If the plate separation is 1 cm, determine the dielectric constant of the medium between plates.

14.4 A $d = 1$ cm parallel-plate guide is made with glass ($n = 1.45$) between plates. If the operating frequency is 32 GHz, which modes will propagate?

14.5 For the guide of Problem 14.4, and at the 32 GHz frequency, determine the difference between the group delays of the highest-order mode (TE or TM) and the TEM mode. Assume a propagation distance of 10 cm.

14.6 The cutoff frequency of the $m = 1$ TE and TM modes in a parallel-plate guide is known to be $f_{c1} = 7.5$ GHz. The guide is used at wavelength $\lambda = 1.5$ cm. Find the group velocity of the $m = 2$ TE and TM modes.

14.7 A parallel-plate guide is partially filled with two lossless dielectrics (Fig. 14.23) where $\epsilon'_{R1} = 4.0$, $\epsilon'_{R2} = 2.1$, and $d = 1$ cm. At a certain frequency, it is found that the TM_1 mode propagates through the guide without suffering any reflective loss at the dielectric interface. (a) Find this frequency. (b) Is the guide operating at a single TM mode at the frequency found in part (a)? Hint: Remember Brewster's angle?

14.8 In the guide of Problem 14.7, it is found that $m = 1$ modes propagating from left to right totally reflect at the interface, so that no power is transmitted into the region of dielectric constant ϵ'_{R2}. (a) Determine the range of frequencies over which this will occur. (b) Does your part (a) answer in any way relate to the cutoff frequency for $m = 1$ modes in either region? Hint: Remember the critical angle?

14.9 A rectangular waveguide has dimensions $a = 6$ cm and $b = 4$ cm. (a) Over what range of frequencies will the guide operate single mode? (b) Over what frequency range will the guide support *both* TE_{10} and TE_{01} modes and no others?

14.10 Two rectangular waveguides are joined end-to-end. The guides have identical dimensions, where $a = 2b$. One guide is air-filled; the other is filled with a lossless dielectric characterized by ϵ'_R. (a) Determine the maximum allowable value of ϵ'_R such that single-mode operation can be simultaneously ensured in *both* guides at some frequency. (b) Write an expression for the frequency range over which single-mode operation will occur in both guides; your answer should be in terms of ϵ'_R, guide dimensions as needed, and other known constants.

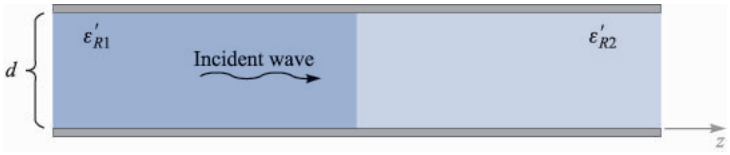


FIGURE 14.23
See Problem 14.7.

14.11 An air-filled rectangular waveguide is to be constructed for single-mode operation at 15 GHz. Specify the guide dimensions, a and b, such that the design frequency is 10% higher than the cutoff frequency for the TE_{10} mode, while being 10% lower than the cutoff frequency for the next higher-order mode.

14.12 Using the relation $\mathcal{P}_{av} = \frac{1}{2}\text{Re}\{\mathbf{E}_s \times \mathbf{H}_s^*\}$, and Eqs. (44) through (46), show that the average power density in the TE_{10} mode in a rectangular waveguide is given by

$$\mathcal{P}_{av} = \frac{\beta_{10}}{2\omega\mu} E_0^2 \sin(\kappa_{10} x)\mathbf{a}_z \qquad \text{W/m}^2$$

14.13 Integrate the result of Problem 14.12 over the guide cross section, $0 < x < a$, $0 < y < b$, to show that the power in Watts transmitted down the guide is given as

$$P = \frac{\beta_{10} ab}{4\omega\mu} E_0^2 = \frac{ab}{4\eta} E_0^2 \sin\theta_{10} \qquad \text{W}$$

where $\eta = \sqrt{\mu\epsilon}$ and θ_{10} is the wave angle associated with the TE_{10} mode. Interpret.

14.14 Show that the group dispersion parameter, $d^2\beta/d\omega^2$, for a given mode in a parallel-plate or rectangular waveguide is given by

$$\frac{d^2\beta}{d\omega^2} = -\frac{n}{\omega c}\left(\frac{\omega_c}{\omega}\right)^2 \left[1 - \left(\frac{\omega_c}{\omega}\right)^2\right]^{-3/2}$$

where ω_c is the radian cutoff frequency for the mode in question (note that the first derivative form was already found, resulting in Eq. (23)).

14.15 Consider a transform-limited pulse of center frequency $f = 10$ GHz, and of full-width $2T = 1.0$ ns. The pulse propagates in a lossless single mode rectangular guide which is air-filled and in which the 10 GHz operating frequency is 1.1 times the cutoff frequency of the TE_{10} mode. Using the result of Problem 14.14, determine the length of guide over which the pulse broadens to twice its initial width. What simple step can be taken to reduce the amount of pulse broadening in this guide, while maintaining the same initial pulse width? Additional background for this problem is found in Sec. 12.6.

14.16 A symmetric dielectric slab waveguide has a slab thickness $d = 10\,\mu$m, with $n_1 = 1.48$ and $n_2 = 1.45$. If the operating wavelength is $\lambda = 1.3\,\mu$m, what modes will propagate?

14.17 A symmetric slab waveguide is known to support only a single pair of TE and TM modes at wavelength $\lambda = 1.55\,\mu$m. If the slab thickness is 5 μm, what is the maximum value of n_1 if $n_2 = 3.3$?

14.18 $n_1 = 1.45$, $n_2 = 1.50$, and $d = 10\,\mu$m in a symmetric slab waveguide. (a) What is the phase velocity of the $m = 1$ TE or TM mode at cutoff? (b) What is the phase velocity of the $m = 2$ TE or TM modes at cutoff?

14.19 An *asymmetric* slab waveguide is shown in Fig. 14.24. In this case, the regions above and below the slab have unequal refractive indices, where $n_1 < n_3 < n_2$. (a) Write, in terms of the appropriate indices, an expression for the minimum possible wave angle, θ_1, that a guided mode may have. (b) Write an expression for the maximum phase velocity a guided mode may have in this structure, using given or known parameters.

14.20 A step index optical fiber is known to be single mode at wavelengths $\lambda > 1.2\,\mu$m. Another fiber is to be fabricated from the same materials, but it is to be single mode at wavelengths $\lambda > 0.63\,\mu$m. By what percentage must the core radius of the new fiber differ from the old one, and should it be larger or smaller?

14.21 A short dipole carrying current $I_0 \cos \omega t$ in the $\mathbf{a}_z$ direction is located at the origin in free space. (a) If $\beta = 1$ rad/m, $r = 2$ m, $\theta = 45°$, $\phi = 0$, and $t = 0$, give a unit vector in rectangular components that shows the instantaneous direction of $\mathbf{E}$. (b) What fraction of the total average power is radiated in the belt, $80° < \theta < 100°$?

14.22 Prepare a curve, r vs. θ in polar coordinates, showing the locus in the $\phi = 0$ plane where: (a) the radiation field $|E_{\theta s}|$ is one-half of its value at $r = 10^4$ m, $\theta = \pi/2$; (b) the average radiated power density $\mathcal{P}_{r,av}$ is one-half of its value at $r = 10^4$ m, $\theta = \pi/2$.

14.23 Two short antennas at the origin in free space carry identical currents of $5 \cos \omega t$ A, one in the $\mathbf{a}_z$ direction, one in the $\mathbf{a}_y$ direction. Let $\lambda = 2\pi$ m and $d = 0.1$ m. Find $\mathbf{E}_s$ at the distant point: (a) $(x = 0, y = 1000, z = 0)$; (b) (0,0,1000); (c) (1000,0,0); (d) Find $\mathbf{E}$ at (1000,0,0) at $t = 0$. (e) Find $|\mathbf{E}|$ at (1000,0,0) at $t = 0$.

14.24 A short current element has $d = 0.03\lambda$. Calculate the radiation resistance for each of the following current distributions: (a) uniform, I_0; (b) linear, $I(z) = I_0(0.5d - |z|)/0.5d$; (c) step, I_0 for $0 < |z| < 0.25d$ and $0.5I_0$ for $0.25d < |z| < 0.5d$.

14.25 A dipole antenna in free space has a linear current distribution. If the length d is 0.02λ, what value of I_0 is required to: (a) provide a radiation-field amplitude of 100 mV/m at a distance of 1 mi, at $\theta = 90°$; (b) radiate a total power of 1 W?

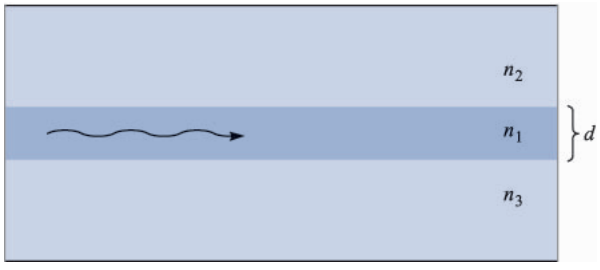


FIGURE 14.24
See Problem 14.19.

14.26 A monopole antenna in free space, extending vertically over a perfectly conducting plane, has a linear current distribution. If the length of the antenna is 0.01λ, what value of I_0 is required to: (a) provide a radiation-field amplitude of 100 mV/m at a distance of 1 mi, at $\theta = 90^\circ$; (b) radiate a total power of 1 W?

14.27 The radiation field of a certain short vertical current element is $E_{\theta s} = (20/r)\sin\theta\, e^{-j10\pi r}$ V/m if it is located at the origin in free space. (a) Find $E_{\theta s}$ at $P(r = 100,\ \theta = 90^\circ,\ \phi = 30^\circ)$. (b) Find $E_{\theta s}$ at $P(100, 90^\circ, 30^\circ)$ if the vertical element is located at $A(0.1, 90^\circ, 90^\circ)$. (c) Find $E_{\theta s}$ at $P(100, 90^\circ, 30^\circ)$ if identical vertical elements are located at $A(0.1, 90^\circ, 90^\circ)$ and $B(0.1, 90^\circ, 270^\circ)$.