Rectangular and Triangular Waveguides

To illustrate the general waveguide theory thus far developed, we shall determine the mode functions and associated eigenvalues for those few guide shapes that permit exact analytical treatment. In this chapter we will discuss guides constructed from plane surfaces. Circular boundaries will be the subject of Chapter 6.

5.1 Rectangular Waveguide

The cross section of a rectangular guide with dimensions $a$ and $b$ is shown in Fig. 5.1, together with a convenient coordinate system. The positive sense of the $z$ axis is assumed to be out of the plane of the page. The wave equations (4.22a) and (4.23a) can be separated in rectangular coordinates; that is, particular solutions for both E-mode and H-mode functions can be found in the form $X(x)Y(y)$, where the two functions thus introduced satisfy one-dimensional wave equations.

Fig. 5.1. Cross section of a rectangular waveguide, with sides $a$ and $b$, and coordinates labelled $(x, y)$
\[
\left( \frac{d^2}{dx^2} + \gamma_x^2 \right) X(x) = 0 , \\
\left( \frac{d^2}{dy^2} + \gamma_y^2 \right) Y(y) = 0 ,
\]

(5.1a)

(5.1b)

where

\[
\gamma^2 = \gamma_x^2 + \gamma_y^2 .
\]

(5.1c)

The E-mode functions are characterized by the boundary conditions

\[
X(0) = X(a) = 0 , \quad Y(0) = Y(b) = 0 ,
\]

(5.2a)

while the H-mode functions are determined by

\[
\frac{d}{dx} X(0) = \frac{d}{dx} X(a) = 0 , \quad \frac{d}{dy} Y(0) = \frac{d}{dy} Y(b) = 0 .
\]

(5.2b)

The scalar mode functions that satisfy these requirements are \((m, n = 0, 1, 2, \ldots)\)

E mode: \(\varphi(x, y) = \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b} ,\)

H mode: \(\psi(x, y) = \cos \frac{m\pi x}{a} \cos \frac{n\pi y}{b} ,\)

(5.3a)

(5.3b)

and for both types of modes:

\[
\gamma^2 = \left( \frac{m\pi}{a} \right)^2 + \left( \frac{n\pi}{b} \right)^2 .
\]

(5.4)

The latter result can also be written

\[
\frac{1}{\lambda^2_c} = \left( \frac{m}{2a} \right)^2 + \left( \frac{n}{2b} \right)^2 ,
\]

(5.5)

or

\[
\lambda_c = \frac{2ab}{\sqrt{(mb)^2 + (na)^2}} .
\]

(5.6)

Observe that neither of the integers \(m\) or \(n\) can be zero for an E mode, or the scalar function vanishes. Although no H-mode corresponding to \(m = n = 0\) exists, since a constant scalar function generates no electromagnetic field, one of the integers can be zero. Thus, there exists a doubly infinite set of E modes characterized by the integers \(m, n = 1, 2, 3, \ldots\); the general member of the set is designated as the \(E_{mn}\) mode, the two integers replacing the index \(a\) of the general theory. Similarly, a doubly infinite set of H modes exists corresponding to the various combinations of integers \(m, n = 0, 1, 2, \ldots\), with \(m = n = 0\) excluded. A particular H mode characterized by \(m\) and \(n\) is called the \(H_{mn}\) mode. The double subscript notation is extended to the various quantities describing a mode; thus, the critical wavenumber of an \(E_{mn}\) or \(H_{mn}\) mode is written \(\gamma_{mn}\).
The scalar function associated with an \( E_{mn} \) mode, normalized in accordance with (4.55), is
\[
\phi(x, y) = \frac{2}{\gamma_{mn}\sqrt{ab}} \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b} = \frac{2}{\pi} \frac{1}{\sqrt{m^2 \frac{b}{a} + n^2 \frac{a}{b}}} \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b}, \tag{5.7}
\]
and the various field components for the \( E_{mn} \) modes are from (4.32a)–(4.32g)
\[
\begin{align*}
E_x &= -\frac{2}{a} \frac{m}{\sqrt{m^2 \frac{b}{a} + n^2 \frac{a}{b}}} \cos \frac{m\pi x}{a} \sin \frac{n\pi y}{b} V(z), \tag{5.8a} \\
E_y &= -\frac{2}{b} \frac{n}{\sqrt{m^2 \frac{b}{a} + n^2 \frac{a}{b}}} \sin \frac{m\pi x}{a} \cos \frac{n\pi y}{b} V(z), \tag{5.8b} \\
H_x &= \frac{2}{b} \frac{n}{\sqrt{m^2 \frac{b}{a} + n^2 \frac{a}{b}}} \sin \frac{m\pi x}{a} \cos \frac{n\pi y}{b} I(z), \tag{5.8c} \\
H_y &= -\frac{2}{a} \frac{m}{\sqrt{m^2 \frac{b}{a} + n^2 \frac{a}{b}}} \cos \frac{m\pi x}{a} \sin \frac{n\pi y}{b} I(z), \tag{5.8d} \\
E_z &= \frac{i\zeta}{k} \frac{2\pi}{ab} \frac{m^2 \frac{b}{a} + n^2 \frac{a}{b}}{\sqrt{m^2 \frac{b}{a} + n^2 \frac{a}{b}}} \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b} I(z), \tag{5.8e} \\
H_z &= 0. \tag{5.8f}
\end{align*}
\]
Of this set of waveguide fields, the \( E_{11} \) mode has the smallest cutoff wavenumber,
\[
\gamma_{11} = \pi \sqrt{\frac{1}{a^2} + \frac{1}{b^2}} = \pi \frac{\sqrt{a^2 + b^2}}{ab}, \tag{5.9}
\]
or, equivalently, the largest cutoff wavelength,
\[
(\lambda_c)_{11} = 2 \frac{ab}{\sqrt{a^2 + b^2}}. \tag{5.10}
\]
This wavelength may be designated as the absolute E-mode cutoff wavelength, in the sense that if the intensive wavelength in the guide exceeds \( (\lambda_c)_{11} \), no E modes can be propagated. The \( E_{11} \) mode is called the dominant E mode, for in the frequency range between the absolute E-mode cutoff frequency and the next smallest E-mode cutoff frequency (that of the \( E_{21} \) mode if \( a \) is the larger dimension), E-mode wave propagation in the guide is restricted to the \( E_{11} \) mode.

It is convenient to consider separately the \( H_{mn} \) modes for which neither integer is zero, and the set of modes for which one integer vanishes, \( H_{m0} \) and \( H_{0n} \). The scalar function associated with a member of the former set is
\[ \psi_{mn}(x, y) = \frac{2}{\pi} \frac{1}{\sqrt{m^2 \frac{b}{a} + n^2 \frac{a}{b}}} \cos \frac{m \pi x}{a} \cos \frac{n \pi y}{b}, \quad m, n \neq 0, \quad (5.11) \]

normalized according to (4.59); the field components are, according to (4.33a)–(4.33g)

\[ E_x = \frac{2}{b} \frac{n}{\sqrt{m^2 \frac{b}{a} + n^2 \frac{a}{b}}} \cos \frac{m \pi x}{a} \sin \frac{n \pi y}{b} V(z), \quad (5.12a) \]
\[ E_y = -\frac{2}{a} \frac{m}{\sqrt{m^2 \frac{b}{a} + n^2 \frac{a}{b}}} \sin \frac{m \pi x}{a} \cos \frac{n \pi y}{b} V(z), \quad (5.12b) \]
\[ H_x = \frac{2}{a} \frac{m}{\sqrt{m^2 \frac{b}{a} + n^2 \frac{a}{b}}} \sin \frac{m \pi x}{a} \cos \frac{n \pi y}{b} I(z), \quad (5.12c) \]
\[ H_y = \frac{2}{b} \frac{n}{\sqrt{m^2 \frac{b}{a} + n^2 \frac{a}{b}}} \cos \frac{m \pi x}{a} \sin \frac{n \pi y}{b} I(z), \quad (5.12d) \]
\[ H_z = \frac{i \eta}{k} \frac{2}{ab} \frac{m^2 \frac{b}{a} + n^2 \frac{a}{b}}{ab} \cos \frac{m \pi x}{a} \cos \frac{n \pi y}{b} V(z), \quad (5.12e) \]
\[ E_z = 0. \quad (5.12f) \]

If \( n = 0 \), the appropriately normalized scalar function if

\[ \psi_{m0}(x, y) = \frac{1}{m \pi} \sqrt{\frac{2a}{b}} \cos \frac{m \pi x}{a}, \quad (5.13) \]

which differs by a factor of \( 1/\sqrt{2} \) from the result obtained on placing \( n = 0 \) in (5.11). The field components of the \( H_{m0} \) mode are

\[ E_x = 0, \quad (5.14a) \]
\[ E_y = -\frac{2}{ab} \sin \frac{m \pi x}{a} V(z), \quad (5.14b) \]
\[ H_x = \frac{2}{ab} \sin \frac{m \pi x}{a} I(z), \quad (5.14c) \]
\[ H_y = 0, \quad (5.14d) \]
\[ E_z = 0, \quad (5.14e) \]
\[ H_z = \frac{i \eta}{ka} \frac{m \pi}{ab} \cos \frac{m \pi x}{a} V(z), \quad (5.14f) \]

of which the most notable feature is the absence of all electric field components save \( E_y \). We also observe that \( H_y = 0 \), whence the \( H_{m0} \) modes behave like E modes with respect to the \( y \) axis. The field structure of the \( H_{0n} \) modes is analogous, with the \( x \) axis as the preferred direction. The smallest H-mode cutoff wavenumber is that of the \( H_{10} \) mode if \( a \) is the larger dimension:
5.1 Rectangular Waveguide

\[ \gamma_{10} = \frac{\pi}{a}. \]  

Thus, the cutoff wavelength is simply

\[ (\lambda_c)_{10} = 2a, \]

and is independent of the dimension \( b \), the height of the guide. We may designate \((\lambda_c)_{10}\) both as the absolute H-mode cutoff wavelength, and as the absolute cutoff wavelength of the guide, for the \( H_{10} \) has the smallest critical frequency of all the waveguide modes. Thus the \( H_{10} \) mode is called the dominant H mode, and the dominant mode of the guide, for in the frequency range between the absolute cutoff frequency of the guide and the next smallest cutoff frequency (which mode this represents depends upon the ratio \( b/a \)), only \( H_{10} \) wave propagation can occur. It is evident from the result (5.6) that a hollow waveguide is intrinsically suited to microwave frequencies, if metal tubes of convenient dimensions are to be employed. Furthermore, the frequency range over which a rectangular guide can be operated as a simple transmission line (only dominant mode propagation) is necessarily less than an octave, for the restrictions thereby imposed on the wavelength are \( 2a > \lambda > a \) (expressing the gap between the \( m = 1, n = 0 \) and the \( m = 2, n = 0 \) modes) and \( \lambda > 2b \) (which marks the beginning of the \( m = 0, n = 1 \) mode). (If \( b > a/\sqrt{3} \) the \( m = 1, n = 1 \) mode sets in before \( \lambda \) gets as small as \( a \).)

The field components of the \( H_{10} \) mode contained in (5.14a)–(5.14f) involve voltages and currents defined with respect to the field impedance choice of characteristic impedance. If the definition discussed in (4.85) is adopted, \( N = 2b/a \) (because the + subscript on the integral in (4.85) means that \( x \) ranges only from 0 to \( a/2 \)), and the nonvanishing \( H_{10} \) mode field components read

\[ E_y = -\frac{1}{b} \sin \frac{\pi x}{a} V(z), \]  

\[ H_x = \frac{2}{a} \sin \frac{\pi x}{a} I(z), \]  

\[ H_z = i \frac{\pi}{\kappa a} \cos \frac{\pi x}{a} Y V(z), \]

where the characteristic admittance is

\[ Y = \frac{\kappa a}{\eta K 2b}. \]

Thus the voltage at a given cross section is defined as the line integral of the electric field between the top and bottom faces of the guide taken along the line of maximum field intensity, \( x = a/2 \). Some important properties of the dominant mode emerge from a study of the surface current flowing in the various guide walls. On the side walls, \( x = 0, a \), the only tangential magnetic field component is \( H_z \) and therefore the surface current flows entirely in the \( y \) direction. The surface current density on the \( x = 0 \) face is \( K_y = -H_z \), while
that on the $x = a$ face is $K_y = H_z$. However, since $H_z$ is of opposite sign on the two surfaces, the current flows in the same direction on both side surfaces, with the density

$$x = 0, a : \quad K_y = -\frac{i\pi 2}{\kappa a} Y V(z) . \quad (5.19)$$

On the top and bottom surfaces, $y = b, 0$, there are two tangential field components, $H_x$ and $H_z$, and correspondingly two components of surface current, $K_z = \pm H_x$ and $K_x = \mp H_z$. The upper and lower signs refer to the top and bottom surface, respectively. On the top surface ($y = b$), the lateral and longitudinal components of surface current are

$$K_x = -\frac{i\pi}{\kappa a} \frac{2}{a} \cos \frac{\pi x}{a} V(z) , \quad (5.20a)$$

$$K_z = \frac{2}{a} \sin \frac{\pi x}{a} I(z) . \quad (5.20b)$$

Equal and opposite currents flow on the bottom surface $y = 0$. Since $\cos \frac{\pi x}{a}$ has reversed signs on opposite sides of the center line $x = a/2$, and vanishes at the latter point, it is clear that the transverse current flowing up the side walls continues to move toward the center of the top surface, but with diminishing magnitude, and vanishes at $x = a/2$. Thus the lines of current flow must turn as the center is approached, the flow becoming entirely longitudinal at $x = a/2$. In agreement with this, we observe that $K_z$ vanishes at the boundaries of the top surface $x = 0, a$, and is a maximum at the center line. As a function of position along the guide, the direction of current flow reverses every half-guide wavelength in consequence of the corresponding behavior of the current and voltage. The essential character of the current distribution can also be seen by noting that the surface charge density $\tau = -\mathbf{n} \cdot \mathbf{D} = -\varepsilon \mathbf{n} \cdot \mathbf{E}$ is confined to the top and bottom surfaces where the charge densities are $\tau = \mp \varepsilon E_y$, respectively. Thus, at $y = b$

$$\tau = \frac{\varepsilon}{b} \sin \frac{\pi x}{a} V(z) . \quad (5.21)$$

Hence, current flows around the circumference of the guide between the charges of opposite sign residing on the top and bottom surfaces, and current flows longitudinally on each surface between the oppositely charged regions separated by half the guide wavelength. Lines of current flow illustrating these general remarks are depicted in Fig. ???. The most immediate practical conclusion to be drawn from this analysis follows from the fact that no current crosses the center line of the top and bottom surfaces. Thus, if the metal were cut along the center line no disturbance of the current flow or of the field in the guide would result. In practice, a slot of appreciable dimensions can be cut in the guide wall without appreciable effect, which permits the insertion of a probe to determine the state of the field, without thereby markedly altering the field to be measured. Further discussion of such a slotted guide will be given in Chapter ???.
The field of a propagating mode in a rectangular waveguide can be regarded as a superposition of elementary plane waves arising from successive reflection at the various inner guide surfaces. This is illustrated most simply for a progressive $H_{10}$ wave, where the single component of the electric field has the form

$$E_y = -\frac{V}{b} \sin \frac{\pi x}{a} e^{i\kappa z} = \frac{i}{2} \frac{V}{b} \left[ e^{i(\kappa z + \pi x/a)} - e^{i(\kappa z - \pi x/a)} \right],$$

which is simply a superposition of the two plane waves $\exp[ik(\cos \theta z \pm \sin \theta x)]$ that travel in the $x$-$z$ plane at an angle $\theta$ with respect to the $z$ axis. Here

$$\cos \theta = \frac{\kappa}{k}, \quad \sin \theta = \frac{1}{k} \frac{\pi}{a},$$

which defines a real angle since $\kappa < k$, and in fact

$$k^2 - \kappa^2 = \gamma_{10}^2 = \left(\frac{\pi}{a}\right)^2.$$ 

Each plane wave results from the other on reflection at the two surfaces $x = 0$ and $a$, and each component is an elementary free-space wave with its electric vector directed along the $y$ axis and its magnetic vector contained in the plane of propagation. As an obvious generalization to be drawn from this simple situation, we may regard the various propagating modes in a waveguide as the result of the coherent interference of the secondary plane waves produced by the successive reflection of an elementary wave at the guide walls. From this point of view, the difference between $E$ and $H$ modes is just that of the polarizations of the plane-wave components. The plane-wave point of view also affords a simple picture of the phase and group velocities associated with a guide mode. A study of Fig. 5.2, which depicts a plane wave moving with speed $c$ at an angle $\theta$ relative to the $z$-axis, shows that during the time interval $dt$ the $z$ projection of a point on an equiphase surface advances a distance $c \cos \theta \, dt$, while the $z$ intercept of the equiphase surface advances a distance $\frac{\dot{z}}{\cos \theta} \, dt$. Thus the group velocity is

**Fig. 5.2.** Elementary wave propagating with angle $\theta$ with respect to $z$ axis.
92 5 Rectangular and Triangular Waveguides

\[ v = c \cos \theta = c \frac{\kappa}{k}. \]  \hspace{1cm} (5.25a)

while the phase velocity is

\[ u = \frac{c}{\cos \theta} = c \frac{k}{\kappa}. \]  \hspace{1cm} (5.25b)

in agreement with our previous results. The different nature of these velocities in particularly apparent at the cutoff frequency, where the group velocity is zero and the phase velocity infinite. From the plane wave viewpoint, the field consists of elementary wave moving perpendicularly to the guide axis. There is then no progression of waves along the guide – zero group velocity – while all points on a line parallel to the \( z \)-axis have the same phase – infinite phase velocity. Thus the essential difference between the two velocities is contained in the statement that the former is a physical velocity, the latter a geometrical velocity.

It has been remarked that the \( H_{m0} \) modes behave like \( E \) modes with respect to the \( y \) axis. A complete set of waveguide fields with \( E \) or \( H \) character relative to the \( y \) axis can be constructed from those already obtained. The \( E_{mn} \) and \( H_{mn} \) modes associated with the same nonvanishing integers are degenerate, since they possess equal critical frequencies and propagation constants. Furthermore, the respective transverse field components manifest the same dependence upon \( x \) and \( y \) – cf. (5.8a)–(5.8d) and (5.12a)–(5.12d). Hence, by a suitable linear combination of these modes, it is possible to construct fields for which either \( E_y \) or \( H_y \) vanishes. Similar remarks apply to the \( x \) axis. Thus, a decomposition into \( E \) and \( H \) modes can be performed with any of the three axes as preferred directions. The result is to be anticipated from the general analysis of Sec. 4.1, since a rectangular waveguide has cylindrical symmetry with respect to all three axes.

### 5.2 Isosceles Right Triangle

A square waveguide, \( a = b \), has a further type of degeneracy, since the \( E_{mn} \) and \( E_{nm} \) modes have the same critical frequency, as do the \( H_{mn} \) and \( H_{nm} \) modes. By suitable linear combinations of these degenerate modes, it is possible to construct the mode functions appropriate to a guide with a cross section in the form of an isosceles right triangle. The (unnormalized) mode function,

\[ \varphi_{mn}(x, y) = \sin \frac{m \pi x}{a} \sin \frac{n \pi y}{a} - \sin \frac{n \pi x}{a} \sin \frac{m \pi y}{a}, \]  \hspace{1cm} (5.26)

describes a possible \( E \) mode in a square guide, with the cutoff wavenumber

\[ \gamma_{mn} = \frac{\pi}{a} \sqrt{m^2 + n^2}. \]  \hspace{1cm} (5.27)

The function thus constructed vanishes on the line \( y = x \) as well as on the boundaries \( y = 0, x = a \), and therefore satisfies all the boundary conditions
for an E mode in an isosceles right triangular guide, as shown in Fig. 5.3. The linearly independent square guide E-mode function

$$\sin \frac{m \pi x}{a} \sin \frac{n \pi y}{a} + \sin \frac{n \pi x}{a} \sin \frac{m \pi y}{a}, \quad (5.28)$$

does not vanish on the line $y = x$, and therefore does not describe a possible triangular mode. Note that the function (5.26) vanishes if $m = n$, and that therefore an interchange of the integers produces a trivial change in sign of the function. Hence the possible E modes of an isosceles right triangular guide are obtained from (5.26) with the integers restricted by $0 < m < n$. Thus the dominant E mode corresponds to $m = 1, n = 2$, and has the cutoff wavelength $\lambda_c = \frac{2}{\sqrt{5}a}$. The mode function

$$\psi_{mn}(x, y) = \cos \frac{m \pi x}{a} \cos \frac{n \pi y}{a} + \cos \frac{n \pi x}{a} \cos \frac{m \pi y}{a}, \quad (5.29)$$

describing an H mode in the square guide, has a vanishing derivative normal to the line $y = x$:

$$\frac{\partial}{\partial n} \psi_{mn} = \frac{1}{\sqrt{2}} \left( -\frac{\partial}{\partial x} + \frac{\partial}{\partial y} \right) \psi_{mn} = 0, \quad y = x, \quad (5.30)$$

and therefore satisfies all boundary conditions for an H mode in the triangular guide under consideration. The linearly independent function

$$\cos \frac{m \pi x}{a} \cos \frac{n \pi y}{a} - \cos \frac{n \pi x}{a} \cos \frac{m \pi y}{a}, \quad (5.31)$$

is not acceptable for this purpose. The function (5.29) is symmetrical in the integers $m$ and $n$, and therefore the possible H modes of an isosceles right triangular guide are derived from (5.29) with the integers restricted by $0 \leq m \leq n$, but with $m = n = 0$ excluded. Thus the dominant H mode, and the dominant mode of the guide, corresponds to $m = 0, n = 1$, and has the cutoff wavelength $\lambda_c = 2a$, which is identical with the dominant mode cutoff wavelength of a rectangular guide with the maximum dimension $a$. 

**Fig. 5.3.** Isosceles right triangular waveguide (shown in cross section) obtained by bisecting a square waveguide by a plane diagonal to the square.
The discussion just presented may appear incomplete, for although we have constructed a set of triangular modes from the modes of the square guide, the possibility remains that there exist other modes not so obtainable. We shall demonstrate that all the modes of the isosceles right triangle are contained among those of the square, thereby introducing a method that will prove fruitful in the derivation of the modes of an equilateral triangle. It will be convenient to use a coordinate system in which the $x$ axis coincides with the diagonal side of the triangle, the latter then occupying a space below the $x$ axis. Let $\psi(x, y)$ be an $E$-mode function of the triangle, which therefore satisfies the wave equation (4.22a) and vanishes on the three triangular boundaries. We now define the function $\varphi(x, y)$ for positive $y$, in terms of its known values within the triangle for negative $y$, by

$$\varphi(x, y) = -\varphi(x, -y), \quad y > 0.$$ (5.32)

The definition of the function is thereby extended to the triangle obtained from the original triangle by reflection in the $x$ axis, the two regions together forming a square, as seen in Fig. 5.3. The two parts of the extended function are continuous and have continuous normal derivative across the line $y = 0$, since $\varphi(x, +0) = 0$, and

$$\varphi_y(x, y) = \frac{\partial}{\partial y} \varphi(x, y) = -\frac{\partial}{\partial(-y)} \varphi(x, -y) \equiv \varphi_y(x, -y),$$ (5.33)

whence $\varphi_y(x, +0) = \varphi_y(x, -0)$. Furthermore, $\varphi(x, y), y > 0,$ satisfies the wave equation:

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \gamma^2\right) \varphi(x, y) = -\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial(-y)^2} + \gamma^2\right) \varphi(x, -y) = 0,$$ (5.34)

and clearly vanishes on the orthogonal sides of the triangle produced by reflection. Hence, the extended function satisfies all the requirements for an $E$ mode function of the square, and thus our contention is proved, since every triangular $E$-mode function generates an $E$ mode of the square. The analogous $H$-mode discussion employs the reflection

$$\psi(x, y) = \psi(x, -y), \quad y > 0$$ (5.35)

to extend the definition of the function. All the necessary continuity and boundary conditions are easily verified.

### 5.3 Equilateral Triangle

The method just discussed may be used to great advantage in the construction of a complete set of $E$ and $H$ modes associated with a guide that has a
cross section in the form of an equilateral triangle. A mode function within the triangle can be extended to three neighboring equilateral triangles by reflection about the three lines, parallel to the sides of the triangle, that intersect the opposite vertices, as shown in Fig. 5.4. We now suppose that the function thus defined in these new regions is further extended by similar reflection processes, and so on, indefinitely, as sketched in Fig. 5.5. It is apparent from Fig. 5.5 that the infinite system of equilateral triangles so formed uniformly cover the entire $x$-$y$ plane. Hence, a function has been defined which is finite, continuous, has continuous derivatives, and satisfies the wave equation at every point of the $x$-$y$ plane. Such a function must be composed of uniform plane waves, and therefore the modes of the equilateral triangle must be
constructible from such plane waves. To proceed further, we observe that the extended mode function has special periodicity properties. Consider a row of triangles parallel to one of the three sides of the triangle. The value of the function at a point within a second neighboring parallel row is obtained by two successive reflections from the value at a corresponding point in the original row, and is therefore identical for both E and H modes. Hence the extended mode function must be periodic with respect to the three dimensions normal to the sides of the triangle, the periodicity interval being $2h$, where $h$ is the length of the perpendicular from an apex to the adjacent side. In terms of $a$, the base length of the triangle, $h = \frac{\sqrt{3}}{2}a$. To supply a mathematical proof for these statements we need merely write the equations that provide the successive definitions of the extended function in one of the directions normal to the sides of the triangle. Thus, for an E mode,

$$\varphi(x, y) = -\varphi(x, 2h - y), \quad h \leq y \leq 2h,$$

$$\varphi(x, y) = -\varphi(x, 4h - y), \quad 2h \leq y \leq 3h,$$

which relates the values of the function in three successive rows parallel to the $x$ axis. On eliminating the value of the function in the center row, we obtain

$$\varphi(x, y) = \varphi(x, y + 2h), \quad 0 \leq y \leq h,$$

which establishes the stated periodicity. The same proof is immediately applicable to translation in the other two directions, if a corresponding coordinate system is employed. The verification for H modes is similar. To construct a function having the desired periodicity properties, we introduce the three unit vectors normal to the sides of the triangle, and oriented in an inward sense (see Fig. 5.6)

$$e_1 = j, \quad e_2 = -\frac{\sqrt{3}}{2}i - \frac{1}{2}j, \quad e_3 = \frac{\sqrt{3}}{2}i - \frac{1}{2}j,$$

and remark that the function of vectorial position in the plane, $r$

$$F(r) = \sum_{\lambda, \mu, \nu} f(r - \lambda 2he_1 - \mu 2he_2 - \nu 2he_3)$$

is unaltered by a translation of magnitude $2h$ in any of the directions specified by $e_1$, $e_2$, and $e_3$. The summation is to be extended over all integral values of $\lambda$, $\mu$, $\nu$. If $f(r)$ is a solution of the wave equation everywhere, $F(r)$ will also be a solution, with the proper periodicity. Furthermore, if $f(r)$ is a uniform plane wave, $F(r)$ will also be of this form, and by a proper combination of such elementary fields, we can construct the modes of the equilateral triangle. Thus, assuming that

$$f(r) = e^{i\gamma r},$$

where the magnitude of the real vector $\gamma$ is the cutoff wavenumber of the mode, we find
Each summation is of the form
\begin{equation}
\sum_{\mu=-\infty}^{\infty} e^{i\mu x},
\end{equation}
which is a periodic function of \( x \), with the periodicity 2\( \pi \). Within the range \(-\pi \leq x < \pi\), the series may be recognized as the Fourier series expansion of 2\( \pi \delta(x) \):
\begin{equation}
\delta(x) = \frac{1}{2\pi} \sum_{\mu=-\infty}^{\infty} e^{i\mu x} \int_{-\pi}^{\pi} dx' e^{-i\mu x'} \delta(x') = \frac{1}{2\pi} \sum_{\mu=-\infty}^{\infty} e^{i\mu x}.
\end{equation}
We can therefore write, for all \( x \),
\begin{equation}
\sum_{\mu=-\infty}^{\infty} e^{i\mu x} = 2\pi \sum_{m=-\infty}^{\infty} \delta(x - 2\pi n),
\end{equation}
since this equation reduces to (5.43) in the range \(-\pi \leq x < \pi\), all delta function terms but the \( m = 0 \) one being nowhere different from zero in this range, and since both sides of the equation are periodic functions with periodicity 2\( \pi \). The relation (5.44) is known as the Poisson sum formula. Applying this result to the three summations contained in (5.41), we obtain
\begin{equation}
F(r) = e^{i\gamma \cdot r} (2\pi)^3 \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(2\hbar e_1 \cdot \gamma - 2\pi l) \times \delta(2\hbar e_2 \cdot \gamma - 2\pi m) \delta(2\hbar e_3 \cdot \gamma - 2\pi n).
\end{equation}
Hence, the function vanishes entirely unless the arguments of all three delta functions vanish simultaneously, which requires that
\begin{equation}
e_1 \cdot \gamma = \frac{\pi l}{\hbar}, \quad e_2 \cdot \gamma = \frac{\pi m}{\hbar}, \quad e_3 \cdot \gamma = \frac{\pi n}{\hbar},
\end{equation}
where \( l, m, \) and \( n \) are integers, positive, negative, or zero. The three integers may not be assigned independently, however, for
\begin{equation}
(e_1 + e_2 + e_3) \cdot \gamma = \frac{\pi}{\hbar} (l + m + n) = 0,
\end{equation}
since
\begin{equation}
e_1 + e_2 + e_3 = 0,
\end{equation}
which is evident from the geometry, or from the explicit form of the vectors, (5.38). Therefore,
With the relations (5.46) and (5.49) we have determined the cutoff wavenumbers of all E and H modes in the triangular guide. Thus

\[(e_1 \cdot \gamma)^2 + (e_2 \cdot \gamma)^2 + (e_3 \cdot \gamma)^2 = \gamma \cdot (e_1 e_1 + e_2 e_2 + e_3 e_3) \cdot \gamma = \frac{\pi^2}{h^2} (l^2 + m^2 + n^2) \cdot \gamma = \frac{\pi^2}{h^2} (l^2 + m^2 + n^2) . \quad (5.50)\]

However, it follows from the explicit form of the vectors that the dyadic

\[e_1 e_1 + e_2 e_2 + e_3 e_3 = \frac{3}{2} (ii + jj) \quad (5.51)\]
a multiple of the unit dyadic in two dimensions\(^1\) whence

\[\gamma^2 = \frac{2}{3} \frac{\pi^2}{h^2} (l^2 + m^2 + n^2) = \frac{8}{9} \frac{\pi^2}{a^2} (l^2 + m^2 + n^2) . \quad (5.52)\]

If we wish, \(l\) can be eliminated by means of (5.49), with the result

\[\gamma = \frac{4}{3} \frac{\pi}{a} \sqrt{m^2 + mn + n^2} . \quad (5.53)\]

We have shown that the elementary functions from which the triangular mode functions are to be constructed have the form

\[e^{i\gamma \cdot r} , \quad (5.54)\]

with the components of \(\gamma\) determined by (5.46). In view of the latter relations, and of (5.51), the quantity \(\gamma \cdot r\) is conveniently rewritten as follows:

\[\gamma \cdot r = \gamma \cdot (ii + jj) \cdot r = \frac{2}{3} \gamma \cdot (e_1 e_1 + e_2 e_2 + e_3 e_3) \cdot r = \frac{2}{3} \frac{\pi}{h} (e_1 \cdot r + e_2 \cdot r + e_3 \cdot r) = \frac{4}{3} \frac{\pi}{a} (lu + mv + nw) . \quad (5.55)\]

The three new variables,

\[u = e_1 \cdot r = y , \quad (5.56a)\]
\[v = e_2 \cdot r = -\frac{\sqrt{3}}{2} x - \frac{1}{2} y , \quad (5.56b)\]
\[w = e_3 \cdot r = \frac{\sqrt{3}}{2} x - \frac{1}{2} y , \quad (5.56c)\]

known as trilinear coordinates, are the projections of the position vector on the tree directions specified by the unit vectors \(e_1, e_2, e_3\), and are related by

\(^1\) The fact that the multiple is not unity signifies the overcompleteness of the vectors \(e_i\).
If the origin of coordinates is placed at the intersection of the three perpendiculars from each vertex to the opposite side, the trilinear coordinates of a point are the perpendicular distances from the origin to the three lines drawn through the point parallel to the sides of the triangle. A coordinate is considered negative if the line lies between the origin and the corresponding side; thus, the equations for the sides of the triangle are

$$u = -r, \quad v = -r, \quad w = -r,$$

where $r$ is the radius of the inscribed circle (see Fig. 5.6)

$$r = \frac{1}{3} h = \frac{a}{2\sqrt{3}}. \quad (5.58)$$

To construct the mode functions, we must attempt to satisfy the boundary conditions by combining all functions of the form (5.45) that correspond to the same value of $\gamma$. There are twelve such functions, of which six are obtained from (5.45) by permutation of the integers $l, m,$ and $n$, and another six from these by a common sign reversal of $l, m,$ and $n$, or equivalently, by taking the complex conjugate, cf. (5.41). It would not be possible to reverse the sign of $l$ alone, for example, since the condition (5.49) would be violated. We first consider $E$ modes and construct pairs of functions that vanish on the boundary $u = -r$. Thus

$$e^{\frac{2\pi i}{3} (lu + mv + nw - lh)} - e^{-\frac{2\pi i}{3} (lu + nv + mw - lh)} \quad (5.59)$$

has this property, for in consequence of the relations (5.49) and (5.57) it can be written

$$e^{\frac{2\pi i}{3} [l(u - h) + (m - n) \frac{u - w}{2}]} - e^{\frac{2\pi i}{3} [-l(u - h) + (m - n) \frac{u - w}{2}]}$$

$$= 2i \sin \frac{ln}{h} \left( u - \frac{2}{3} h \right) e^{\frac{2\pi i}{3} (m - n) \frac{u - w}{2}}, \quad (5.60)$$

which clearly vanishes when $u = -\frac{1}{3} h$. The two additional functions obtained by a cyclic permutation of $l, m, n$ also vanish on the boundary $u = -r$, as
do the complex conjugates of the three pairs of functions. Of course, the six functions and their conjugates could also be grouped in pairs so that they vanish on the boundary \( v = -r \), or \( w = -r \). It is now our task to construct one or more functions that have the proper grouping relative to all three boundaries, and are thus the desired general E-mode functions. It is easy to verify that the required functions are

\[
\varphi(x, y) = e^{i \frac{2\pi}{3} (lu + mw + nw - lh)} - e^{-i \frac{2\pi}{3} (lu + mw + lw - lh)} \\
+ e^{i \frac{2\pi}{3} (mu + lv + lw - mh)} - e^{-i \frac{2\pi}{3} (mu + lv + lw - mh)} \\
+ e^{i \frac{2\pi}{3} (nu + lv + lw - nh)} - e^{-i \frac{2\pi}{3} (nu + lv + lw - nh)},
\]

(5.61)

and its complex conjugate. Each E mode is therefore two-fold degenerate, and the real and imaginary parts of (5.61) are equally admissible mode functions. Note that permutations of the integers \( l, m, n \) yield no new functions, for (5.61) is unchanged under the three cyclic permutations, and the other three permutations convert (5.61) into its negative complex conjugate. On combining the pairs of functions in the matter of (5.60), and taking real and imaginary parts, we obtain the two E-mode functions associated with the integers \( l, m, n \) in the form

\[
\varphi(x, y) = \sin \frac{l\pi}{h} \left( y - \frac{2h}{3} \right) \cos \frac{\pi}{\sqrt{3}h} (m - n)x \\
+ \sin \frac{m\pi}{h} \left( y - \frac{2h}{3} \right) \cos \frac{\pi}{\sqrt{3}h} (n - l)x \\
+ \sin \frac{n\pi}{h} \left( y - \frac{2h}{3} \right) \cos \frac{\pi}{\sqrt{3}h} (l - m)x,
\]

(5.62)

where we have replaced \( u \) and \( \frac{w - w}{2} \) by \( y \) and \( -\frac{\sqrt{3}}{2}x \), respectively, according to (5.56a)–(5.56c).

To find the dominant E mode, we note that (5.62) vanishes identically if any integer is zero, and therefore the E mode of lowest cutoff wavenumber corresponds to \( m = n = 1, l = -2 \), and thus from (5.53) has the cutoff wavelength \( \lambda_c = \frac{\sqrt{3}}{2}a = h \). It is important to observe that the E-mode function constructed from the sine functions of \( x \) vanishes if two integers are equal. Hence the dominant E mode is nondegenerate. The H-mode function analogous to (5.61) is

\[
\psi(x, y) = e^{i \frac{2\pi}{3} (lu + nv + lw - lh)} + e^{-i \frac{2\pi}{3} (lu + nv + lw - lh)} \\
+ e^{i \frac{2\pi}{3} (mu + lv + lw - mh)} + e^{-i \frac{2\pi}{3} (mu + lv + lw - mh)} \\
+ e^{i \frac{2\pi}{3} (nu + lv + lw - nh)} + e^{-i \frac{2\pi}{3} (nu + lv + lw - nh)}.
\]

(5.63)

To verify this statement, it must be shown that \( \psi(x, y) \) has vanishing derivatives normal to each of the triangular sides. Now

\[
e_1 \cdot \nabla = \frac{\partial}{\partial y} = \frac{\partial}{\partial u} - \frac{1}{2} \frac{\partial}{\partial v} - \frac{1}{2} \frac{\partial}{\partial w} = \frac{3}{2} \frac{\partial}{\partial u},
\]

(5.64)
since the operator $\partial / \partial u + \partial / \partial v + \partial / \partial w$ annihilates all of the twelve plane waves in consequence of $l = m = n = 0$. The symmetry of this relation is sufficient assurance for the validity of

$$e_2 \cdot \nabla = \frac{3}{2} \frac{\partial}{\partial v}, \quad e_3 \cdot \nabla = \frac{3}{2} \frac{\partial}{\partial w}. \quad (5.65)$$

Thus, each normal derivative is proportional to the derivative with respect to the corresponding trilinear coordinate, and the application of any of these operators to (5.63) produces a form in which the same cancellation occurs on the boundaries as for the E modes. The two-fold degenerate mode functions obtained as the real and imaginary parts of (5.63) are

$$\psi(x, y) = \cos \frac{l\pi}{h} \left( y - \frac{2h}{3} \right) \cos \frac{\pi}{\sqrt{3}h} (m - n)x$$

$$+ \cos \frac{m\pi}{h} \left( y - \frac{2h}{3} \right) \cos \frac{\pi}{\sqrt{3}h} (n - l)x$$

$$+ \cos \frac{m\pi}{h} \left( y - \frac{2h}{3} \right) \cos \frac{\pi}{\sqrt{3}h} (l - m)x. \quad (5.66)$$

It should again be noted that a permutation of the integers produces no new mode, the function (5.66) either remaining invariant or reversing sign under such an operation. It is permissible for one integer to be zero, and therefore the dominant H mode, which is also the dominant mode of the guide, corresponds to $m = 1, n = -1, l = 0$, and has the cutoff wavelength $\lambda_c = \frac{3}{2}a$. Unlike the dominant E mode, the dominant H mode is doubly degenerate.

The mode functions of the equilateral triangle have interesting properties relative to coordinate transformations that leave the triangle invariant. If, for example, the coordinate system indicated in Fig. 5.6 is rotated counterclockwise through $120^\circ$, the triangle has the same aspect relative to the new coordinate system that it had in the original system. The relationship between the unit vectors and trilinear coordinates in the two coordinate systems is expressed by

$$e_1 = e'_3, \quad e_2 = e'_1, \quad e_3 = e'_2, \quad (5.67a)$$

$$u = w', \quad v = u', \quad w = v', \quad (5.67b)$$

where primes indicate that the new reference system is involved. Now, the wave equation is invariant in form under a rotation of coordinates and the boundary conditions have the same form in the two systems. If therefore a mode function $f(u, v, w)$ is expressed in the new coordinates:

$$f(u, v, w) = f'(u', v', w'), \quad (5.68)$$

the new function $f'(u, v, w)$ must also be a possible mode function corresponding to the same eigenvalue, and can therefore be expressed as a linear
combination of the degenerate mode functions associated with that eigenvalue. This conclusion is most easily verified by considering the complex E and H mode functions (5.61) and (5.63). It is not difficult to show that the effect of the substitution (5.67b) is to produce a function of the primed coordinates that has exactly the same form, save that each function is multiplied by the constant

$$C = e^{i \frac{2\pi}{3} (n-m)} = e^{i \frac{2\pi}{3} (l-n)} = e^{i \frac{2\pi}{3} (m-l)} .$$

The equivalence of these expressions follows from

$$(n - m) - (l - n) = 3n ,$$

$$(m - l) - (n - m) = 3m ,$$

$$(l - n) - (m - l) = 3l .$$

(5.70)

Of course, the complex conjugate of a mode function is multiplied by $C^* = e^{i \frac{2\pi}{3} (m-n)/3}$. If two such rotations are applied in succession, implying that the coordinate system is rotated through 240°, each mode function is multiplied by $C^2$. If a rotation through 360° is performed, we must expect that the mode function preserve their original form, which requires that

$$C^3 = 1 ,$$

as indeed it is. Hence all modes can be divided into three classes, depending on whether the mode function is multiplied by 1, $e^{2\pi i/3}$, or $e^{4\pi i/3} = e^{-2\pi i/3}$, under the influence of a rotation through 120°. The corresponding classification of $n - m$ is that it is divisible by 3 with a remainder that is either 0, 1, or 2. Thus, the dominant E-mode functions ($m = n = 1$) are invariant under rotation, while the dominant H-mode function ($m = -n = 1$) and its complex conjugate are multiplied by $e^{-2\pi i/3}$ and $e^{2\pi i/3}$, respectively, under a rotation through 120°. It should be noted that although the complex functions (5.61) and (5.63) are transformed into multiples of themselves by a rotation, the real and imaginary parts will not have this simple behavior, save for the class that is invariant under rotation.

Another coordinate transformation that leaves the triangle unaltered is reflection. If, in the coordinate system of Fig. 5.6, the positive sense of the $x$ axis is reversed, the triangle has the same aspect relative to the new system. Again the wave equation is unaltered in form by the reflection transformation

$$x = -x' , \quad y = y' ,$$

(5.72a)

or

$$u = u' , \quad v = w' , \quad w = v' .$$

(5.72b)

and we conclude that a mode function expressed in the new coordinates must be a linear combination of the mode functions associated with the same eigenvalue. Indeed, the H-mode function (5.63) is converted into its complex conjugate, while the E-mode function becomes its negative complex conjugate.
However, it should be observed that now the real functions (5.62) and (5.66) change into multiples of themselves under the transformation, either remaining unaltered in form or reversing sign. It follows that two successive reflections produce no change in the mode function, as must be required. Two other reflection operations exist, associated with the remaining perpendiculars from the vertices to opposite sides. They are described by

\[ u = w', \quad v = v', \quad w = u', \]  

(5.73a)

and

\[ u = v', \quad v = u', \quad w = w'. \]  

(5.73b)

Thus, the entire set of symmetry operations of the triangle, consisting of three rotations and three reflections, are described by the six permutations of the trilinear coordinates \( u, v, \) and \( w. \) Only two of these transformations are independent, in the sense that successive application of them will generate all the other transformations. Thus, a rotation through 120°, \( u = w', \) \( v = v', \) \( w = v', \) (5.67b), followed by the reflection (5.72b), \( u' = u'', \) \( v' = w'', \) \( w' = v'', \) is equivalent to \( u = v'', \) \( v = u'', \) \( w = w'', \) the reflection (5.73b).

Similarly, a rotation through 240°, which is the transformation (5.67b) applied twice, followed by the reflection (5.72b), generates the reflection (5.73a). As a particular consequence of these statements, note that a mode function which is invariant with respect to rotations and one of the reflections, is also invariant under the other two reflections.

In constructing the complete set of E and H modes for a guide with a cross section in the form of an equilateral triangle, we have also obtained a complete solution for the modes of a right angle triangle with angles of 30° and 60°. The connection between the two problems is the same as that between the square and the isosceles right angle triangle. The set of equilateral E modes (5.62) that involve sine functions of \( x \) vanish on the line \( x = 0 \) and therefore satisfy all the boundary conditions for an E mode of the 30°, 60° triangles thus obtained. To find the dominant E mode we must recall that no integer can be zero and that no two integers can be equal for the mode that is an odd function of \( x. \) Hence the dominant mode corresponds to \( m = 1, \) \( n = 2, \) \( l = -3, \) and has from (5.53) the cutoff wavelength \( \lambda_c = \frac{3}{2\sqrt{7}} a, \) where \( a \) is now the length of the diagonal side. The set of equilateral triangular H modes (5.66) that contain cosine functions of \( x \) have vanishing derivatives normal to the y axis and therefore yield the H modes of a 30°, 60° triangle. The dominant H mode of the latter triangle is the same as that of the equilateral triangle \( (m = 1, n = -1, l = 0) \) and has the same cutoff wavelength \( \lambda_c = \frac{3}{2} a. \) The modes of the equilateral triangle also furnish us with modes for a guide that has a cross section formed by any closed curve drawn along the lines of Fig. 5.5. In particular, we obtain modes for a cross section in the form of a regular hexagon, but only some of the modes are obtained in this way.
5.4 Problems for Chapter 5

1. Verify that the H modes that satisfy (5.35) fulfill all the necessary requirements to be H modes for the square, and that therefore all the H modes of the isosceles right triangle are contained within those for the square.