2 Waves guided by perfectly conducting boundaries

2.1 TEM TRANSMISSION LINES

The simplest systems for guiding electromagnetic waves are the two-wire lines. Figure 2.1 illustrates some common types. These lines all have at least two conductors which are electrically insulated from each other. In addition the whole space where the electric field is not zero is filled with a uniform dielectric material. A number of other types of two-wire line exist which have more than one dielectric around them. These lines which have rather more complicated behaviour are discussed in the next chapter.

Elementary treatments of the theory of lines such as those shown in Fig. 2.1 assume that the fields are transverse. It is straightforward to show that this is consistent with the circuit approach for the special case of the coaxial cable (Carter, 1986). We now must prove that this result holds for all lines of this type.

The fields around the conductors obey Maxwell's equations (1.8) and (1.16). Taking the curl of (1.8) and using (1.16) gives

$$\begin{aligned} \nabla \wedge (\nabla \wedge \boldsymbol{E}) &= -\frac{\partial}{\partial t}(\nabla \wedge \boldsymbol{B}) \\ &= -\mu\frac{\partial}{\partial t}(\nabla \wedge \boldsymbol{H}) \\ &= -\varepsilon\mu\frac{\partial^2 \boldsymbol{E}}{\partial t^2}. \end{aligned} \tag{2.1}$$

The left hand side of this equation can be written

$$\nabla \wedge (\nabla \wedge E) = \nabla(\nabla \cdot \boldsymbol{E}) - \nabla^2 \boldsymbol{E}. \tag{2.2}$$

This relationship can be proved by evaluating the derivatives in terms of the Cartesian components of $\boldsymbol{E}$ though the task is a bit laborious. Now, since there is no free charge between the conductors $\varrho = 0$ and therefore $\nabla \cdot \boldsymbol{E} = 0$ from (1.5). Thus from (2.1) and (2.2) we have

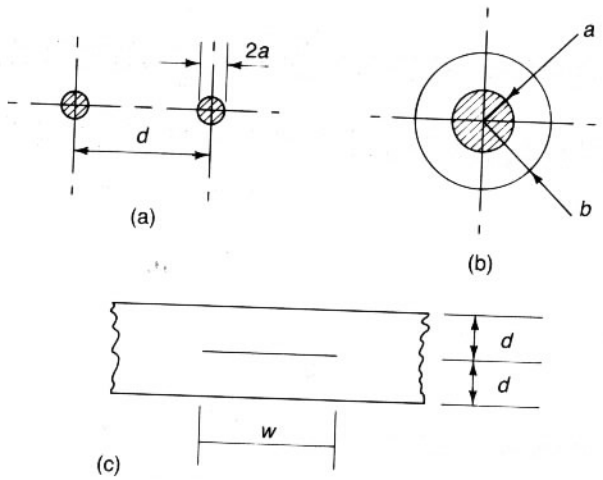

Fig. 2.1 Two-wire transmission lines: (a) parallel-wire line, (b) coaxial cable, and (c) triplate line.

$$\nabla^2 \boldsymbol{E} = \varepsilon\mu \frac{\partial^2 \boldsymbol{E}}{\partial t^2}. \tag{2.3}$$

This is the three-dimensional form of the wave equation. In Cartesian coordinates it can be written

$$\frac{\partial^2 E_x}{\partial x^2} + \frac{\partial^2 E_x}{\partial y^2} + \frac{\partial^2 E_x}{\partial z^2} = \varepsilon\mu \frac{\partial^2 E_x}{\partial t^2} \tag{2.4}$$

together with similar equations for E_y and E_z. Now suppose that $\boldsymbol{E}$ has only x and y components and that these vary as $\exp j(\omega t - kz)$. Substitution into (2.4) produces

$$\frac{\partial^2 E_x}{\partial x^2} + \frac{\partial^2 E_x}{\partial y^2} - k^2 E_x = -\varepsilon\mu\omega^2 E_x. \tag{2.5}$$

If we assume that

$$k^2 = \varepsilon\mu\omega^2. \tag{2.6}$$

Then

$$\frac{\partial^2 E_x}{\partial x^2} + \frac{\partial^2 E_x}{\partial y^2} = 0. \tag{2.7}$$

But $\boldsymbol{E}$ can be written in terms of a scalar potential V

$$E_x = -\frac{\partial V}{\partial x} \tag{2.8}$$

so that

$$\frac{\partial^2}{\partial x^2}\left(\frac{\partial V}{\partial x}\right) + \frac{\partial^2}{\partial y^2}\left(\frac{\partial V}{\partial x}\right) = 0$$

$$\frac{\partial}{\partial x}\left(\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2}\right) = 0. \tag{2.9}$$

Similarly by considering the y component of $\boldsymbol{E}$ we obtain

$$\frac{\partial}{\partial y}\left(\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2}\right) = 0. \tag{2.10}$$

Equations (2.9) and (2.10) can be satisfied simultaneously if the expression in the brackets is zero, that is, if V satisfies Laplace's equation (Carter, 1986, p. 16). Thus for the mode of propagation which satisfies (2.6) the electric field distribution is identical to the electrostatic field between the electrodes. That is why it is possible to compute the capacitance and inductance per unit length of such a line from the static field solutions. It should be noted that, whilst this TEM wave is a possible solution, there could be other solutions for which (2.6) is not satisfied and the fields differ from the static fields. This point will be examined further in Section 2.8.

The use of field theory here, as very often in electromagnetism, is to provide a way of calculating the circuit parameters. The phase velocity of a TEM wave on a two-wire line is constant and equal to that of an unbounded plane TEM wave propagating through the same medium as that which separates the conductors. Expressions for the capacitance and inductance per unit length and the characteristic impedance for the two-wire lines shown in Fig. 2.1 are given in Table 2.1.

Table 2.1

Line	C (F m^{-1})	L (H m^{-1})	Z (Ω)
Coaxial	$\dfrac{2\pi\varepsilon}{\ln(b/a)}$	$\dfrac{\mu}{2\pi}\ln\left(\dfrac{b}{a}\right)$	$\dfrac{1}{2\pi}\sqrt{\left(\dfrac{\mu}{\varepsilon}\right)}\ln\left(\dfrac{b}{a}\right)$
Two wire	$\dfrac{\pi\varepsilon}{\cosh^{-1}\left(\dfrac{d}{2a}\right)}$	$\dfrac{\mu}{\pi}\cosh^{-1}\left(\dfrac{d}{2a}\right)$	$\dfrac{1}{\pi}\sqrt{\left(\dfrac{\mu}{\varepsilon}\right)}\cosh^{-1}\left(\dfrac{d}{2a}\right)$
Triplate $w \gg d$	$\dfrac{2\varepsilon w}{d}$	$\dfrac{\mu d}{2w}$	$\dfrac{d}{2w}\sqrt{\left(\dfrac{\mu}{\varepsilon}\right)}$

2.2 REFLECTION OF WAVES BY A CONDUCTING PLANE

The simplest case of reflection of waves by a conducting plane occurs when the direction of propagation is normal to the plane. Equation (1.45) shows

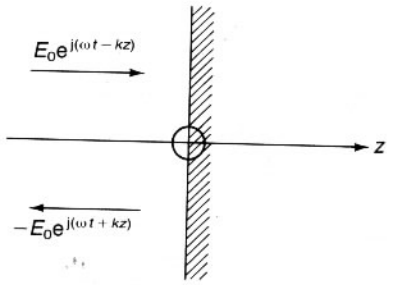

Fig. 2.2 Reflection of electromagnetic waves from a conducting surface at normal incidence.

that the wave impedance in a good conductor tends to zero as ω tends to infinity. The boundary conditions at the surface of the conductor are therefore given by (1.86). If the conducting surface lies in the (x, y) plane as shown in Fig. 2.2 the electric field in the region to the left of the plane is given by

$$E_0 \exp j(\omega t - kz) - E_0 \exp j(\omega t + kz) = -2jE_0 \sin kz \exp(j\omega t). \quad (2.11)$$

The field is therefore a standing wave as shown in Fig. 2.3. The amplitude of the electric field at each point in space is fixed and the actual field values vary sinusoidally with time between the limits indicated by the solid and broken curves shown in the figure.

The corresponding magnetic field is

$$H_0 \exp j(\omega t - kz) + H_0 \exp j(\omega t + kz) = -\frac{2E_0}{Z_0} \cos kx \exp j\omega t. \quad (2.12)$$

This field is also a standing wave but with a maximum rather than a zero of field at the surface of the plane. The factor j which appears in (2.11) but not in (2.12) shows that the electric and magnetic fields are in phase quadrature so that there is no flow of energy. The energy is stored alternately in the electric and magnetic fields. This situation is analogous to the transfer of energy between kinetic and potential energy in a pendulum, or between electric and magnetic stored energy in a resonant circuit. A simple electromagnetic resonator can be constructed by putting a second plane parallel to the first. If the second plane is located at one of the zeroes of the electric field, shown in Fig. 2.4, then the boundary conditions will be satisfied upon it. Taking the separation between the planes as d the condition for resonance is

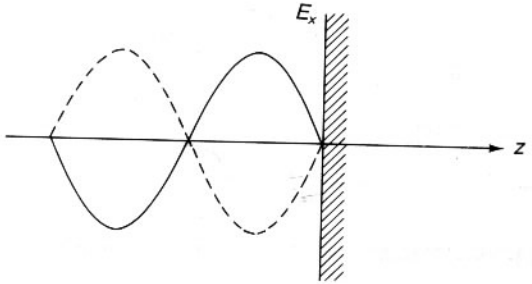

Fig. 2.3 Standing wave produced by the reflection of electromagnetic waves by a conducting surface at normal incidence.

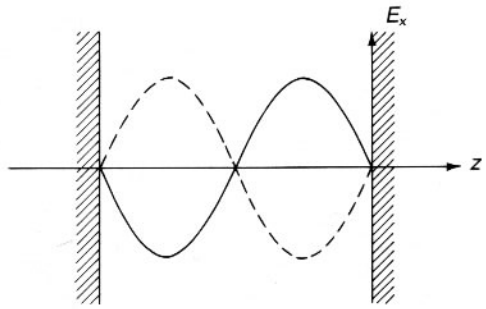

Fig. 2.4 Standing wave produced by reflection of an electromagnetic wave between two parallel conducting planes.

$$\sin kd = 0$$

$$k = \frac{\pi}{d}, \frac{2\pi}{d}, \frac{3\pi}{d}, \text{ etc.} \quad (2.13)$$

We shall return to this subject in chapter 7.

2.3 TRANSVERSE ELECTRIC WAVES

When the waves are incident obliquely upon the boundary the situation is rather more complicated. The cases when the electric and magnetic field vectors are parallel to the plane must be treated separately. Any general case can then be regarded as a superposition of these two particular cases.

Consider first the case with the electric field vector parallel to the conducting plane. Equation (1.95) shows that the incident and reflected waves must be in antiphase at the boundary because the tangential component of the electric field must always be zero there. It is helpful to discuss the wave pattern produced by this reflection by thinking in terms of wavefronts. Figure 2.5 shows the electric fields of the incident and reflected waves diagrammatically at an instant of time. The solid lines represent places where the electric field is a maximum and directed out of the paper and the broken lines places where it is maximum but directed into the paper. The directions of motion of these wavefronts are shown by the arrows. At the

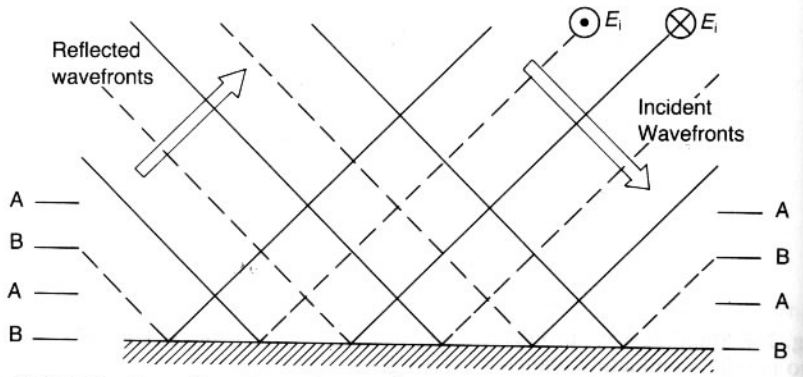

Fig. 2.5 Reflection of a plane electromagnetic wave by a conducting plane for oblique incidence with the electric field vectors parallel to the plane.

conducting surface the incident and reflected waves are in antiphase as required by the boundary conditions.

On planes such as A–A the two waves are in phase with each other producing a maximum of the electric field. This maximum is just twice the amplitude of the incident wave. Similarly on planes such as B–B the waves are in antiphase so that the electric field is zero. These planes are parallel to the conducting surface, are evidently equally spaced, and their positions are independent of the instantaneous positions of the wavefronts. Thus in the direction normal to the plane there is a standing wave exactly as in the case of normal incidence.

The separation between the planes can be found by considering Fig. 2.6. PP′ and QQ′ are successive wavefronts in a wave with angle of incidence θ. The perpendicular distance between them is equal to the wavelength of the incident wave λ_0. The wavelength along the plane is therefore

$$\mathrm{PQ} = \lambda_0/\sin\theta. \tag{2.14}$$

If Q–Q″ is a reflected wavefront then PP′ and QQ″ intersect on the line

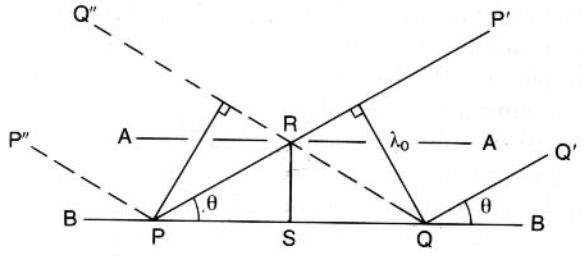

Fig. 2.6 Geometry of the reflection of an electromagnetic wave by a conducting plane for oblique incidence.

AA at R and RS is perpendicular to PQ. Then from the triangle PRS we have

$$\begin{aligned}\mathrm{RS} &= \mathrm{PS}\tan\theta \\ &= \lambda_0/(2\cos\theta)\end{aligned} \tag{2.15}$$

which is the separation between the planes A–A and B–B.

Looking again at Fig. 2.5 we see that the pattern of wavefronts is moving parallel to the plane. The wavelength in this direction is $\lambda_0/\sin\theta$ so that the phase velocity is given by

$$v_p = f\lambda_0/\sin\theta = v_{p0}/\sin\theta \tag{2.16}$$

where v_{p0} is the phase velocity of the incident wave. If the space in which the waves are propagating is empty then (2.16) shows that the phase velocity of the waves exceeds the velocity of light. This apparently contradicts the assumption in the theory of relativity that nothing can move faster than the speed of light. There is, in fact, no contradiction as will be shown later on.

The behaviour of the magnetic field can be deduced by considering Fig. 2.8 which shows wavefronts in terms of the magnetic field vectors at a moment in time. The arrows show the directions of the field. The component of $\boldsymbol{H}$ parallel to the surface is not reversed by the reflection so that it is consistent with equation (1.96). The planes A–A and B–B are identical to those in Fig. 2.5. The pattern of the magnetic field is the superposition of the fields of the incident and reflected waves. Using the Cartesian coordinate axes shown in Fig. 2.8 we can resolve the magnetic field vectors parallel and perpendicular to the plane to give

The behaviour of the magnetic field can be deduced by considering Fig. 2.8 which shows wavefronts in terms of the magnetic field vectors at a moment in time. The arrows show the directions of the field. The component of $\boldsymbol{H}$ parallel to the surface is not reversed by the reflection so that it is consistent with equation (1.96). The planes A–A and B–B are identical to those in Fig. 2.5. The pattern of the magnetic field is the

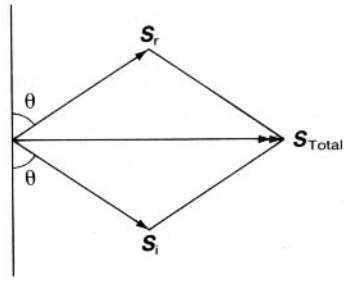

Fig. 2.7 Vector addition of the Poynting vectors of the incident and reflected waves.

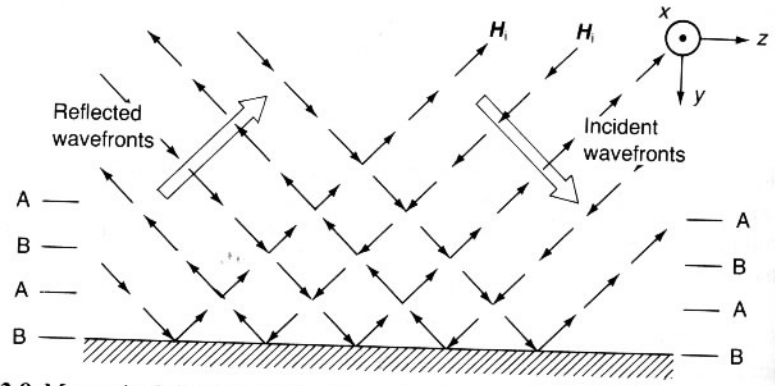

Fig. 2.8 Magnetic field vectors for the reflection of electromagnetic waves by a plane as illustrated in Fig. 2.5.

superposition of the fields of the incident and reflected waves. Using the Cartesian coordinate axes shown in Fig. 2.8 we can resolve the magnetic field vectors parallel and perpendicular to the plane to give

$$H_y = |\boldsymbol{H}_i| \sin\theta \quad \text{and} \quad H_z = -|\boldsymbol{H}_i| \cos\theta \tag{2.18}$$

for the incident wave, and

$$H_y = -|\boldsymbol{H}_i| \sin\theta \quad \text{and} \quad H_z = -|\boldsymbol{H}_i| \cos\theta \tag{2.19}$$

for the reflected wave. Then, on planes (A–A) on which the magnetic fields are in antiphase, the z components cancel and

$$H_y = 2|\boldsymbol{H}_i| \sin\theta. \tag{2.20}$$

Similarly on planes where they are in phase (B–B) the y components cancel so that

$$H_z = 2|\boldsymbol{H}_i| \cos\theta. \tag{2.21}$$

The local directions of the magnetic field are therefore as shown in Fig. 2.9. The magnetic flux lines are seen to form closed loops as required by equation (1.2).

We are now in a position to write down the equations which define the electric and magnetic field vectors at every point in space. These are

$$E_x = 2|\boldsymbol{E}_i| \sin k_y y \cos(\omega t - k_z z) \tag{2.22}$$

$$H_y = \frac{2|\boldsymbol{E}_i|}{Z_0} \sin\theta \sin k_y y \cos(\omega t - k_z z) \tag{2.23}$$

$$H_z = \frac{2|\boldsymbol{E}_i|}{Z_0} \cos\theta \cos k_y y \sin(\omega t - k_z z), \tag{2.24}$$

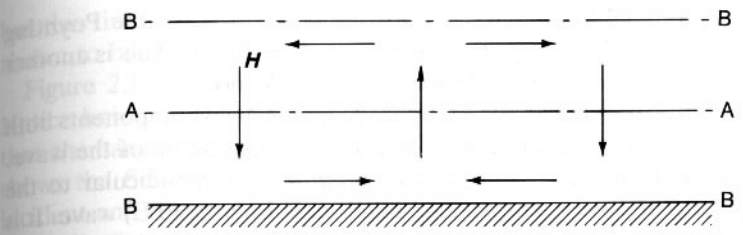

Fig. 2.9 Magnetic field vectors at various positions close to the conducting plane of Fig. 2.5.

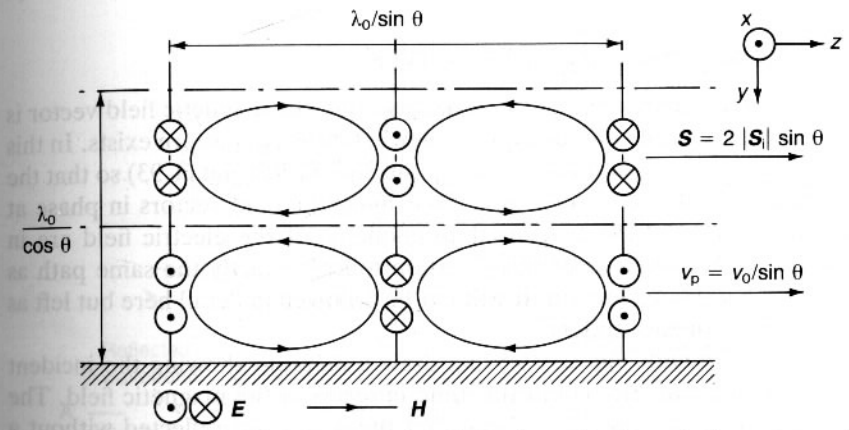

Fig. 2.10 Complete field pattern produced by the reflection of electromagnetic waves by a conducting plane when the electric field vectors are parallel to the plane.

where $$k_y = \frac{2\pi}{\lambda_0} \cos\theta = k_0 \cos\theta \tag{2.25}$$

and $$k_z = \frac{2\pi}{\lambda_0} \sin\theta = k_0 \sin\theta. \tag{2.26}$$

These equations show that we can think of a wave travelling at an angle to the coordinate axes as propagating as

$$\exp j(\omega t - k_y y - k_z z), \tag{2.27}$$

where k_y and k_z are the components of the vector propagation constant $\boldsymbol{k}_0$ in the y and z directions. $\boldsymbol{k}_0$ is a vector in the direction of propagation whose magnitude is $2\pi/\lambda_0$.

Figure 2.10 summarizes these results in the form of a diagram. The field pattern shown is repeated periodically in both the y and z directions. It is worth spending some time studying Fig. 2.10 in relation to equations (2.22)

to (2.24). Notice that H_z is in phase quadrature with E_x so the Poynting vector derived from them has a time average value of zero. This is another way of showing that the power flow is in the z direction.

The wave pattern shown in Fig. 2.10 has magnetic field components both parallel to and perpendicular to the direction of propagation of the wave. The electric field, however, only has a component perpendicular to the z axis. This kind of wave is known as a Transverse Electric (TE) wave. It is useful to distinguish it as a separate kind of wave although it is, in reality, just the superposition of two TEM waves travelling at an angle to each other. Because the magnetic field has a component in the z direction this kind of wave is sometimes referred to as an $\boldsymbol{H}$ wave.

2.4 TRANSVERSE MAGNETIC WAVES

When the incident waves are polarized so that the magnetic field vector is parallel to the reflecting surface a somewhat different pattern exists. In this case the boundary conditions to be applied are (1.92) and (1.93) so that the incident and reflected waves have their magnetic field vectors in phase at the boundary and the tangential components of the electric field are in antiphase. The analysis of this situation follows exactly the same path as that in the previous section. It will not be followed in detail here but left as an exercise for the student.

By analogy with Fig. 2.5 we can draw a diagram showing the incident and reflected wavefronts but this time in terms of the magnetic field. The result is shown in Fig. 2.11. Note that the waves are reflected without a change of phase. As before there is a standing wave in the direction normal to the plane with planes A–A and B–B on which the fields cancel or add

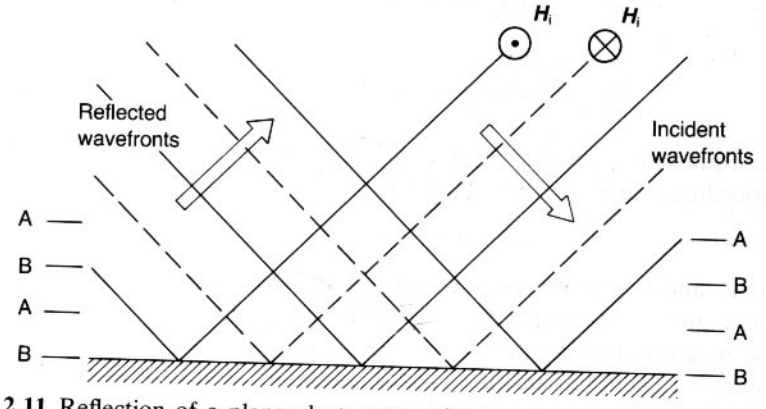

Fig. 2.11 Reflection of a plane electromagnetic wave by a conducting plane for oblique incidence with the magnetic field vectors parallel to the plane.

but this time A–A is a null plane. The separations of the planes can be shown to be exactly as before.

Figure 2.12 shows the electric field vectors by analogy with Fig. 2.8. Notice how the directions of the vectors at the reflecting surface ensure that the tangential component of the electric field is zero. Figure 2.13 shows the field vectors at certain points obtained by superimposing the incident and relected waves.

Finally Fig. 2.14 shows the complete field pattern for this case. Notice how the electric field lines form closed loops except immediately adjacent to the conducting plane where they terminate on surface charges. The equations describing the fields are found by inspection to be

$$H_x = \frac{2|\boldsymbol{E}_i|}{Z_0} \cos(k_y y) \cos(\omega t - k_z z) \tag{2.28}$$

$$E_y = -2|\boldsymbol{E}_i| \sin\theta \cos(k_y y) \cos(\omega t - k_z z) \tag{2.29}$$

$$E_z = 2|\boldsymbol{E}_i| \cos\theta \sin(k_y y) \sin(\omega t - k_z z), \tag{2.30}$$

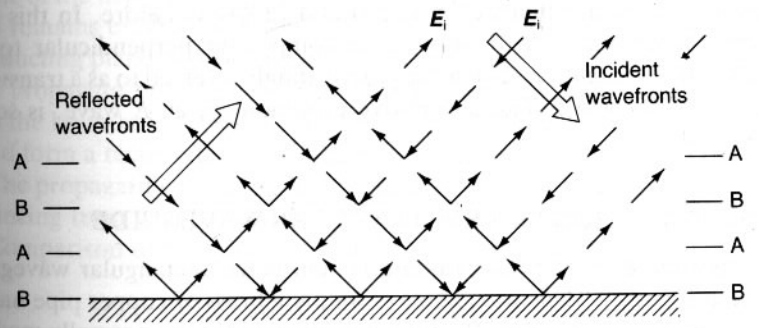

Fig. 2.12 Electric field vectors for the reflection of electromagnetic waves by a plane as illustrated in Fig. 2.11.

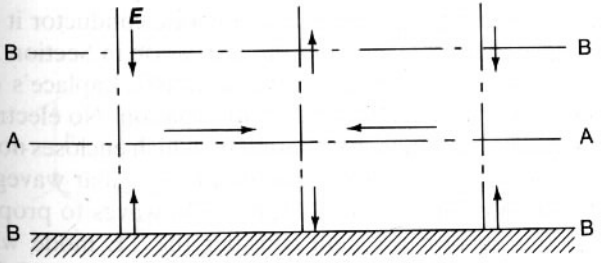

Fig. 2.13 Electric field vectors at various positions close to the conducting plane of Fig. 2.11.

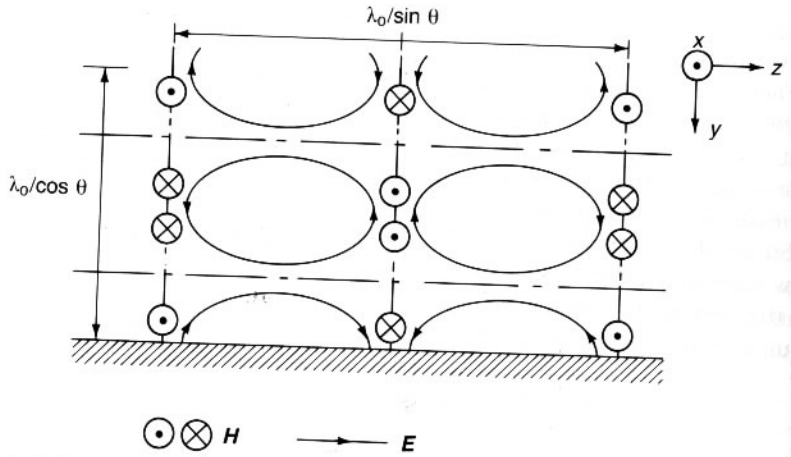

Fig. 2.14 Complete field pattern produced by the reflection of electromagnetic waves by a conducting plane when the magnetic field vectors are parallel to the plane.

where k_y and k_z are defined by (2.25) and (2.26) as before. In this field pattern the magnetic field only has a component perpendicular to the direction of propagation. The wave is accordingly referred to as a transverse magnetic (TM) wave. The alternative terminology, an ***E*** wave, is sometimes used.

2.5 PROPAGATION IN A RECTANGULAR WAVEGUIDE

An important practical waveguiding system is the rectangular waveguide shown in Fig. 2.15. This waveguide is simply a brass or copper pipe having a rectangular cross section. Rectangular waveguides are normally made in standard sizes with width a approximately twice the height b. Table 2.2 lists the sizes in common use with their designations, inside dimensions and recommended frequency bands. The reasoning behind the choice of dimensions is discussed in Section 2.7.

Because a rectangular waveguide has only one conductor it cannot support a TEM wave. This is clear from the discussion in Section 2.1 where it was shown that the fields of such a wave satisfy Laplace's equation in planes perpendicular to the direction of propagation. No electrostatic field can exist within a closed conducting boundary which encloses no free charge and therefore TEM waves cannot exist in a rectangular waveguide.

It is, however, possible for both TE and TM waves to propagate down the guide. In this section we shall consider the TE mode which has its electric field vector parallel to the narrow wall of the guide. Other possible modes of propagation will be considered in Section 2.7. To consider how

Table 2.2 Rectangular waveguides in common use

Designation UK	Designation USA	a (mm)	b (mm)	Frequency (GHz)	Power (MW)
WG6	WR650	165.1	82.6	1.14 to 1.73	13.5
WG8	WR430	109.2	54.6	1.72 to 2.61	5.9
WG10	WR284	72.14	34.04	2.60 to 3.95	2.4
WG12	WR187	47.6	22.1	3.94 to 5.99	1.0
WG14	WR137	34.85	15.80	5.38 to 8.17	0.54
WG16	WR90	22.86	10.16	8.20 to 12.50	0.23
WG18	WR62	15.80	7.90	11.9 to 18.00	0.12
WG20	WR42	10.67	4.32	17.6 to 26.7	0.048
WG22	WR28	7.11	3.56	26.4 to 40.1	0.025

this TE mode propagates we refer to Fig. 2.10. In this figure the waves are propagating from left to right parallel to the conducting plane and the electric field is normal to the plane of the paper. If another conducting plane is placed parallel to the first so that it coincides with the first null plane of the field pattern then the field between the two planes is unaffected and remains exactly as shown in Fig. 2.10. If, in addition, another pair of conducting planes is placed parallel to the plane of the paper the wave can still propagate because the electric field is perpendicular to these planes and the boundary conditions can be satisfied. The set of four planes so defined form a rectangular conducting pipe exactly as illustrated in Fig. 2.15.

The propagating wave can be thought of as a combination of TEM waves bouncing from wall to wall down the waveguide.

Comparison of Fig. 2.10 and 2.15 shows that

$$\cos\theta = \lambda_0/2a. \tag{2.31}$$

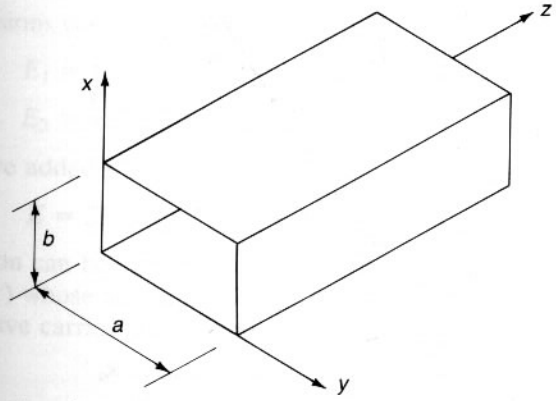

Fig. 2.15 A rectangular waveguide.

It is convenient to define the guide wavelength of the TE wave by

$$\lambda_g = \lambda_0/\sin\theta \tag{2.32}$$

(see Fig. 2.10) so that

$$\sin\theta = \lambda_0/\lambda_g. \tag{2.33}$$

The angle θ is not apparent to the user of the waveguide. It can be eliminated by squaring and adding (2.31) and (2.33) to give

$$\frac{1}{\lambda_0^2} = \frac{1}{(2a)^2} + \frac{1}{\lambda_g^2} \tag{2.34}$$

or

$$\frac{1}{\lambda_g^2} = \frac{1}{\lambda_0^2} - \frac{1}{(2a)^2}. \tag{2.35}$$

Multiplying this equation by $4\pi^2$, recalling that the propagation constant is given by $k = 2\pi/\lambda$, and that in free space $k = \omega/c$, where c is the velocity of light, we obtain

$$k_g = \frac{1}{c}(\omega^2 - \omega_c^2)^{\frac{1}{2}} \tag{2.36}$$

where $\omega_c = (\pi c/a)$. To understand the significance of this equation it is helpful to consider the graph of ω against k derived from it shown in Fig. 2.16. When $\omega = \omega_c$, $k_g = 0$ so that the wave does not propagate. The guide is then said to be at cut-off and ω_c is referred to as the cut-off frequency. The physical significance of this result can be shown by considering equation (2.33). At cut-off λ_g tends to infinity so that $\theta = 0$ and the wave bounces backwards and forwards across the guide without progressing down it.

At frequencies above cut-off the relationship between the guide wavelength and the frequency is given by the curve in Fig. 2.16. At any frequency

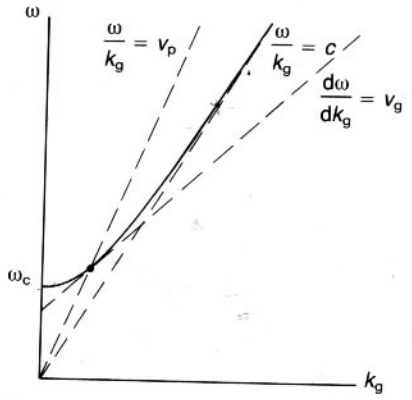

Fig. 2.16 Dispersion diagram of a waveguide.

the phase velocity of the wave is given by the ratio of the frequency to the propagation constant, that is by the slope of the line from a point on the curve to the origin of the graph. Mathematically

$$v_p = \frac{\omega}{k_g} = c\left[1 + \left(\frac{\pi}{k_g a}\right)^2\right]^{\frac{1}{2}}. \tag{2.37}$$

Thus the phase velocity of a wave in an empty waveguide is always greater than the velocity of light and tends to c as k_g tends to infinity. Unlike the TEM transmission lines discussed earlier the phase velocity of a wave varies with frequency. This means that, if an electromagnetic pulse is injected into the guide, the different Fourier components from which it is made up travel with different velocities and the shape of the pulse which emerges from the end of the guide differs from that injected. The pulse is said to be *dispersed* by the guide and the guide itself is described as dispersive. This property is not limited to rectangular waveguides, it occurs in any wave propagating medium in which the phase velocity is a function of frequency. Figure 2.16 is called the dispersion diagram of the waveguide.

Equation (2.37) apparently violates the axiom of the theory of relativity that nothing can travel faster than the velocity of light. There is in fact no contradiction. The axiom is more accurately stated as: 'Information cannot travel faster than the velocity of light.' A continuous sine wave carries no information, to become a carrier of information it must be modulated in some way. The effects of modulation are most easily illustrated by considering the superposition of two signals having the same amplitude and slightly different frequencies. Technically this is double-sideband suppressed-carrier modulation (J.J. O'Reilly, 1984). If the two waves are

$$E_1 = E_0 \exp \mathrm{j}[(\omega + \delta\omega)t - (k + \delta k)z]$$

and

$$E_2 = E_0 \exp \mathrm{j}[(\omega - \delta\omega)t - (k - \delta k)z]. \tag{2.38}$$

These expressions can be rearranged to give

$$E_1 = E_0 \exp \mathrm{j}(\omega t - kz) \exp \mathrm{j}(\delta\omega t - \delta kz)$$

and

$$E_2 = E_0 \exp \mathrm{j}(\omega t - kz) \exp [-\mathrm{j}(\delta\omega t - \delta kz)]. \tag{2.39}$$

When they are added together the result is

$$E = 2E_0 \exp \mathrm{j}(\omega t - kz) \cos(\delta\omega t - \delta kz). \tag{2.40}$$

This expression can be interpreted as a carrier wave which propagates as $\exp \mathrm{j}(\omega t - kz)$ whose amplitude is $2E_0 \cos(\delta\omega t - \delta kz)$. The envelope of the carrier wave carries information at a velocity given by

$$v_g = \lim_{\delta\omega \to 0}\left(\frac{\delta\omega}{\delta k}\right) = \frac{d\omega}{dk}. \tag{2.41}$$

This velocity is known as the *group velocity* of the wave. On the dispersion diagram (Fig. 2.16) it is represented by the slope of the tangent to the dispersion curve at a point. It is evidently always less that the velocity of light and it tends to zero at the cut-off frequency.

By differentiating (2.36) it can be shown that the group velocity in a rectangular waveguide is

$$v_g = c(\lambda_0/\lambda_g) = c \sin \theta \tag{2.42}$$

which is just the z component of the phase velocity of the wave bouncing down the guide.

2.6 POWER FLOW IN A RECTANGULAR WAVEGUIDE

The fields of the TE mode discussed in the previous section are readily derived from equations (2.22) to (2.24) since

$$k_y = \pi/a \tag{2.43}$$

from (2.25) and (2.31), and

$$k_z = k_g. \tag{2.44}$$

So, making use of (2.31) and (2.33)

$$E_x = E_0 \sin\left(\frac{\pi y}{a}\right) \cos(\omega t - k_g z) \tag{2.45}$$

$$H_y = \frac{E_0}{Z_0}\frac{\lambda_0}{\lambda_g} \sin\left(\frac{\pi y}{a}\right) \cos(\omega t - k_g z) \tag{2.46}$$

$$H_z = \frac{E_0}{Z_0}\frac{\lambda_0}{\lambda_c} \cos\left(\frac{\pi y}{a}\right) \sin(\omega t - k_g z). \tag{2.47}$$

From these equations it can be seen that E_x and H_y are in phase with one another whilst H_z has a phase difference from them of 90°. The time average Poynting vector is therefore

$$S_z = \tfrac{1}{2}|E_x|\,|H_y| \tag{2.48}$$

and the average power flow is obtained by integrating S_z across the cross section of the guide

$$W = \frac{1}{2}\int_0^b dx \int_0^a |E_x|\,|H_y|\,dy$$
$$= \frac{ab}{4}\frac{\lambda_0}{\lambda_g}\frac{|E_0|^2}{Z_0}. \tag{2.49}$$

This equation shows that the power flowing in a guide depends upon the electric field strength at the centre of the guide and the cross-sectional

area. In a guide filled with air at atmospheric pressure the electric field is limited by dielectric breakdown. This sets a limit on the power density in the guide. The maximum power handling capabilities of waveguides decrease with the square of the cut-off frequency. Table 2.2 shows the maximum recommended power-handling capabilities of standard waveguides. Higher powers can be dealt with by pressurizing the guide either with air or with a gas such as freon or sulphur hexafluoride. The power which can be transmitted can also be increased by evacuating the guide because that also raises the breakdown field.

By calculating the total stored energy in one wavelength of the wave and comparing it with (2.48) it is possible to show that the velocity of propagation of energy is equal to the group velocity (see Exercise 2.3).

If the signal frequency is below the cut-off frequency then k is imaginary as can be seen from (2.36). The wave no longer propagates down the guide and the fields decay as exp $-k_g z$. Equations (2.45) and (2.46) show that E_x and H_y are in phase quadrature (because λ_g is imaginary) so that the Poynting vector is purely imaginary. There is then no flow of energy but only a reactive storage of it. A cut-off wave of this kind is known as an evanescent wave.

It is often useful to apply transmission line theory to waveguide problems. To do this we require the characteristic impedance of the guide as well as the guide wavelength. The characteristic impedance is defined by

$$Z_g = |V_0|^2/2W, \tag{2.50}$$

where $V_0 = bE_0$ is the magnitude of the potential difference between the centres of the broad walls of the guide. Substituting for the power flow from (2.49) yields

$$Z_g = \frac{2b}{a}\frac{\lambda_g}{\lambda_0} Z_0. \tag{2.51}$$

Note that the characteristic impedance of a waveguide varies with frequency.

The theory of transmission lines is commonly based upon an equivalent circuit for the line. Our discussion of rectangular waveguides has, so far, adopted a field approach. It is, however, possible to represent a waveguide by an equivalent circuit. The derivation of that circuit provides a useful example of the way in which such circuits can be derived for microwave components.

Figure 2.17(a) shows a sectioned view of a waveguide including the electric and magnetic field patterns. To satisfy the boundary conditions there must be charge concentrations and currents flowing in the walls as shown in Fig. 2.17(b). To derive the equivalent circuit of a short length dz of the guide we include shunt capacitance to represent the displacement current paths and shunt and series inductance to represent the conduction current paths. The result is the circuit shown in Fig. 2.18. In order to derive

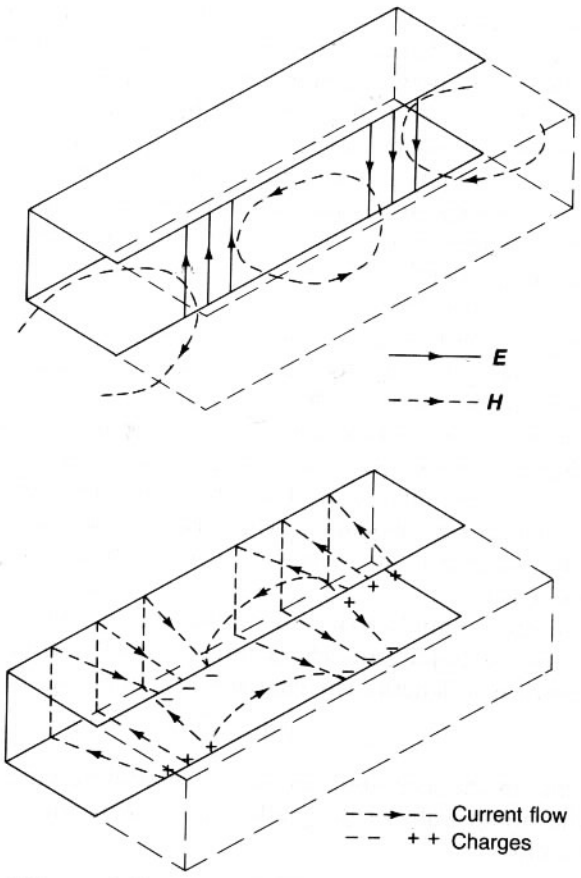

Fig. 2.17 The TE_{01} mode in a rectangular waveguide: (a) Electric and magnetic field patterns, and (b) wall charges and currents.

expressions for the component values in terms of the dimensions of the guide we require three conditions, namely:

1. The equivalent circuit must have the same cut-off frequency as the guide;
2. The phase velocity calculated from the equivalent circuit must tend to the velocity of light in the limit of high frequencies; and
3. The equivalent circuit must have the same characteristic impedance as the guide.

Assuming that the phase change produced by the section is $k\,dz$ the analysis of the circuit shows that

$$k = \left[\frac{L_2}{L_1}(\omega^2 L_1 C - 1)\right]^{\frac{1}{2}}. \tag{2.52}$$

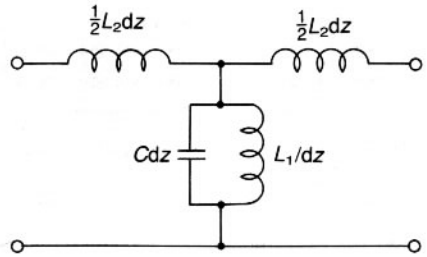

Fig. 2.18 Equivalent circuit for the TE_{01} mode in a rectangular waveguide.

This equation fits exactly the curve shown in Fig. 2.16 provided that

$$L_1 C = 1/\omega_c^2 \quad \text{and} \quad \frac{L_2}{L_1} = \frac{\omega_c^2}{c^2} \tag{2.53}$$

which satisfy conditions 1 and 2. The iterative impedance of the network tends to

$$Z_g = \frac{\omega L_1}{\omega^2 L_1 C - 1} k \tag{2.54}$$

as dz tends to zero. Equating the expressions for Z_g given by (2.51) and (2.54) to satisfy condition 3 gives

$$L_1 = \frac{2ab}{\pi^2}\mu_0 \tag{2.55}$$

and, from (2.53) we obtain

$$C = \frac{a}{2b}\varepsilon_0 \tag{2.56}$$

and

$$L_2 = \frac{2b}{a}\mu_0. \tag{2.57}$$

It is important to note that the dimensions of these expressions are henry metres, farads per metre and henries per metre, respectively. It is also important to note that (2.56) is not equal to the capacitance per unit length between the broad walls of the guide regarded as a parallel plate capacitor. This underlines the point that the fields in a waveguide cannot be obtained by solving Laplace's equation.

2.7 HIGHER-ORDER MODES IN A RECTANGULAR WAVEGUIDE

In Section 2.5 we saw that a rectangular waveguide could be constructed by inserting a conducting sheet along the null plane closest to the conducting

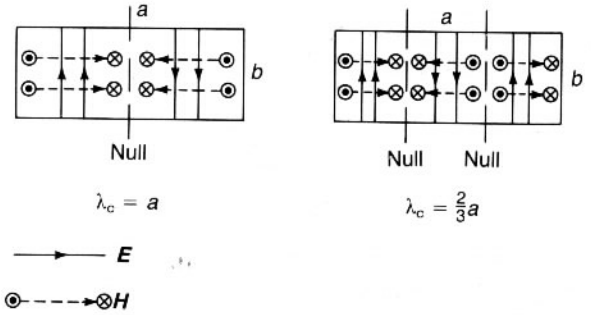

Fig. 2.19 Higher-order modes in rectangular waveguide: (a) TE_{02} and, (b) TE_{03}.

plane in the field pattern shown in Fig. 2.10. We could just as easily have chosen any of the other null planes for this purpose. This would have led to other possible solutions for the propagation of waves in the guide. Figure 2.19 shows the next two such modes with their cut-off wavelengths ($\lambda_c = 2\pi c/\omega_c$). Each has a longitudinal null plane and a transverse electric field described by

$$E_x = E_0 \sin\left(\frac{n\pi y}{a}\right) \cos(\omega t - k_g z), \tag{2.58}$$

where $n = 1, 2, 3$, etc. We could, instead, have chosen to study modes with their electric fields in the y direction so that

$$E_y = E_0 \sin\left(\frac{m\pi x}{b}\right) \cos(\omega t - k_g z), \tag{2.59}$$

where $m = 1, 2, 3$, etc. With such a proliferation of modes it is desirable to have a way of referring to them. The usual notation is to call them TE_{mn} modes where m and n have integer values. On this basis the mode discussed in Sections 2.5 and 2.6 is the TE_{01} mode and those shown in Fig. 2.19 are the TE_{02} and TE_{03} modes. It is natural to ask whether modes can exist for which m and n are both non-zero. A complete solution of the electromagnetic wave equation for transverse electric waveguide modes (Ramo *et al.*, 1965) shows that this is indeed the case. Figure 2.20 shows, for example, the field pattern for the TE_{11} mode. It is useful to be able to sketch field patterns for the different modes because it is then possible to write down the equations which describe them by inspection. Thus from Fig. 2.20 we get

$$E_x = E_1 \cos\left(\frac{\pi x}{b}\right) \sin\left(\frac{\pi y}{a}\right) \cos(\omega t - k_g z) \tag{2.60}$$

and

$$E_y = E_2 \sin\left(\frac{\pi x}{b}\right) \cos\left(\frac{\pi y}{a}\right) \cos(\omega t - k_g z). \tag{2.61}$$

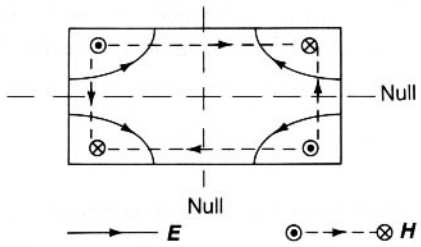

Fig. 2.20 The TE_{11} mode in rectangular waveguide.

This field pattern must satisfy the wave equation (2.4) so that

$$\left(\frac{\pi}{b}\right)^2 + \left(\frac{\pi}{a}\right)^2 + k_g^2 = k_0^2. \tag{2.62}$$

In general for a TE_{mn} mode

$$\left(\frac{m\pi}{b}\right)^2 + \left(\frac{n\pi}{a}\right)^2 + k_g^2 = k_0^2 \tag{2.63}$$

so that all the modes obey the equation

$$k_g^2 = k_0^2 - k_c^2, \tag{2.64}$$

where

$$k_c^2 = \left(\frac{2\pi}{\lambda_c}\right)^2 = \left(\frac{m\pi}{b}\right)^2 + \left(\frac{n\pi}{a}\right)^2. \tag{2.65}$$

Equation (2.64) is a generalization of (2.35). It can be seen from (2.65) that λ_c decreases and therefore the cut-off frequency increases as m and n increase. The mode which has the lowest cut-off frequency of all is the TE_{01} mode discussed in Sections 2.5 and 2.6. Table 2.3 shows the cut-off frequencies for some of the lower modes in standard WG16 waveguide.

Table 2.3 Cut-off frequencies of TE modes in waveguide WG16

Mode		Cut-off frequency
m	n	(GHz)
0	1	6.55
0	2	13.10
1	0	14.71
1	1	16.10
0	3	19.65
1	2	19.69
1	3	24.54
2	1	30.13
2	2	32.20

The relationship between the amplitudes of the x and y components of the electric field can be obtained by making use of Maxwell's equation (1.5) which is

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} = 0 \tag{2.66}$$

since there is no free charge within the waveguide and $E_z = 0$. Substitution from (2.60) and (2.61) yields

$$\left[\frac{\pi E_1}{b} + \frac{\pi E_2}{a}\right]\left[\sin\left(\frac{\pi x}{b}\right)\sin\left(\frac{\pi y}{a}\right)\cos\left(\omega t - k_g z\right)\right] = 0 \tag{2.67}$$

so that

$$E_2 = -\frac{a}{b}E_1. \tag{2.68}$$

A little thought shows that this result is equivalent to saying that the electric field lines must all start and finish on the walls of the guide. This means that the amplitudes of the fields at the centres of the walls must be in inverse ratio to the widths of the walls.

The magnetic field pattern of a mode can be obtained in a similar way by substituting the electric field into Maxwell's equation (1.8). For the TE_{11} mode the result is

$$-j\omega\boldsymbol{B} = \begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ E_x & E_y & 0 \end{vmatrix}$$

$$= \left[-\frac{\partial E_y}{\partial z}\right]\hat{\boldsymbol{x}} + \left[\frac{\partial E_x}{\partial z}\right]\hat{\boldsymbol{y}} + \left[\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial z}\right]\hat{\boldsymbol{z}}. \tag{2.69}$$

Substituting the expressions for the electric fields gives

$$H_x = \frac{jk_g}{\omega\mu_0}\frac{a}{b}E_1\sin\left(\frac{\pi x}{b}\right)\cos\left(\frac{\pi y}{a}\right)\sin\left(\omega t - k_g z\right) \tag{2.70}$$

$$H_y = \frac{jk_g}{\omega\mu_0}E_1\cos\left(\frac{\pi x}{b}\right)\sin\left(\frac{\pi y}{a}\right)\sin\left(\omega t - k_g z\right) \tag{2.71}$$

and

$$H_z = \frac{j}{\omega\mu_0}\left[\frac{\pi E_1}{a} - \frac{\pi a E_1}{b^2}\right]\cos\left(\frac{\pi x}{b}\right)\cos\left(\frac{\pi y}{a}\right)\cos\left(\omega t - k_g z\right). \tag{2.72}$$

These expressions can be derived by solving the wave equation (2.3) subject to the appropriate boundary conditions but the approach given here has the advantage of being tied more closely to an intuitive understanding of electromagnetic phenomena.

So far we have only considered transverse electric modes. Section 2.4 demonstrated the existence of transverse magnetic modes when plane waves are incident obliquely on a conducting plane. Can such waves pro-

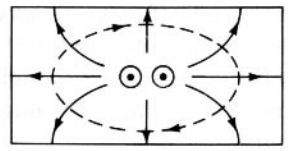

Fig. 2.21 The TM_{11} mode in rectangular waveguide.

pagate in a waveguide? The answer is that they can. Like the TE waves they are actually the result of the superposition of plane TEM waves. Unfortunately there is no simple construction like that used for TE waves by which the TM mode patterns can be deduced. The problem is that although a conducting sheet can be introduced parallel to the conducting plane in Fig. 2.14 it is not possible simply to add a further pair of planes at right angles to the first. This is because the tangential component of the electric field cannot be zero on them.

The full solution to the wave equation for the TM modes can be found in standard texts (Ramo *et al.*, 1965). Here we shall proceed by sketching the field patterns and observing that the fields can be deduced from them as before. Figure 2.21 shows the field pattern for the TM_{11} mode. This is, in fact, the lowest frequency TM mode because modes with m or $n = 0$ cannot satisfy the boundary conditions. It is easy to show that the TM modes must also satisfy equations (2.64) and (2.65) so that their behaviour is just like that of the TE modes.

We are now in a position to explain the operating bands quoted for standard waveguides in Table 2.2. Operation in the TE_{01} mode ensures that that is the only mode which can propagate; all the others are cut off. As the frequency approaches the cut-off frequency the guide becomes increasingly dispersive. From equation (2.42) we find that the group velocity is

$$v_g = [1 - (\omega_c/\omega)^2]^{\frac{1}{2}}c. \tag{2.73}$$

At the bottom of the frequency band quoted for a waveguide the group velocity is 60% of the velocity of light.

The upper frequency limit is set by the need to ensure that all the higher order modes are well beyond cut-off. Table 2.3 shows that cut-off frequency of the TE_{02} mode in waveguide 16 is a factor of 1.05 above the highest working frequency of the guide. From equation (2.64) we find that the propagation constant is $82j\,m^{-1}$. At the same frequency the propagation constant of the TE_{01} mode is $223\,m^{-1}$. The attenuation of the evanescent TE_{02} mode signal in one wavelength of the TE_{01} mode is 46 dB.

The useful bandwidth of a rectangular waveguide is therefore governed

by the separation between the cut-off frequencies of the lowest and next to lowest modes of propagation. The aspect ratio of 2 : 1 which is commonly used makes the TE_{10} cut-off frequency a little higher than that of the TE_{02} mode. It is therefore close to the ratio of dimensions which gives the maximum possible bandwidth.

2.8 OTHER WAVEGUIDES

Although all the discussion so far in this chapter has been about rectangular waveguides it is obvious that any conducting pipe can act as a waveguide with properties similar to those of the rectangular guides. Two which are commonly encountered are the circular and ridge waveguides.

The circular waveguide has the special property that no plane containing the axis is distinguishable from any other. As a result different polarizations of a TE mode have the same cut-off and guide wavelengths. This property is employed in the rotating joints used in the waveguides feeding radar antennas. It is also used to make accurate attenuators and phase shifters as we shall see in Chapter 4.

The analysis of modes in circular waveguide requires the use of cylindrical polar coordinates (see Appendix B). The suffixes which describe the modes of propagation refer to the number of nodes in the tangential (θ) and radial (r) directions, respectively. Figure 2.22 shows a few of the modes which exist in a cylindrical waveguide. Details of the fields and the cut-off wavelengths for the different modes can be found in Ramo *et al.* (1965).

In the previous section we saw that the useful bandwidth of a rectangular waveguide is limited to around 1.5 : 1. For some purposes this is inconvenient. In order to increase the useful bandwidth we need to increase the

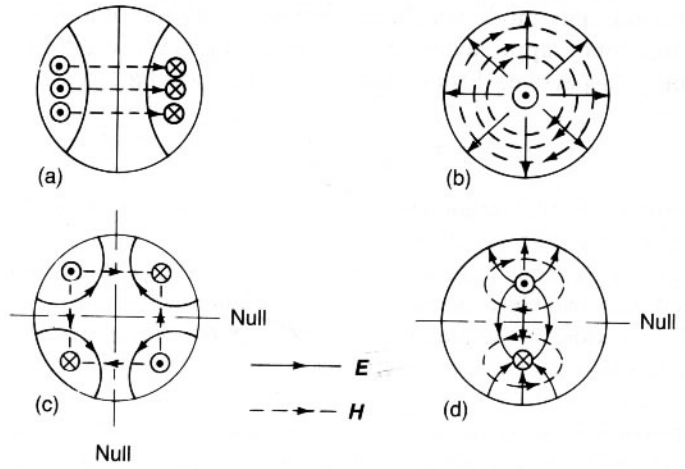

Fig. 2.22 Modes in circular waveguide: (a) TE_{11}, (b) TM_{01}, (c) TE_{21} and, (d) TM_{11}.

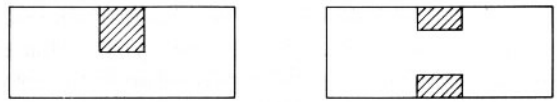

Fig. 2.23 Single- and double-ridge waveguides.

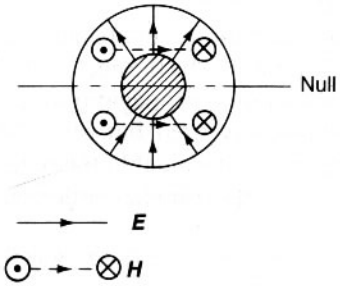

Fig. 2.24 The TE_{11} waveguide mode in a coaxial line.

separation between the cut-off frequencies of the TE_{01} and TE_{02} modes. The equivalent circuit shown in Fig. 2.18 suggests a possible way. If a ridge is added to the centre of one, or both, of the broad walls of the guide, as shown in Fig. 2.23, then the effect on the TE_{01} mode will be to increase both the shunt capacitance and the shunt inductance shown in Fig. 2.18. The consequence is that the cut-off frequency of this mode is lowered. The TE_{02} mode, however, will be little affected because the ridges are in regions of weak electric field. The price paid for the increase in bandwidth is a reduction in power handling capability. At low power levels two-wire lines provide a more compact means of broadband signal transmission, but ridge waveguides are useful where a combination of broad-bandwidth and moderately high power is required. The notation for modes in ridge waveguide is the same as that used for rectangular waveguides.

It is not always appreciated that two-wire lines can also support higher-order modes of propagation. Figure 2.24 shows the TE_{11} mode for a coaxial line. Like any other waveguide mode this has a lower cut-off frequency and a guide wavelength which obeys equation (2.64). The useful bandwidth of a coaxial line is much greater than that of a waveguide because the lower cut-off frequency is zero. The presence of higher-order modes provides a limit on the highest frequency for which a line can be used.

2.9 CONCLUSION

In this chapter we have seen how electromagnetic waves are guided by conducting surfaces. The TEM waves which propagate in free space combine when reflected off a conductor to give waves which can be classified as

either TE or TM. Waves of both these types can propagate down metallic tubes at frequencies above some cut-off frequency. The phase velocities of these waves vary with frequency so that pulses formed of groups of waves of different frequencies are dispersed as they travel down the waveguide. The phase velocity is always greater than the velocity of light but it is found that information and power propagate with the group velocity which is less than the velocity of light.

Waveguides can support an infinite set of higher-order modes. The useful bandwidth of a waveguide is limited by the need to avoid excessive dispersion on the one hand and multi-mode propagation on the other. Rectangular waveguides are most commonly used but other cross sections including circular and ridge waveguides are valuable for special purposes. Two-wire transmission lines can also support higher-order modes.

EXERCISES

2.1 Calculate the dimensions of the following 50 Ω transmission lines:

1. Coaxial cable with polythene dielectric ($\varepsilon_r = 2.25$) and an inner conductor diameter of 1.5 mm.
2. An air-spaced parallel-wire line with a conductor diameter of 5 mm.
3. A triplate line with alumina dielectric ($\varepsilon_r = 8.9$) and centre conductor width 2.5 mm.

2.2 Calculate the cut-off wavelength of the WG10 waveguide and the guide wavelength at 2.5, 3.0, 3.5 and 4.0 GHz. Plot a dispersion diagram for the waveguide and compute the phase and group velocities at 3.0 GHz.

2.3 Find an expression for the stored energy per wavelength in a rectangular waveguide and show, by comparison with (2.49) that energy propagates down the guide at the group velocity.

2.4 Calculate the cut-off frequencies of the TE_{02}, TE_{11} and TM_{11} modes in a WG10 waveguide.

2.5 Calculate the characteristic impedances of waveguides whose heights are 2.0, 4.0, 6.0 and 7.9 mm and whose width is 15.8 mm.

2.6 Calculate the cut-off frequency and characteristic impedance of WG12 waveguide when it is filled with: 1. air at atmospheric pressure and, 2. paraffin wax ($\varepsilon_r = 2.25$).