Copyright 1990 - R.G. Carter

6 Coupling between wave-guiding systems

6.1 INTRODUCTION

We have already seen that transmission lines and waveguides can support a variety of modes of propagation. Many important microwave devices incorporate transitions between regions with different propagation characteristics. In this chapter we shall examine the basic principles underlying coupling between these different regions and show how they are related to the theory and design of microwave components. We shall see that the coupling can often be thought of in terms of either the electric field or the magnetic field. For convenience all kinds of waveguiding structures will be referred to as waveguides. Such a huge variety of examples of coupling exists that it will be necessary to study only a representative sample and concentrate on the fundamental principles in operation. The various devices are usually represented by equivalent circuits and these are the basis of design calculations. Many of the equivalent circuits have been derived empirically and throw no light on the basic physical principles at work. For this reason a detailed discussion is not given here and the reader is referred to the books by Marcuvitz (1986), and Edwards (1981) for further details.

6.2 DISCONTINUITIES

All couplings between waveguides involve some kind of discontinuity in the waveguiding structure. It is therefore useful to begin by establishing what the effects of these discontinuities are.

Figure 6.1(a) shows a diaphragm which partially obstructs the width of a waveguide. At the plane of the diaphragm the fields must satisfy the boundary conditions given in Section 1.9. The field distribution at the plane of the diaphragm is a compressed TE_{01} mode as shown in Fig. 6.1(b) with $E_x = 0$ outside the gap. The field in the waveguide adjacent to the diaphragm must match that at the diaphragm in every respect. The field of

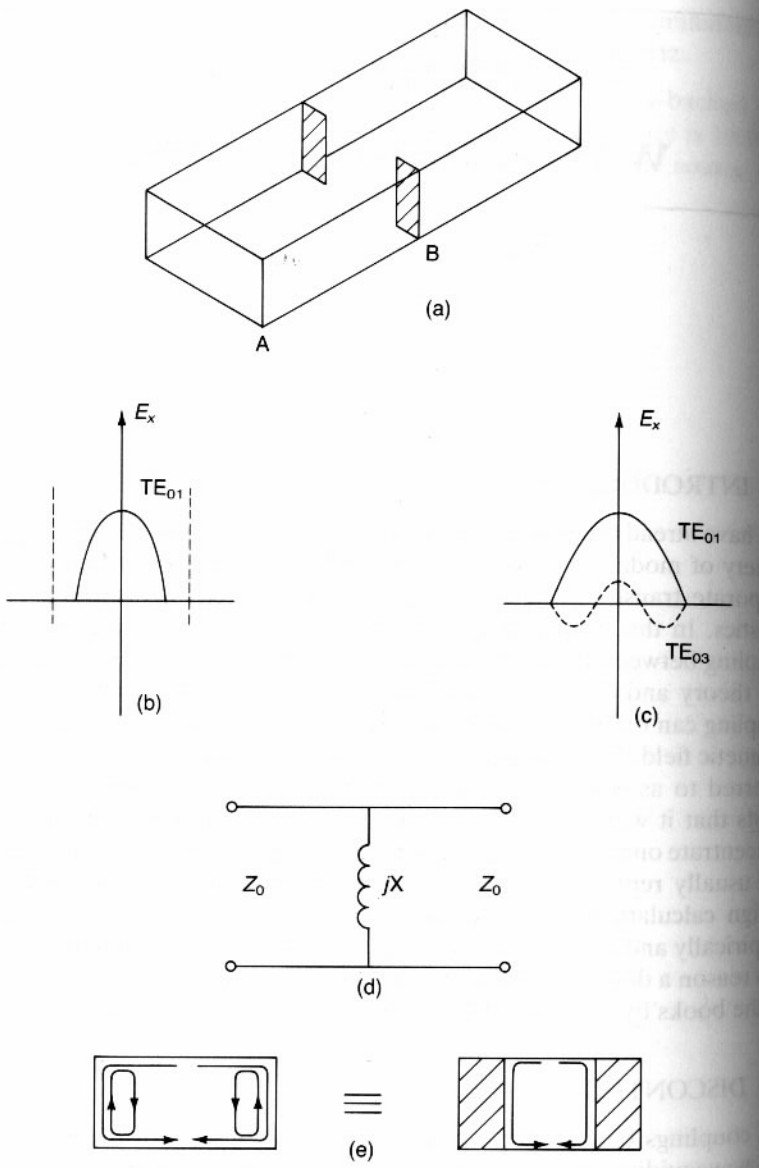

Fig. 6.1 An inductive waveguide iris: (a) general arrangement, (b) the field within the iris, (c) superposition of the waveguide TE_{01} mode and higher-order modes to satisfy the boundary conditions, (d) equivalent circuit, and (e) showing how the current in the plane of the iris can be represented by the superposition of inductive current loops on the current flow in the walls of the waveguide.

the incident TE_{01} wave shown in Fig. 6.1(c) plainly does not obey this requirement. In order to match the boundary conditions correctly we recall that an infinite set of higher-order modes can exist in the guide. The boundary conditions can thus be satisfied by a superposition of all these modes with appropriate amplitudes. Figure 6.1(c) shows the TE_{03} mode in the waveguide. Clearly the addition of the field of this mode to that of the TE_{01} mode produces a field distribution which is more nearly the same as that at B. In this case the amplitude of the TE_{02} would be zero because of the symmetry of the problem.

We have, therefore, established the basic principle that any discontinuity in a waveguide results in the transfer of energy to higher-order modes. Now the working frequency bands of waveguides are normally chosen so that the higher-order modes are cut off. The power coupled into them, therefore, does not propagate but is stored in evanescent waves close to the discontinuity. For many purposes the effect of the discontinuity can be represented as a lumped reactance at its plane. In the case of the diaphragm shown in Fig. 6.1(a) the effect of the discontinuity is to alter the paths of the conduction currents in the walls of the waveguide. Figure 6.1(e) shows how the effect of the diaphragm in altering the current paths can be represented as the superposition of a pair of current loops on the current flow in the unmodified waveguide. The diaphragm therefore introduces a shunt inductance and the discontinuity can be represented by the equivalent circuit shown in Fig. 6.1(d).

If the diaphragm changed the height of the guide, instead of its width, as shown in Fig. 6.2(a) a different set of higher-order modes would be excited. Once again these would normally be cut off. The discontinuity affects the electric field distribution so the shunt reactance is capacitive. An alternative way of introducing a shunt capacitance is to put a screw in the broad wall of the waveguide as shown in Fig. 6.2(b). This arrangement is useful for matching purposes because the capacitance is easily adjusted. An important special case arises when the diaphragm constricts both the height and the width of the guide as shown in Fig. 6.2(c). This arrangement presents a parallel combination of inductance and capacitance to the incident wave. The equivalent circuit is shown in Fig. 6.2(d). The arrangement is known as an iris and its resonant properties make it an important component in certain types of waveguide filter.

A comprehensive discussion of discontinuities in microstrip is given by Edwards (1981). Two are considered here to illustrate other ways in which discontinuities can be represented by equivalent circuits. We have already noted that it is easier to make open circuits than short circuits in microstrip. Thus matching stubs are normally open circuited at their free ends. The calculation of the lengths of such stubs is an important part of circuit design. Figure 6.3(a) shows the electric field in the region of an open circuit termination. It is clear that the field does not stop abruptly at the end of the

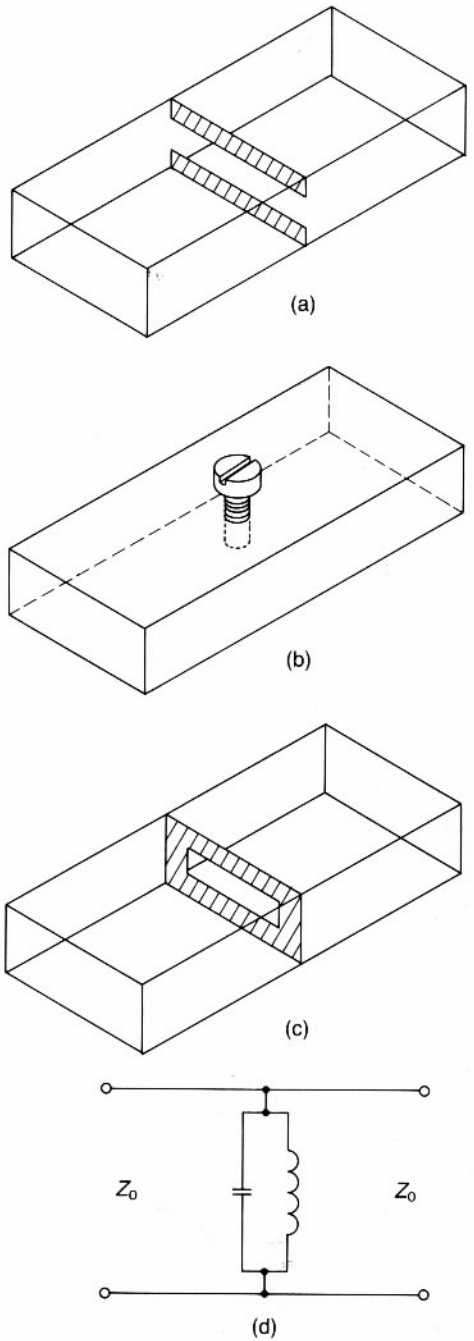

Fig. 6.2 Waveguide obstacles: (a) capacitative iris, (b) capacitative tuning screw, (c) resonant iris, and (d) equivalent circuit for the resonant iris.

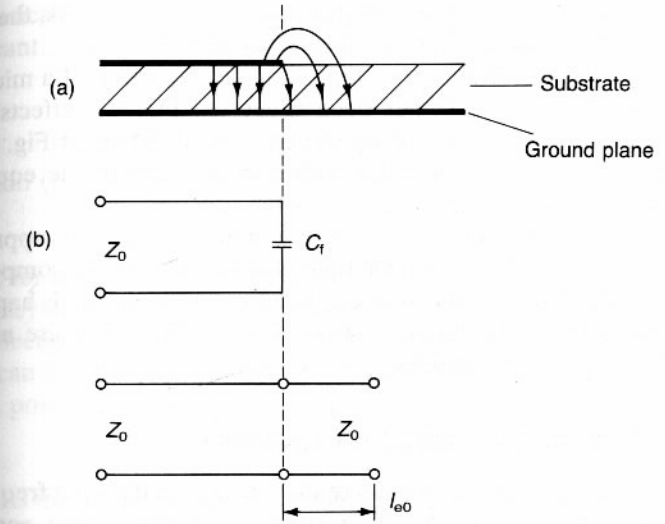

Fig. 6.3 Microstrip open circuit: (a) electric fringing field, (b) equivalent capacitance, and (c) equivalent transmission line.

line but fringes into the region beyond. This fringing field can be allowed for by a capacitor as shown in Fig. 6.3(b). The capacitance can be calculated accurately enough from a static field solution. Edwards quotes a formula for C_f.

An alternative way of looking at this effect is to consider that a short length of open-circuited transmission line will look like a capacitance. Thus the fringing can be represented by an additional length l_{e0} of the same

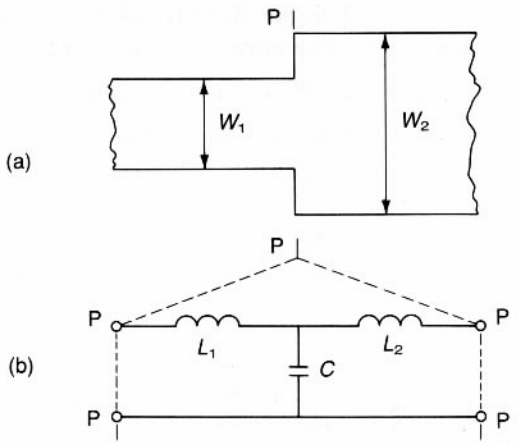

Fig. 6.4 (a) Step change in the width of a microstrip line, and (b) an equivalent circuit for the step.

transmission line as shown in Fig. 6.3(c). In design calculations, therefore, the line should be shorter than the theoretical length by l_{e0}.

As a final example consider the step change in the width of a microstrip line shown in Fig. 6.4(a). It has been suggested that the effects of the discontinuity can be represented by the equivalent circuit of Fig. 6.4(b). Here there is no clear physical link between the form of the equivalent circuit and the field patterns occurring at the junction.

The three examples given illustrate between them the various approaches which are used to produce equivalent circuits for microwave components. The first two are based upon a physical understanding of what is happening at the discontinuity. The last is a network whose elements are adjusted empirically to give useful results.

6.3 BROADBAND MATCHING TECHNIQUES

Mismatches in transmission lines can readily be matched at spot frequencies by arranging other mismatches to provide a reflected wave equal and opposite to the one to be eliminated. The matching elements take the form of shunt stubs in coaxial line and microstrip and inductive or capacitive irises in waveguide. A full description of the use of the Smith chart to design matching networks of this kind can be found in Dunlop and Smith (1984).

For many purposes this kind of narrow-band matching is not satisfactory and some way of producing a broad-band match is needed. To show how this can be achieved consider Fig. 6.5 which shows a transmission line with a matched termination which has four equally spaced reflecting objects with reflection coefficients ϱ_1 and ϱ_2 as shown. The electrical length of the line between the objects is ϕ. We will assume that the reflection coefficients are small so that the amplitude of the incident wave is hardly affected by the reflections. Now suppose that D is the obstacle to be matched and A, B and C are the matching elements. The apparent reflection coefficient at A is

$$\begin{aligned}\varrho &= \varrho_1 + \varrho_2 e^{-2j\phi} + \varrho_2 e^{-4j\phi} + \varrho_1 e^{-6j\phi} \\ &= 2\varrho_1 e^{-3j\phi}\cos 3\phi + 2\varrho_2 e^{-3j\phi}\cos\phi. \end{aligned} \tag{6.1}$$

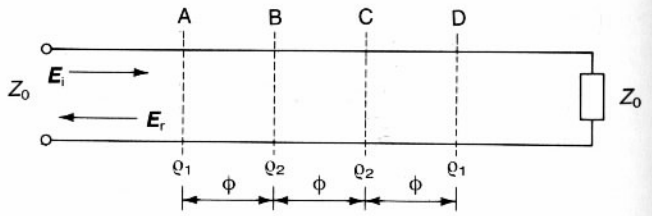

Fig. 6.5 Arrangement of four regularly spaced discontinuities on a transmission line.

In practical terms it is usually the magnitude of the reflection which is significant, that is

$$|\varrho| = |8\varrho_1 \cos^3\phi + 2(\varrho_2 - 3\varrho_1)\cos\phi| \tag{6.2}$$

since $\cos 3\phi = 4\cos^3\phi - 3\cos\phi.$

Equation (6.2) may be written

$$|\varrho| = |8\varrho_1 x^3 + 2(\varrho_2 - 3\varrho_1)x|, \tag{6.3}$$

where, for convenience, we have written $x = \cos\phi$. Now ϕ is a function of frequency and, therefore, so is x. Equation (6.3) describes the variation of the reflection coefficient with frequency. Clearly the shape of this cubic curve can be changed by making different choices of ϱ_2.

One possibility is to make $\varrho_2 = 3\varrho_1$. The result is

$$|\varrho| = |8\varrho_1 x^3| \tag{6.4}$$

shown as the continuous curve in Fig. 6.6. A polynomial for which only the coefficient of the highest power of x is non-zero is known as a maximally flat polynomial. It has the property that the highest possible number of derivatives of the function are zero at the origin and that the curve is therefore as flat as possible close to the origin. The match having this frequency dependence is called a maximally flat or Butterworth response. It is the best possible match over a narrow band using a given number of matching elements.

For a given maximum reflection coefficient ϱ_m the bandwidth is the distance between the two points on the curve at which $\varrho = \varrho_m$ so that, from (6.4)

$$B_B = 2x = 2\left(\frac{\varrho_m}{8\varrho_1}\right)^{\frac{1}{3}} = \left(\frac{\varrho_m}{\varrho_1}\right)^{\frac{1}{3}}. \tag{6.5}$$

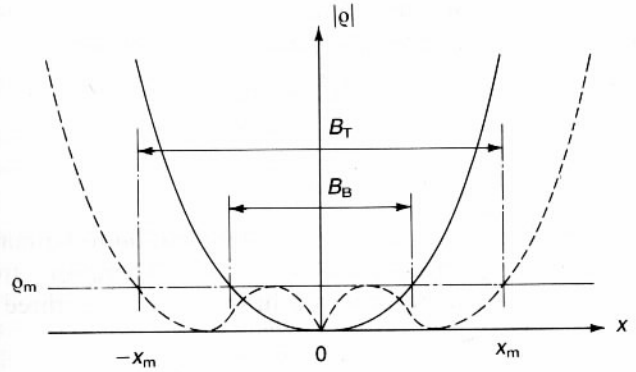

Fig. 6.6 Third-order Butterworth (maximally flat) and Tchebychev (equal-ripple) reflection characteristics.

This demonstrates that, as one would expect, it is possible to achieve a greater bandwidth by accepting a poorer match.

For broad-band matching it is better to sacrifice some of the excellence of the match close to the origin in order to gain greater bandwidth. The broken curve in Fig. 6.6 shows such a possibility. There are, of course, an infinite number of possible polynomials but of special interest are the Tchebychev polynomials. The first six polynominals are listed in Table 6.1.

Table 6.1 Tchebychev polynomials

n	$T_n(x)$
0	1
1	x
2	$2x^2 - 1$
3	$4x^3 - 3x$
4	$8x^4 - 8x^2 + 1$
5	$16x^5 - 20x^3 + 5x$

These polynomials have the property that over the range $-1 < x < 1$ the magnitude is always less than or equal to 1. Of all possible polynomials of a given order and specified in-band ripple the Tchebychev polynomial has the broadest bandwidth.

To design a broad-band match using Tchebychev polynomials we introduce the scaling factors ϱ_m and x_m. The third-order polynomial can then be written

$$\left(\frac{\varrho}{\varrho_m}\right) = 4\left(\frac{x}{x_m}\right)^3 - 3\left(\frac{x}{x_m}\right). \tag{6.6}$$

The response curve has an in-band ripple equal to ϱ_m and the band edges are at $x = \pm x_m$ as shown in Fig. 6.6.

Equating coefficients of x^3 in (6.3) and (6.6) we obtain

$$4\varrho_m/x_m^3 = 8\varrho_1 \tag{6.7}$$

so that

$$x_m^3 = \varrho_m/2\varrho_1. \tag{6.8}$$

Now ϱ_1 is the reflection coefficient of the termination to be matched and ϱ_m is the maximum acceptable in-band match. Therefore, from (6.8), the bandwidth which can be achieved in this way using three matching elements is

$$B_T = 2x_m = 2\left(\frac{\varrho_m}{2\varrho_1}\right)^{\frac{1}{3}} = 2^{\frac{2}{3}}\left(\frac{\varrho_m}{\varrho_1}\right)^{\frac{1}{3}}. \tag{6.9}$$

Comparing (6.9) with (6.5) we see that the bandwidth of the Tchebychev match is wider than that of the Butterworth match by a factor of $2^{2/3} = 1.59$ when three matching elements are used.

To complete the design of the Tchebychev matching network we equate the coefficients of x in (6.3) and (6.6) to give

$$2(\varrho_2 - 3\varrho_1) = -3\varrho_m/x_m \tag{6.10}$$

so that

$$\varrho_2 = 3\varrho_1 - 3\varrho_m/2x_m. \tag{6.11}$$

Since ϱ_1, ϱ_m and x_m are all known ϱ_2 can be calculated.

Example

Design third-order Butterworth and Tchebychev matching networks to give a matched reflection coefficient of 0.001 for an obstacle which has a reflection coefficient of 0.01.

Solution

For a Butterworth match the value of ϱ_2 is 0.03 and the bandwidth is 0.47 from (6.5).

For a Tchebychev match $x_m = 0.368$, the bandwidth is 0.74 from (6.9) and the value of ϱ_2 is 0.0259 from (6.11).

It must be remembered that the theory developed above assumes that all the reflection coefficients are small. It can illustrate the principles of broad-band matching but cannot be applied to the problem of designing matching networks for large mismatches. However methods have been developed for designing both Butterworth and Tchebychev matching networks for large mismatches. (Matthaei *et al.*, 1964)

6.4 COUPLING WITHOUT CHANGE OF MODE

Microwave circuits involve bends and junctions in the waveguides which introduce discontinuities into the system. Some of these junctions do not involve any change in the mode of propagation. These are discussed here and those in which a change of mode is involved are discussed in the next section.

One of the simplest waveguide components is the bend. Provided that the bend is gradual the mismatch is slight and no problems arise. This is normally the case with coaxial lines. Sometimes, however, it is necessary to arrange a change of direction in rather a small space and the design of the bend then becomes important.

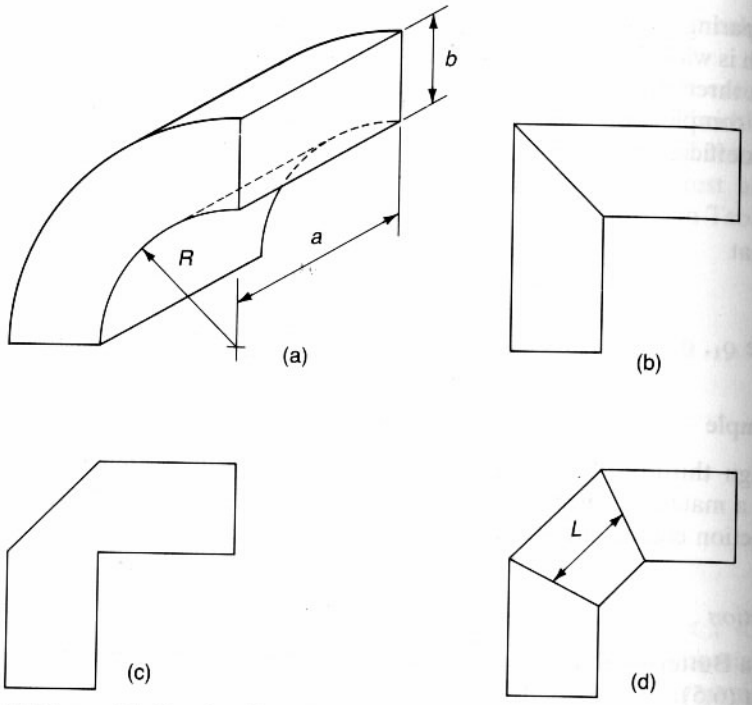

Fig. 6.7 Waveguide bends: (a) radiussed bend, (b) mitred bend, (c) chamfered bend, and (d) double mitred bend.

Waveguide bends are often made in the form of circular arcs as shown in Fig. 6.7(a). The direction of the bend can be either in the ***E*** plane or the ***H*** plane. Southworth (see Harvey, 1963) has shown that the minimum radius for a satisfactory match is 1.5 times the width of the guide in the plane of the bend. Thus for the bend shown in Fig. 6.7(a) the minimum value of R is $1.5b$. Very accurate manufacturing is necessary to avoid reflections and tight bends are usually electro-formed. More gentle bends can be made by bending a straight waveguide using special equipment.

Radiussed waveguide bends are expensive so it is sometimes better to use fabricated mitred bends instead. Figure 6.7(b) shows a simple right-angle bend. Clearly the field patterns within the bend do not match those in the connecting waveguides so an appreciable reactive mismatch can be expected. An improvement (Fig. 6.7(c)) would be to chamfer the outside of the bend to make the effective height of the waveguide the same as that of the connecting guides. Better still if space allows is to use a double-mitred bend as shown in Fig. 6.7d. If the separation L between the two joints is made equal to an odd number of half wavelengths then the mismatches caused by the junctions cancel each other out at one frequency

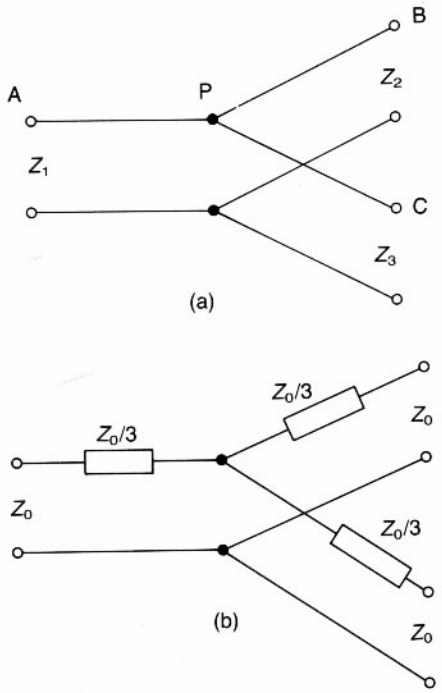

Fig. 6.8 Transmission line tee junctions: (a) simple tee, and (b) a matched tee.

and produce near cancellation over a useful band of frequencies. Similar techniques are used in microstrip.

At low frequencies we are accustomed to join three or more wires together at a point without thinking about the details of what we are doing. At microwave frequencies more care is needed. Figure 6.8(a) shows three transmission lines meeting at a point. If the characteristic impedances of the lines are as shown, A is the input port, and B and C are matched then the impedance presented at P is Z_2 in parallel with Z_3. Very often the impedances of the three lines will be the same and there is then clearly an appreciable mismatch at P. One solution to the problem is to incorporate a matching network of some kind between A and P. This approach leaves B and C incorrectly matched. If it is essential that all the ports are matched then the solution is to add series resistors as shown in Fig. 6.8(b). A little thought shows that this arrangement is correctly matched at all the ports but at the expense of dissipating half the input power within the device.

The networks shown in Fig. 6.8 are idealized, in practice we must take account of the discontinuity in the lines at the junction. Figure 6.9(a) shows a microstrip T junction. By analogy with the step change in width shown in Fig. 6.4(a) we expect that the equivalent circuit will need to have a shunt

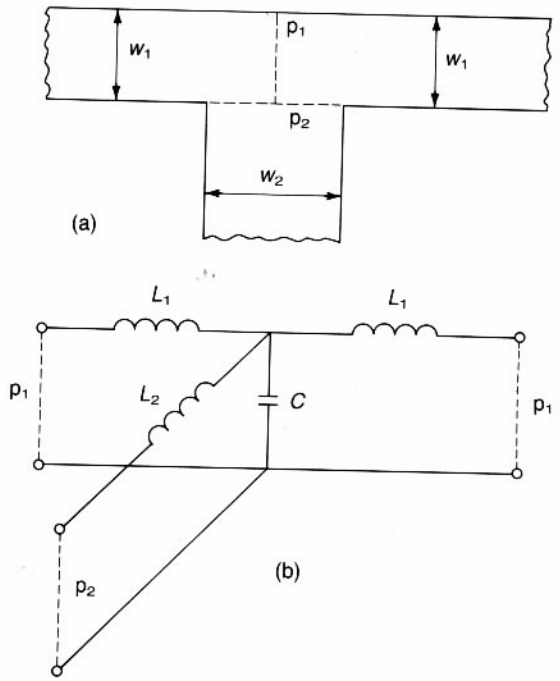

Fig. 6.9 Microstrip tee junction: (a) arrangement of the junction, and (b) equivalent circuit.

capacitance at the junction and series inductances in the three arms. The result is shown in Fig. 6.9(b). One important matter is the definition of the planes on the microstrip which are represented by the terminals of the equivalent circuit. For the main line it is natural to select the plane of symmetry (p_1) whilst for the side arm a plane such as p_2 might be used. It turns out that at frequencies above about 1 GHz it is necessary to add a transformer in the side arm whose turns ratio is a function of frequency. Similar equivalent circuits apply to other types of waveguide.

As a final example of this kind of coupling consider the four-port waveguide junction shown in Fig. 6.10(a). This arrangement combines an ***E***-plane tee with an ***H***-plane tee and is known as a hybrid tee or, sometimes, a magic tee. It has the property that any power fed in at port 1 is divided equally between ports 2 and 3, and any power fed in at port 4 is similarly divided. There is a very high isolation between ports 1 and 4. The reason why it works in this fashion can be understood from sketches of the fields for the two possible excitations. Figure 6.10(b) shows the field patterns when port 1 is the input. The signals emerge from ports 2 and 3 in phase with each other. The electric field at the entrance to the fourth arm is attempting to set up a TM_{11} wave, but this is cut off so the result is a

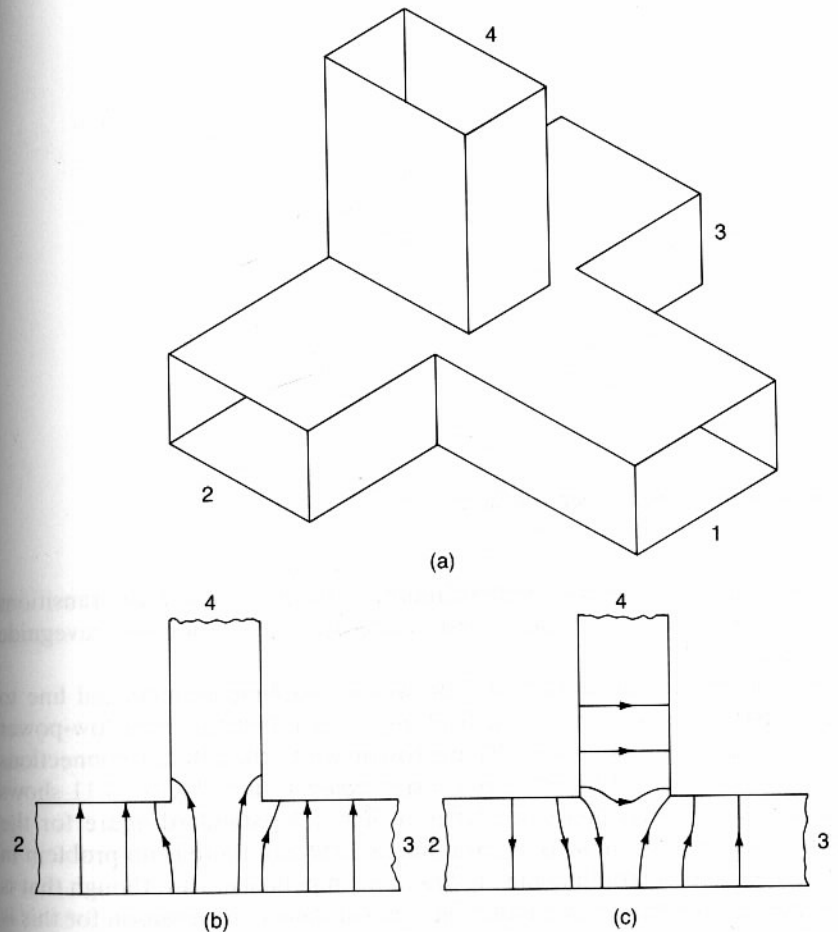

Fig. 6.10 Waveguide hybrid tee junction ('magic tee'): (a) arrangement of the junction, (b) electric fields when the input is at port 1, and (c) electric fields when the input is at port 4.

reactive loading of the junction. In the same way Fig. 6.10(c) shows the field patterns when port 4 is the input. This time the signals emerging from ports 2 and 3 are in antiphase and the wave excited in the remaining arm is the cut-off TE_{02} mode. The reactive elements of the junction are matched by a capacitive boss at its centre. It is difficult to match this device over a wide band.

6.5 COUPLING WITH CHANGE OF MODE

In the previous section we considered examples of coupling between waveguides which were propagating the same mode. Very often it is necessary

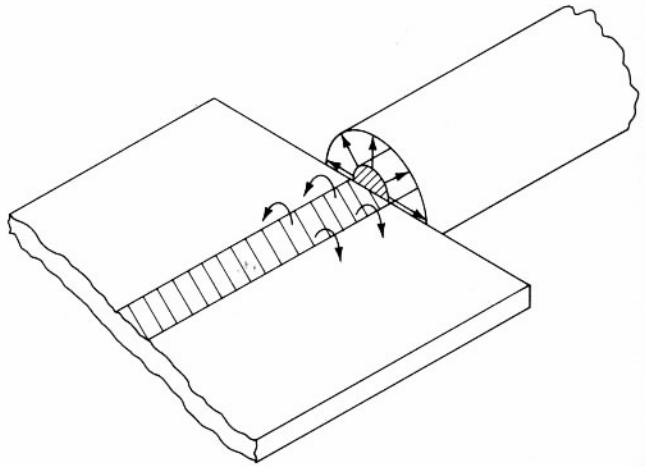

Fig. 6.11 Junction between coaxial line and microstrip.

to couple power from one mode to another. Examples of this are transitions from coaxial line to waveguide and to microstrip and from one waveguide mode to another.

The most straightforward of these is the coupling from coaxial line to microstrip. This is of great practical importance because most low-power microwave circuits are realized in microstrip whilst the external connections between them are normally made using coaxial lines. Figure 6.11 shows a junction between these two types of line. The standard figure for the characteristic impedance of coaxial line is 50 Ω and there is no problem in designing a microstrip line having the same impedance. But, though that is a necessary condition for a match, it is not sufficient. The reason for this is clear from Fig. 6.11 which shows the difference in the electric field patterns of the two modes. It is therefore necessary to provide some matching elements to compensate for the reactance of the junction.

In many other cases there is a very poor match between the field patterns of the two modes which are to be coupled. When that is so the approach is to try to find some way of matching the field patterns as nearly as possible. In that way a high proportion of the power is coupled into the desired mode and relatively little into other modes. An example of this is the transition from the TE_{01} rectangular waveguide mode to the TM_{01} circular waveguide mode required to make the rotating waveguide joints used to feed power to radar antennas. Figure 6.12(a) shows the field patterns in the two guides. It is immediately evident that the magnetic field patterns are similar if the guides are arranged at right angles to each other as shown. The electric fields are not so well matched because the removal of part of the broad wall of the rectangular guide means that the field lines have to be

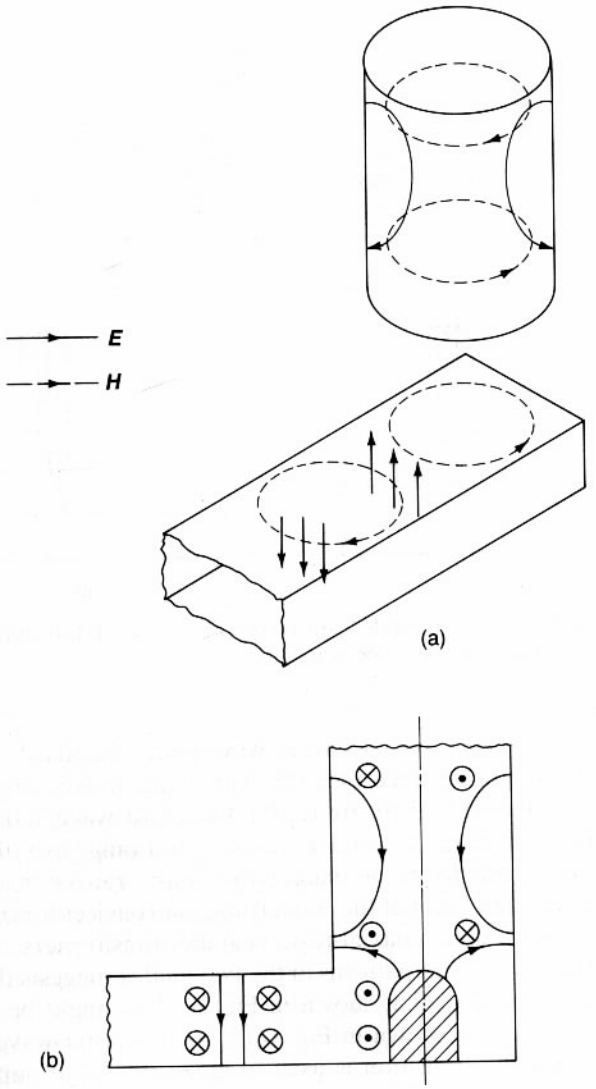

Fig. 6.12 TE_{10} rectangular waveguide to TM_{01} circular waveguide junction: (a) arrangements of the fields in the waveguides, and (b) a cross-sectional view of a 'door-knob' transition showing the field patterns.

radically redistributed. The solution is to put a boss (A) at the centre of the transition as shown in Fig. 6.12(b). Additional irises may be needed to match the transition completely over a band of frequencies. This transition is known as a 'door-knob' transition.

As a final example we will consider the problems of making a good

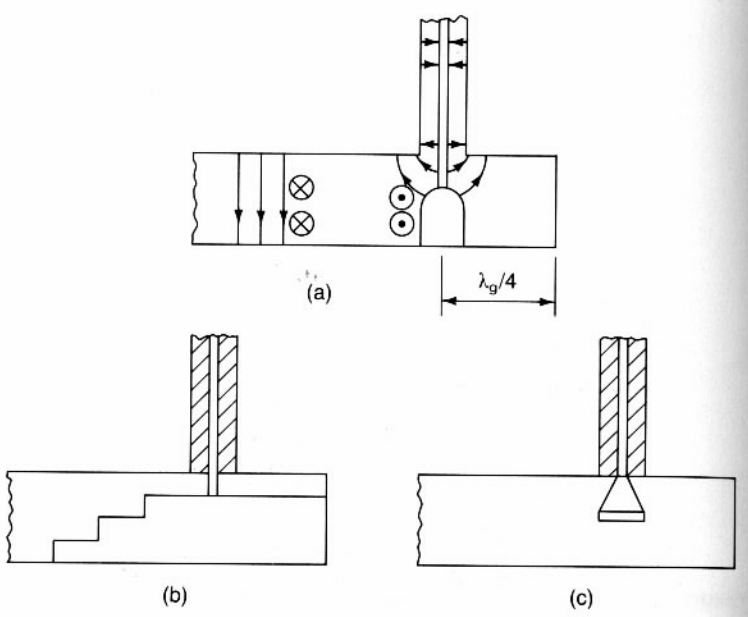

Fig. 6.13 Coaxial line to waveguide transitions: (a) door-knob transition, (b) ridge-waveguide transition, and (c) probe transition.

broad-band match from coaxial line to waveguide. Standard coaxial line has a characteristic impedance of 50 Ω. The characteristic impedance of rectangular waveguide is given by (2.51). Standard waveguides have an aspect ratio (a/b) of about 2 : 1 and are used over a range of λ_g/λ_0 of about 1.7 to 1.2. The impedance to be matched therefore ranges from 630 Ω to 440 Ω, roughly ten times that of the coaxial line, and devices for this purpose are therefore often called 'coaxial-to-waveguide transformers'.

Consideration of the field patterns in the two guides suggests that a door-knob transition similar to that shown in Fig. 6.12(b) might be used. The field patterns arising are shown in Fig. 6.13(a). A quarter-wave length of waveguide behind the transition is used to make the shunt impedance of the end of the waveguide an open circuit at the transition. It can be seen that the field patterns are not very well matched.

One of the difficulties of the door-knob transition is that it has to cope with the change from a low-voltage, high-current, wave in the coaxial line to a high-voltage, low-current, wave in the waveguide. The existence of a direct current path makes this difficult. Two solutions to this problem suggest themselves: one is to make use of a lower impedance waveguide, the other is to avoid a direct current path.

To reduce the waveguide impedance we must reduce its height (see (2.51)). But the reduction to something like one tenth of the height of

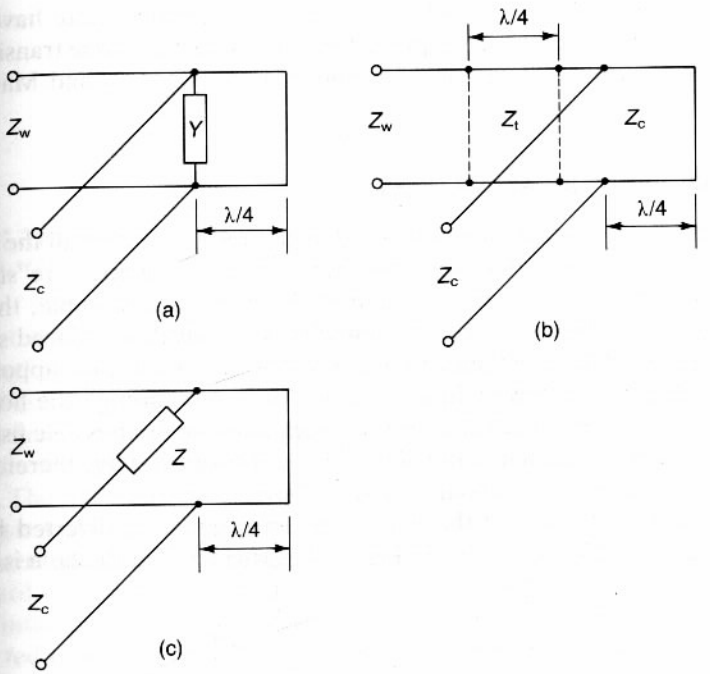

Fig. 6.14 Equivalent circuits for the coaxial line to waveguide transitions shown in Fig. 6.13.

standard waveguide presents new problems for the transformation between the two. A better solution is to use ridge waveguide which can have an impedance of 50 Ω without needing such narrow gaps. Figure 6.13(b) shows a transition using a ridge waveguide with a stepped impedance transformer to standard waveguide and, once again, a quarter-wave section of short-circuited line behind it.

The third approach is to couple into the waveguide using displacement current rather than conduction current. Figure 6.13(c) shows how this is achieved. The free end of the centre conductor is, in effect, an antenna radiating into the waveguide. The use of an expanded end to it provides for the necessary impedance transformation. It also has the effect of providing a good match between the field patterns as can be seen from Fig. 6.13(c). These two statements are really just saying the same thing, one in circuit terms and the other in field terms.

The three transitions described above can be represented by the equivalent circuits shown in Figs. 6.14(a), (b) and (c), respectively. The first is unsatisfactory and is seldom used. The second works well but is bulky. The third can be made to have a good broadband match and has the advantage of being very compact. Many other kinds of transition have been invented,

some of which are in common use. The three described here have been included to illustrate the principles involved in making a mode transition of this kind. Further details can be found in Harvey (1963) and Marcuvitz (1986).

6.6 COUPLING BY APERTURES

So far we have discussed cases where it is desired to transfer all the power flowing in one waveguide into another so that the guides are strongly coupled; sometimes this is not required. Consider, for example, the situation shown in Fig. 6.15. Two rectangular waveguides are placed side by side with a small hole in their common wall. It is reasonable to suppose that some of the power flowing in waveguide 1 will pass through the hole and excite waves in waveguide 2 as shown. Such a device has possible uses as a way of sampling the signal in waveguide 1. It is interesting, therefore, to consider the theory of coupling by small holes.

The current flowing in the wall of the waveguide is diverted by the presence of the hole as shown in Fig. 6.16. Provided that the hole is small,

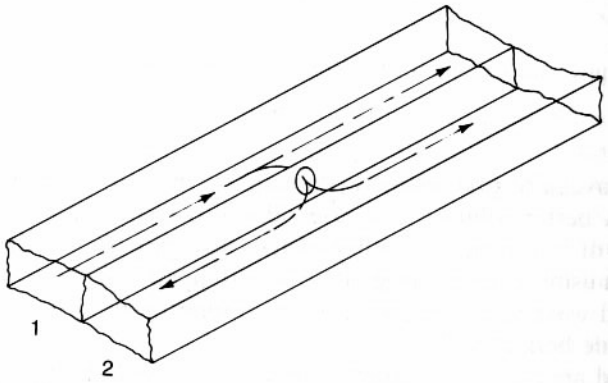

Fig. 6.15 Coupling between waveguides by a small hole in a common wall.

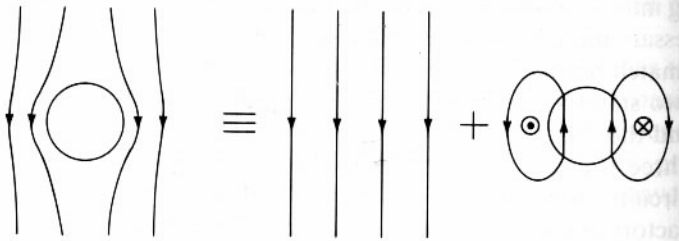

Fig. 6.16 Representation of the current flow around a hole in a waveguide wall by the superposition of the field of a magnetic dipole on a uniform field.

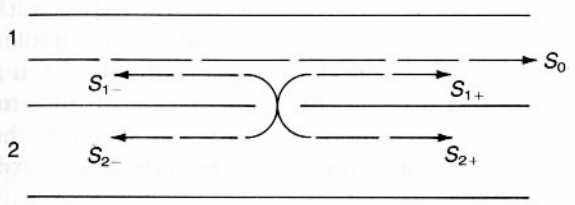

Fig. 6.17 The waves radiated into two waveguides by a magnetic dipole in their common wall.

so that the current patterns away from the hole are undisturbed and the hole itself is small compared with the free-space wavelength, we can think of its effect as the superposition of a small current loop on the undisturbed current flow. This current loop generates an alternating magnetic field and so acts as a magnetic dipole antenna which radiates power into both waveguides. The radiated power excites four waves as shown in Fig. 6.17 besides exciting cut-off higher-order modes. S_{1+} must be roughly in antiphase with S_0 because the hole is extracting some power from the incident wave. The backward wave S_{1-} shows that the presence of the hole introduces a mismatch into guide 1. The signals S_{2+} and S_{2-} provide samples of S_0 which can be fed to a detector or power meter. The whole system is linear so the sampled power is proportional to the incident power.

When the two guides have the broad wall as their common boundary the situation is a little more complicated. This is because there are now two components of the tangential magnetic field and one of the normal electric field. An argument similar to the one in the previous paragraph can be used. Figure 6.18(a) shows how the fringing of the normal electric field through a small hole can be represented by the superposition of a small electric dipole on the unperturbed field. Similarly, Fig. 6.18(b) shows the representation of the fringing of magnetic field lines through a small hole by a small magnetic dipole. The latter is an alternative way of looking at

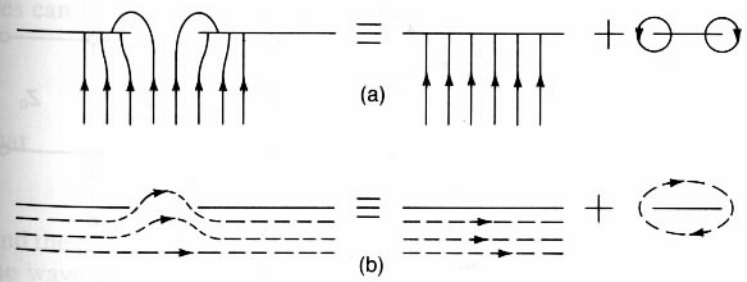

Fig. 6.18 Representation of the fringing electric and magnetic fields around a hole in a conducting wall by the superposition of a dipole field on a uniform field: (a) electric fringing field, and (b) magnetic fringing field.

the situation shown in Fig. 6.16. A little thought shows that the forward and backward waves radiated into a guide by an electric dipole are in phase with each other whilst those due to a magnetic dipole are in antiphase.

The dipole moments are evidently proportional to the excitation and may be expected to depend upon the size and shape of the hole. The dependence of the dipole moment on the size of the hole is expressed as its polarizability, so that

$$p = \varepsilon_0 \alpha_e |\boldsymbol{E}_i| \quad \text{and} \quad j = \mu_0 \alpha_m |\boldsymbol{H}_i|, \tag{6.12}$$

where p and j are the electric and magnetic dipole moments, α_e and α_m are the electric and magnetic polarizabilities and $|\boldsymbol{E}_i|$ and $|\boldsymbol{H}_i|$ are the exciting fields. For a circular hole of radius r the polarizabilities are given by

$$\alpha_e = \tfrac{2}{3} r^3 \quad \text{and} \quad \alpha_m = \tfrac{4}{3} r^3. \tag{6.13}$$

The mathematical theory of coupling by small holes is rather complicated so it will not be reproduced here. Details can be found in Collin (1966). In practice the design of devices which use coupling holes generally proceeds of the basis of measured parameters (see Harvey, 1963).

6.7 EFFECT OF HOLES IN SCREENS ON SCREENING EFFECTIVENESS

In Chapter 4 we saw that a conducting enclosure can be a very effective screen against electromagnetic interference provided that there are no holes in it. We are now in a position to consider how the presence of holes might affect the screening effectiveness of an enclosure.

Figure 6.19(a) shows a plane TEM wave incident normally on a conducting sheet with a square hole in it. The thickness of the sheet is d and

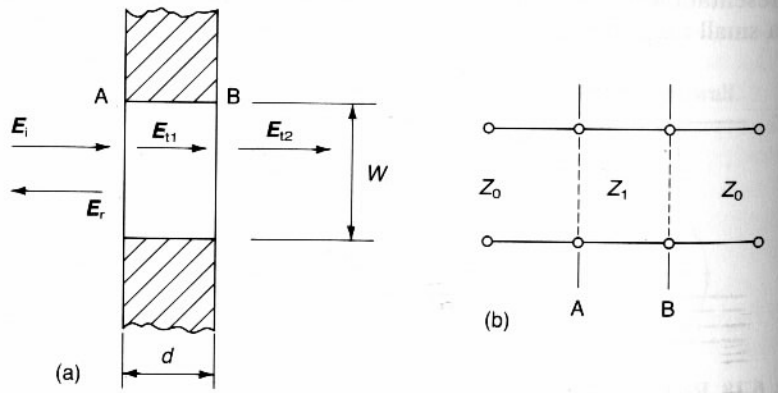

Fig. 6.19 Transmission of waves through a small hole in a conducting wall: (a) general arrangement, and (b) transmission-line equivalent circuit.

the width of the hole is W. Figure 6.19(b) shows the equivalent circuit for this problem. The hole normally has a width much less than the free-space wavelength so it behaves as a strongly cut-off waveguide and the signal is heavily attenuated as it passes through. The impedance mismatches at A and B are large giving rise to strong reflections and the possibility of multiple reflections if the attenuation of the hole is not very great. Thus the screening effectiveness of the sheet with the hole in it can be described by the equation

$$S = A_a + R_a + B_a \tag{6.14}$$

where A_a is the attenuation of the signal as it passes through the aperture, R_a is the reflection of the signal at the transition from the aperture to free space, and B_a is a term to account for multiple reflections.

The propagation constant in a rectangular waveguide is given by (2.64) and, for the lowest TE mode $k_c = \pi/W$. Now, if $k_0 \ll k_c$, $k_g = \mathrm{j}k_c$. The signal therefore propagates through the aperture as

$$\boldsymbol{E} = \boldsymbol{E}_0 \exp - (\pi z/W), \tag{6.15}$$

so that

$$\begin{aligned} A_a &= -20 \log_{10} [\exp - (\pi d/W)] \\ &= 27.3(d/W)\,\mathrm{dB}. \end{aligned} \tag{6.16}$$

The wave impedance of the wave in the aperture is

$$\begin{aligned} Z_1 &= \frac{E_x}{H_y} = \frac{\lambda_g}{\lambda_0} Z_0 \\ &= -\frac{2\mathrm{j}W}{\lambda_0} Z_0 \end{aligned} \tag{6.17}$$

from (2.45) and (2.46). This impedance is typically much less than the wave impedance of free space so there is a substantial mismatch. Assuming that the mismatch of impedances dominates so that the effects of higher-order modes can be neglected we have, from (1.85)

$$\frac{E_{t1}}{E_i} = \frac{2Z_1}{Z_0 + Z_1} \simeq -\frac{2\mathrm{j}W}{\lambda_0} \tag{6.18}$$

so that

$$R_a \simeq -20 \log_{10} (2W/\lambda_0). \tag{6.19}$$

To find the effects of multiple reflections consider Fig. 6.20. If the amplitude of the wave entering the hole at A is $\boldsymbol{E}_0$ then in the absence of multiple reflections the amplitude of the wave arriving at B is

$$\boldsymbol{E}_f = \boldsymbol{E}_0 \mathrm{e}^{-\alpha d}. \tag{6.20}$$

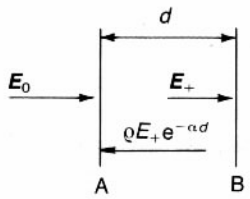

Fig. 6.20 Multiple reflection of waves within an aperture in a conducting sheet.

Now the voltage-reflection coefficients at A and B for waves travelling within the hole are both

$$\varrho = \frac{Z_0 - Z_1}{Z_0 + Z_1} \tag{6.21}$$

from (1.84). So, if the wave is multiply reflected at A and B, we must have

$$\boldsymbol{E}_+ = \boldsymbol{E}_0 e^{-\alpha d} + \varrho^2 \boldsymbol{E}_+ e^{-2\alpha d} \tag{6.22}$$

or

$$\boldsymbol{E}_+ = \frac{\boldsymbol{E}_0 e^{-\alpha d}}{(1 - \varrho^2 e^{-2\alpha d})}. \tag{6.23}$$

The bottom line of this expression represents the effects of multiple reflections so

$$B_a = 20 \log_{10} (1 - \varrho^2 e^{-2\alpha d}). \tag{6.24}$$

Now $Z_0 \gg Z_1$ so $\varrho \simeq 1$ and (6.24) can be written

$$B_a \simeq 20 \log_{10} (1 - 10^{-A_a/10}). \tag{6.25}$$

To see the implications of this theory in a practical case let us consider the effect of making a small hole in the aluminium screen whose screening effectiveness was estimated in Chapter 4.

Example

Estimate the leakage of signals through a square hole 5 mm × 5 mm in an aluminium sheet 0.2 mm thick over the frequency range 10 Hz to 10 MHz.

Solution

Substituting these figures into (6.16), (6.19) and (6.24) we get the figures in Table 6.2.

Comparison of these figures with those computed in the last chapter shows that even a small hole can produce a dramatic reduction in the screening effectiveness of an enclosure. As might be expected the degredation gets worse as the frequency increases.

Table 6.2

f	A(dB)	R(dB)	B(dB)	S(dB)
10 Hz	1.1	190	−13	178
100 Hz	1.1	170	−13	158
1 kHz	1.1	150	−13	138
10 kHz	1.1	130	−13	118
100 kHz	1.1	110	−13	98
1 MHz	1.1	90	−13	78
10 MHz	1.1	70	−13	58

So far we have only considered the case of a single square hole. If the hole had been some other shape the same general principles would apply but the figures would have been slightly different. When there are several holes it is necessary to include corrections for the proportion of the surface area occupied by the holes and for the coupling between them. In the limit of a wire mesh an additional correction is needed to take account of the skin effects in the wires. For further discussion of this subject see Keiser (1979).

6.8 WAVEGUIDE DIRECTIONAL COUPLERS

We have seen in Section 6.6 that it is possible to couple power from one waveguide to another through small holes in a common wall. By a suitable choice of the positions of those holes it is possible to make the coupling directional. Figure 6.21 shows schematically such an arrangement. In ideal coupler if a signal P_1 were injected at port 1 signals P_2 and P_3 would appear at ports 2 and 3 with nothing at port 4. Then the coupling of the coupler is defined as

$$C = 10 \log_{10} \left(\frac{P_1}{P_3} \right). \tag{6.26}$$

From the symmetry of the device the coupling would be the same if the power were injected at any other port with a cyclical permutation of the port numbers.

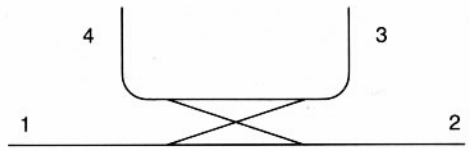

Fig. 6.21 Schematic diagram of a directional coupler.

In practice no directional coupler is perfect and some of the incident power at port 1 is coupled to port 4. The closeness of the coupler to the ideal is measured by its directivity defined by

$$D = 10 \log \left(\frac{P_4}{P_3}\right). \tag{6.27}$$

Notice carefully that the directivity is the ratio of the forward and backward powers in the second guide and not the ratio of the backward power in the second guide to the incident power in the first guide.

A device of this kind has many uses in microwave systems. It can be used to measure the forward and backward powers in the primary guide with very little disturbance to the quantities being measured provided that the coupling is weak enough.

To illustrate the way in which it is possible to make a directional coupler consider the two-hole side-wall coupler shown in Fig. 6.22. Let the electrical distance between the two holes be ϕ and let each radiate forward and backward waves with amplitudes $\boldsymbol{E}$ for an incident wave amplitude $\boldsymbol{E}_1$. The coupling is magnetic and we recall that the forward and backward waves radiated from each hole are in antiphase. Taking the plane A as the reference plane the amplitude of the forward wave in the secondary guide is

$$\boldsymbol{E}_3 = \boldsymbol{E} + (\boldsymbol{E}e^{-j\phi})e^{j\phi} \tag{6.28}$$

whilst the backward wave is

$$\boldsymbol{E}_4 = -\boldsymbol{E} - (\boldsymbol{E}e^{-j\phi})e^{-j\phi}. \tag{6.29}$$

The forward-wave components add for all values of the electrical length. The backward wave components cancel each other out if the electrical length between the holes is a quarter wavelength. This condition only exists at certain frequencies. In between, the backward power will be non-zero and could be equal to the forward power. Note that if the coupler is

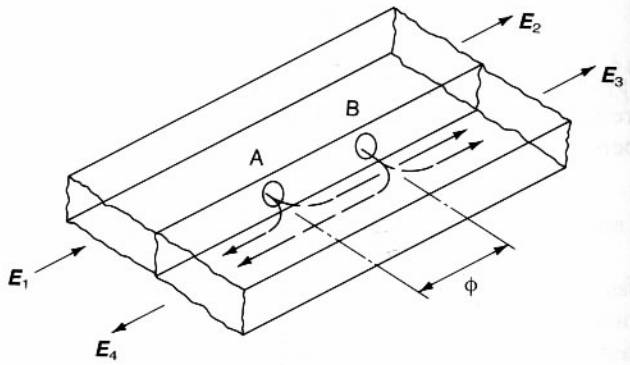

Fig. 6.22 Simple waveguide directional coupler.

symmetrical the backward wave will not quite be zero because the amplitude of the signal at B is less than that at A as a result of the transfer of some power at A from the first guide to the second. The bandwidth of the coupler can be increased by adding more holes to give either a Butterworth or a Tchebychev response (Levy, 1959). The coupling could equally well have been made through holes in a common broad wall though then the situation would have been complicated by the combination of electric and magnetic coupling.

Multi-hole directional couplers are commonly used where good directivity and a flat frequency resonse are required. The coupling is usually in the range 10 dB to 30 dB and a good instrumentation coupler has a directivity exceeding 40 dB throughout its band of operation.

Multi-hole directional couplers can be obtained with either three ports or four. The only difference is that in a three-port design the fourth port has been terminated by a built-in waveguide load. This can give a better match than if a separate external load were used. The reason why this is important can be seen from Fig. 6.21. Suppose that port 1 is connected to the signal generator whilst port 2 is connected to a load with a small reflection coefficient. Port 3 is terminated by a load and the power at port 4 is used to measure the reflection coefficient of the load in the main guide. The actual signal detected at port 4 will be made up of three components: first the backward power coupled from the main guide, second the sample of the forward power coupled to port 4 because of the finite directivity of the coupler, and third the reflection of the forward-coupled power by the imperfect load at port 3.

Example

A directional coupler has a coupling of 20 dB, directivity of 40 dB and an internal load whose return loss is 40 dB. Find the possible range of error if it is used to measure the return loss of a load whose actual return loss is 30 dB.

Solution

If the signal supplied by the signal generator is 1 mW (power levels are sometimes expressed in decibels referred to 1 milliwatt, written 'dBm', so this is 0 dBm) then the three components of the signal detected are:

power at port 2 coupled to port 4 = −50 dBm (10 nW);
power at port 1 coupled to port 4 = −60 dBm (1 nW); and
power at port 1 coupled to port 3 and reflected = −60 dBm (1 nW).

The power levels of the two smaller signals are each 1/10 of that of the

signal to be measured so their fields are smaller by a factor of $1/\sqrt{10}$. The total signal level detected is

$$\boldsymbol{E} = \boldsymbol{E}_0 \pm \frac{2}{\sqrt{10}}\boldsymbol{E}_0,$$

where the two signs represent the cases where the smaller waves are in phase or in antiphase with the signal to be measured. Thus the maximum and minimum signal levels are

$$1.632\boldsymbol{E}_0 \quad \text{and} \quad 0.368\boldsymbol{E}_0,$$

giving limits of +4.25 dB and −8.68 dB.

Many other different configurations of coupler have been designed. One which is particularly useful because of its compactness is the cross-guide coupler shown in Fig. 6.23. The use of offset holes in a common broad wall means that each hole must be represented by one electric dipole and two magnetic dipoles. Because of the different phase relationships for these dipoles it is possible to make each hole have directional properties. As with the coupler shown in Fig. 6.22 the correct choice of hole separation can also give directional properties. Cross-guide couplers typically have couplings in the range 10 dB to 30 dB, but with directivities only a little better than 20 dB.

It is possible to make couplers with couplings as strong as 3 dB but it is correspondingly more difficult to achieve good directivity or a flat frequency response. The hybrid tee junction discussed earlier can be thought of as a 3 dB coupler. A pair of these couplers can be used as power combiners. Figure 6.24 shows one application. The incoming power is split at A into two branches one of which contains a variable phase shifter B. The signals are then recombined at C. As the phase shifter is adjusted the output power changes from zero up to the full input power. The couplers and the

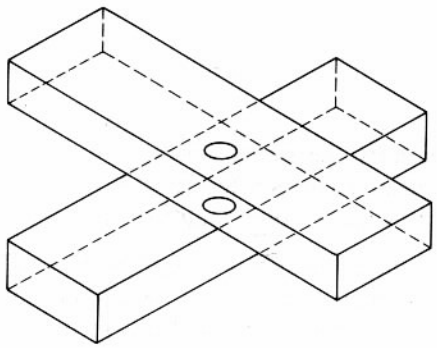

Fig. 6.23 Cross-guide directional coupler.

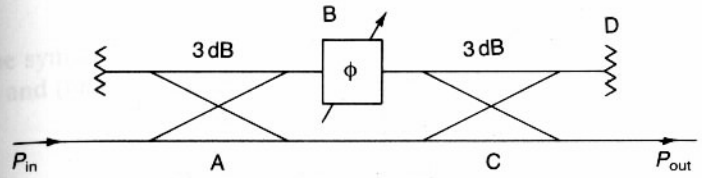

Fig. 6.24 Construction of a high-power attenuator using 3 dB couplers, a phase shifter and a high-power load.

phase shifter are lossless devices so the remaining power must end up in the load D. If D is a high-power load then this circuit can be used as a high-power attenuator.

6.9 DISTRIBUTED COUPLING

In the preceding section we have considered examples of coupling where the coupling is achieved by a number of discrete coupling elements. This is the kind of coupling normally encountered in waveguide devices. In microstrip and optical-fibre lines the coupling is usually distributed over a length of line. Figure 6.25 shows two parallel microstrip lines and the two possible ways in which they can be coupled together. In the first case (Fig. 6.25(a)) the signals on the lines are in phase with each other and the coupling is via the magnetic field. In the second case (Fig. 6.25(b)) the coupling is via the electric field and the signals on the lines are in antiphase. Notice carefully that in both these cases the signals on both lines are propagating into the paper as can be seen from the directions of the Poynting vectors. Note also that the two lines could be parallel tracks on a printed circuit board. Thus the theory given in this section applies also to cross-talk in high-speed

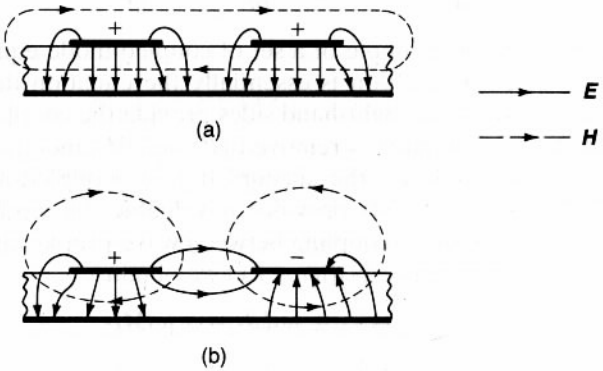

Fig. 6.25 Coupling between microstrip lines: (a) even-mode coupling, and (b) odd-mode coupling.

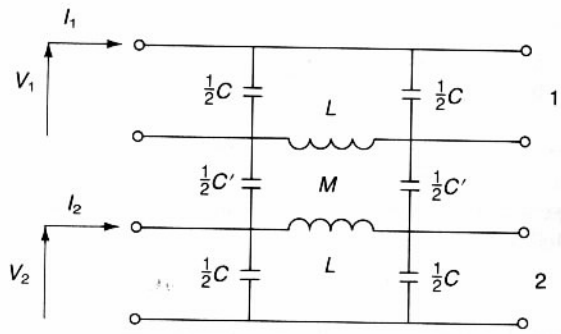

Fig. 6.26 Equivalent circuit for distributed coupling between transmission lines.

digital circuits. The coupled transmission lines can be represented by the equivalent circuit shown in Fig. 6.26. This shows pi sections of two identical transmission lines coupled by capacitance C' and mutual inductance M per unit length.

Using the same method of analysis as is used for single TEM lines (Carter, 1986) we can write down the set of equations which describe the coupled system.

$$\frac{\partial V_1}{\partial x} + L\frac{\partial I_1}{\partial t} = -M\frac{\partial I_2}{\partial t} \tag{6.30}$$

$$\frac{\partial V_2}{\partial x} + L\frac{\partial I_2}{\partial t} = -M\frac{\partial I_1}{\partial t} \tag{6.31}$$

$$\frac{\partial I_1}{\partial x} + (C + C')\frac{\partial V_1}{\partial t} = C'\frac{\partial V_2}{\partial t} \tag{6.32}$$

$$\frac{\partial I_2}{\partial x} + (C + C')\frac{\partial V_2}{\partial t} = C'\frac{\partial V_1}{\partial t} \tag{6.33}$$

These equations are an example of a set of coupled-mode equations. The left-hand side of each equation is essentially the equation describing an uncoupled mode whilst the right-hand sides provide the coupling terms. It is evident that if the coupling is removed (C' and M tending to zero) the equations reduce to those for the uncoupled lines. Coupled-mode theory (Pierce, 1954; Louisell, 1960) provides a valuable conceptual tool for studying problems involving coupling between wave-propagating systems.

Assuming propagation as exp j($\omega t - kx$) we obtain

$$-\mathrm{j}kV_1 + \mathrm{j}\omega LI_1 = -\mathrm{j}\omega MI_2 \tag{6.34}$$

$$-\mathrm{j}kV_2 + \mathrm{j}\omega LI_2 = -\mathrm{j}\omega MI_1 \tag{6.35}$$

$$-\mathrm{j}kI_1 + \mathrm{j}\omega(C + C')V_1 = \mathrm{j}\omega C'V_2 \tag{6.36}$$

$$-\mathrm{j}kI_2 + \mathrm{j}\omega(C + C')V_2 = \mathrm{j}\omega C'V_1. \tag{6.37}$$

For the symmetric mode (Fig. 6.25(a)) $V_1 = V_2$ and $I_1 = I_2$ so that from (6.34) and (6.36) we obtain

$$-\mathrm{j}kV_1 + \mathrm{j}\omega(L + M)I_1 = 0 \tag{6.38}$$

$$-\mathrm{j}kI_1 + \mathrm{j}\omega CV_1 = 0, \tag{6.39}$$

whence

$$k^2 - \omega^2(L + M)C = 0 \tag{6.40}$$

so that the propagation constant for the symmetric mode is

$$k_+ = \pm\omega\sqrt{[(L + M)C]}. \tag{6.41}$$

By substitution into (6.39) we find that the characteristic impedance of this mode is

$$Z_+ = \sqrt{\left(\frac{L + M}{C}\right)}. \tag{6.42}$$

Notice that C' does not appear in either (6.41) or (6.42), showing that the coupling is pure magnetic.

For the antisymmetric mode $V_1 = -V_2$ and $I_1 = -I_2$. Analysis similar to that given above leads to the following expressions for the propagation constant and the characteristic impedance

$$k_- = \pm\omega\sqrt{[(L - M)(C + 2C')]} \tag{6.43}$$

$$Z_- = \sqrt{\left(\frac{L - M}{C + 2C'}\right)}. \tag{6.44}$$

The propagation constants of the two modes can be written

$$k_+ = \pm k_0\sqrt{\left(1 + \frac{M}{L}\right)} \tag{6.45}$$

$$k_- = \pm k_0\sqrt{\left[\left(1 - \frac{M}{L}\right)\left(1 + \frac{2C'}{C}\right)\right]}, \tag{6.46}$$

where $k_0 = \omega\sqrt{(LC)}$ is the propagation constant of the uncoupled lines. Note that k_+ is always greater than k_0. The product of k_+ and k_- is

$$k_+k_- = k_0^2\left[\left(1 - \frac{M^2}{L^2}\right)\left(1 + \frac{2C'}{C}\right)\right]^{\frac{1}{2}}. \tag{6.47}$$

For the special case of weak coupling the terms involving M and C' can be neglected and

$$k_+k_- \simeq k_0^2. \tag{6.48}$$

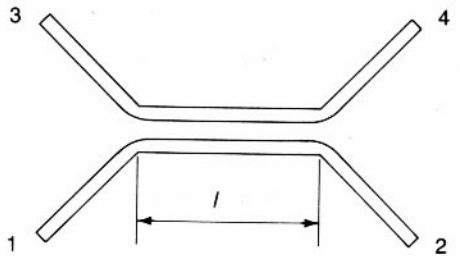

Fig. 6.27 Arrangement of a simple microstrip directional coupler.

Similarly, the product of the coupling impedances is

$$Z_+ Z_- = \sqrt{\left(\frac{L^2 - M^2}{C^2 + 2CC'}\right)} \tag{6.49}$$

and for weak coupling

$$Z_+ Z_- \simeq Z_0^2. \tag{6.50}$$

Now consider a microstrip directional coupler made by arranging two lines as shown in Fig. 6.27. An input signal at port 1 may be expected to excite both even and odd modes in the coupler. The analysis of this coupler proceeds by assuming forward and backward symmetric and antisymmetric waves in each arm (eight waves in all) and then applying the boundary conditions to obtain the amplitudes of each.

The analysis (Edwards, 1981) is tedious rather than difficult so it will not be reproduced here. The results may be expressed in terms of the signal amplitudes at the ports as

$$\frac{V_2}{V_1} = \frac{\sqrt{(1 - C^2)}}{\sqrt{(1 - C^2)}\cos\theta + \mathrm{j}\sin\theta} \tag{6.51}$$

$$\frac{V_3}{V_1} = \frac{\mathrm{j}C\sin\theta}{\sqrt{(1 - C^2)}\cos\theta + \mathrm{j}\sin\theta} \tag{6.52}$$

$$V_4 = 0, \tag{6.53}$$

where $\theta = k_0 l$ (the difference between k_+ and k_- being neglected) and the coupling constant C is defined by

$$C = \frac{Z_+ - Z_-}{Z_+ + Z_-}. \tag{6.54}$$

The interesting thing about these results is that the coupled power appears at port 3 not at port 4. A coupler with this property is known as a contradirectional coupler as opposed to a co-directional coupler. Maximum coupling is obtained when the coupled lines are a quarter of a wavelength long and then the coupling is

$$20 \log_{10}\left(\frac{Z_+ - Z_-}{Z_+ + Z_-}\right). \tag{6.55}$$

Because the electrical length of the lines varies with frequency the coupling is frequency dependent and the design is relatively narrow band.

It would appear from (6.53) that this kind of coupler must have infinite directivity. This is not the case because the difference between the propagation constants of the two modes has been neglected in the analysis. There are also, inevitably, reflections at the ends of the coupler where the modes are mismatched to the connecting lines. We have also assumed that the uncoupled modes can be regarded as pure TEM modes. In practice it is difficult to make a coupler of this kind with a directivity better than 15 dB.

The bandwidth of the coupler shown in Fig. 6.27 can be improved by the addition of lumped capacitors at its ends as shown in Fig. 6.28(a). The

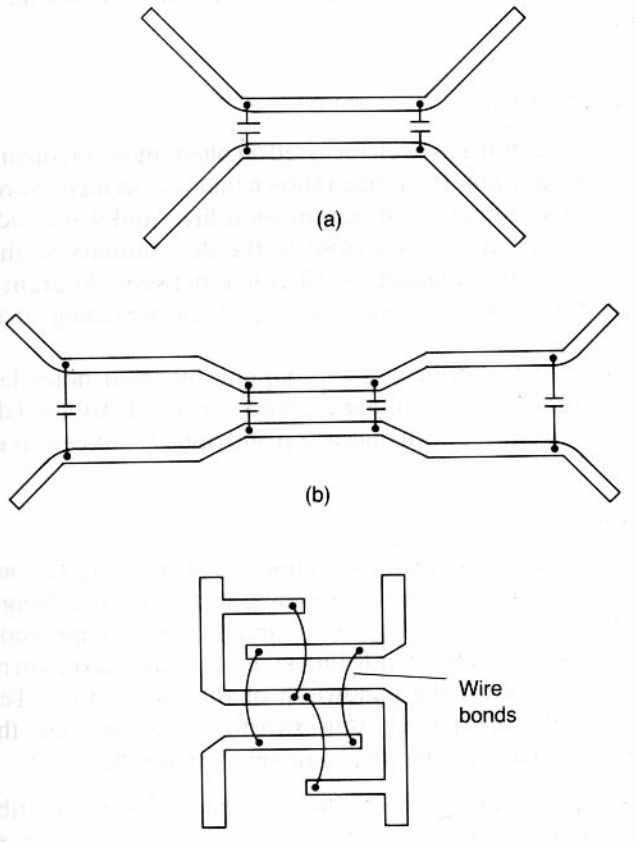

Fig. 6.28 Microstrip directional couplers: (a) narrow-band coupler, (b) broad-band coupler, and (c) Lange coupler.

capacitors have the effect of equalizing the phase velocities of the odd and even modes. Another approach (shown in Fig. 6.28(b)) is to make use of a number of sections to achieve a Tchebychev response.

An alternative type of microstrip directional coupler which is very commonly used is the Lange coupler shown in Fig. 6.28(c). With this arrangement couplers can be made with 10 dB coupling and 20 dB directivity over a frequency band of an octave. The design is largely empirical. For further details see Edwards (1981) and the references therein.

The discussion so far in this section has concentrated on microstrip couplers as examples of distributed directional couplers. It is possible for coupling between optical fibres to occur in a similar way. We have seen in Chapter 3 that there is an evanescent wave just outside the core of an optical fibre cable. If a second core is placed close enough to the first it will be coupled to the evanescent wave. The result is a 'leaky-wall' coupler which has distributed coupling and properties similar to the microstrip example discussed above. This coupling could result in crosstalk in bundles of optical fibres.

6.10 CONCLUSION

In this chapter attention has been focussed on phenomena involving coupling of modes of propagation. It has been shown that cut-off higher-order modes are excited at discontinuities in transmission lines and waveguides. These modes store energy in a region close to the discontinuity so they can be represented by lumped reactances. Coupling between different modes of propagation has been discussed in terms of the matching of their field patterns.

A discussion of the principles of coupling by small holes led to consideration of multi-hole directional couplers. Finally distributed directional couplers were introduced with microstrip and optical couplers as examples.

EXERCISES

6.1 Use the method described in Section 6.3 to investigate fourth-order Butterworth and Tchebychev matching using four matching elements with reflection coefficients ϱ_1, ϱ_2, ϱ_3 and ϱ_2. Derive expressions similar to (6.5) and (6.9) for the bandwidth in terms of the maximum mismatch. What is the ratio of the bandwidth of the fourth-order Tchebychev match to the fourth-order Butterworth match and the third-order Tchebychev match having the same maximum reflection?

6.2 Estimate the screening effectiveness of the enclosure described in the worked example on p. 146 if the hole is enlarged to 5 mm × 10 mm.

6.3 Design a transition at 6 GHz from WG14 waveguide to 50 Ω coaxial line

by coupling the line to a 50 Ω section of reduced-height waveguide and using a single-stage quarter-wave transformer to match the reduced-height guide to the standard guide. Why is this technique not used?

6.4 Two microstrip lines each 2 mm wide are constructed with their centres 5 mm apart on an alumina substrate ($\varepsilon_r = 10$) 1.0 mm thick. Estimate the magnitudes of the equivalent circuit parameters shown in Fig. 6.26 and, hence, the even- and odd-mode characteristic impedances and the propagation constants at 900 MHz.