

conductor are not perturbed by the sidewalls. We then have a closed finite region in which the potential $\Phi(x, y)$ satisfies Laplace's equation,

$$\nabla_t^2 \Phi(x, y) = 0 \quad \text{for } |x| \leq a/2, 0 \leq y \leq b, \quad (3.182)$$

with the boundary conditions

$$\Phi(x, y) = 0, \quad \text{at } x = \pm a/2, \quad (3.183a)$$

$$\Phi(x, y) = 0, \quad \text{at } y = 0, b. \quad (3.183b)$$

Laplace's equation can be solved by the method of separation of variables. Because the center conductor at $y = b/2$ will contain a surface charge density, the potential $\Phi(x, y)$ will have a slope discontinuity there because $\bar{D} = -\epsilon_0 \epsilon_r \nabla_t \Phi$ is discontinuous at $y = b/2$. Therefore, separate solutions for $\Phi(x, y)$ must be found for $0 < y < b/2$ and $b/2 < y < b$. The general solutions for $\Phi(x, y)$ in these two regions can be written as

$$\Phi(x, y) = \begin{cases} \sum_{\substack{n=1 \\ \text{odd}}}^{\infty} A_n \cos \frac{n\pi x}{a} \sinh \frac{n\pi y}{a} & \text{for } 0 \leq y \leq b/2 \\ \sum_{\substack{n=1 \\ \text{odd}}}^{\infty} B_n \cos \frac{n\pi x}{a} \sinh \frac{n\pi}{a}(b - y) & \text{for } b/2 \leq y \leq b. \end{cases} \quad (3.184)$$

Only the odd- n terms are needed in (3.184) because the solution is an even function of x . The reader can verify by substitution that (3.184) satisfies Laplace's equation in the two regions and satisfies the boundary conditions of (3.183).

The potential must be continuous at $y = b/2$, which from (3.184) leads to

$$A_n = B_n. \quad (3.185)$$

The remaining set of unknown coefficients, A_n , can be found by solving for the charge density on the center strip. Because $E_y = -\partial\Phi/\partial y$, we have

$$E_y = \begin{cases} -\sum_{\substack{n=1 \\ \text{odd}}}^{\infty} A_n \left(\frac{n\pi}{a}\right) \cos \frac{n\pi x}{a} \cosh \frac{n\pi y}{a} & \text{for } 0 \leq y \leq b/2 \\ \sum_{\substack{n=1 \\ \text{odd}}}^{\infty} A_n \left(\frac{n\pi}{a}\right) \cos \frac{n\pi x}{a} \cosh \frac{n\pi}{a}(b - y) & \text{for } b/2 \leq y \leq b. \end{cases} \quad (3.186)$$

The surface charge density on the strip at $y = b/2$ is

$$\begin{aligned} \rho_s &= D_y(x, y = b/2^+) - D_y(x, y = b/2^-) \\ &= \epsilon_0 \epsilon_r [E_y(x, y = b/2^+) - E_y(x, y = b/2^-)] \\ &= 2\epsilon_0 \epsilon_r \sum_{\substack{n=1 \\ \text{odd}}}^{\infty} A_n \left(\frac{n\pi}{a}\right) \cos \frac{n\pi x}{a} \cosh \frac{n\pi b}{2a}, \end{aligned} \quad (3.187)$$

which is seen to be a Fourier series in x for the surface charge density, ρ_s , on the strip at $y = b/2$. If we know the surface charge density we could easily find the unknown constants, A_n , and then the capacitance. We do not know the exact surface charge density, but we can make a good guess by approximating it as a constant over the width of the strip,

$$\rho_s(x) = \begin{cases} 1 & \text{for } |x| < W/2 \\ 0 & \text{for } |x| > W/2. \end{cases} \quad (3.188)$$

Equating this to (3.187) and using the orthogonality properties of the $\cos(n\pi x/a)$ functions gives the constants A_n as

$$A_n = \frac{2a \sin(n\pi W/2a)}{(n\pi)^2 \epsilon_0 \epsilon_r \cosh(n\pi b/2a)}. \quad (3.189)$$

The voltage of the strip conductor relative to the bottom conductor can be found by integrating the vertical electric field from $y = 0$ to $b/2$. Because the solution is approximate, this voltage is not constant over the width of the strip but varies with position, x . Rather than choosing the voltage at an arbitrary position, we can obtain an improved result by averaging the voltage over the width of the strip:

$$V_{\text{avg}} = \frac{1}{W} \int_{-W/2}^{W/2} \int_0^{b/2} E_y(x, y) dy dx = \sum_{\substack{n=1 \\ \text{odd}}}^{\infty} A_n \left(\frac{2a}{n\pi W} \right) \sin \frac{n\pi W}{2a} \sinh \frac{n\pi b}{2a}. \quad (3.190)$$

The total charge per unit length on the center conductor is

$$Q = \int_{-W/2}^{W/2} \rho_s(x) dx = W \text{ Coul/m}, \quad (3.191)$$

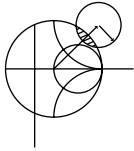
so the capacitance per unit length of the stripline is

$$C = \frac{Q}{V_{\text{avg}}} = \frac{W}{\sum_{\substack{n=1 \\ \text{odd}}}^{\infty} A_n \left(\frac{2a}{n\pi W} \right) \sin \frac{n\pi W}{2a} \sinh \frac{n\pi b}{2a}} \text{ F/m}. \quad (3.192)$$

Finally, the characteristic impedance is given by

$$Z_0 = \sqrt{\frac{L}{C}} = \frac{\sqrt{LC}}{C} = \frac{1}{v_p C} = \frac{\sqrt{\epsilon_r}}{c C},$$

where $c = 3 \times 10^8$ m/sec.



EXAMPLE 3.6 NUMERICAL CALCULATION OF STRIPLINE IMPEDANCE

Evaluate the above expressions for a stripline having $\epsilon_r = 2.55$ and $a = 100b$ to find the characteristic impedance for $W/b = 0.25$ to 5.0. Compare with the results from (3.179).

Solution

A computer program was written to evaluate (3.192). The series was truncated after 500 terms, and the results for Z_0 are as follows.

W/b	Z_0, Ω		
	Numerical, Eq. (3.192)	Formula, Eq. (3.179)	Commercial CAD
0.25	90.9	86.6	85.3
0.50	66.4	62.7	61.7
1.0	43.6	41.0	40.2
2.0	25.5	24.2	24.4
5.0	11.1	10.8	11.9

We see that the results are in reasonable agreement with the closed-form equations of (3.179) and the results from a commercial CAD package, particularly for wider strips where the charge density is closer to uniform. Better results could be obtained if more sophisticated estimates were used for the charge density. ■

3.8 MICROSTRIP LINE

Microstrip line is one of the most popular types of planar transmission lines primarily because it can be fabricated by photolithographic processes and is easily miniaturized and integrated with both passive and active microwave devices. The geometry of a microstrip line is shown in Figure 3.25a. A conductor of width W is printed on a thin, grounded dielectric substrate of thickness d and relative permittivity ϵ_r ; a sketch of the field lines is shown in Figure 3.25b.

If the dielectric substrate were not present ($\epsilon_r = 1$), we would have a two-wire line consisting of a flat strip conductor over a ground plane, embedded in a homogeneous medium (air). This would constitute a simple TEM transmission line with phase velocity $v_p = c$ and propagation constant $\beta = k_0$.

The presence of the dielectric, particularly the fact that the dielectric does not fill the region above the strip ($y > d$), complicates the behavior and analysis of microstrip line. Unlike stripline, where all the fields are contained within a homogeneous dielectric region, microstrip has some (usually most) of its field lines in the dielectric region between the strip conductor and the ground plane and some fraction in the air region above the substrate. For this reason microstrip line cannot support a pure TEM wave since the phase velocity of TEM fields in the dielectric region would be $c/\sqrt{\epsilon_r}$, while the phase velocity of TEM fields in the air region would be c , so a phase-matching condition at the dielectric–air interface would be impossible to enforce.

In actuality, the exact fields of a microstrip line constitute a hybrid TM-TE wave and require more advanced analysis techniques than we are prepared to deal with here. In most practical applications, however, the dielectric substrate is electrically very thin ($d \ll \lambda$), and so the fields are quasi-TEM. In other words, the fields are essentially the same as those of the static (DC) case. Thus, good approximations for the phase velocity, propagation constant, and characteristic impedance can be obtained from static, or *quasi-static*, solutions. Then the phase velocity and propagation constant can be expressed as

$$v_p = \frac{c}{\sqrt{\epsilon_e}}, \quad (3.193)$$

$$\beta = k_0 \sqrt{\epsilon_e}, \quad (3.194)$$

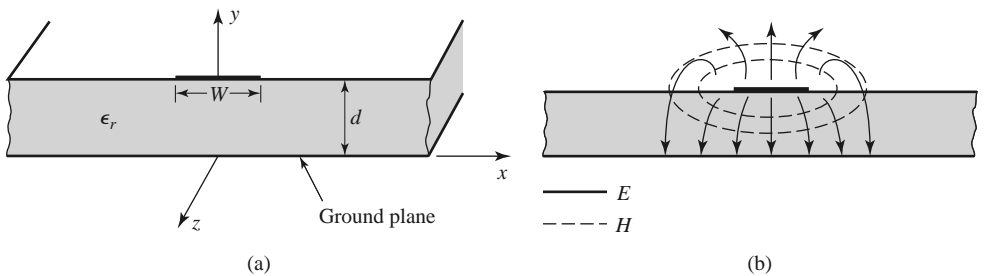


FIGURE 3.25 Microstrip transmission line. (a) Geometry. (b) Electric and magnetic field lines.

where ϵ_e is the *effective dielectric constant* of the microstrip line. Because some of the field lines are in the dielectric region and some are in air, the effective dielectric constant satisfies the relation

$$1 < \epsilon_e < \epsilon_r$$

and depends on the substrate dielectric constant, the substrate thickness, the conductor width, and the frequency.

We will present approximate design formulas for the effective dielectric constant, characteristic impedance, and attenuation of microstrip line; these results are curve-fit approximations to rigorous quasi-static solutions [8, 9]. Then we will discuss additional aspects of microstrip lines, including frequency-dependent effects, higher order modes, and parasitic effects.

Formulas for Effective Dielectric Constant, Characteristic Impedance, and Attenuation

The effective dielectric constant of a microstrip line is given approximately by

$$\epsilon_e = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \frac{1}{\sqrt{1 + 12d/W}}. \quad (3.195)$$

The effective dielectric constant can be interpreted as the dielectric constant of a homogeneous medium that equivalently replaces the air and dielectric regions of the microstrip line, as shown in Figure 3.26. The phase velocity and propagation constant are then given by (3.193) and (3.194).

Given the dimensions of the microstrip line, the characteristic impedance can be calculated as

$$Z_0 = \begin{cases} \frac{60}{\sqrt{\epsilon_e}} \ln \left(\frac{8d}{W} + \frac{W}{4d} \right) & \text{for } W/d \leq 1 \\ \frac{120\pi}{\sqrt{\epsilon_e} [W/d + 1.393 + 0.667 \ln (W/d + 1.444)]} & \text{for } W/d \geq 1. \end{cases} \quad (3.196)$$

For a given characteristic impedance Z_0 and dielectric constant ϵ_r , the W/d ratio can be found as

$$\frac{W}{d} = \begin{cases} \frac{8e^A}{e^{2A} - 2} & \text{for } W/d < 2 \\ \frac{2}{\pi} \left[B - 1 - \ln(2B - 1) + \frac{\epsilon_r - 1}{2\epsilon_r} \left\{ \ln(B - 1) + 0.39 - \frac{0.61}{\epsilon_r} \right\} \right] & \text{for } W/d > 2, \end{cases} \quad (3.197)$$

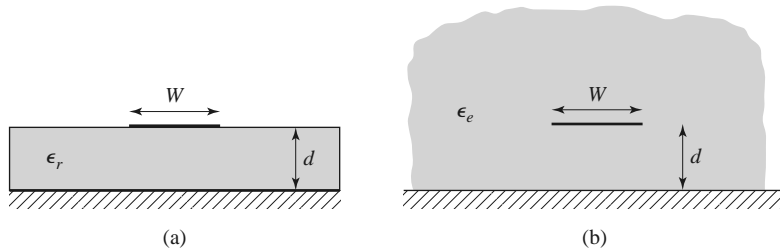


FIGURE 3.26 Equivalent geometry of a quasi-TEM microstrip line. (a) Original geometry. (b) Equivalent geometry, where the dielectric substrate of relative permittivity ϵ_r is replaced with a homogeneous medium of effective relative permittivity ϵ_e .

where

$$A = \frac{Z_0}{60} \sqrt{\frac{\epsilon_r + 1}{2}} + \frac{\epsilon_r - 1}{\epsilon_r + 1} \left(0.23 + \frac{0.11}{\epsilon_r} \right)$$

$$B = \frac{377\pi}{2Z_0\sqrt{\epsilon_r}}.$$

Considering a microstrip line as a quasi-TEM line, we can determine the attenuation due to dielectric loss as

$$\alpha_d = \frac{k_0\epsilon_r(\epsilon_e - 1) \tan \delta}{2\sqrt{\epsilon_e}(\epsilon_r - 1)} \text{ Np/m}, \quad (3.198)$$

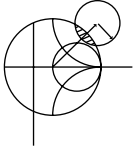
where $\tan \delta$ is the loss tangent of the dielectric. This result is derived from (3.30) by multiplying by a “filling factor,”

$$\frac{\epsilon_r(\epsilon_e - 1)}{\epsilon_e(\epsilon_r - 1)},$$

which accounts for the fact that the fields around the microstrip line are partly in air (lossless) and partly in the dielectric (lossy). The attenuation due to conductor loss is given approximately by [8]

$$\alpha_c = \frac{R_s}{Z_0 W} \text{ Np/m}, \quad (3.199)$$

where $R_s = \sqrt{\omega\mu_0/2\sigma}$ is the surface resistivity of the conductor. For most microstrip substrates, conductor loss is more significant than dielectric loss; exceptions may occur, however, with some semiconductor substrates.



EXAMPLE 3.7 MICROSTRIP LINE DESIGN

Design a microstrip line on a 0.5 mm alumina substrate ($\epsilon_r = 9.9$, $\tan \delta = 0.001$) for a 50Ω characteristic impedance. Find the length of this line required to produce a phase delay of 270° at 10 GHz, and compute the total loss on this line, assuming copper conductors. Compare the results obtained from the approximate formulas of (3.195)–(3.199) with those from a microwave CAD package.

Solution

First find W/d for $Z_0 = 50 \Omega$, and initially guess that $W/d < 2$. From (3.197),

$$A = 2.142, \quad W/d = 0.9654.$$

So the condition that $W/d < 2$ is satisfied; otherwise we would use the expression for $W/d > 2$. Then the required line width is $W = 0.9654d = 0.483 \text{ mm}$. From (3.195) the effective dielectric constant is $\epsilon_e = 6.665$. The line length, ℓ , for a 270° phase shift is found as

$$\phi = 270^\circ = \beta\ell = \sqrt{\epsilon_e}k_0\ell,$$

$$k_0 = \frac{2\pi f}{c} = 209.4 \text{ m}^{-1},$$

$$\ell = \frac{270^\circ(\pi/180^\circ)}{\sqrt{\epsilon_e}k_0} = 8.72 \text{ mm}.$$

Attenuation due to dielectric loss is found from (3.198) as $\alpha_d = 0.255 \text{ Np/m} = 0.022 \text{ dB/cm}$. The surface resistivity for copper at 10 GHz is 0.026Ω , and the attenuation due to conductor loss is, from (3.199), $\alpha_c = 0.0108 \text{ Np/cm} = 0.094 \text{ dB/cm}$. The total loss on the line is then 0.101 dB .

A commercial microwave CAD package gives the following results: $W = 0.478 \text{ mm}$, $\epsilon_e = 6.83$, $\ell = 8.61 \text{ mm}$, $\alpha_d = 0.022 \text{ dB/cm}$, and $\alpha_c = 0.054 \text{ dB/cm}$. The approximate formulas give results that are within a few percent of the CAD data for linewidth, effective dielectric constant, line length, and dielectric attenuation. The greatest discrepancy occurs for the attenuation constant for conductor loss. ■

Frequency-Dependent Effects and Higher Order Modes

The results for the parameters of microstrip line presented in the previous section were based on the quasi-static approximation and are strictly valid only at DC (or very low frequencies). At higher frequencies a number of effects can occur that lead to variations from the quasi-static results for effective dielectric constant, characteristic impedance, and attenuation of microstrip line. In addition, new effects can arise, such as higher order modes and parasitic reactances.

Because microstrip line is not a true TEM line, its propagation constant is not a linear function of frequency, meaning that the effective dielectric constant varies with frequency. The electromagnetic field that exists on microstrip line involves a hybrid coupling of TM and TE modes, complicated by the boundary condition imposed by the air and dielectric substrate interface. In addition, the current on the strip conductor is not uniform across the width of the strip, and this distribution varies with frequency. The thickness of the strip conductor also has an effect on the current distribution and hence affects the line parameters (especially the conductor loss).

The variation with frequency of the parameters of a transmission line is important for several reasons. First, if the variation is significant it becomes important to know and use the parameters at the particular frequency of interest to avoid errors in design or analysis. Typically, for microstrip line, the frequency variation of the effective dielectric constant is more significant than the variation of characteristic impedance, both in terms of relative change and the relative effect on performance. A change in the effective dielectric constant may have a substantial effect on the phase delay through a long section of line, while a small change in characteristic impedance has the primary effect of introducing a small impedance mismatch. Second, a variation in line parameters with frequency means that different frequency components of a broadband signal will propagate differently. A variation in phase velocity, for example, means that different frequency components will arrive at the output of the line at different times, leading to *signal dispersion* and distortion of the input signal. Third, because of the complexity of modeling these effects, approximate formulas are generally useful only for a limited range of frequency and line parameters, and numerical computer models are usually more accurate and useful.

There are a number of approximate formulas, developed from numerical computer solutions and/or experimental data, that have been suggested for predicting the frequency variation of microstrip line parameters [8, 9]. A popular frequency-dependent model for the effective dielectric constant has a form similar to the following formula [8]:

$$\epsilon_e(f) = \epsilon_r - \frac{\epsilon_r - \epsilon_e(0)}{1 + G(f)}, \quad (3.200)$$

where $\epsilon_e(f)$ represents the frequency-dependent effective dielectric constant, ϵ_r is the relative permittivity of the substrate, and $\epsilon_e(0)$ is the effective dielectric constant of the line at

DC, as given by (3.195). The function $G(f)$ can take various forms, but one suggested in reference [8] is that $G(f) = g(f/f_p)^2$, with $g = 0.6 + 0.009 Z_0$ and $f_p = Z_0/8\pi d$ (Z_0 is in ohms, f is in GHz, and d is in cm). It can be seen from the form of (3.200) that $\epsilon_e(f)$ reduces to the DC value $\epsilon_e(0)$ when $f = 0$ and increases toward ϵ_r as frequency increases.

Approximate formulas like the above were primarily developed in the years before computer-aided design tools for RF and microwave engineering became commonly available (see the Point of Interest on computer-aided design in Chapter 4). Such tools usually give accurate results for a wide range of line parameters and today are usually preferred over closed-form approximations.

Another potential difficulty with microstrip line is that it may support several types of higher order modes, particularly at higher frequencies. Some of these are directly related to the TM and TE surface waves modes that were discussed in Section 3.6, while others are related to waveguide-type modes in the cross section of the line.

The TM_0 surface wave mode for a grounded dielectric substrate has a zero cutoff frequency, as we know from (3.167). Because some of the field lines of this mode are aligned with the field lines of the quasi-TEM mode of a microstrip line, it is possible for coupling to occur from the desired microstrip mode to a surface wave, leading to excess power loss and possibly undesired coupling to adjacent microstrip elements. Because the fields of the TM_0 surface wave are zero at DC, there is little coupling to the quasi-TEM microstrip mode until a critical frequency is reached. Studies have shown that this threshold frequency is greater than zero and less than the cutoff frequency of the TM_1 surface wave mode. A commonly used approximation is [8]

$$f_{T1} \simeq \frac{c}{2\pi d} \sqrt{\frac{2}{\epsilon_r - 1}} \tan^{-1} \epsilon_r. \quad (3.201)$$

For ϵ_r ranging from 1 to 10, (3.201) gives a frequency that is 35% to 66% of f_{c1} , the cutoff frequency of the TM_1 surface wave mode.

When a microstrip circuit has transverse discontinuities (such as bends, junctions, or even step changes in width), the transverse currents on the conductors that are generated may allow coupling to TE surface wave modes. Most practical microstrip circuits involve such discontinuities, so this type of coupling is often important. The minimum threshold frequency where such coupling becomes important is given by the cutoff of the TE_1 surface wave, from (3.174):

$$f_{T2} \simeq \frac{c}{4d\sqrt{\epsilon_r - 1}}. \quad (3.202)$$

For wide microstrip lines, it is possible to excite a transverse resonance along the x axis of the microstrip line below the strip in the dielectric region because the sides below the strip conductor appear approximately as magnetic walls. This condition occurs when the width is about $\lambda/2$ in the dielectric, but because of field fringing the effective width of the strip is somewhat larger than the physical width. A rough approximation for the effective width is $W + d/2$, so the approximate threshold frequency for transverse resonance is

$$f_{T3} \simeq \frac{c}{\sqrt{\epsilon_r} (2W + d)}. \quad (3.203)$$

It is rare that a microstrip line is wide enough to approach this limit in practice.

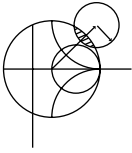
Finally, a parallel plate-type waveguide mode may propagate when the vertical spacing between the strip conductor and ground plane approaches $\lambda/2$ in the dielectric. Thus, an

approximation for the threshold frequency for this mode (valid for wide microstrip lines) can be given as

$$f_{T4} \simeq \frac{c}{2d\sqrt{\epsilon_r}}. \quad (3.204)$$

Thinner microstrip lines will have more fringing field that effectively lengthens the path between the strip and ground plane, thus reducing the threshold frequency by as much as 50%.

The net effect of the threshold frequencies given in (3.201)–(3.204) is to impose an upper frequency limit of operation for a given microstrip geometry. This limit is a function of the substrate thickness, dielectric constant, and strip width.



EXAMPLE 3.8 FREQUENCY DEPENDENCE OF EFFECTIVE DIELECTRIC CONSTANT

Use the approximate formula of (3.200) to plot the change in effective dielectric constant over frequency for a 25 Ω microstrip line on a substrate having a relative permittivity of 10.0 and a thickness of 0.65 mm. Compare the approximate data with results from a CAD model for frequencies up to 20 GHz. Compare the calculated phase delay at 10 GHz through a 1.093 cm length of line when using $\epsilon_e(0)$ versus $\epsilon_e(10 \text{ GHz})$.

Solution

The required linewidth for a 25 Ω impedance is $w = 2.00 \text{ mm}$. The effective dielectric constant for this line at low frequencies can be found from (3.195) to be $\epsilon_e(0) = 7.53$. A short computer program was used to calculate the effective dielectric constant as a function of frequency using (3.200), and the result is shown in Figure 3.27. Comparison with a commercial microwave CAD package shows that the approximate model is reasonably accurate up to about 10 GHz but gives an overestimate at higher frequencies.

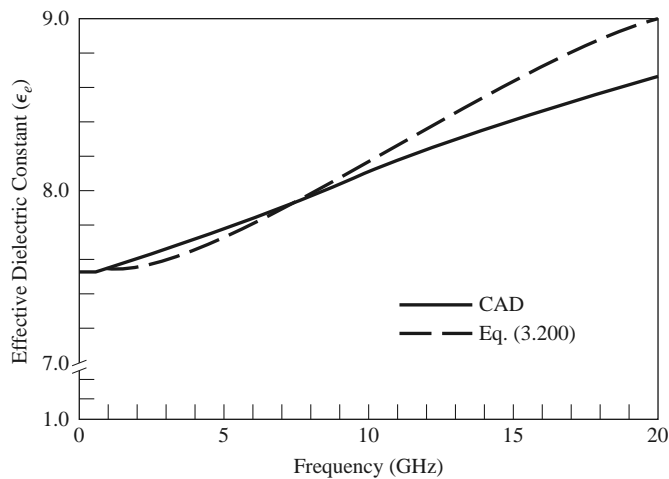


FIGURE 3.27 Effective dielectric constant versus frequency for the microstrip line of Example 3.8, comparing the approximate model of (3.200) with data from a computer-aided design package.

Using an effective dielectric constant of $\epsilon_e(0) = 7.53$, we find the phase delay through a 1.093 cm length of line to be $\phi_0 = \sqrt{\epsilon_e(0)}k_0\ell = 360^\circ$. The effective dielectric constant at 10 GHz is 8.120 (CAD), with a corresponding phase delay of $\phi_{10} = \sqrt{\epsilon_e(10 \text{ GHz})}k_0\ell = 374^\circ$ —an error of about 14° . ■

3.9

THE TRANSVERSE RESONANCE TECHNIQUE

According to the general solutions of Maxwell's equations for TE or TM waves given in Section 3.1, a uniform waveguide structure always has a propagation constant of the form

$$\beta = \sqrt{k^2 - k_c^2} = \sqrt{k^2 - k_x^2 - k_y^2}, \quad (3.205)$$

where $k_c = \sqrt{k_x^2 + k_y^2}$ is the cutoff wave number of the guide and, for a given mode, is a fixed function of the cross-sectional geometry of the guide. Thus, if we know k_c we can determine the propagation constant of the guide. In previous sections we determined k_c by solving the wave equation in the guide, subject to the appropriate boundary conditions. Although this technique is very powerful and general, it can be complicated for complex waveguides, especially if dielectric layers are present. In addition, the wave equation solution gives a complete field description inside the waveguide, which is often more information than we really need if we are only interested in the propagation constant of the guide. The *transverse resonance technique* employs a transmission line model of the transverse cross section of the waveguide and gives a much simpler and more direct solution for the cutoff frequency. This is another example where circuit and transmission line theory offers a simplified alternative to a field theory solution.

The transverse resonance procedure is based on the fact that in a waveguide at cutoff, the fields form standing waves in the transverse plane of the guide, as can be inferred from the “bouncing plane wave” interpretation of waveguide modes discussed in Section 3.2. This situation can be modeled with an equivalent transmission line circuit operating at resonance. One of the conditions of such a resonant line is the fact that, at any point on the line, the sum of the input impedances seen looking to either side must be zero. That is,

$$Z_{\text{in}}^r(x) + Z_{\text{in}}^\ell(x) = 0 \quad \text{for all } x, \quad (3.206)$$

where $Z_{\text{in}}^r(x)$ and $Z_{\text{in}}^\ell(x)$ are the input impedances seen looking to the right and left, respectively, at any point x on the resonant line.

The transverse resonance technique only gives results for the cutoff frequency of the guide. If fields or attenuation due to conductor loss are needed, the complete field theory solution will be required. The procedure will now be illustrated with an example.

TE_{0n} Modes of a Partially Loaded Rectangular Waveguide

The transverse resonance technique is particularly useful when the guide contains dielectric layers because the boundary conditions at the dielectric interfaces, which require the solution of simultaneous algebraic equations in the field theory approach, can be easily handled as junctions of different transmission lines. As an example, consider a rectangular waveguide partially filled with dielectric, as shown in Figure 3.28. To find the cutoff frequencies for the TE_{0n} modes, the equivalent transverse resonance circuit shown in the figure can be used. The line for $0 < y < t$ represents the dielectric-filled part of the guide

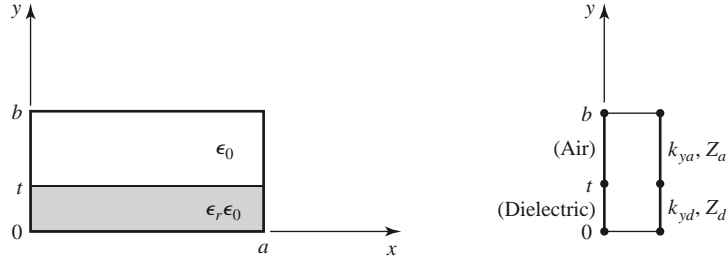


FIGURE 3.28 A rectangular waveguide partially filled with dielectric and the transverse resonance equivalent circuit.

and has a transverse propagation constant k_{yd} and a characteristic impedance for TE modes given by

$$Z_d = \frac{k\eta}{k_{yd}} = \frac{k_0\eta_0}{k_{yd}}, \quad (3.207a)$$

where $k_0 = \omega\sqrt{\mu_0\epsilon_0}$ and $\eta_0 = \sqrt{\mu_0/\epsilon_0}$. For $t < y < b$, the guide is air filled and has a transverse propagation constant k_{ya} and an equivalent characteristic impedance given by

$$Z_a = \frac{k_0\eta_0}{k_{ya}}. \quad (3.207b)$$

Applying condition (3.206) yields

$$k_{ya} \tan k_{yd}t + k_{yd} \tan k_{ya}(b-t) = 0. \quad (3.208)$$

This equation contains two unknowns, k_{ya} and k_{yd} . An additional equation is obtained from the fact that the longitudinal propagation constant, β , must be the same in both regions for phase matching of the tangential fields at the dielectric interface. Thus, with $k_x = 0$,

$$\beta = \sqrt{\epsilon_r k_0^2 - k_{yd}^2} = \sqrt{k_0^2 - k_{ya}^2},$$

or

$$\epsilon_r k_0^2 - k_{yd}^2 = k_0^2 - k_{ya}^2. \quad (3.209)$$

Equations (3.208) and (3.209) can be solved (numerically or graphically) to obtain k_{yd} and k_{ya} . There will be an infinite number of solutions, corresponding to the n dependence (number of variations in y) of the TE_{0n} mode.

3.10 WAVE VELOCITIES AND DISPERSION

We have so far encountered two types of velocities related to the propagation of electromagnetic waves:

- The speed of light in a medium ($1/\sqrt{\mu\epsilon}$)
- The phase velocity ($v_p = \omega/\beta$)

The speed of light in a medium is the velocity at which a plane wave would propagate in that medium, while the phase velocity is the speed at which a constant phase point travels. For a TEM plane wave, these two velocities are identical, but for other types of guided wave propagation the phase velocity may be greater or less than the speed of light.

If the phase velocity and attenuation of a line or guide are constants that do not change with frequency, then the phase of a signal that contains more than one frequency component will not be distorted. If the phase velocity is different for different frequencies, then the individual frequency components will not maintain their original phase relationships as they propagate down the transmission line or waveguide, and signal distortion will occur. Such an effect is called *dispersion* since different phase velocities allow the “faster” waves to lead in phase relative to the “slower” waves, and the original phase relationships will gradually be dispersed as the signal propagates down the line. In such a case, there is no single phase velocity that can be attributed to the signal as a whole. However, if the bandwidth of the signal is relatively small or if the dispersion is not too severe, a *group velocity* can be defined in a meaningful way. This velocity can be used to describe the speed at which the signal propagates.

Group Velocity

As discussed earlier, the physical interpretation of group velocity is the velocity at which a narrowband signal propagates. We will derive the relation of group velocity to the propagation constant by considering a signal $f(t)$ in the time domain. The Fourier transform of this signal is defined as

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt, \quad (3.210a)$$

and the inverse transform is

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j\omega t} d\omega. \quad (3.210b)$$

Now consider the transmission line or waveguide on which the signal $f(t)$ is propagating as a linear system, with a transfer function $Z(\omega)$ that relates the output, $F_o(\omega)$, of the line to the input, $F(\omega)$, of the line, as shown in Figure 3.29. Thus,

$$F_o(\omega) = Z(\omega) F(\omega). \quad (3.211)$$

For a lossless matched transmission line or waveguide, the transfer function $Z(\omega)$ can be expressed as

$$Z(\omega) = A e^{-j\beta z} = |Z(\omega)| e^{-j\psi}, \quad (3.212)$$

where A is a constant and β is the propagation constant of the line or guide.

The time domain representation of the output signal, $f_o(t)$, can then be written as

$$f_o(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) |Z(\omega)| e^{j(\omega t - \psi)} d\omega. \quad (3.213)$$

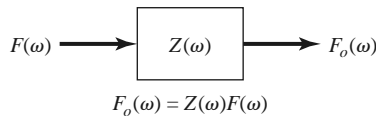


FIGURE 3.29

A transmission line or waveguide represented as a linear system with transfer function $Z(\omega)$.

If $|Z(\omega)| = A$ is a constant and the phase ψ of $Z(\omega)$ is a linear function of ω , say $\psi = a\omega$, the output can be expressed as

$$f_o(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} AF(\omega)e^{j\omega(t-a)}d\omega = Af(t-a), \quad (3.214)$$

which is seen to be a replica of $f(t)$, except for an amplitude factor A and time shift a . Thus, a transfer function of the form $Z(\omega) = Ae^{-j\omega a}$ does not distort the input signal. A lossless TEM wave has a propagation constant $\beta = \omega/c$, which is of this form, so a TEM line is *dispersionless* and does not lead to signal distortion. If the TEM line is lossy, however, the attenuation may be a function of frequency, which could lead to signal distortion.

Now consider a narrowband input signal of the form

$$s(t) = f(t) \cos \omega_0 t = \text{Re} \left\{ f(t)e^{j\omega_0 t} \right\}, \quad (3.215)$$

which represents an amplitude-modulated carrier wave of frequency ω_0 . Assume that the highest frequency component of $f(t)$ is ω_m , where $\omega_m \ll \omega_0$. The Fourier transform, $S(\omega)$, of $s(t)$, is

$$S(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega_0 t} e^{j\omega t} dt = F(\omega - \omega_0), \quad (3.216)$$

where we have used the complex form of the input signal as expressed in (3.215). We will need to take the real part of the output inverse transform to obtain the time domain output signal. The spectra of $F(\omega)$ and $S(\omega)$ are depicted in Figure 3.30.

The output signal spectrum is

$$S_o(\omega) = AF(\omega - \omega_0)e^{-j\beta z}, \quad (3.217)$$

and in the time domain,

$$\begin{aligned} s_o(t) &= \frac{1}{2\pi} \text{Re} \int_{-\infty}^{\infty} S_o(\omega)e^{j\omega t} d\omega \\ &= \frac{1}{2\pi} \text{Re} \int_{\omega_0 - \omega_m}^{\omega_0 + \omega_m} AF(\omega - \omega_0)e^{j(\omega t - \beta z)} d\omega. \end{aligned} \quad (3.218)$$

In general, the propagation constant β may be a complicated function of ω . However, if $F(\omega)$ is narrowband ($\omega_m \ll \omega_0$), then β can often be linearized by using a Taylor series expansion about ω_0 :

$$\beta(\omega) = \beta(\omega_0) + \left. \frac{d\beta}{d\omega} \right|_{\omega=\omega_0} (\omega - \omega_0) + \frac{1}{2} \left. \frac{d^2\beta}{d\omega^2} \right|_{\omega=\omega_0} (\omega - \omega_0)^2 + \dots \quad (3.219)$$

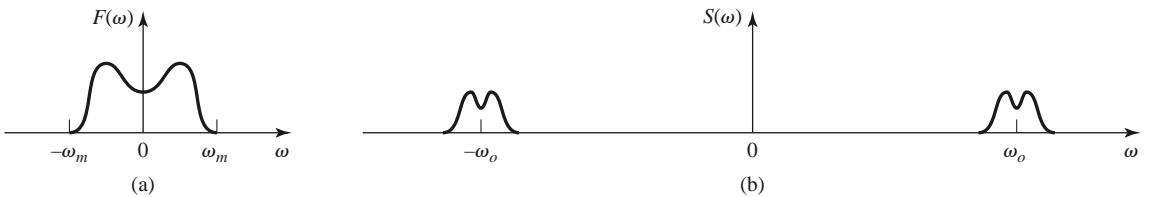


FIGURE 3.30 Fourier spectra of the signals (a) $f(t)$ and (b) $s(t)$.

Retaining the first two terms gives

$$\beta(\omega) \simeq \beta_o + \beta'_o(\omega - \omega_o), \quad (3.220)$$

where

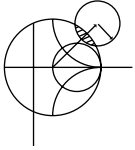
$$\begin{aligned} \beta_o &= \beta(\omega_o), \\ \beta'_o &= \left. \frac{d\beta}{d\omega} \right|_{\omega=\omega_o}. \end{aligned}$$

After a change of variables to $y = \omega - \omega_o$, the expression for $s_o(t)$ becomes

$$\begin{aligned} s_o(t) &= \frac{A}{2\pi} \operatorname{Re} \left\{ e^{j(\omega_o t - \beta_o z)} \int_{-\omega_m}^{\omega_m} F(y) e^{j(t - \beta'_o z)y} dy \right\} \\ &= A \operatorname{Re} \left\{ f(t - \beta'_o z) e^{j(\omega_o t - \beta_o z)} \right\} \\ &= A f(t - \beta'_o z) \cos(\omega_o t - \beta_o z), \end{aligned} \quad (3.221)$$

which is a time-shifted replica of the original modulation envelope, $f(t)$, of (3.215). The velocity of this envelope is the group velocity, v_g :

$$v_g = \frac{1}{\beta'_o} = \left(\frac{d\beta}{d\omega} \right)^{-1} \bigg|_{\omega=\omega_o}. \quad (3.222)$$



EXAMPLE 3.9 WAVEGUIDE WAVE VELOCITIES

Calculate the group velocity for a waveguide mode propagating in an air-filled guide. Compare this velocity to the phase velocity and speed of light.

Solution

The propagation constant for a mode in an air-filled waveguide is

$$\beta = \sqrt{k_0^2 - k_c^2} = \sqrt{(\omega/c)^2 - k_c^2}.$$

Taking the derivative with respect to frequency gives

$$\frac{d\beta}{d\omega} = \frac{\omega/c^2}{\sqrt{(\omega/c)^2 - k_c^2}} = \frac{k_o}{c\beta},$$

so from (3.234) the group velocity is

$$v_g = \left(\frac{d\beta}{d\omega} \right)^{-1} = \frac{c\beta}{k_o}.$$

The phase velocity is $v_p = \omega/\beta = ck_o/\beta$. Since $\beta < k_o$, we have that $v_g < c < v_p$, which indicates that the phase velocity of a waveguide mode may be greater than the speed of light, but the group velocity (the velocity of a narrow-band signal) will be less than the speed of light. ■

3.11

SUMMARY OF TRANSMISSION LINES AND WAVEGUIDES

We have discussed a variety of transmission lines and waveguides in this chapter, and here we will summarize some of the basic properties of these transmission media and their relative advantages in a broader context.

TABLE 3.6 Comparison of Common Transmission Lines and Waveguides

Characteristic	Coax	Waveguide	Stripline	Microstrip
Modes: Preferred	TEM	TE ₁₀	TEM	Quasi-TEM
Other	TM,TE	TM,TE	TM,TE	Hybrid TM,TE
Dispersion	None	Medium	None	Low
Bandwidth	High	Low	High	High
Loss	Medium	Low	High	High
Power capacity	Medium	High	Low	Low
Physical size	Large	Large	Medium	Small
Ease of fabrication	Medium	Medium	Easy	Easy
Integration with	Hard	Hard	Fair	Easy

We made a distinction between TEM, TM, and TE waves and saw that transmission lines and waveguides can be categorized according to which type of waves they can support. We saw that TEM waves are nondispersive, with no cutoff frequency, while TM and TE waves exhibit dispersion and generally have nonzero cutoff frequencies. Other electrical considerations include bandwidth, attenuation, and power-handling capacity. Mechanical factors are also very important, however, and include such considerations as physical size (volume and weight), ease of fabrication (cost), and the ability to be integrated with other devices (active or passive). Table 3.6 compares several types of transmission media with regard to these considerations; this table only gives general guidelines, as specific cases may give better or worse results than those indicated.

Other Types of Lines and Guides

Although we have discussed the most common types of waveguides and transmission lines, there are many other guides and lines (and many variations) that we are not able to present in detail. A few of the more popular types are briefly mentioned here.

Ridge waveguide: The practical bandwidth of rectangular waveguide is slightly less than an octave (a 2:1 frequency range). This is because the TE₂₀ mode begins to propagate at a frequency equal to twice the cutoff frequency of the TE₁₀ mode. The ridge waveguide, shown in Figure 3.31, consists of a rectangular waveguide loaded with conducting ridges on the top and/or bottom walls. This loading tends to lower the cutoff frequency of the dominant mode, leading to increased bandwidth and better (more constant) impedance characteristics. Ridge waveguides are often used for impedance matching purposes, where

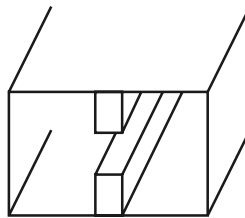


FIGURE 3.31 Cross section of a ridge waveguide.

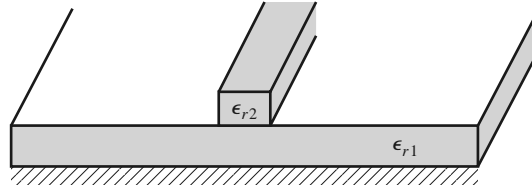


FIGURE 3.32 Dielectric waveguide geometry.

the ridge may be tapered along the length of the guide. The presence of the ridge, however, reduces the power-handling capacity of the waveguide.

Dielectric waveguide: As we have seen from our study of surface waves, metallic conductors are not necessary to confine and support a propagating electromagnetic field. The dielectric waveguide shown in Figure 3.32 is another example of such a guide, where ϵ_{r2} , the dielectric constant of the ridge, is usually greater than ϵ_{r1} , the dielectric constant of the substrate. The fields are thus mostly confined to the ridge and the surrounding area. This type of guide supports TM and TE modes, and is convenient for miniaturization and integration with active devices. Its small size makes it useful for millimeter wave to optical frequencies, although it can be very lossy at bends or junctions in the ridge line. Many variations in this basic geometry are possible.

Slotline: Slotline is another one of the many possible types of planar transmission lines. The geometry of a slotline is shown in Figure 3.33. It consists of a thin slot in the ground plane on one side of a dielectric substrate. Thus, like microstrip line, the two conductors of slotline lead to a quasi-TEM type of mode. The width of the slot controls the characteristic impedance of the line.

Coplanar waveguide: The coplanar waveguide, shown in Figure 3.34, is similar to the slotline, and can be viewed as a slotline with a third conductor centered in the slot region. Because of the presence of this additional conductor, this type of line can support even or odd quasi-TEM modes, depending on whether the electric fields in the two slots are in the opposite direction or the same direction. Coplanar waveguides are particularly useful for fabricating active circuitry due to the presence of the center conductor and the close proximity of the ground planes.

Covered microstrip: Many variations of the basic microstrip line geometry are possible, but one of the more common is the covered microstrip, shown in Figure 3.35. The metallic cover plate is often used for electrical shielding and physical protection of the microstrip circuitry and is usually situated several substrate thicknesses away from the circuit. Its presence, however, can perturb the operation of the circuit enough so that its effect must be taken into account during design.

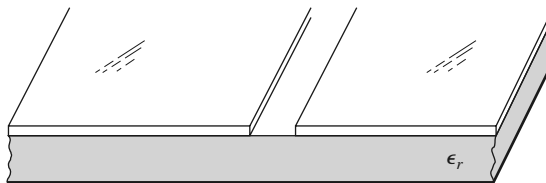


FIGURE 3.33 Geometry of a printed slotline.

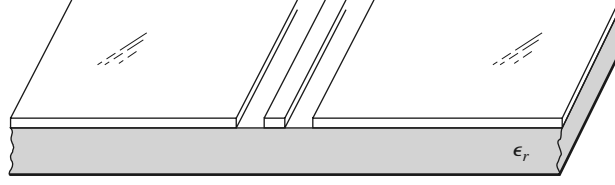
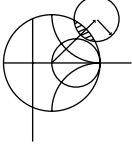


FIGURE 3.34 Coplanar waveguide geometry.



POINT OF INTEREST: Power Capacity of Transmission Lines

The power-handling capacity of an air-filled transmission line or waveguide is usually limited by voltage breakdown, which occurs at a field strength of about $E_d = 3 \times 10^6$ V/m for room temperature air at sea level pressure. Thermal effects may also serve to limit the power capacity of some types of lines.

In an air-filled coaxial line the electric field varies as $E_\rho = V_o/(\rho \ln b/a)$, which has a maximum at $\rho = a$ (at the inner conductor). Thus the maximum voltage before breakdown is

$$V_{\max} = E_d a \ln \frac{b}{a} \quad (\text{peak-to-peak}),$$

and the maximum power capacity is then

$$P_{\max} = \frac{V_{\max}^2}{2Z_0} = \frac{\pi a^2 E_d^2}{\eta_0} \ln \frac{b}{a}.$$

As might be expected, this result shows that power capacity can be increased by using a larger coaxial cable (larger a, b with fixed b/a for the same characteristic impedance). However, propagation of higher order modes limits the maximum operating frequency for a given cable size. Thus, there is an upper limit on the power capacity of a coaxial line for a given maximum operating frequency, f_{\max} , which can be shown to be given by

$$P_{\max} = \frac{0.025}{\eta_0} \left(\frac{c E_d}{f_{\max}} \right)^2 = 5.8 \times 10^{12} \left(\frac{E_d}{f_{\max}} \right)^2.$$

As an example, at 10 GHz the maximum peak power capacity of any coaxial line with no higher order modes is about 520 kW.

In an air-filled rectangular waveguide the electric field varies as $E_y = E_o \sin(\pi x/a)$, which has a maximum value of E_o at $x = a/2$ (the middle of the guide). Thus the maximum power capacity before breakdown is

$$P_{\max} = \frac{ab E_o^2}{4Z_w} = \frac{ab E_d^2}{4Z_w},$$

which shows that power capacity increases with guide size. For most standard waveguides, $b \simeq 2a$. To avoid propagation of the TE_{20} mode we must have $a < c/f_{\max}$, where f_{\max} is the

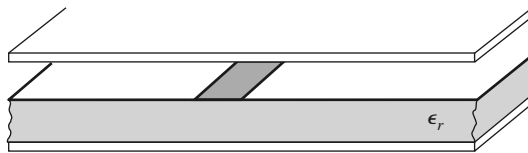


FIGURE 3.35 Covered microstrip line.

maximum operating frequency. Then the maximum power capacity of the guide can be shown to be

$$P_{\max} = \frac{0.11}{\eta_0} \left(\frac{cE_d}{f_{\max}} \right)^2 = 2.6 \times 10^{13} \left(\frac{E_d}{f_{\max}} \right)^2.$$

As an example, at 10 GHz the maximum peak power capacity of a rectangular waveguide operating in the TE_{10} mode is about 2300 kW, which is considerably higher than the power capacity of a coaxial cable at the same frequency.

Because arcing and voltage breakdown are high-speed transient effects, these voltage and power limits are peak values; average power capacity is lower. In addition, it is good engineering practice to provide a safety factor of at least two, so the maximum powers that can be safely transmitted should be limited to about half of the above values. If there are reflections on the line or guide, the power capacity is further reduced. In the worst case, a reflection coefficient magnitude of unity will double the maximum voltage on the line, so the power capacity will be reduced by a factor of four.

The power capacity of a line can be increased by pressurizing the line with air or an inert gas or by using a dielectric. The dielectric strength (E_d) of most dielectric materials is greater than that of air, but the power capacity may be further limited by the heating of the dielectric due to ohmic loss.

Reference: P. A. Rizzi, *Microwave Engineering—Passive Circuits*, Prentice-Hall, Englewood Cliffs, N.J., 1988.

REFERENCES

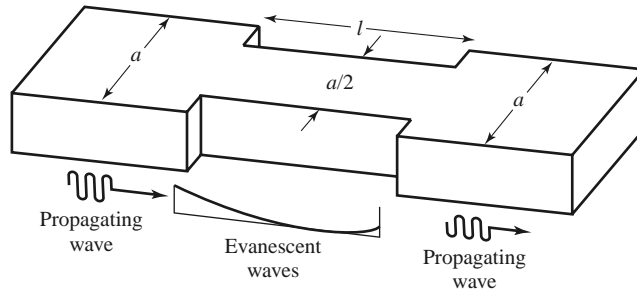
- [1] O. Heaviside, *Electromagnetic Theory*, Vol. 1, 1893. Reprinted by Dover, New York, 1950.
- [2] Lord Rayleigh, "On the Passage of Electric Waves through Tubes," *Philosophical Magazine*, vol. 43, pp. 125–132, 1897. Reprinted in *Collected Papers*, Cambridge University Press, Cambridge, 1903.
- [3] K. S. Packard, "The Origin of Waveguides: A Case of Multiple Rediscovery," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-32, pp. 961–969, September 1984.
- [4] R. M. Barrett, "Microwave Printed Circuits—An Historical Perspective," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-32, pp. 983–990, September 1984.
- [5] D. D. Grieg and H. F. Englemann, "Microstrip—A New Transmission Technique for the Kilomega-cycle Range," *Proceedings of the IRE*, vol. 40, pp. 1644–1650, December 1952.
- [6] H. Howe, Jr., *Stripline Circuit Design*, Artech House, Dedham, Mass., 1974.
- [7] I. J. Bahl and R. Garg, "A Designer's Guide to Stripline Circuits," *Microwaves*, January 1978, pp. 90–96.
- [8] I. J. Bahl and D. K. Trivedi, "A Designer's Guide to Microstrip Line," *Microwaves*, May 1977, pp. 174–182.
- [9] K. C. Gupta, R. Garg, and I. J. Bahl, *Microstrip Lines and Slotlines*, Artech House, Dedham, Mass., 1979.

PROBLEMS

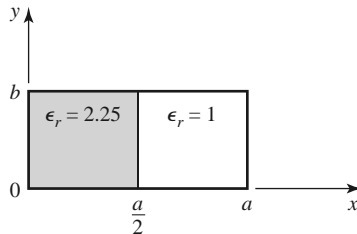
- 3.1 Devise at least two variations of the basic coaxial transmission line geometry of Section 3.5, and discuss the advantages and disadvantages of your proposed lines in terms of size, loss, cost, higher order modes, dispersion, or other considerations. Repeat this exercise for the microstrip line geometry of Section 3.8.
- 3.2 Derive equations (3.5a)–(3.5d) from equations (3.3) and (3.4).
- 3.3 Calculate the attenuation due to conductor loss for the TE_n mode of a parallel plate waveguide.
- 3.4 Consider a section of air-filled K-band waveguide. From the dimensions given in Appendix I, determine the cutoff frequencies of the first two propagating modes. From the recommended operating range given in Appendix I for this guide, determine the percentage reduction in bandwidth

that this operating range represents, relative to the theoretical bandwidth for a single propagating mode.

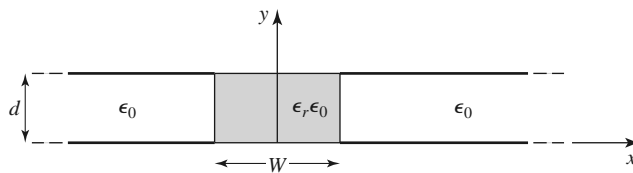
- 3.5** A 10 cm length of a K-band copper waveguide is filled with a dielectric material with $\epsilon_r = 2.55$ and $\tan \delta = 0.0015$. If the operating frequency is 15 GHz, find the total loss through the guide and the phase delay from the input to the output of the guide.
- 3.6** An attenuator can be made using a section of waveguide operating below cutoff, as shown in the accompanying figure. If $a = 2.286$ cm and the operating frequency is 12 GHz, determine the required length of the below-cutoff section of waveguide to achieve an attenuation of 100 dB between the input and output guides. Ignore the effect of reflections at the step discontinuities.



- 3.7** Find expressions for the electric surface current density on the walls of a rectangular waveguide for a TE_{10} mode. Why can a narrow slot be cut along the centerline of the broad wall of a rectangular waveguide without perturbing the operation of the guide? (Such a slot is often used in a slotted line for a probe to sample the standing wave field inside the guide.)
- 3.8** Derive the expression for the attenuation of the TM_{mn} mode of a rectangular waveguide due to imperfectly conducting walls.
- 3.9** For the partially loaded rectangular waveguide shown in the accompanying figure, solve (3.109) with $\beta = 0$ to find the cutoff frequency of the TE_{10} mode. Assume $a = 2.286$ cm, $t = a/2$, and $\epsilon_r = 2.25$.

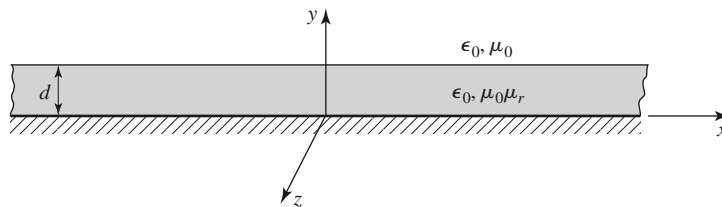


- 3.10** Consider the partially filled parallel plate waveguide shown in the accompanying figure. Derive the solution (fields and cutoff frequency) for the lowest order TE mode of this structure. Assume the metal plates are infinitely wide. Can a TEM wave propagate on this structure?

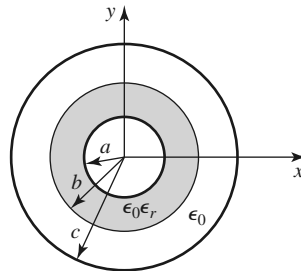


- 3.11** Derive equations (3.110a)–(3.110d) for the transverse field components in terms of longitudinal fields, in cylindrical coordinates.

- 3.12** Derive the expression for the attenuation of the TM_{nm} mode in a circular waveguide with finite conductivity.
- 3.13** A circular copper waveguide has a radius of 0.4 cm and is filled with a dielectric material having $\epsilon_r = 1.5$ and $\tan \delta = 0.0002$. Identify the first four propagating modes and their cutoff frequencies. For the dominant mode, calculate the total attenuation at 20 GHz.
- 3.14** Derive the \vec{E} and \vec{H} fields of a coaxial line from the expression for the potential given in (3.153). Also find expressions for the voltage and current on the line and the characteristic impedance.
- 3.15** Derive a transcendental equation for the cutoff frequency of the TM modes of a coaxial waveguide. Using tables, obtain an approximate value of $k_c a$ for the TM_{01} mode if $b/a = 2$.
- 3.16** Derive an expression for the attenuation of a TE surface wave on a grounded dielectric substrate when the ground plane has finite conductivity.
- 3.17** Consider the grounded magnetic substrate shown in the accompanying figure. Derive a solution for the TM surface waves that can propagate on this structure.



- 3.18** Consider the partially filled coaxial line shown in the accompanying figure. Can a TEM wave propagate on this line? Derive the solution for the TM_{0m} (no azimuthal variation) modes of this geometry.



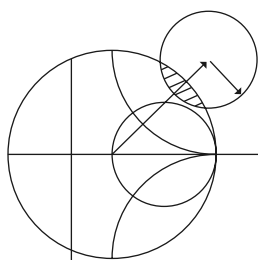
- 3.19** A copper stripline transmission line is to be designed for a $100\ \Omega$ characteristic impedance. The ground plane separation is 1.02 mm and the dielectric constant is 2.20, with $\tan \delta = 0.001$. At 5 GHz, find the guide wavelength on the line and the total attenuation.
- 3.20** A copper microstrip transmission line is to be designed for a $100\ \Omega$ characteristic impedance. The substrate is 0.51 mm thick, with $\epsilon_r = 2.20$ and $\tan \delta = 0.001$. At 5 GHz, find the guide wavelength on the line and the total attenuation. Compare these results with those for the similar stripline case of the preceding problem.
- 3.21** A $100\ \Omega$ microstrip line is printed on a substrate of thickness 0.0762 cm with a dielectric constant of 2.2. Ignoring losses and fringing fields, find the shortest length of this line that appears at its input as a capacitor of 5 pF at 2.5 GHz. Repeat for an inductance of 5 nH. Using a microwave CAD package with a physical model for the microstrip line, compute the actual input impedance seen when losses are included (assume copper conductors and $\tan \delta = 0.001$).
- 3.22** A microwave antenna feed network operating at 5 GHz requires a $50\ \Omega$ printed transmission line that is 16λ long. Possible choices are (1) copper microstrip, with $d = 0.16$ cm, $\epsilon_r = 2.20$, and $\tan \delta = 0.001$, or (2) copper stripline, with $b = 0.32$ cm, $\epsilon_r = 2.20$, $t = 0.01$ mm, and $\tan \delta = 0.001$. Which line should be used if attenuation is to be minimized?

- 3.23** Consider the TE modes of an arbitrary uniform waveguiding structure in which the transverse fields are related to H_z as in (3.19). If H_z is of the form $H_z(x, y, z) = h_z(x, y)e^{-j\beta z}$, where $h_z(x, y)$ is a real function, compute the Poynting vector and show that real power flow occurs only in the z direction. Assume that β is real, corresponding to a propagating mode.
- 3.24** A piece of rectangular waveguide is air filled for $z < 0$ and dielectric filled for $z > 0$. Assume that both regions can support only the dominant TE₁₀ mode and that a TE₁₀ mode is incident on the interface from $z < 0$. Using a field analysis, write general expressions for the transverse field components of the incident, reflected, and transmitted waves in the two regions and enforce the boundary conditions at the dielectric interface to find the reflection and transmission coefficients. Compare these results to those obtained with an impedance approach, using Z_{TE} for each region.
- 3.25** Use the transverse resonance technique to derive a transcendental equation for the propagation constant of the TM modes of a rectangular waveguide that is air filled for $0 < x < d$ and dielectric filled for $d < x < a$.
- 3.26** Apply the transverse resonance technique to find the propagation constants for the TE surface waves that can be supported by the structure of Problem 3.17.
- 3.27** An X-band waveguide filled with Rexolite is operating at 9.0 GHz. Calculate the speed of light in this material and the phase and group velocities in the waveguide.
- 3.28** As discussed in the Point of Interest on the power-handling capacity of transmission lines, the maximum power capacity of a coaxial line is limited by voltage breakdown and is given by

$$P_{\max} = \frac{\pi a^2 E_d^2}{\eta_0} \ln \frac{b}{a},$$

where E_d is the field strength at breakdown. Find the value of b/a that maximizes the maximum power capacity and show that the corresponding characteristic impedance is about 30 Ω .

- 3.29** A microstrip circuit is fabricated on an alumina substrate having a dielectric constant of 9.9, a thickness of 2.0 mm, and a 50 Ω linewidth of 1.93 mm. Find the threshold frequencies of the four higher order modes discussed in Section 3.8, and recommend the maximum operating frequency for this microstrip circuit.



Microwave Network Analysis

Circuits operating at low frequencies, for which the circuit dimensions are small relative to the wavelength, can be treated as an interconnection of lumped passive or active components with unique voltages and currents defined at any point in the circuit. In this situation the circuit dimensions are small enough such that there is negligible phase delay from one point in the circuit to another. In addition, the fields can be considered as TEM fields supported by two or more conductors. This leads to a quasi-static type of solution to Maxwell's equations and to the well-known Kirchhoff voltage and current laws and impedance concepts of circuit theory [1]. As the reader is aware, there is a powerful and useful set of techniques for analyzing low-frequency circuits. In general, these techniques cannot be directly applied to microwave circuits, but it is the purpose of the present chapter to show how basic circuit and network concepts can be extended to handle many microwave analysis and design problems of practical interest.

The main reason for doing this is that it is usually much easier to apply the simple and intuitive ideas of circuit analysis to a microwave problem than it is to solve Maxwell's equations for the same problem. In a way, field analysis gives us much more information about the particular problem under consideration than we really want or need. That is, because the solution to Maxwell's equations for a given problem is complete, it gives the electric and magnetic fields at all points in space. However, usually we are only interested in the voltage or current at a set of terminals, the power flow through a device, or some other type of "terminal" quantity, as opposed to a minute description of the fields at all points in space. Another reason for using circuit or network analysis is that it is then very easy to modify the original problem, or combine several elements together and find the response, without having to reanalyze in detail the behavior of each element in combination with its neighbors. A field analysis using Maxwell's equations for such problems would be hopelessly difficult. There are situations, however, in which such circuit analysis techniques are an oversimplification and may lead to erroneous results. In such cases one must resort to a field analysis approach, using Maxwell's equations. Fortunately, there are a number of commercially available computer-aided design packages that can model RF and microwave problems using both field theory analysis and network analysis. It is part of the education of a microwave engineer to be able to determine when network analysis concepts apply and when they should be cast aside in favor of more rigorous analysis.

The basic procedure for microwave network analysis is as follows. We first treat a set of basic, canonical problems rigorously, using field analysis and Maxwell's equations (as we have done in Chapters 2 and 3, for a variety of transmission line and waveguide problems). When so doing, we try to obtain quantities that can be directly related to a circuit or transmission line parameter. For example, when we treated various transmission lines and waveguides in Chapter 3 we derived the propagation constant and characteristic impedance of the line. This allowed the transmission line or waveguide to be treated as an idealized distributed component characterized by its length, propagation constant, and characteristic impedance. At this point, we can interconnect various components and use network and/or transmission line theory to analyze the behavior of the entire system of components, including effects such as multiple reflections, loss, impedance transformations, and transitions from one type of transmission medium to another (e.g., coax to microstrip). As we will see, a transition between different transmission lines, or a discontinuity on a transmission line, generally cannot be treated as a simple junction between two transmission lines, but typically includes some type of equivalent circuit to account for reactances associated with the transition or discontinuity.

Microwave network theory was originally developed in the service of radar system and component development at the MIT Radiation Lab in the 1940s. This work was continued at the Polytechnic Institute of Brooklyn and other locations by researchers such as E. Weber, N. Marcuvitz, A. A. Oliner, L. B. Felsen, A. Hessel, and many others [2].

4.1 IMPEDANCE AND EQUIVALENT VOLTAGES AND CURRENTS

Equivalent Voltages and Currents

At microwave frequencies the measurement of voltage or current is difficult (or impossible), unless a clearly defined terminal pair is available. Such a terminal pair may be present in the case of TEM-type lines (such as coaxial cable, microstrip line, or stripline), but does not strictly exist for non-TEM lines (such as rectangular, circular, or surface waveguides).

Figure 4.1 shows the electric and magnetic field lines for an arbitrary two-conductor TEM transmission line. As in Chapter 3, the voltage, V , of the + conductor relative to the – conductor can be found as

$$V = \int_{+}^{-} \vec{E} \cdot d\vec{\ell}, \quad (4.1)$$

where the integration path begins on the + conductor and ends on the – conductor. It is important to realize that, because of the electrostatic nature of the transverse fields between the two conductors, the voltage defined in (4.1) is unique and does not depend on the shape of the integration path. The total current flowing on the + conductor can be determined from an application of Ampere's law as

$$I = \oint_{C^{+}} \vec{H} \cdot d\vec{\ell}, \quad (4.2)$$

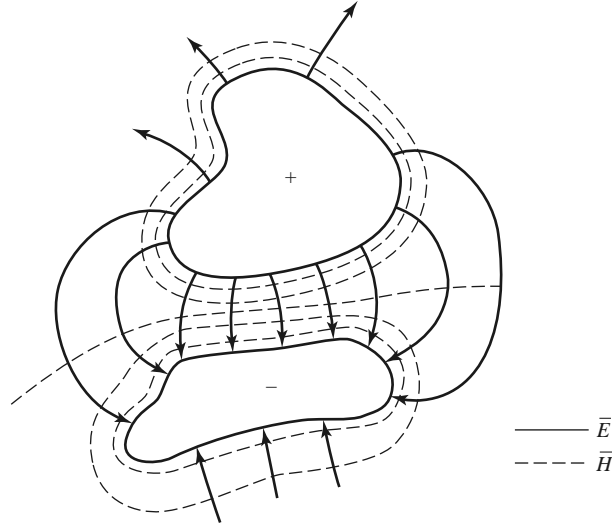


FIGURE 4.1 Electric and magnetic field lines for an arbitrary two-conductor TEM line.

where the integration contour is any closed path enclosing the + conductor (but not the - conductor). A characteristic impedance Z_0 can then be defined for traveling waves as

$$Z_0 = \frac{V}{I}. \quad (4.3)$$

At this point, after having defined and determined a voltage, current, and characteristic impedance (and assuming we know the propagation constant for the line), we can proceed to apply the circuit theory for transmission lines developed in Chapter 2 to characterize this line as a circuit element.

The situation is more difficult for waveguides. To see why, we will look at the case of a rectangular waveguide, as shown in Figure 4.2. For the dominant TE_{10} mode, the transverse fields can be written, from Table 3.2, as

$$E_y(x, y, z) = \frac{j\omega\mu a}{\pi} A \sin \frac{\pi x}{a} e^{-j\beta z} = A e_y(x, y) e^{-j\beta z}, \quad (4.4a)$$

$$H_x(x, y, z) = \frac{j\beta a}{\pi} A \sin \frac{\pi x}{a} e^{-j\beta z} = A h_x(x, y) e^{-j\beta z}. \quad (4.4b)$$

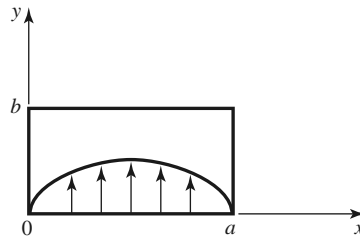


FIGURE 4.2 Electric field lines for the TE_{10} mode of a rectangular waveguide.

Applying (4.1) to the electric field of (4.4a) gives

$$V = \frac{-j\omega\mu a}{\pi} A \sin \frac{\pi x}{a} e^{-j\beta z} \int_y dy. \quad (4.5)$$

Thus it is seen that this voltage depends on the position, x , as well as the length of the integration contour along the y direction. For example, integrating from $y = 0$ to b for $x = a/2$ gives a voltage that is quite different from that obtained by integrating from $y = 0$ to b for $x = 0$. What, then, is the correct voltage? The answer is that there is no “correct” voltage in the sense of being unique or pertinent for all applications. A similar problem arises with current, and also impedance. We will now show how we can define *equivalent* voltages, currents, and impedances that can be useful for non-TEM lines.

There are many ways to define equivalent voltage, current, and impedance for waveguides since these quantities are not unique for non-TEM lines, but the following considerations usually lead to the most useful results [1, 3, 4]:

- Voltage and current are defined only for a particular waveguide mode, and are defined so that the voltage is proportional to the transverse electric field and the current is proportional to the transverse magnetic field.
- In order to be useful in a manner similar to voltages and currents of circuit theory, the equivalent voltages and currents should be defined so that their product gives the power flow of the waveguide mode.
- The ratio of the voltage to the current for a single traveling wave should be equal to the characteristic impedance of the line. This impedance may be chosen arbitrarily, but is usually selected as equal to the wave impedance of the line, or else normalized to unity.

For an arbitrary waveguide mode with both positively and negatively traveling waves, the transverse fields can be written as

$$\bar{E}_t(x, y, z) = \bar{e}(x, y)(A^+ e^{-j\beta z} + A^- e^{j\beta z}) = \frac{\bar{e}(x, y)}{C_1}(V^+ e^{-j\beta z} + V^- e^{j\beta z}), \quad (4.6a)$$

$$\bar{H}_t(x, y, z) = \bar{h}(x, y)(A^+ e^{-j\beta z} - A^- e^{j\beta z}) = \frac{\bar{h}(x, y)}{C_2}(I^+ e^{-j\beta z} - I^- e^{j\beta z}), \quad (4.6b)$$

where \bar{e} and \bar{h} are the transverse field variations of the mode, and A^+ , A^- are the field amplitudes of the traveling waves. Because \bar{E}_t and \bar{H}_t are related by the wave impedance, Z_w , according to (3.22) or (3.26), we also have that

$$\bar{h}(x, y) = \frac{\hat{z} \times \bar{e}(x, y)}{Z_w}. \quad (4.7)$$

Equation (4.6) also defines equivalent voltage and current waves as

$$V(z) = V^+ e^{-j\beta z} + V^- e^{j\beta z}, \quad (4.8a)$$

$$I(z) = I^+ e^{-j\beta z} - I^- e^{j\beta z}, \quad (4.8b)$$

with $V^+/I^+ = V^-/I^- = Z_0$. This definition embodies the idea of making the equivalent voltage and current proportional to the transverse electric and magnetic fields, respectively. The proportionality constants for this relationship are $C_1 = V^+/A^+ = V^-/A^-$ and $C_2 = I^+/A^+ = I^-/A^-$, and can be determined from the remaining two conditions for power and impedance.

The complex power flow for the incident wave is given by

$$P^+ = \frac{1}{2} |A^+|^2 \int_S \bar{\mathbf{e}} \times \bar{\mathbf{h}}^* \cdot \hat{\mathbf{z}} ds = \frac{V^+ I^{+*}}{2C_1 C_2^*} \int_S \bar{\mathbf{e}} \times \bar{\mathbf{h}}^* \cdot \hat{\mathbf{z}} ds. \quad (4.9)$$

Because we want this power to be equal to $(1/2)V^+ I^{+*}$, we have the result that

$$C_1 C_2^* = \int_S \bar{\mathbf{e}} \times \bar{\mathbf{h}}^* \cdot \hat{\mathbf{z}} ds, \quad (4.10)$$

where the surface integration is over the cross section of the waveguide. The characteristic impedance is

$$Z_0 = \frac{V^+}{I^+} = \frac{V^-}{I^-} = \frac{C_1}{C_2}, \quad (4.11)$$

since $V^+ = C_1 A$ and $I^+ = C_2 A$, from (4.6a) and (4.6b). If it is desired to have $Z_0 = Z_w$, the wave impedance (Z_{TE} or Z_{TM}) of the mode, then

$$\frac{C_1}{C_2} = Z_w (Z_{TE} \text{ or } Z_{TM}). \quad (4.12a)$$

Alternatively, it may be desirable to normalize the characteristic impedance to unity ($Z_0 = 1$), in which case we have

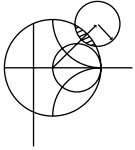
$$\frac{C_1}{C_2} = 1. \quad (4.12b)$$

For a given waveguide mode, (4.10) and (4.12) can be solved for the constants C_1 and C_2 , and equivalent voltages and currents defined. Higher order modes can be treated in the same way, so that a general field in a waveguide can be expressed in the following form:

$$\bar{\mathbf{E}}_t(x, y, z) = \sum_{n=1}^N \left(\frac{V_n^+}{C_{1n}} e^{-j\beta_n z} + \frac{V_n^-}{C_{1n}} e^{j\beta_n z} \right) \bar{\mathbf{e}}_n(x, y), \quad (4.13a)$$

$$\bar{\mathbf{H}}_t(x, y, z) = \sum_{n=1}^N \left(\frac{I_n^+}{C_{2n}} e^{-j\beta_n z} - \frac{I_n^-}{C_{2n}} e^{j\beta_n z} \right) \bar{\mathbf{h}}_n(x, y), \quad (4.13b)$$

where V_n^\pm and I_n^\pm are the equivalent voltages and currents for the n th mode, and C_{1n} and C_{2n} are the proportionality constants for each mode.



EXAMPLE 4.1 EQUIVALENT VOLTAGE AND CURRENT FOR A RECTANGULAR WAVEGUIDE

Find the equivalent voltages and currents for a TE_{10} mode in a rectangular waveguide.

Solution

The transverse field components and power flow of the TE_{10} rectangular waveguide mode and the equivalent transmission line model of this mode can be written as follows:

Waveguide Fields	Transmission Line Model
$E_y = (A^+ e^{-j\beta z} + A^- e^{j\beta z}) \sin \frac{\pi x}{a}$	$V(z) = V^+ e^{-j\beta z} + V^- e^{j\beta z}$
$H_x = \frac{-1}{Z_{TE}} (A^+ e^{-j\beta z} - A^- e^{j\beta z}) \sin \frac{\pi x}{a}$	$I(z) = I^+ e^{-j\beta z} - I^- e^{j\beta z}$ $= \frac{1}{Z_0} (V^+ e^{-j\beta z} - V^- e^{j\beta z})$
$P^+ = \frac{-1}{2} \int_S E_y H_x^* dx dy = \frac{ab}{4Z_{TE}} A^+ ^2$	$P^+ = \frac{1}{2} V^+ I^{+*}$

We now find the constants $C_1 = V^+/A^+ = V^-/A^-$ and $C_2 = I^+/A^+ = I^-/A^-$ that relate the equivalent voltages V^\pm and currents I^\pm to the field amplitudes, A^\pm . Equating incident powers gives

$$\frac{ab |A^+|^2}{4Z_{TE}} = \frac{1}{2} V^+ I^{+*} = \frac{1}{2} |A^+|^2 C_1 C_2^*.$$

If we choose $Z_0 = Z_{TE}$, then we also have that

$$\frac{V^+}{I^+} = \frac{C_1}{C_2} = Z_{TE}.$$

Solving for C_1 , C_2 gives

$$C_1 = \sqrt{\frac{ab}{2}},$$

$$C_2 = \frac{1}{Z_{TE}} \sqrt{\frac{ab}{2}},$$

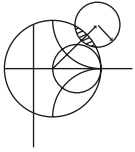
which completes the transmission line equivalence for the TE_{10} mode. ■

The Concept of Impedance

We have used the idea of impedance in several different ways, so it may be useful at this point to summarize this important concept. The term *impedance* was first used by Oliver Heaviside in the nineteenth century to describe the complex ratio V/I in AC circuits consisting of resistors, inductors, and capacitors; the impedance concept quickly became indispensable in the analysis of AC circuits. It was then applied to transmission lines, in terms of lumped-element equivalent circuits and the distributed series impedance and shunt admittance of the line. In the 1930s, S. A. Schelkunoff recognized that the impedance concept could be extended to electromagnetic fields in a systematic way, and noted that impedance should be regarded as characteristic of the type of field, as well as of the medium [2]. In addition, in relation to the analogy between transmission lines and plane wave propagation, impedance may even be dependent on direction. The concept of impedance, then, forms an important link between field theory and transmission line or circuit theory.

We summarize the various types of impedance we have used so far, and their notation:

- $\eta = \sqrt{\mu/\epsilon}$ = intrinsic impedance of the medium. This impedance is dependent only on the material parameters of the medium, and is equal to the wave impedance for plane waves.
- $Z_w = E_t/H_t = 1/Y_w$ = wave impedance. This impedance is a characteristic of the particular type of wave. TEM, TM, and TE waves each have different wave impedances (Z_{TEM} , Z_{TM} , Z_{TE}), which may depend on the type of line or guide, the material, and the operating frequency.
- $Z_0 = 1/Y_0 = V^+/I^+$ = characteristic impedance. Characteristic impedance is the ratio of voltage to current for a traveling wave on a transmission line. Because voltage and current are uniquely defined for TEM waves, the characteristic impedance of a TEM wave is unique. TE and TM waves, however, do not have a uniquely defined voltage and current, so the characteristic impedance for such waves may be defined in different ways.



EXAMPLE 4.2 APPLICATION OF WAVEGUIDE IMPEDANCE

Consider a rectangular waveguide with $a = 2.286$ cm and $b = 1.016$ cm (X-band guide), air filled for $z < 0$ and Rexolite filled ($\epsilon_r = 2.54$) for $z > 0$, as shown in Figure 4.3. If the operating frequency is 10 GHz, use an equivalent transmission line model to compute the reflection coefficient of a TE_{10} wave incident on the interface from $z < 0$.

Solution

The waveguide propagation constants in the air ($z < 0$) and the dielectric ($z > 0$) regions are

$$\beta_a = \sqrt{k_0^2 - \left(\frac{\pi}{a}\right)^2} = 158.0 \text{ m}^{-1},$$

$$\beta_d = \sqrt{\epsilon_r k_0^2 - \left(\frac{\pi}{a}\right)^2} = 304.1 \text{ m}^{-1},$$

where $k_0 = 209.4 \text{ m}^{-1}$.

The reader may verify that the TE_{10} mode is the only propagating mode in either waveguide region. We can set up an equivalent transmission line for the TE_{10} mode in each waveguide, and treat the problem as the reflection of an incident voltage wave at the junction of two infinite transmission lines.

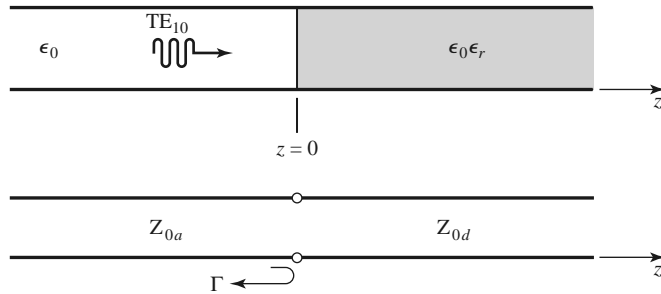


FIGURE 4.3 Geometry of a partially filled waveguide and its transmission line equivalent for Example 4.2.

By Example 4.1 and Table 3.2, the equivalent characteristic impedances for the two lines are

$$Z_{0a} = \frac{k_0 \eta_0}{\beta_a} = \frac{(209.4)(377)}{158.0} = 500.0 \, \Omega,$$

$$Z_{0d} = \frac{k\eta}{\beta_d} = \frac{k_0 \eta_0}{\beta_d} = \frac{(209.4)(377)}{304.1} = 259.6 \, \Omega.$$

The reflection coefficient seen looking into the dielectric filled region is then

$$\Gamma = \frac{Z_{0d} - Z_{0a}}{Z_{0d} + Z_{0a}} = -0.316.$$

With this result, expressions for the incident, reflected, and transmitted waves can be written in terms of fields, or in terms of equivalent voltages and currents. ■

We now consider the arbitrary one-port network shown in Figure 4.4 and derive a general relation between its impedance properties and electromagnetic energy stored in, and the power dissipated by, the network. The complex power delivered to this network is given by (1.91):

$$P = \frac{1}{2} \oint_S \bar{\mathbf{E}} \times \bar{\mathbf{H}}^* \cdot d\bar{\mathbf{s}} = P_\ell + 2j\omega(W_m - W_e), \quad (4.14)$$

where P_ℓ is real and represents the average power dissipated by the network, and W_m and W_e represent the stored magnetic and electric energy, respectively. Note that the unit normal vector in Figure 4.4 is pointing into the volume.

If we define real transverse modal fields $\bar{\mathbf{e}}$ and $\bar{\mathbf{h}}$ over the terminal plane of the network such that

$$\bar{\mathbf{E}}_t(x, y, z) = V(z)\bar{\mathbf{e}}(x, y)e^{-j\beta z}, \quad (4.15a)$$

$$\bar{\mathbf{H}}_t(x, y, z) = I(z)\bar{\mathbf{h}}(x, y)e^{-j\beta z}, \quad (4.15b)$$

with a normalization such that

$$\int_S \bar{\mathbf{e}} \times \bar{\mathbf{h}} \cdot d\bar{\mathbf{s}} = 1,$$

then we can express (4.14) in terms of the terminal voltage and current:

$$P = \frac{1}{2} \int_S V\bar{\mathbf{e}} \times \bar{\mathbf{h}} \cdot d\bar{\mathbf{s}} = \frac{1}{2} V\bar{\mathbf{I}}^*. \quad (4.16)$$

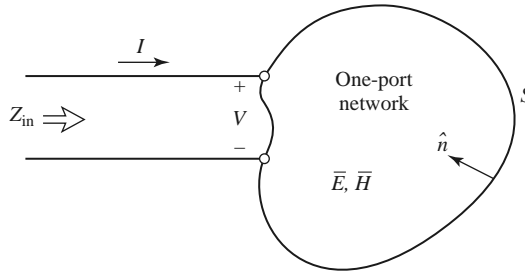


FIGURE 4.4 An arbitrary one-port network.

Then the input impedance is

$$Z_{\text{in}} = R + jX = \frac{V}{I} = \frac{VI^*}{|I|^2} = \frac{P}{\frac{1}{2}|I|^2} = \frac{P_\ell + 2j\omega(W_m - W_e)}{\frac{1}{2}|I|^2}. \quad (4.17)$$

Thus we see that the real part, R , of the input impedance is related to the dissipated power, while the imaginary part, X , is related to the net energy stored in the network. If the network is lossless, then $P_\ell = 0$ and $R = 0$. Then Z_{in} is purely imaginary, with a reactance

$$X = \frac{4\omega(W_m - W_e)}{|I|^2}, \quad (4.18)$$

which is positive for an inductive load ($W_m > W_e$), and negative for a capacitive load ($W_m < W_e$).

Even and Odd Properties of $Z(\omega)$ and $\Gamma(\omega)$

Consider the driving point impedance, $Z(\omega)$, at the input port of an electrical network. The voltage and current at this port are related as $V(\omega) = Z(\omega)I(\omega)$. For an arbitrary frequency dependence, we can find the time-domain voltage by taking the inverse Fourier transform of $V(\omega)$:

$$v(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} V(\omega) e^{j\omega t} d\omega. \quad (4.19)$$

Because $v(t)$ must be real, we have that $v(t) = v^*(t)$, or

$$\int_{-\infty}^{\infty} V(\omega) e^{j\omega t} d\omega = \int_{-\infty}^{\infty} V^*(\omega) e^{-j\omega t} d\omega = \int_{-\infty}^{\infty} V^*(-\omega) e^{j\omega t} d\omega,$$

where the last term was obtained by a change of variable from ω to $-\omega$. This shows that $V(\omega)$ must satisfy the relation

$$V(-\omega) = V^*(\omega), \quad (4.20)$$

which means that $\text{Re}\{V(\omega)\}$ is even in ω , while $\text{Im}\{V(\omega)\}$ is odd in ω . Similar results hold for $I(\omega)$, and for $Z(\omega)$ since

$$V^*(-\omega) = Z^*(-\omega)I^*(-\omega) = Z^*(-\omega)I(\omega) = V(\omega) = Z(\omega)I(\omega).$$

Thus, if $Z(\omega) = R(\omega) + jX(\omega)$, then $R(\omega)$ is even in ω and $X(\omega)$ is odd in ω . These results can also be inferred from (4.17).

Now consider the reflection coefficient at the input port:

$$\Gamma(\omega) = \frac{Z(\omega) - Z_0}{Z(\omega) + Z_0} = \frac{R(\omega) - Z_0 + jX(\omega)}{R(\omega) + Z_0 + jX(\omega)}. \quad (4.21)$$

Then

$$\Gamma(-\omega) = \frac{R(\omega) - Z_0 - jX(\omega)}{R(\omega) + Z_0 - jX(\omega)} = \Gamma^*(\omega), \quad (4.22)$$

which shows that the real and imaginary parts of $\Gamma(\omega)$ are even and odd, respectively, in ω . Finally, the magnitude of the reflection coefficient is

$$|\Gamma(\omega)|^2 = \Gamma(\omega)\Gamma^*(\omega) = \Gamma(\omega)\Gamma(-\omega) = |\Gamma(-\omega)|^2, \quad (4.23)$$

which shows that $|\Gamma(\omega)|^2$ and $|\Gamma(\omega)|$ are even functions of ω . This result implies that only even series of the form $a + b\omega^2 + c\omega^4 + \dots$ can be used to represent $|\Gamma(\omega)|$ or $|\Gamma(\omega)|^2$.

4.2

IMPEDANCE AND ADMITTANCE MATRICES

In the previous section we have seen how equivalent voltages and currents can be defined for TEM and non-TEM waves. Once such voltages and currents have been defined at various points in a microwave network, we can use the impedance and/or admittance matrices of circuit theory to relate these terminal or *port* quantities to each other, and thus to essentially arrive at a matrix description of the network. This type of representation lends itself to the development of equivalent circuits of arbitrary networks, which will be quite useful when we discuss the design of passive components such as couplers and filters. (The term *port* was introduced by H. A. Wheeler in the 1950s to replace the less descriptive and more cumbersome phrase “two-terminal pair” [2, 3].)

We begin by considering an arbitrary N -port microwave network, as depicted in Figure 4.5. The ports in Figure 4.5 may be any type of transmission line or transmission line equivalent of a single propagating waveguide mode. If one of the physical ports of the network is a waveguide supporting more than one propagating mode, additional electrical ports can be added to account for these modes. At a specific point on the n th port, a terminal plane, t_n , is defined along with equivalent voltages and currents for the incident (V_n^+ , I_n^+) and reflected (V_n^- , I_n^-) waves. The terminal planes are important in providing a phase reference for the voltage and current phasors. Now, at the n th terminal plane, the total voltage and current are given by

$$V_n = V_n^+ + V_n^-, \quad (4.24a)$$

$$I_n = I_n^+ - I_n^-, \quad (4.24b)$$

as seen from (4.8) when $z = 0$.

The impedance matrix $[Z]$ of the microwave network then relates these voltages and currents:

$$\begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_N \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{1N} \\ Z_{21} & & & \vdots \\ \vdots & & & \vdots \\ Z_{N1} & \cdots & \cdots & Z_{NN} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_N \end{bmatrix},$$

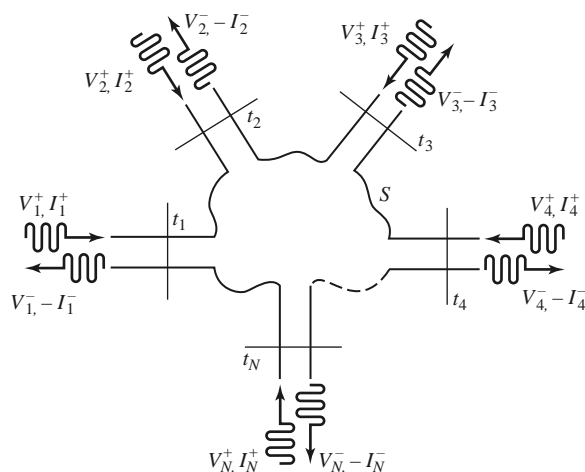


FIGURE 4.5 An arbitrary N -port microwave network.

or in matrix form as

$$[V] = [Z][I]. \quad (4.25)$$

Similarly, we can define an admittance matrix $[Y]$ as

$$\begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_N \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} & \cdots & Y_{1N} \\ Y_{21} & & & \vdots \\ \vdots & & & \vdots \\ Y_{N1} & \cdots & \cdots & Y_{NN} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_N \end{bmatrix},$$

or in matrix form as

$$[I] = [Y][V]. \quad (4.26)$$

Of course, the $[Z]$ and $[Y]$ matrices are the inverses of each other:

$$[Y] = [Z]^{-1}. \quad (4.27)$$

Note that both the $[Z]$ and $[Y]$ matrices relate the total port voltages and currents.

From (4.25), we see that Z_{ij} can be found as

$$Z_{ij} = \left. \frac{V_i}{I_j} \right|_{I_k=0 \text{ for } k \neq j}. \quad (4.28)$$

In words, (4.28) states that Z_{ij} can be found by driving port j with the current I_j , open-circuiting all other ports (so $I_k = 0$ for $k \neq j$), and measuring the open-circuit voltage at port i . Thus, Z_{ii} is the input impedance seen looking into port i when all other ports are open-circuited, and Z_{ij} is the transfer impedance between ports i and j when all other ports are open-circuited.

Similarly, from (4.26), Y_{ij} can be found as

$$Y_{ij} = \left. \frac{I_i}{V_j} \right|_{V_k=0 \text{ for } k \neq j}, \quad (4.29)$$

which states that Y_{ij} can be determined by driving port j with the voltage V_j , short-circuiting all other ports (so $V_k = 0$ for $k \neq j$), and measuring the short-circuit current at port i .

In general, each Z_{ij} or Y_{ij} element may be complex. For an arbitrary N -port network, the impedance and admittance matrices are $N \times N$ in size, so there are $2N^2$ independent quantities or degrees of freedom. In practice, however, many networks are either reciprocal or lossless, or both. If the network is reciprocal (not containing any active devices or nonreciprocal media, such as ferrites or plasmas), we will show that the impedance and admittance matrices are symmetric, so that $Z_{ij} = Z_{ji}$, and $Y_{ij} = Y_{ji}$. If the network is lossless, we can show that all the Z_{ij} or Y_{ij} elements are purely imaginary. Either of these special cases serves to reduce the number of independent quantities or degrees of freedom that an N -port network may have. We now derive the above characteristics for reciprocal and lossless networks.

Reciprocal Networks

Consider the arbitrary network of Figure 4.5 to be reciprocal (no active devices, ferrites, or plasmas), with short circuits placed at all terminal planes except those of ports 1 and 2. Let \bar{E}_a , \bar{H}_a and \bar{E}_b , \bar{H}_b be the fields anywhere in the network due to two independent sources,

a and b , located somewhere in the network. Then the reciprocity theorem of (1.156) states that

$$\oint_S \bar{E}_a \times \bar{H}_b \cdot d\bar{s} = \oint_S \bar{E}_b \times \bar{H}_a \cdot d\bar{s}, \quad (4.30)$$

where S is the closed surface along the boundaries of the network and through the terminal planes of the ports. If the boundary walls of the network and transmission lines are metal, then $\bar{E}_{\text{tan}} = 0$ on these walls (assuming perfect conductors). If the network or the transmission lines are open structures, like microstrip line or slotline, the boundaries of the network can be taken arbitrarily far from the lines so that \bar{E}_{tan} is negligible. Then the only nonzero contribution to the integrals of (4.30) come from the cross-sectional areas of ports 1 and 2.

From Section 4.1, the fields due to sources a and b can be evaluated at the terminal planes t_1 and t_2 as

$$\bar{E}_{1a} = V_{1a}\bar{e}_1, \quad \bar{H}_{1a} = I_{1a}\bar{h}_1, \quad (4.31a)$$

$$\bar{E}_{1b} = V_{1b}\bar{e}_1, \quad \bar{H}_{1b} = I_{1b}\bar{h}_1, \quad (4.31b)$$

$$\bar{E}_{2a} = V_{2a}\bar{e}_2, \quad \bar{H}_{2a} = I_{2a}\bar{h}_2, \quad (4.31c)$$

$$\bar{E}_{2b} = V_{2b}\bar{e}_2, \quad \bar{H}_{2b} = I_{2b}\bar{h}_2, \quad (4.31d)$$

where \bar{e}_1 , \bar{h}_1 and \bar{e}_2 , \bar{h}_2 are the transverse modal fields of ports 1 and 2, respectively, and the V s and I s are the equivalent total voltages and currents. (For instance, \bar{E}_{1b} is the transverse electric field at terminal plane t_1 of port 1 due to source b .) Substituting the fields of (4.31) into (4.30) gives

$$(V_{1a}I_{1b} - V_{1b}I_{1a}) \int_{S_1} \bar{e}_1 \times \bar{h}_1 \cdot d\bar{s} + (V_{2a}I_{2b} - V_{2b}I_{2a}) \int_{S_2} \bar{e}_2 \times \bar{h}_2 \cdot d\bar{s} = 0, \quad (4.32)$$

where S_1 and S_2 are the cross-sectional areas at the terminal planes of ports 1 and 2.

As in Section 4.1, the equivalent voltages and currents have been defined so that the power through a given port can be expressed as $VI^*/2$; then, comparing (4.31) to (4.6) implies that $C_1 = C_2 = 1$ for each port, so that

$$\int_{S_1} \bar{e}_1 \times \bar{h}_1 \cdot d\bar{s} = \int_{S_2} \bar{e}_2 \times \bar{h}_2 \cdot d\bar{s} = 1. \quad (4.33)$$

This reduces (4.32) to

$$V_{1a}I_{1b} - V_{1b}I_{1a} + V_{2a}I_{2b} - V_{2b}I_{2a} = 0. \quad (4.34)$$

Now use the 2×2 admittance matrix of the (effectively) two-port network to eliminate the I s:

$$I_1 = Y_{11}V_1 + Y_{12}V_2,$$

$$I_2 = Y_{21}V_1 + Y_{22}V_2.$$

Substitution into (4.34) gives

$$(V_{1a}V_{2b} - V_{1b}V_{2a})(Y_{12} - Y_{21}) = 0. \quad (4.35)$$

Because the sources a and b are independent, the voltages V_{1a} , V_{1b} , V_{2a} , and V_{2b} can take on arbitrary values. So in order for (4.35) to be satisfied for any choice of sources, we must have $Y_{12} = Y_{21}$, and since the choice of which ports are labeled as 1 and 2 is arbitrary, we have the general result that

$$Y_{ij} = Y_{ji}. \quad (4.36)$$

Then if $[Y]$ is a symmetric matrix, its inverse, $[Z]$, is also symmetric.

Lossless Networks

Now consider a reciprocal lossless N -port junction; we will show that the elements of the impedance and admittance matrices must be pure imaginary. If the network is lossless, then the net real power delivered to the network must be zero. Thus, $\text{Re}\{P_{\text{avg}}\} = 0$, where

$$\begin{aligned} P_{\text{avg}} &= \frac{1}{2}[V]^t[I]^* = \frac{1}{2}([Z][I])^t[I]^* = \frac{1}{2}[I]^t[Z][I]^* \\ &= \frac{1}{2}(I_1 Z_{11} I_1^* + I_1 Z_{12} I_2^* + I_2 Z_{21} I_1^* + \cdots) \\ &= \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N I_m Z_{mn} I_n^*. \end{aligned} \quad (4.37)$$

We have used the result from matrix algebra that $([A][B])^t = [B]^t[A]^t$. Because the I_n are independent, we must have the real part of each self term ($I_n Z_{nn} I_n^*$) equal to zero, since we could set all port currents equal to zero except for the n th current. So,

$$\text{Re}\{I_n Z_{nn} I_n^*\} = |I_n|^2 \text{Re}\{Z_{nn}\} = 0,$$

or

$$\text{Re}\{Z_{nn}\} = 0. \quad (4.38)$$

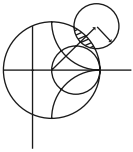
Now let all port currents be zero except for I_m and I_n . Then (4.37) reduces to

$$\text{Re}\{(I_n I_m^* + I_m I_n^*) Z_{mn}\} = 0,$$

since $Z_{mn} = Z_{nm}$. However, $(I_n I_m^* + I_m I_n^*)$ is a purely real quantity that is, in general, nonzero. Thus we must have that

$$\text{Re}\{Z_{mn}\} = 0. \quad (4.39)$$

Then (4.38) and (4.39) imply that $\text{Re}\{Z_{mn}\} = 0$ for any m, n . The reader can verify that this also leads to an imaginary $[Y]$ matrix.



EXAMPLE 4.3 EVALUATION OF IMPEDANCE PARAMETERS

Find the Z parameters of the two-port T-network shown in Figure 4.6.

Solution

From (4.28), Z_{11} can be found as the input impedance of port 1 when port 2 is open-circuited:

$$Z_{11} = \left. \frac{V_1}{I_1} \right|_{I_2=0} = Z_A + Z_C.$$

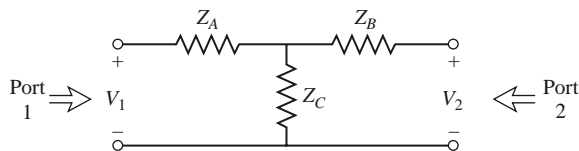


FIGURE 4.6 A two-port T-network.

The transfer impedance Z_{12} can be found measuring the open-circuit voltage at port 1 when a current I_2 is applied at port 2. By voltage division,

$$Z_{12} = \frac{V_1}{I_2} \Big|_{I_1=0} = \frac{V_2}{I_2} \frac{Z_C}{Z_B + Z_C} = Z_C.$$

The reader can verify that $Z_{21} = Z_{12}$, indicating that the circuit is reciprocal. Finally, Z_{22} is found as

$$Z_{22} = \frac{V_2}{I_2} \Big|_{I_1=0} = Z_B + Z_C. \quad \blacksquare$$

4.3

THE SCATTERING MATRIX

We have already discussed the difficulty in defining voltages and currents for non-TEM lines. In addition, a practical problem exists when trying to measure voltages and currents at microwave frequencies because direct measurements usually involve the magnitude (inferred from power) and phase of a wave traveling in a given direction or of a standing wave. Thus, equivalent voltages and currents, and the related impedance and admittance matrices, become somewhat of an abstraction when dealing with high-frequency networks. A representation more in accord with direct measurements, and with the ideas of incident, reflected, and transmitted waves, is given by the scattering matrix.

Like the impedance or admittance matrix for an N -port network, the scattering matrix provides a complete description of the network as seen at its N ports. While the impedance and admittance matrices relate the total voltages and currents at the ports, the scattering matrix relates the voltage waves incident on the ports to those reflected from the ports. For some components and circuits, the scattering parameters can be calculated using network analysis techniques. Otherwise, the scattering parameters can be measured directly with a vector network analyzer; a photograph of a modern network analyzer is shown in Figure 4.7. Once the scattering parameters of the network are known, conversion to other matrix parameters can be performed, if needed.

Consider the N -port network shown in Figure 4.5, where V_n^+ is the amplitude of the voltage wave incident on port n and V_n^- is the amplitude of the voltage wave reflected from port n . The scattering matrix, or $[S]$ matrix, is defined in relation to these incident and reflected voltage waves as

$$\begin{bmatrix} V_1^- \\ V_2^- \\ \vdots \\ V_N^- \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1N} \\ S_{21} & & & \vdots \\ S_{N1} & \cdots & & S_{NN} \\ \vdots & & & \end{bmatrix} \begin{bmatrix} V_1^+ \\ V_2^+ \\ \vdots \\ V_N^+ \end{bmatrix},$$

or

$$[V^-] = [S][V^+]. \quad (4.40)$$

A specific element of the scattering matrix can be determined as

$$S_{ij} = \frac{V_i^-}{V_j^+} \Big|_{V_k^+ = 0 \text{ for } k \neq j}. \quad (4.41)$$

In words, (4.41) says that S_{ij} is found by driving port j with an incident wave of voltage V_j^+ and measuring the reflected wave amplitude V_i^- coming out of port i . The incident

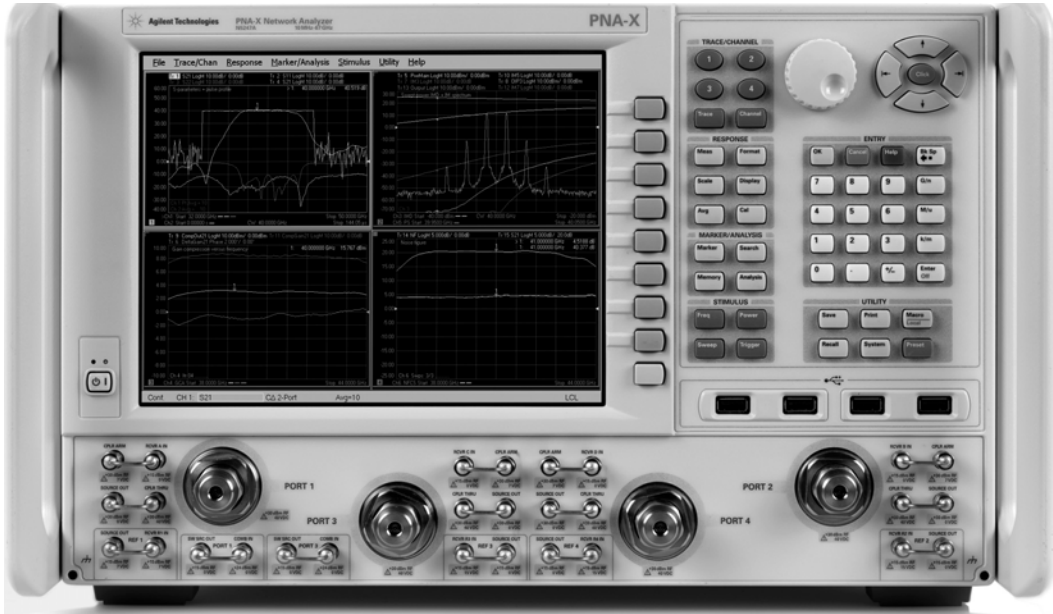
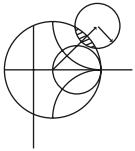


FIGURE 4.7 Photograph of the Agilent N5247A Programmable Network Analyzer. This instrument is used to measure the scattering parameters of RF and microwave networks from 10 MHz to 67 GHz. The instrument is programmable, performs error correction, and has a wide variety of display formats and data conversions.
Courtesy of Agilent Technologies.

waves on all ports except the j th port are set to zero, which means that all ports should be terminated in matched loads to avoid reflections. Thus, S_{ii} is the reflection coefficient seen looking into port i when all other ports are terminated in matched loads, and S_{ij} is the transmission coefficient from port j to port i when all other ports are terminated in matched loads.



EXAMPLE 4.4 EVALUATION OF SCATTERING PARAMETERS

Find the scattering parameters of the 3 dB attenuator circuit shown in Figure 4.8.

Solution

From (4.41), S_{11} can be found as the reflection coefficient seen at port 1 when port 2 is terminated in a matched load ($Z_0 = 50 \Omega$):

$$S_{11} = \left. \frac{V_1^-}{V_1^+} \right|_{V_2^+=0} = \Gamma^{(1)}|_{V_2^+=0} = \frac{Z_{\text{in}}^{(1)} - Z_0}{Z_{\text{in}}^{(1)} + Z_0} \bigg|_{Z_0 \text{ on port 2}},$$

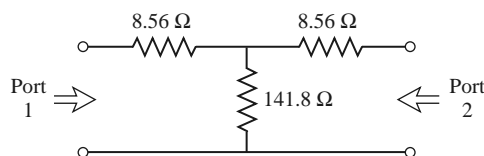


FIGURE 4.8 A matched 3 dB attenuator with a 50Ω characteristic impedance (Example 4.4).

but $Z_{\text{in}}^{(1)} = 8.56 + [141.8(8.56 + 50)]/(141.8 + 8.56 + 50) = 50 \Omega$, so $S_{11} = 0$. Because of the symmetry of the circuit, $S_{22} = 0$.

We can find S_{21} by applying an incident wave at port 1, V_1^+ , and measuring the outgoing wave at port 2, V_2^- . This is equivalent to the transmission coefficient from port 1 to port 2:

$$S_{21} = \left. \frac{V_2^-}{V_1^+} \right|_{V_2^+ = 0}.$$

From the fact that $S_{11} = S_{22} = 0$, we know that $V_1^- = 0$ when port 2 is terminated in $Z_0 = 50 \Omega$, and that $V_2^+ = 0$. In this case we have that $V_1^+ = V_1$ and $V_2^- = V_2$. By applying a voltage V_1 at port 1 and using voltage division twice we find $V_2^- = V_2$ as the voltage across the 50Ω load resistor at port 2:

$$V_2^- = V_2 = V_1 \left(\frac{41.44}{41.44 + 8.56} \right) \left(\frac{50}{50 + 8.56} \right) = 0.707 V_1,$$

where $41.44 = 141.8(58.56)/(141.8 + 58.56)$ is the resistance of the parallel combination of the 50Ω load and the 8.56Ω resistor with the 141.8Ω resistor. Thus, $S_{12} = S_{21} = 0.707$.

If the input power is $|V_1^+|^2/2Z_0$, then the output power is $|V_2^-|^2/2Z_0 = |S_{21}V_1^+|^2/2Z_0 = |S_{21}|^2/2Z_0|V_1^+|^2 = |V_1^+|^2/4Z_0$, which is one-half (-3 dB) of the input power. ■

We now show how the scattering matrix can be determined from the $[Z]$ (or $[Y]$) matrix and vice versa. First, we must assume that the characteristic impedances, Z_{0n} , of all the ports are identical. (This restriction will be removed when we discuss generalized scattering parameters.) Then, for convenience, we can set $Z_{0n} = 1$. From (4.24) the total voltage and current at the n th port can be written as

$$V_n = V_n^+ + V_n^-, \quad (4.42a)$$

$$I_n = I_n^+ - I_n^- = V_n^+ - V_n^-. \quad (4.42b)$$

Using the definition of $[Z]$ from (4.25) with (4.42) gives

$$[Z][I] = [Z][V^+] - [Z][V^-] = [V] = [V^+] + [V^-],$$

which can be rewritten as

$$([Z] + [U])[V^-] = ([Z] - [U])[V^+], \quad (4.43)$$

where $[U]$ is the unit, or identity, matrix defined as

$$[U] = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & \\ 0 & & \cdots & 1 \end{bmatrix}.$$

Comparing (4.43) to (4.40) suggests that

$$[S] = ([Z] + [U])^{-1} ([Z] - [U]), \quad (4.44)$$

giving the scattering matrix in terms of the impedance matrix. Note that for a one-port network (4.44) reduces to

$$S_{11} = \frac{z_{11} - 1}{z_{11} + 1},$$

in agreement with the result for the reflection coefficient seen looking into a load with a normalized input impedance of z_{11} .

To find $[Z]$ in terms of $[S]$, rewrite (4.44) as $[Z][S] + [U][S] = [Z] - [U]$, and solve for $[Z]$ to give

$$[Z] = ([U] + [S]) ([U] - [S])^{-1}. \quad (4.45)$$

Reciprocal Networks and Lossless Networks

As we discussed in Section 4.2, the impedance and admittance matrices are symmetric for reciprocal networks, and are purely imaginary for lossless networks. The scattering matrices for these particular types of networks also have special properties. We will show that the scattering matrix for a reciprocal network is symmetric, and that the scattering matrix for a lossless network is unitary.

By adding (4.42a) and (4.42b) we obtain

$$V_n^+ = \frac{1}{2}(V_n + I_n),$$

or

$$[V^+] = \frac{1}{2}([Z] + [U])[I]. \quad (4.46a)$$

By subtracting (4.42a) and (4.42b) we obtain

$$V_n^- = \frac{1}{2}(V_n - I_n),$$

or

$$[V^-] = \frac{1}{2}([Z] - [U])[I]. \quad (4.46b)$$

Eliminating $[I]$ from (4.46a) and (4.46b) gives

$$[V^-] = ([Z] - [U])([Z] + [U])^{-1}[V^+],$$

so that

$$[S] = ([Z] - [U])([Z] + [U])^{-1}. \quad (4.47)$$

Taking the transpose of (4.47) gives

$$[S]^t = \{([Z] + [U])^{-1}\}^t ([Z] - [U])^t.$$

Now $[U]$ is diagonal, so $[U]^t = [U]$, and if the network is reciprocal, $[Z]$ is symmetric, so that $[Z]^t = [Z]$. The above equation then reduces to

$$[S]^t = ([Z] + [U])^{-1}([Z] - [U]),$$

which is equivalent to (4.44). We have thus shown that

$$[S] = [S]^t, \quad (4.48)$$

so the scattering matrix is symmetric for reciprocal networks.

If the network is lossless, no real power can be delivered to the network. Thus, if the characteristic impedances of all the ports are identical and assumed to be unity, the average power delivered to the network is

$$\begin{aligned} P_{\text{avg}} &= \frac{1}{2} \text{Re}\{[V]^t [I]^*\} = \frac{1}{2} \text{Re}\{([V^+]^t + [V^-]^t)([V^+]^* - [V^-]^*)\} \\ &= \frac{1}{2} \text{Re}\{[V^+]^t [V^+]^* - [V^+]^t [V^-]^* + [V^-]^t [V^+]^* - [V^-]^t [V^-]^*\} \\ &= \frac{1}{2} [V^+]^t [V^+]^* - \frac{1}{2} [V^-]^t [V^-]^* = 0, \end{aligned} \quad (4.49)$$

since the terms $-[V^+]^t [V^-]^* + [V^-]^t [V^+]^*$ are of the form $A - A^*$, and so are purely imaginary. Of the remaining terms in (4.49), $(1/2)[V^+]^t [V^+]^*$ represents the total incident power, while $(1/2)[V^-]^t [V^-]^*$ represents the total reflected power. So, for a lossless junction, we have the intuitive result that the incident and reflected powers are equal:

$$[V^+]^t [V^+]^* = [V^-]^t [V^-]^*. \quad (4.50)$$

Using $[V^-] = [S][V^+]$ in (4.50) gives

$$[V^+]^t [V^+]^* = [V^+]^t [S]^t [S]^* [V^+]^*,$$

so that, for nonzero $[V^+]$,

$$[S]^t [S]^* = [U], \quad (4.51)$$

$$\text{or } [S]^* = \{[S]^t\}^{-1}.$$

A matrix that satisfies the condition of (4.51) is called a *unitary matrix*.

The matrix equation of (4.51) can be written in summation form as

$$\sum_{k=1}^N S_{ki} S_{kj}^* = \delta_{ij}, \text{ for all } i, j, \quad (4.52)$$

where $\delta_{ij} = 1$ if $i = j$, and $\delta_{ij} = 0$ if $i \neq j$, is the Kronecker delta symbol. Thus, if $i = j$, (4.52) reduces to

$$\sum_{k=1}^N S_{ki} S_{ki}^* = 1, \quad (4.53a)$$

while if $i \neq j$, (4.52) reduces to

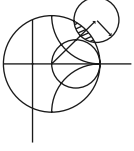
$$\sum_{k=1}^N S_{ki} S_{kj}^* = 0, \text{ for } i \neq j. \quad (4.53b)$$

In words, (4.53a) states that the dot product of any column of $[S]$ with the conjugate of that same column gives unity, while (4.53b) states that the dot product of any column with the

conjugate of a different column gives zero (the columns are orthonormal). From (4.51) we also have that

$$[S][S]^* = [U],$$

so the same statements can be made about the rows of the scattering matrix.



EXAMPLE 4.5 APPLICATION OF SCATTERING PARAMETERS

A two-port network is known to have the following scattering matrix:

$$[S] = \begin{bmatrix} 0.15 \angle 0^\circ & 0.85 \angle -45^\circ \\ 0.85 \angle 45^\circ & 0.2 \angle 0^\circ \end{bmatrix}$$

Determine if the network is reciprocal and lossless. If port 2 is terminated with a matched load, what is the return loss seen at port 1? If port 2 is terminated with a short circuit, what is the return loss seen at port 1?

Solution

Because $[S]$ is not symmetric, the network is not reciprocal. To be lossless, the scattering parameters must satisfy (4.53). Taking the first column [$i = 1$ in (4.53a)] gives

$$|S_{11}|^2 + |S_{21}|^2 = (0.15)^2 + (0.85)^2 = 0.745 \neq 1,$$

so the network is not lossless.

When port 2 is terminated with a matched load, the reflection coefficient seen at port 1 is $\Gamma = S_{11} = 0.15$. So the return loss is

$$\text{RL} = -20 \log |\Gamma| = -20 \log(0.15) = 16.5 \text{ dB}.$$

When port 2 is terminated with a short circuit, the reflection coefficient seen at port 1 can be found as follows. From the definition of the scattering matrix and the fact that $V_2^+ = -V_2^-$ (for a short circuit at port 2), we can write

$$V_1^- = S_{11} V_1^+ + S_{12} V_2^+ = S_{11} V_1^+ - S_{12} V_2^-,$$

$$V_2^- = S_{21} V_1^+ + S_{22} V_2^+ = S_{21} V_1^+ - S_{22} V_2^-.$$

The second equation gives

$$V_2^- = \frac{S_{21}}{1 + S_{22}} V_1^+.$$

Dividing the first equation by V_1^+ and using the above result gives the reflection coefficient seen at port 1 as

$$\begin{aligned} \Gamma = \frac{V_1^-}{V_1^+} &= S_{11} - S_{12} \frac{V_2^-}{V_1^+} = S_{11} - \frac{S_{12} S_{21}}{1 + S_{22}} \\ &= 0.15 - \frac{(0.85 \angle -45^\circ)(0.85 \angle 45^\circ)}{1 + 0.2} = -0.452. \end{aligned}$$

So the return loss is $\text{RL} = -20 \log |\Gamma| = -20 \log(0.452) = 6.9 \text{ dB}$. ■

An important point to understand about scattering parameters is that the reflection coefficient looking into port n is not equal to S_{nn} unless all other ports are matched (this

is illustrated in the above example). Similarly, the transmission coefficient from port m to port n is not equal to S_{nm} unless all other ports are matched. The scattering parameters of a network are properties only of the network itself (assuming the network is linear), and are defined under the condition that all ports are matched. Changing the terminations or excitations of a network does not change its scattering parameters, but may change the reflection coefficient seen at a given port, or the transmission coefficient between two ports.

A Shift in Reference Planes

Because scattering parameters relate amplitudes (magnitude and phase) of traveling waves incident on and reflected from a microwave network, phase reference planes must be specified for each port of the network. We now show how scattering parameters are transformed when the reference planes are moved from their original locations.

Consider the N -port microwave network shown in Figure 4.9, where the original terminal planes are assumed to be located at $z_n = 0$ for the n th port, where z_n is an arbitrary coordinate measured along the transmission line feeding the n th port. The scattering matrix for the network with this set of terminal planes is denoted by $[S]$. Now consider a new set of reference planes defined at $z_n = \ell_n$ for the n th port, and let the new scattering matrix be denoted as $[S']$. Then in terms of the incident and reflected port voltages we have that

$$[V^-] = [S][V^+], \quad (4.54a)$$

$$[V'^-] = [S'] [V'^+], \quad (4.54b)$$

where the unprimed quantities are referenced to the original terminal planes at $z_n = 0$, and the primed quantities are referenced to the new terminal planes at $z_n = \ell_n$.

From the theory of traveling waves on lossless transmission lines we can relate the new wave amplitudes to the original ones as

$$V_n'^+ = V_n^+ e^{j\theta_n}, \quad (4.55a)$$

$$V_n'^- = V_n^- e^{-j\theta_n}, \quad (4.55b)$$

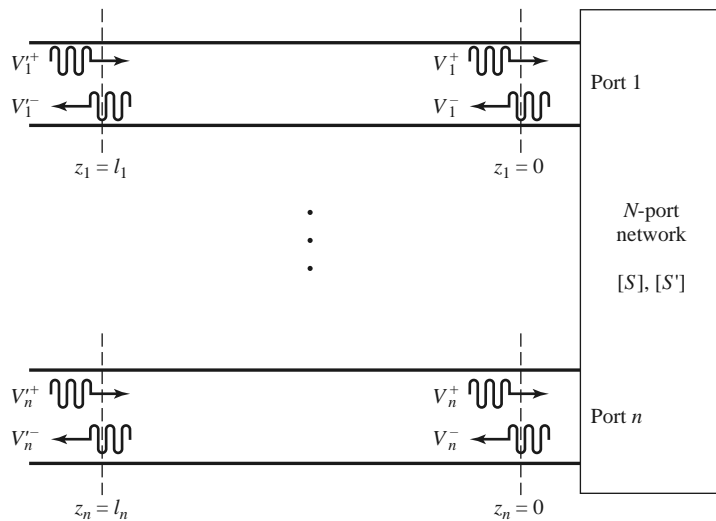


FIGURE 4.9 Shifting reference planes for an N -port network.

where $\theta_n = \beta_n \ell_n$ is the electrical length of the outward shift of the reference plane of port n . Writing (4.55) in matrix form and substituting into (4.54a) gives

$$\begin{bmatrix} e^{j\theta_1} & & 0 \\ & e^{j\theta_2} & \\ & \ddots & \\ 0 & & e^{j\theta_N} \end{bmatrix} [V'^-] = [S] \begin{bmatrix} e^{-j\theta_1} & & 0 \\ & e^{-j\theta_2} & \\ & \ddots & \\ 0 & & e^{-j\theta_N} \end{bmatrix} [V'^+].$$

Multiplying by the inverse of the first matrix on the left gives

$$[V'^-] = \begin{bmatrix} e^{-j\theta_1} & & 0 \\ & e^{-j\theta_2} & \\ & \ddots & \\ 0 & & e^{-j\theta_N} \end{bmatrix} [S] \begin{bmatrix} e^{-j\theta_1} & & 0 \\ & e^{-j\theta_2} & \\ & \ddots & \\ 0 & & e^{-j\theta_N} \end{bmatrix} [V'^+].$$

Comparing with (4.54b) shows that

$$[S'] = \begin{bmatrix} e^{-j\theta_1} & & 0 \\ & e^{-j\theta_2} & \\ & \ddots & \\ 0 & & e^{-j\theta_N} \end{bmatrix} [S] \begin{bmatrix} e^{-j\theta_1} & & 0 \\ & e^{-j\theta_2} & \\ & \ddots & \\ 0 & & e^{-j\theta_N} \end{bmatrix}, \quad (4.56)$$

which is the desired result. Note that $S'_{nn} = e^{-2j\theta_n} S_{nn}$, meaning that the phase of S_{nn} is shifted by twice the electrical length of the shift in terminal plane n because the wave travels twice over this length upon incidence and reflection. This result is consistent with (2.42), which gives the change in the reflection coefficient on a transmission line due to a shift in the reference plane.

Power Waves and Generalized Scattering Parameters

We previously expressed the total voltage and current on a transmission line in terms of the incident and reflected voltage wave amplitudes, as in (2.34) or (4.42):

$$V = V_0^+ + V_0^-, \quad (4.57a)$$

$$I = \frac{1}{Z_0} (V_0^+ - V_0^-), \quad (4.57b)$$

with Z_0 being the characteristic impedance of the line. Inverting (4.57) gives the incident and reflected voltage wave amplitudes in terms of the total voltage and current:

$$V_0^+ = \frac{V + Z_0 I}{2}, \quad (4.58a)$$

$$V_0^- = \frac{V - Z_0 I}{2}. \quad (4.58b)$$

The average power delivered to a load can be expressed as

$$\begin{aligned} P_L &= \frac{1}{2} \operatorname{Re} \{ V I^* \} = \frac{1}{2Z_0} \operatorname{Re} \left\{ |V_0^+|^2 - V_0^+ V_0^{-*} + V_0^{+*} V_0^- - |V_0^-|^2 \right\} \\ &= \frac{1}{2Z_0} (|V_0^+|^2 - |V_0^-|^2), \end{aligned} \quad (4.59)$$

where the last step follows because the quantity $V_0^{+*} V_0^- - V_0^+ V_0^{-*}$ is pure imaginary. This is a physically satisfying result since it expresses the net power delivered to the load as the difference between the incident and reflected powers. Unfortunately, this result is only

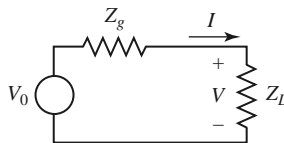


FIGURE 4.10 A generator with impedance Z_g connected to a load impedance Z_L .

valid when the characteristic impedance is real; it does not apply when Z_0 is complex, as in the case of a lossy line. In addition, these results are not useful when no transmission line is present between the generator and load, as in the circuit shown in Figure 4.10.

In the circuit of Figure 4.10 there is no defined characteristic impedance, nor is there a voltage reflection coefficient, or incident and reflected voltage or current waves. It is possible, however, to define a new set of waves, called *power waves*, which have useful properties when dealing with power transfer between a generator and a load, and can be applied to circuits like that of Figure 4.10, as well as to problems with lossless or lossy transmission lines. We will also see how power waves lead to a generalization of scattering parameters.

The incident and reflected power wave amplitudes a and b are defined as the following linear transformations of the total voltage and current:

$$a = \frac{V + Z_R I}{2\sqrt{R_R}}, \quad (4.60a)$$

$$b = \frac{V - Z_R^* I}{2\sqrt{R_R}}, \quad (4.60b)$$

where $Z_R = R_R + jX_R$ is known as the *reference impedance*, and may be complex. Note that the power wave amplitudes of (4.60) are similar in form to the voltage waves of (4.58), but do not have units of power, voltage, or current.

Inverting (4.60) gives the total voltage and current in terms of the power wave amplitudes:

$$V = \frac{Z_R^* a + Z_R b}{\sqrt{R_R}}, \quad (4.61a)$$

$$I = \frac{a - b}{\sqrt{R_R}}. \quad (4.61b)$$

Then the power delivered to the load can be expressed as

$$\begin{aligned} P_L &= \frac{1}{2} \text{Re} \{ V I^* \} = \frac{1}{2R_R} \text{Re} \left\{ Z_R^* |a|^2 - Z_R^* a b^* + Z_R a^* b - Z_R |b|^2 \right\} \\ &= \frac{1}{2} |a|^2 - \frac{1}{2} |b|^2, \end{aligned} \quad (4.62)$$

since the quantity $Z_R a^* b - Z_R^* a b^*$ is pure imaginary. Once again we have the satisfying result that the load power is the difference between the powers of the incident and reflected power waves. It is important to note that this result is valid for any reference impedance Z_R .

The reflection coefficient, Γ_p , for the reflected power wave can be found by using (4.60) and the fact that $V = Z_L I$ at the load:

$$\Gamma_p = \frac{b}{a} = \frac{V - Z_R^* I}{V + Z_R I} = \frac{Z_L - Z_R^*}{Z_L + Z_R}. \quad (4.63)$$

Observe that this reflection coefficient reduces to our usual voltage reflection coefficient of (2.35) when $Z_R = Z_0$ is a real characteristic impedance. Equation (4.63) suggests that

choosing the reference impedance as the conjugate of the load impedance [5],

$$Z_R = Z_L^*, \quad (4.64)$$

will have the useful effect of making the reflected power wave amplitude go to zero.¹

From basic circuit theory, the voltage, current, and load power for the circuit of Figure 4.10 are

$$V = V_0 \frac{Z_L}{Z_L + Z_g}, \quad I = \frac{V_0}{Z_L + Z_g}, \quad P_L = \frac{V_0^2}{2} \frac{R_L}{|Z_L + Z_g|^2}, \quad (4.65a, b, c)$$

where $Z_L = R_L + jX_L$. Then the power wave amplitudes can be found from (4.60), with $Z_R = Z_L^*$, as

$$a = \frac{V + Z_R I}{2\sqrt{R_R}} = V_0 \frac{\frac{Z_L}{Z_L + Z_g} + \frac{Z_L^*}{Z_L + Z_g}}{2\sqrt{R_R}} = V_0 \frac{\sqrt{R_L}}{Z_L + Z_g}, \quad (4.66a)$$

$$b = \frac{V - Z_R^* I}{2\sqrt{R_R}} = V_0 \frac{\frac{Z_L}{Z_L + Z_g} - \frac{Z_L}{Z_L + Z_g}}{2\sqrt{R_R}} = 0. \quad (4.66b)$$

From (4.62) the power delivered to the load is

$$P_L = \frac{1}{2} |a|^2 = \frac{V_0^2}{2} \frac{R_L}{|Z_L + Z_g|^2},$$

in agreement with (4.65c).

When the load is conjugately matched to the generator, so that $Z_g = Z_L^*$, we have $P_L = V_0^2/8R_L$. Note that selecting the reference impedance as $Z_R = Z_L^*$ results in the condition that $b = 0$ (and $\Gamma_p = 0$), but this does not necessarily mean that the load is conjugately matched to the generator, nor that maximum power is delivered to the load. The incident power wave amplitude of (4.66a) depends on Z_L and Z_g , and is maximum only when $Z_g = Z_L^*$.

To define the scattering matrix for power waves for an N -port network, we assume the reference impedance for port i is Z_{Ri} . Then, analogous to (4.60), we define the power wave amplitude vectors in terms of the total voltage and current vectors:

$$[a] = [F] ([V] + [Z_R] [I]), \quad (4.67a)$$

$$[b] = [F] ([V] - [Z_R]^* [I]), \quad (4.67b)$$

where $[F]$ is a diagonal matrix with elements $1/2\sqrt{\text{Re}\{Z_{Ri}\}}$ and $[Z_R]$ is a diagonal matrix with elements Z_{Ri} . By the impedance matrix relation that $[V] = [Z][I]$, (4.67) can be written as

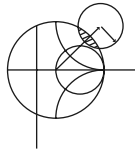
$$[b] = [F] ([Z] - [Z_R]^*) ([Z] + [Z_R])^{-1} [F]^{-1} [a].$$

Because the scattering matrix for power waves, $[S_p]$, should relate $[b]$ to $[a]$, we have

$$[S_p] = [F] ([Z] - [Z_R]^*) ([Z] + [Z_R])^{-1} [F]^{-1}. \quad (4.68)$$

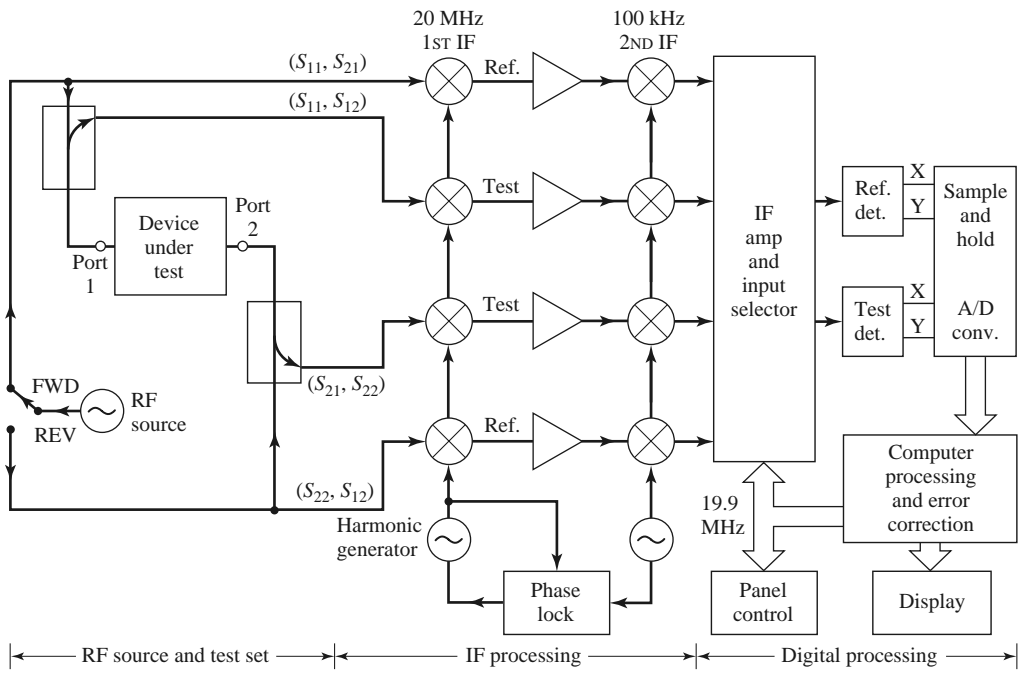
¹ Some authors choose the reference impedance equal to the generator impedance. This has the same effect as (4.64) when the generator and load are conjugately matched, but the choice of (4.64) leads to a zero reflected wave even when the conjugate matching condition is not satisfied, and so can be more useful in general.

The ordinary scattering matrix for a network can first be converted to an impedance matrix, using a relation similar to (4.45), then converted to the generalized power wave scattering matrix using (4.68). The generalized scattering matrix has the useful property that the diagonal elements can be made to be zero by proper selection of the reference impedances.



POINT OF INTEREST: The Vector Network Analyzer

The scattering parameters of passive and active networks can be measured with a *vector network analyzer*, which is a two-channel (or four-channel) microwave receiver designed to process the magnitude and phase of the transmitted and reflected waves from the network. A simplified block diagram of a network analyzer is shown in the accompanying figure. In operation, the RF source is usually set to sweep over a specified bandwidth. A four-port reflectometer samples the incident, reflected, and transmitted RF waves; a switch allows the network to be driven from either port 1 or port 2. Four dual-conversion channels convert these signals to 100-kHz IF frequencies, which are then detected and converted to digital form. An internal computer is used to calculate and display the magnitude and phase of the scattering parameters or other quantities that can be derived from these data, such as SWR, return loss, group delay, impedance, etc. An important feature of the network analyzer is the substantial improvement in accuracy made possible with error-correcting software. Errors caused by directional coupler mismatch, imperfect directivity, loss, and variations in the frequency response of the analyzer system are accounted for by using a 12-term error model and a calibration procedure. Another useful feature is the ability to determine the time-domain response of the network by calculating the inverse Fourier transform of the frequency-domain data.



4.4 THE TRANSMISSION (ABCD) MATRIX

The Z , Y , and S parameter representations can be used to characterize a microwave network with an arbitrary number of ports, but in practice many microwave networks consist of a cascade connection of two or more two-port networks. In this case it is convenient

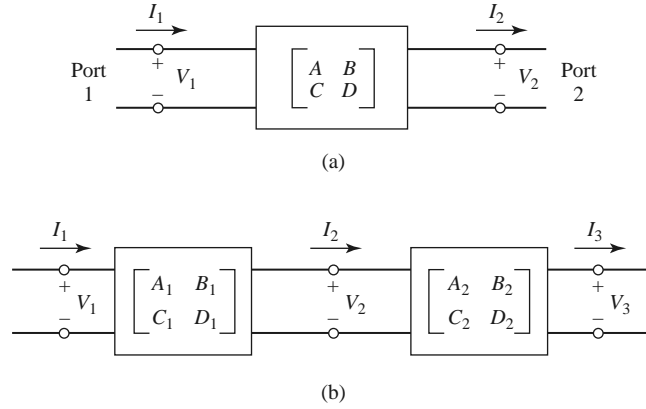


FIGURE 4.11 (a) A two-port network; (b) a cascade connection of two-port networks.

to define a 2×2 *transmission*, or $ABCD$, *matrix*, for each two-port network. We will see that the $ABCD$ matrix of the cascade connection of two or more two-port networks can be easily found by multiplying the $ABCD$ matrices of the individual two-ports.

The $ABCD$ matrix is defined for a two-port network in terms of the total voltages and currents as shown in Figure 4.11a and the following:

$$V_1 = AV_2 + BI_2,$$

$$I_1 = CV_2 + DI_2,$$

or in matrix form as

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} V_2 \\ I_2 \end{bmatrix}. \quad (4.69)$$

It is important to note from Figure 4.11a that a change in the sign convention of I_2 has been made from our previous definitions, which had I_2 as the current flowing *into* port 2. The convention that I_2 flows *out* of port 2 will be used when dealing with $ABCD$ matrices so that in a cascade network I_2 will be the same current that flows into the adjacent network, as shown in Figure 4.11b. Then the left-hand side of (4.69) represents the voltage and current at port 1 of the network, while the column on the right-hand side of (4.69) represents the voltage and current at port 2.

In the cascade connection of two two-port networks shown in Figure 4.11b we have that

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix} \begin{bmatrix} V_2 \\ I_2 \end{bmatrix}, \quad (4.70a)$$

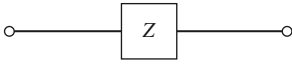
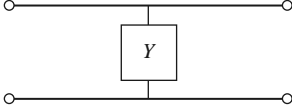
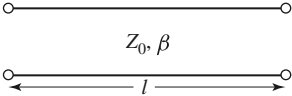
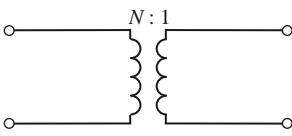
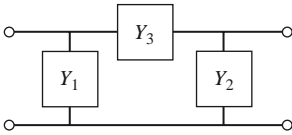
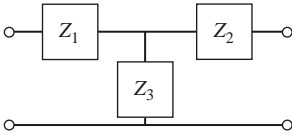
$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix} \begin{bmatrix} V_3 \\ I_3 \end{bmatrix}.$$

Substituting (4.70b) into (4.70a) gives

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix} \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix} \begin{bmatrix} V_3 \\ I_3 \end{bmatrix}, \quad (4.71)$$

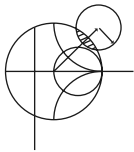
which shows that the $ABCD$ matrix of the cascade connection of the two networks is equal to the product of the $ABCD$ matrices representing the individual two-ports. Note that the

TABLE 4.1 *ABCD* Parameters of Some Useful Two-Port Circuits

Circuit	<i>ABCD</i> Parameters	
	$A = 1$ $C = 0$	$B = Z$ $D = 1$
	$A = 1$ $C = Y$	$B = 0$ $D = 1$
	$A = \cos \beta \ell$ $C = jY_0 \sin \beta \ell$	$B = jZ_0 \sin \beta \ell$ $D = \cos \beta \ell$
	$A = N$ $C = 0$	$B = 0$ $D = \frac{1}{N}$
	$A = 1 + \frac{Y_2}{Y_3}$ $C = Y_1 + Y_2 + \frac{Y_1 Y_2}{Y_3}$	$B = \frac{1}{Y_3}$ $D = 1 + \frac{Y_1}{Y_3}$
	$A = 1 + \frac{Z_1}{Z_3}$ $C = \frac{1}{Z_3}$	$B = Z_1 + Z_2 + \frac{Z_1 Z_2}{Z_3}$ $D = 1 + \frac{Z_2}{Z_3}$

order of multiplication of the matrix must be the same as the order in which the networks are arranged since matrix multiplication is not, in general, commutative.

The usefulness of the *ABCD* matrix representation lies in the fact that a library of *ABCD* matrices for elementary two-port networks can be built up, and applied in building-block fashion to more complicated microwave networks that consist of cascades of these simpler two-ports. Table 4.1 lists a number of useful two-port networks and their *ABCD* matrices.



EXAMPLE 4.6 EVALUATION OF *ABCD* PARAMETERS

Find the *ABCD* parameters of a two-port network consisting of a series impedance *Z* between ports 1 and 2 (the first entry in Table 4.1).

Solution

From the defining relations of (4.69), we have that

$$A = \left. \frac{V_1}{V_2} \right|_{I_2=0},$$

which indicates that A is found by applying a voltage V_1 at port 1, and measuring the open-circuit voltage V_2 at port 2. Thus, $A = 1$. Similarly,

$$B = \left. \frac{V_1}{I_2} \right|_{V_2=0} = \frac{V_1}{V_1/Z} = Z,$$

$$C = \left. \frac{I_1}{V_2} \right|_{I_2=0} = 0,$$

$$D = \left. \frac{I_1}{I_2} \right|_{V_2=0} = \frac{I_1}{I_1} = 1. \quad \blacksquare$$

Relation to Impedance Matrix

The impedance parameters of a network can be easily converted to $ABCD$ parameters. Thus, from the definition of the $ABCD$ parameters in (4.69), and from the defining relations for the Z parameters of (4.25) for a two-port network with I_2 to be consistent with the sign convention used with $ABCD$ parameters,

$$V_1 = I_1 Z_{11} - I_2 Z_{12}, \quad (4.72a)$$

$$V_2 = I_1 Z_{21} - I_2 Z_{22}, \quad (4.72b)$$

we have that

$$A = \left. \frac{V_1}{V_2} \right|_{I_2=0} = \frac{I_1 Z_{11}}{I_1 Z_{21}} = Z_{11}/Z_{21}, \quad (4.73a)$$

$$\begin{aligned} B &= \left. \frac{V_1}{I_2} \right|_{V_2=0} = \left. \frac{I_1 Z_{11} - I_2 Z_{12}}{I_2} \right|_{V_2=0} = Z_{11} \left. \frac{I_1}{I_2} \right|_{V_2=0} - Z_{12} \\ &= Z_{11} \frac{I_1 Z_{22}}{I_1 Z_{21}} - Z_{12} = \frac{Z_{11} Z_{22} - Z_{12} Z_{21}}{Z_{21}}, \end{aligned} \quad (4.73b)$$

$$C = \left. \frac{I_1}{V_2} \right|_{I_2=0} = \frac{I_1}{I_1 Z_{21}} = 1/Z_{21}, \quad (4.73c)$$

$$D = \left. \frac{I_1}{I_2} \right|_{V_2=0} = \frac{I_2 Z_{22}/Z_{21}}{I_2} = Z_{22}/Z_{21}. \quad (4.73d)$$

If the network is reciprocal, then $Z_{12} = Z_{21}$ and (4.73) can be used to show that $AD - BC = 1$.

Equivalent Circuits for Two-Port Networks

The special case of a two-port microwave network occurs so frequently in practice that it deserves further attention. Here we will discuss the use of equivalent circuits to represent an arbitrary two-port network. Useful conversions between two-port network parameters are given in Table 4.2.

Figure 4.12a shows a transition between a coaxial line and a microstrip line, and is an example of a two-port network. Terminal planes can be defined at arbitrary points on the two transmission lines; a convenient choice might be as shown in the figure. However, because of the physical discontinuity in the transition from a coaxial line to a microstrip line, electric and/or magnetic energy can be stored in the vicinity of the junction, leading to reactive effects. Characterization of such effects can be obtained by measurement or by numerical analysis (such analysis may be quite complicated), and represented by the

TABLE 4.2 Conversions Between Two-Port Network Parameters

	S	Z	Y	ABCD
S_{11}	S_{11}	$\frac{(Z_{11} - Z_0)(Z_{22} + Z_0) - Z_{12}Z_{21}}{\Delta Z}$	$\frac{(Y_0 - Y_{11})(Y_0 + Y_{22}) + Y_{12}Y_{21}}{\Delta Y}$	$\frac{A + B/Z_0 - CZ_0 - D}{A + B/Z_0 + CZ_0 + D}$
S_{12}	S_{12}	$\frac{2Z_{12}Z_0}{\Delta Z}$	$\frac{-2Y_{12}Y_0}{\Delta Y}$	$\frac{2(AD - BC)}{A + B/Z_0 + CZ_0 + D}$
S_{21}	S_{21}	$\frac{2Z_{21}Z_0}{\Delta Z}$	$\frac{-2Y_{21}Y_0}{\Delta Y}$	$\frac{2}{A + B/Z_0 + CZ_0 + D}$
S_{22}	S_{22}	$\frac{(Z_{11} + Z_0)(Z_{22} - Z_0) - Z_{12}Z_{21}}{\Delta Z}$	$\frac{(Y_0 + Y_{11})(Y_0 - Y_{22}) + Y_{12}Y_{21}}{\Delta Y}$	$\frac{-A + B/Z_0 - CZ_0 + D}{A + B/Z_0 + CZ_0 + D}$
Z_{11}	$Z_0 \frac{(1 + S_{11})(1 - S_{22}) + S_{12}S_{21}}{(1 - S_{11})(1 - S_{22}) - S_{12}S_{21}}$	Z_{11}	$\frac{Y_{22}}{ Y }$	$\frac{A}{C}$
Z_{12}	$Z_0 \frac{2S_{12}}{(1 - S_{11})(1 - S_{22}) - S_{12}S_{21}}$	Z_{12}	$\frac{-Y_{12}}{ Y }$	$\frac{AD - BC}{C}$
Z_{21}	$Z_0 \frac{2S_{21}}{(1 - S_{11})(1 - S_{22}) - S_{12}S_{21}}$	Z_{21}	$\frac{-Y_{21}}{ Y }$	$\frac{1}{C}$
Z_{22}	$Z_0 \frac{(1 - S_{11})(1 + S_{22}) + S_{12}S_{21}}{(1 - S_{11})(1 - S_{22}) - S_{12}S_{21}}$	Z_{22}	$\frac{Y_{11}}{ Y }$	$\frac{D}{C}$
Y_{11}	$Y_0 \frac{(1 - S_{11})(1 + S_{22}) + S_{12}S_{21}}{(1 + S_{11})(1 + S_{22}) - S_{12}S_{21}}$	$\frac{Z_{22}}{ Z }$	Y_{11}	$\frac{D}{B}$
Y_{12}	$Y_0 \frac{-2S_{12}}{(1 + S_{11})(1 + S_{22}) - S_{12}S_{21}}$	$\frac{-Z_{12}}{ Z }$	Y_{12}	$\frac{BC - AD}{B}$
Y_{21}	$Y_0 \frac{-2S_{21}}{(1 + S_{11})(1 + S_{22}) - S_{12}S_{21}}$	$\frac{-Z_{21}}{ Z }$	Y_{21}	$\frac{-1}{B}$
Y_{22}	$Y_0 \frac{(1 + S_{11})(1 - S_{22}) + S_{12}S_{21}}{(1 + S_{11})(1 + S_{22}) - S_{12}S_{21}}$	$\frac{Z_{11}}{ Z }$	Y_{22}	$\frac{A}{B}$
A	$\frac{(1 + S_{11})(1 - S_{22}) + S_{12}S_{21}}{2S_{21}}$	$\frac{Z_{11}}{Z_{21}}$	$\frac{-Y_{22}}{Y_{21}}$	A
B	$Z_0 \frac{(1 + S_{11})(1 + S_{22}) - S_{12}S_{21}}{2S_{21}}$	$\frac{ Z }{Z_{21}}$	$\frac{-1}{Y_{21}}$	B
C	$\frac{1}{Z_0} \frac{(1 - S_{11})(1 - S_{22}) - S_{12}S_{21}}{2S_{21}}$	$\frac{1}{Z_{21}}$	$\frac{- Y }{Y_{21}}$	C
D	$\frac{(1 - S_{11})(1 + S_{22}) + S_{12}S_{21}}{2S_{21}}$	$\frac{Z_{22}}{Z_{21}}$	$\frac{-Y_{11}}{Y_{21}}$	D
$ Z = Z_{11}Z_{22} - Z_{12}Z_{21}; \quad Y = Y_{11}Y_{22} - Y_{12}Y_{21}; \quad \Delta Y = (Y_{11} + Y_0)(Y_{22} + Y_0) - Y_{12}Y_{21}; \quad \Delta Z = (Z_{11} + Z_0)(Z_{22} + Z_0) - Z_{12}Z_{21}; \quad Y_0 = 1/Z_0.$				

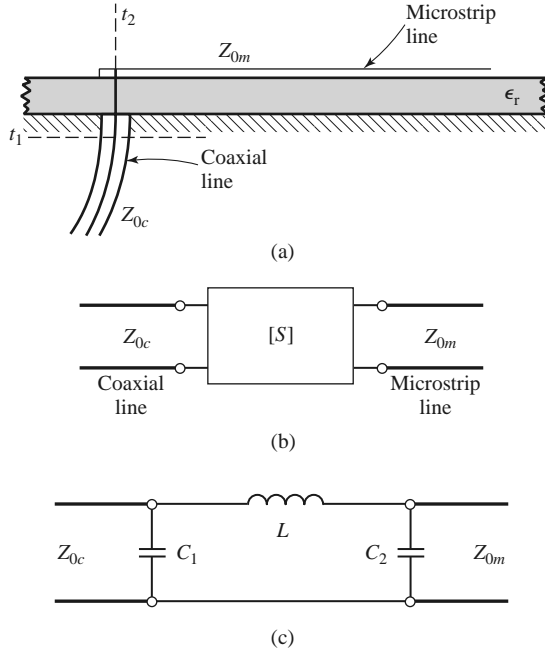


FIGURE 4.12 A coax-to-microstrip transition and equivalent circuit representations. (a) Geometry of the transition. (b) Representation of the transition by a “black box.” (c) A possible equivalent circuit for the transition [6].

two-port “black box” shown in Figure 4.12b. The properties of the transition can then be expressed in terms of the network parameters (Z , Y , S , or $ABCD$) of the two-port network. This type of treatment can be applied to a variety of two-port junctions, such as transitions from one type of transmission line to another, transmission line discontinuities such as step changes in width or bends, etc. When modeling a microwave junction in this way, it is often useful to replace the two-port “black box” with an equivalent circuit containing a few idealized components, as shown in Figure 4.12c. This is particularly useful if the component values can be related to some physical features of the actual junction. There is an unlimited number of ways in which such equivalent circuits can be defined; we will discuss some of the most common and useful types below.

As we have seen, an arbitrary two-port network can be described in terms of impedance parameters as

$$\begin{aligned} V_1 &= Z_{11}I_1 + Z_{12}I_2, \\ V_2 &= Z_{21}I_1 + Z_{22}I_2, \end{aligned} \quad (4.74a)$$

or in terms of admittance parameters as

$$\begin{aligned} I_1 &= Y_{11}V_1 + Y_{12}V_2, \\ I_2 &= Y_{21}V_1 + Y_{22}V_2. \end{aligned} \quad (4.74b)$$

If the network is reciprocal, then $Z_{12} = Z_{21}$ and $Y_{12} = Y_{21}$. These representations lead naturally to the T and π equivalent circuits shown in Figures 4.13a and 4.13b. The relations in Table 4.2 can be used to relate the component values to other network parameters.

Other equivalent circuits can also be used to represent a two-port network. If the network is reciprocal, there are six degrees of freedom (the real and imaginary parts of three matrix elements), so the equivalent circuit should have six independent parameters.

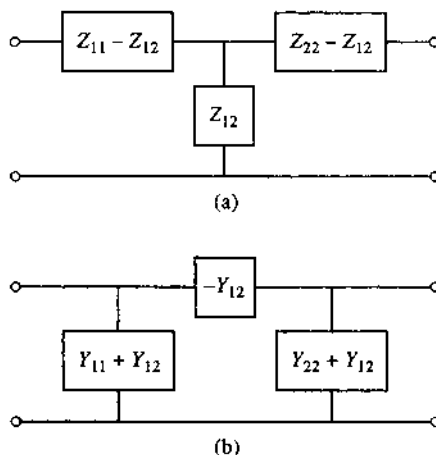


FIGURE 4.13 Equivalent circuits for a reciprocal two-port network. (a) T equivalent. (b) π equivalent.

A nonreciprocal network cannot be represented by a passive equivalent circuit using reciprocal elements.

If the network is lossless, which is a good approximation for many practical two-port junctions, some simplifications can be made in the equivalent circuit. As was shown in Section 4.2, the impedance or admittance matrix elements are purely imaginary for a lossless network. This reduces the degrees of freedom for such a network to three, and implies that the T and π equivalent circuits of Figure 4.13 can be constructed from purely reactive elements.

4.5

SIGNAL FLOW GRAPHS

We have seen how transmitted and reflected waves can be represented by scattering parameters, and how the interconnection of sources, networks, and loads can be treated with various matrix representations. In this section we discuss the *signal flow graph*, which is an additional technique that is very useful for the analysis of microwave networks in terms of transmitted and reflected waves. We first discuss the features and the construction of the flow graph itself, and then present a technique for the reduction, or solution, of the flow graph.

The primary components of a signal flow graph are nodes and branches:

- **Nodes:** Each port i of a microwave network has two nodes, a_i and b_i . Node a_i is identified with a wave entering port i , while node b_i is identified with a wave reflected from port i . The voltage at a node is equal to the sum of all signals entering that node.
- **Branches:** A branch is a directed path between two nodes representing signal flow from one node to another. Every branch has an associated scattering parameter or reflection coefficient.

At this point it is useful to consider the flow graph of an arbitrary two-port network, as shown in Figure 4.14. Figure 4.14a shows a two-port network with incident and reflected waves at each port, and Figure 4.14b shows the corresponding signal flow graph representation. The flow graph gives an intuitive graphical illustration of the network behavior.

For example, a wave of amplitude a_1 incident at port 1 is split, with part going through S_{11} and out port 1 as a reflected wave, and part transmitted through S_{21} to node b_2 .

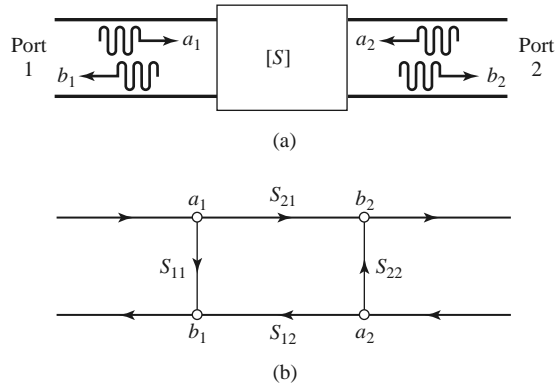


FIGURE 4.14 The signal flow graph representation of a two-port network. (a) Definition of incident and reflected waves. (b) Signal flow graph.

At node b_2 , the wave goes out port 2; if a load with nonzero reflection coefficient is connected at port 2, this wave will be at least partly reflected and reenter the two-port network at node a_2 . Part of this wave can be reflected back out port 2 via S_{22} , and part can be transmitted out port 1 through S_{12} .

Two other special networks—a one-port network and a voltage source—are shown in Figure 4.15, along with their signal flow graph representations. Once a microwave network has been represented in signal flow graph form, it is a relatively easy matter to solve for the ratio of any combination of wave amplitudes. We will discuss how this can be done using four basic decomposition rules, but the same results can also be obtained using Mason's rule from control system theory.

Decomposition of Signal Flow Graphs

A signal flow graph can be reduced to a single branch between two nodes using the following four basic decomposition rules to obtain any desired wave amplitude ratio.

- **Rule 1 (Series Rule).** Two branches, whose common node has only one incoming and one outgoing wave (branches in series), may be combined to form a single branch whose coefficient is the product of the coefficients of the original branches.

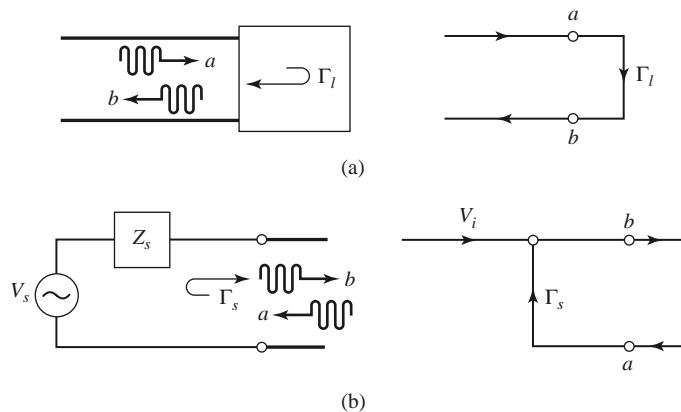


FIGURE 4.15 The signal flow graph representations of a one-port network and a source. (a) A one-port network and its flow graph. (b) A source and its flow graph.

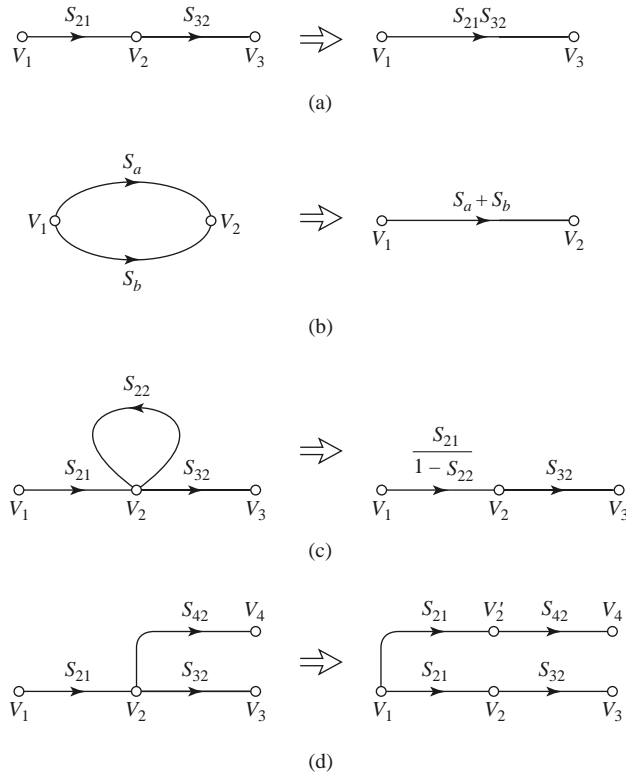


FIGURE 4.16 Decomposition rules. (a) Series rule. (b) Parallel rule. (c) Self-loop rule. (d) Splitting rule.

Figure 4.16a shows the flow graphs for this rule. Its derivation follows from the basic relation

$$V_3 = S_{32} V_2 = S_{32} S_{21} V_1. \quad (4.75)$$

- **Rule 2 (Parallel Rule).** Two branches from one common node to another common node (branches in parallel) may be combined into a single branch whose coefficient is the sum of the coefficients of the original branches. Figure 4.16b shows the flow graphs for this rule. The derivation follows from the obvious relation

$$V_2 = S_a V_1 + S_b V_1 = (S_a + S_b) V_1. \quad (4.76)$$

- **Rule 3 (Self-Loop Rule).** When a node has a self-loop (a branch that begins and ends on the same node) of coefficient S , the self-loop can be eliminated by multiplying coefficients of the branches feeding that node by $1/(1 - S)$. Figure 4.16c shows the flow graphs for this rule, which can be derived as follows. From the original network we have

$$V_2 = S_{21} V_1 + S_{22} V_2, \quad (4.77a)$$

$$V_3 = S_{32} V_2. \quad (4.77b)$$

Eliminating V_2 gives

$$V_3 = \frac{S_{32} S_{21}}{1 - S_{22}} V_1, \quad (4.78)$$

which is seen to be the transfer function for the reduced graph of Figure 4.16c.

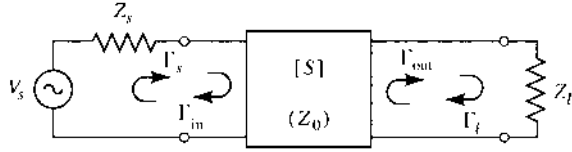


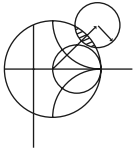
FIGURE 4.17 A terminated two-port network.

- **Rule 4 (Splitting Rule).** A node may be split into two separate nodes as long as the resulting flow graph contains, once and only once, each combination of separate (not self-loops) input and output branches that connect to the original node. This rule is illustrated in Figure 4.16d and follows from the observation that

$$V_4 = S_{42}V_2 = S_{21}S_{42}V_1 \quad (4.79)$$

in both the original flow graph and the flow graph with the split node.

We now illustrate the use of each of these rules with an example.



EXAMPLE 4.7 APPLICATION OF SIGNAL FLOW GRAPH

Use signal flow graphs to derive expressions for Γ_{in} and Γ_{out} for the microwave network shown in Figure 4.17.

Solution

The signal flow graph for the circuit of Figure 4.17 is shown in Figure 4.18. In terms of node voltages, Γ_{in} is given by the ratio b_1/a_1 . The first two steps of the required decomposition of the flow graph are shown in Figures 4.19a and 4.19b, from which the desired result follows by inspection:

$$\Gamma_{in} = \frac{b_1}{a_1} = S_{11} + \frac{S_{12}S_{21}\Gamma_L}{1 - S_{22}\Gamma_L}.$$

Next, Γ_{out} is given by the ratio b_2/a_2 . The first two steps for this decomposition are shown in Figures 4.19c and 4.19d. The desired result is

$$\Gamma_{out} = \frac{b_2}{a_2} = S_{22} + \frac{S_{12}S_{21}\Gamma_s}{1 - S_{11}\Gamma_s} \quad \blacksquare$$

Application to Thru-Reflect-Lin Network Analyzer Calibration

As a further application of signal flow graphs we consider the calibration of a network analyzer using the *Thru-Reflect-Line* (TRL) technique [7]. The general problem is shown in Figure 4.20, where it is intended to measure the scattering parameters of a two-port device

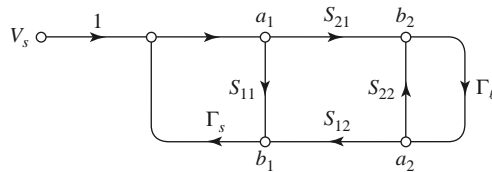


FIGURE 4.18 Signal flow graph for the two-port network with general source and load impedances of Figure 4.17.

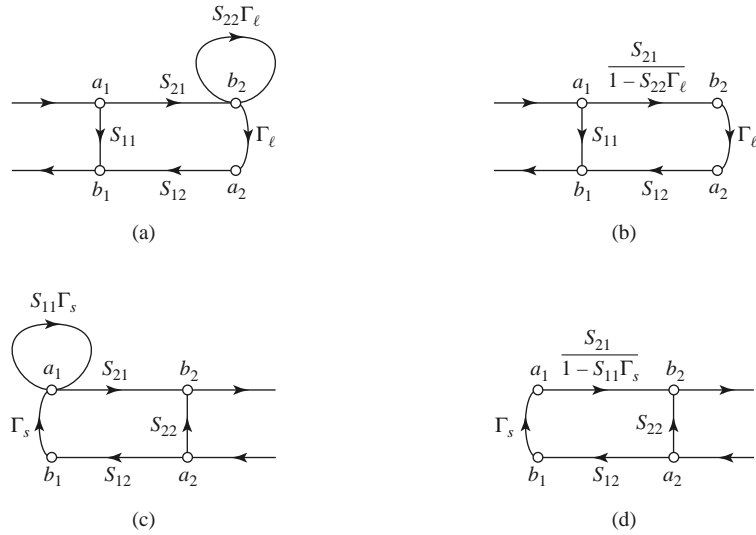


FIGURE 4.19 Decompositions of the flow graph of Figure 4.18 to find $\Gamma_{\text{in}} = b_1/a_1$ and $\Gamma_{\text{out}} = b_2/a_2$. (a) Using Rule 4 on node a_2 . (b) Using Rule 3 for the self-loop at node b_2 . (c) Using Rule 4 on node b_1 . (d) Using Rule 3 for the self-loop at node a_1 .

at the indicated reference planes. As discussed in the previous Point of Interest, a network analyzer measures scattering parameters as ratios of complex voltage amplitudes. The primary reference plane for such measurements is generally at some point within the analyzer itself, so the measurement will include losses and phase delays caused by the effects of the connectors, cables, and transitions that must be used to connect the device under test (DUT) to the analyzer. In the block diagram of Figure 4.20 these effects are lumped together in a two-port *error box* placed at each port between the actual measurement reference plane and the desired reference plane for the two-port DUT. A calibration procedure is used to characterize the error boxes before measurement of the DUT; then the actual error-corrected scattering parameters of the DUT can be calculated from the measured data. Measurement of a one-port network can be considered as a reduced version of the two-port case.

The simplest way to calibrate a network analyzer is to use three or more known loads, such as shorts, opens, and matched loads. The problem with this approach is that such standards are always imperfect to some degree, and therefore introduce errors into the measurement. These errors become increasingly significant at higher frequencies and as the quality of the measurement system improves. The TRL calibration scheme does not

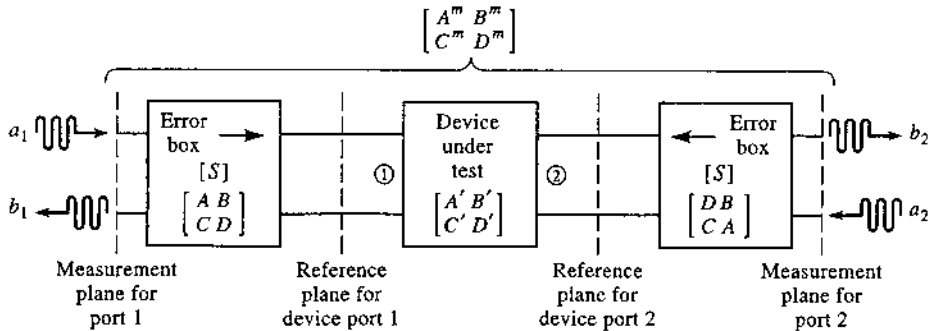


FIGURE 4.20 Block diagram of a network analyzer measurement of a two-port device.

rely on known standard loads, but uses three simple connections to allow the error boxes to be characterized completely. These three connections are shown in Figure 4.21. The *Thru* connection is made by directly connecting port 1 to port 2 at the desired reference planes. The *Reflect* connection uses a load having a large reflection coefficient, Γ_L , such as a nominal open or short. It is not necessary to know the exact value of Γ_L , as this will be determined by the TRL calibration procedure. The *Line* connection involves connecting ports 1 and 3 together through a length of matched transmission line. It is not necessary to know the length of the line, and it is not required that the line be lossless; these parameters will be determined by the TRL procedure.

We can use signal flow graphs to derive the set of equations necessary to find the scattering parameters for the error boxes in the TRL calibration procedure. With reference to Figure 4.20, we will apply the *Thru*, *Reflect*, and *Line* connections at the reference plane for the DUT, and measure the scattering parameters for these three cases at the measurement planes. For simplicity, we assume the same characteristic impedance for ports 1 and 2, and that the error boxes are reciprocal and identical for both ports. The error boxes are characterized by a scattering matrix $[S]$ and, alternatively, by an $ABCD$ matrix. Thus $S_{21} = S_{12}$ for both error boxes. Also note that ports 1 and 2 of the error boxes are in opposite positions since they are symmetrically connected, as shown in the figure. To avoid confusion in notation we will denote the measured scattering parameters for the *Thru*, *Reflect*, and *Line* connections as the $[T]$, $[R]$, and $[L]$ matrices, respectively.

Figure 4.21a shows the arrangement for the *Thru* connection and the corresponding signal flow graph. Observe that we have made use of the fact that $S_{21} = S_{12}$ and that the error boxes are identical and symmetrically arranged. The signal flow graph can be easily reduced using the decomposition rules to give the measured scattering parameters at the measurement planes in terms of the scattering parameters of the error boxes as

$$T_{11} = \left. \frac{b_1}{a_1} \right|_{a_2=0} = S_{11} + \frac{S_{22}S_{12}^2}{1 - S_{22}^2} \quad (4.80a)$$

$$T_{12} = \left. \frac{b_1}{a_2} \right|_{a_1=0} = \frac{S_{12}^2}{1 - S_{22}^2} \quad (4.80b)$$

By symmetry we have $T_{22} = T_{11}$, and by reciprocity we have $T_{21} = T_{12}$.

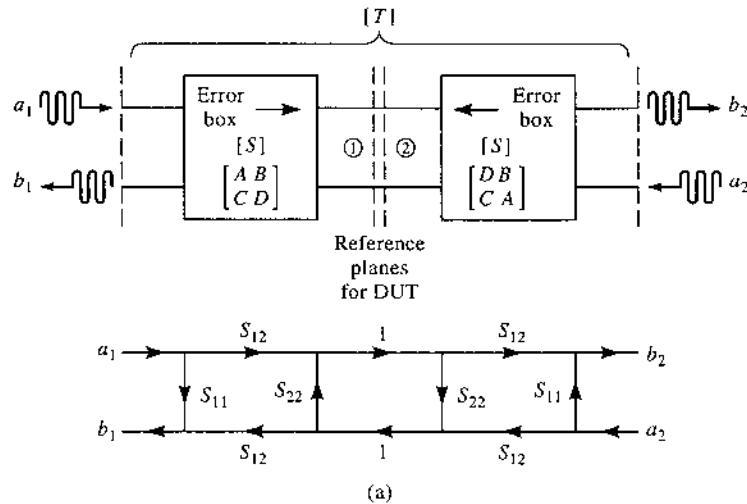


FIGURE 4.21a Block diagram and signal flow graph for the *Thru* connection.

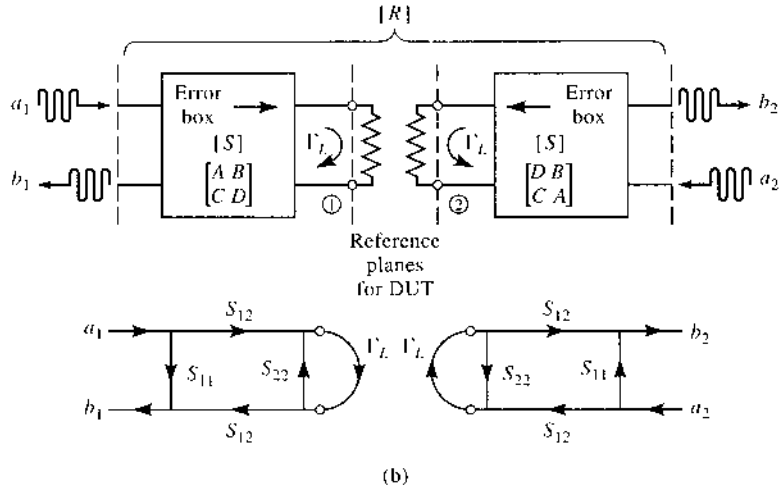


FIGURE 4.21b Block diagram and signal flow graph for the *Reflect* connection.

The *Reflect* connection is shown in Figure 4.21b, with the corresponding signal flow graph. Note that this arrangement effectively decouples the two measurement ports, so $R_{12} = R_{21} = 0$. The signal flow graph can be easily reduced to show that

$$R_{11} = \left. \frac{b_1}{a_1} \right|_{a_2=0} = S_{11} + \frac{S_{12}^2 \Gamma_L}{1 - S_{22} \Gamma_L}. \quad (4.81)$$

By symmetry we have $R_{22} = R_{11}$.

The *Line* connection is shown in Figure 4.21c, with its corresponding signal flow graph. A reduction similar to that used for the *Thru* case gives

$$L_{11} = \left. \frac{b_1}{a_1} \right|_{a_2=0} = S_{11} + \frac{S_{22} S_{12}^2 e^{-2\gamma l}}{1 - S_{22}^2 e^{-2\gamma l}}, \quad (4.82a)$$

$$L_{12} = \left. \frac{b_1}{a_2} \right|_{a_1=0} = \frac{S_{12}^2 e^{-\gamma l}}{1 - S_{22}^2 e^{-2\gamma l}}. \quad (4.82b)$$

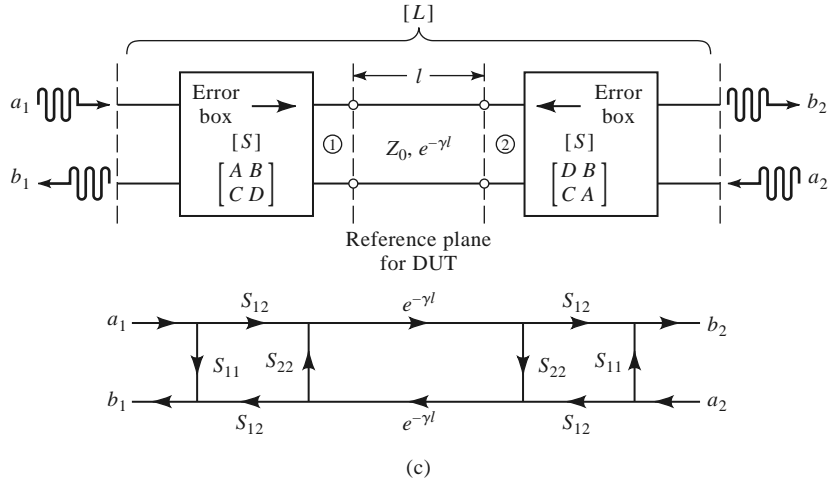


FIGURE 4.21c Block diagram and signal flow graph for the *Line* connection.

By symmetry and reciprocity we have $L_{22} = L_{11}$ and $L_{21} = L_{12}$.

We now have five equations (4.80)–(4.82) for the five unknowns S_{11} , S_{12} , S_{22} , Γ_L , and e^{γ^ℓ} ; the solution is straightforward but lengthy. Because (4.81) is the only equation that contains Γ_L , we can first solve the four equations in (4.80) and (4.82) for the other four unknowns. Equation (4.80b) can be used to eliminate S_{12} from (4.80a) and (4.82), and then S_{11} can be eliminated from (4.80a) and (4.82a). This leaves two equations for S_{22} and e^{γ^ℓ} :

$$L_{12}e^{2\gamma^\ell} - L_{12}S_{22}^2 = T_{12}e^{\gamma^\ell} - T_{12}S_{22}^2e^{\gamma^\ell}, \quad (4.83a)$$

$$e^{2\gamma^\ell}(T_{11} - S_{22}T_{12}) - T_{11}S_{22}^2 = L_{11}(e^{2\gamma^\ell} - S_{22}^2) - S_{22}T_{12}. \quad (4.83b)$$

Equation (4.83a) can be solved for S_{22} and substituted into (4.83b) to give a quadratic equation for e^{γ^ℓ} . Application of the quadratic formula then gives the solution for e^{γ^ℓ} in terms of the measured TRL scattering parameters as

$$e^{\gamma^\ell} = \frac{L_{12}^2 + T_{12}^2 - (T_{11} - L_{11})^2 \pm \sqrt{[L_{12}^2 + T_{12}^2 - (T_{11} - L_{11})^2]^2 - 4L_{12}^2T_{12}^2}}{2L_{12}T_{12}}. \quad (4.84)$$

The choice of sign can be determined by the requirement that the real and imaginary parts of γ be positive, or by knowing the phase of Γ_L [as determined from (4.83)] to within 180° .

Now multiply (4.80b) by S_{22} and subtract from (4.80a) to get

$$T_{11} = S_{11} + S_{22}T_{12}, \quad (4.85a)$$

and similarly multiply (4.82b) by $S_{22}e^{-\gamma^\ell}$ and subtract from (4.82a) to get

$$L_{11} = S_{11} + S_{22}L_{12}e^{-\gamma^\ell}. \quad (4.85b)$$

Eliminating S_{11} from these two equations gives S_{22} in terms of $e^{-\gamma^\ell}$ as

$$S_{22} = \frac{T_{11} - L_{11}}{T_{12} - L_{12}e^{-\gamma^\ell}}. \quad (4.86)$$

Solving (4.85a) for S_{11} gives

$$S_{11} = T_{11} - S_{22}T_{12}, \quad (4.87)$$

and solving (4.80b) for S_{12} gives

$$S_{12}^2 = T_{12}(1 - S_{22}^2). \quad (4.88)$$

Finally, (4.81) can be solved for Γ_L to give

$$\Gamma_L = \frac{R_{11} - S_{11}}{S_{12}^2 + S_{22}(R_{11} - S_{11})}. \quad (4.89)$$

Equations (4.84) and (4.86)–(4.89) give the scattering parameters for the error boxes, as well as the unknown reflection coefficient Γ_L (to within the sign), and the propagation factor $e^{-\gamma^\ell}$. This completes the calibration procedure for the TRL method.

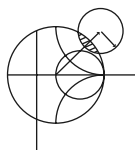
The scattering parameters of the DUT can now be measured at the measurement reference planes shown in Figure 4.20, and corrected using the above TRL error box parameters

to give the scattering parameters at the reference planes of the DUT. Because we are working with a cascade of three two-port networks, it is convenient to use $ABCD$ parameters. Thus, we convert the error box scattering parameters to the corresponding $ABCD$ parameters, and convert the measured scattering parameters of the cascade to the corresponding $A^m B^m C^m D^m$ parameters. If we use $A' B' C' D'$ to denote the parameters for the DUT, then we have

$$\begin{bmatrix} A^m & B^m \\ C^m & D^m \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix} \begin{bmatrix} D & B \\ C & A \end{bmatrix},$$

where the change in the elements of the last matrix account for the reversal of ports for the error box at port 2 of the DUT (see Problem 4.25). Then the $ABCD$ parameters for the DUT can be determined as

$$\begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} \begin{bmatrix} A^m & B^m \\ C^m & D^m \end{bmatrix} \begin{bmatrix} D & B \\ C & A \end{bmatrix}^{-1}. \quad (4.90)$$



POINT OF INTEREST: Computer-Aided Design for Microwave Circuits

Computer-aided design (CAD) software packages have become essential tools for the analysis, design, and optimization of RF and microwave circuits and systems. Several microwave CAD products are commercially available, including Microwave Office (Applied Wave Research), ADS (Agilent Technologies), Microwave Studio (CST), Designer (Ansoft), and many others. RF and microwave CAD packages can be divided into two types: those that use “physics-based” solutions, where Maxwell’s equations are numerically solved for physical geometries such as printed circuit geometries or waveguides, and “circuit-based” solutions, which use equivalent circuits for various elements, including distributed elements, discontinuities, coupled lines, and active devices. Some packages combine these two approaches. Both linear and nonlinear modeling, as well as circuit optimization, are generally possible. Although such computer programs can be fast, powerful, and accurate, they cannot serve as a substitute for engineering experience and a good understanding of microwave principles.

A typical design process usually begins with specifications or design goals for the circuit or system. Based on previous designs and his or her experience, an engineer can develop an initial design, including specific components and a circuit layout. CAD can then be used to model and analyze the design, using data for each of the components and including effects such as loss and discontinuities. The software can be used to optimize the design by adjusting some of the circuit parameters to achieve the best performance. If the specifications are not met, the design may have to be revised. CAD tools can also be used to study the effects of component tolerances and errors to improve circuit reliability and robustness. When the design meets the specifications, an engineering prototype can be built and tested. If the measured results satisfy the specifications, the design process is completed. Otherwise the design will need to be revised and the procedure repeated.

Without CAD tools the design process would require the construction and measurement of laboratory prototypes at each iteration, which is expensive and time consuming. Thus, CAD can greatly decrease the time and cost of a design while enhancing its quality. The simulation and optimization process is especially important for monolithic microwave integrated circuits because these circuits cannot easily be tuned or trimmed after fabrication.

CAD techniques are not without limitations, however. Of primary importance is the fact that any computer model is only an approximation to a “real-world” physical circuit and cannot completely account for the inevitable differences due to component and fabrication tolerances, surface roughness, spurious coupling, higher order modes, junction discontinuities, thermal effects, and a number of other practical issues that can occur with a physical circuit or device.

4.6

DISCONTINUITIES AND MODAL ANALYSIS

By either necessity or design, microwave circuits and networks often consist of transmission lines with various types of discontinuities. In some cases discontinuities are an unavoidable result of mechanical or electrical transitions from one medium to another (e.g., a junction between two waveguides, or a coax-to-microstrip transition), and the discontinuity effect is unwanted but may be significant enough to warrant characterization. In other cases discontinuities may be deliberately introduced into the circuit to perform a certain electrical function (e.g., reactive diaphragms in waveguide, or stubs on a microstrip line for matching or filter circuits). In any event, a transmission line discontinuity can be represented as an equivalent circuit at some point on the transmission line. Depending on the type of discontinuity, the equivalent circuit may be a simple shunt or series element across the line or, in the more general case, a T- or π -equivalent circuit may be required. The component values of an equivalent circuit depend on the parameters of the line and the discontinuity, as well as on the frequency of operation. In some cases the equivalent circuit involves a shift in the phase reference planes on the transmission lines. Once the equivalent circuit of a given discontinuity is known, its effect can be incorporated into the analysis or design of the network using the theory developed previously in this chapter.

The purpose of the present section is to discuss how equivalent circuits are obtained for transmission line discontinuities; we will see that one approach is to start with a field theory solution to a canonical discontinuity problem and develop a circuit model with component values. This is thus another example of our objective of replacing complicated field analyses with circuit concepts. In other cases, it may be easier to measure the network parameters of an isolated discontinuity.

Figures 4.22 and 4.23 show some common transmission line discontinuities and their equivalent circuits. As shown in Figures 4.22a–4.22c, thin metallic diaphragms (or “irises”) can be placed in the cross section of a waveguide to yield equivalent shunt inductance, capacitance, or a resonant combination. Similar effects occur with step changes in the height or width of the waveguide, as shown in Figures 4.22d and 4.22e. Similar discontinuities can also be made in circular waveguide. The classic reference for waveguide discontinuities and their equivalent circuits is the *Waveguide Handbook* [8].

Some typical microstrip discontinuities and transitions are shown in Figure 4.23; similar geometries exist for stripline and other printed transmission lines such as slotline, covered microstrip, coplanar waveguide, etc. Although approximate equivalent circuits have been developed for some printed transmission line discontinuities [9], many do not lend themselves to easy or accurate modeling, and must be treated by numerical analysis. Modern CAD tools are usually capable of accurately modeling such problems.

Modal Analysis of an H -Plane Step in Rectangular Waveguide

The field analysis of most transmission line discontinuity problems is difficult, and beyond the scope of this book. The technique of waveguide *modal analysis*, however, is relatively straightforward and similar in principle to the reflection/transmission problems that were discussed in Chapters 1 and 2. In addition, modal analysis is a rigorous and versatile technique that can be applied to a number of waveguide and coax discontinuity problems, and lends itself well to computer implementation. We will illustrate the technique by applying it to the problem of finding the equivalent circuit of an H -plane step (change in width) in a rectangular waveguide.

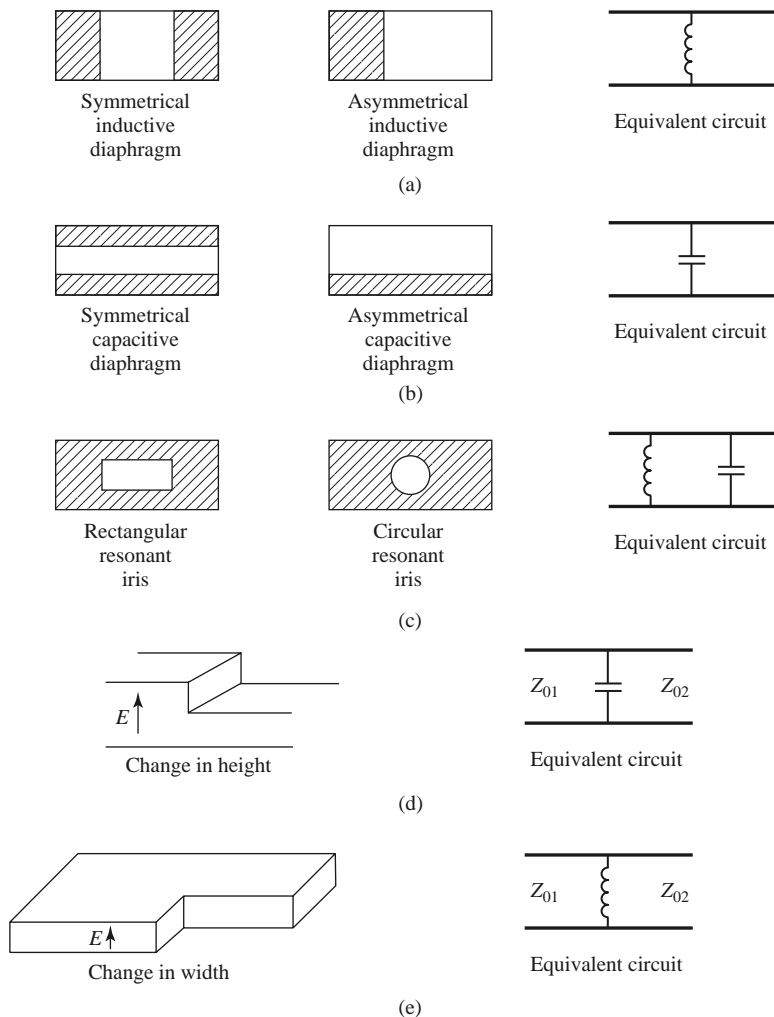


FIGURE 4.22 Rectangular waveguide discontinuities.

The geometry of the H -plane waveguide step is shown in Figure 4.24. It is assumed that only the dominant TE_{10} mode is propagating in guide 1 ($z < 0$) and is incident on the junction from $z < 0$. It is also assumed that no modes are propagating in guide 2, although the analysis to follow is still valid if propagation can occur in guide 2. From Section 3.3, the transverse components of the incident TE_{10} mode can be written, for $z < 0$, as

$$E_y^i = \sin \frac{\pi x}{a} e^{-j\beta_1^a z}, \quad (4.91a)$$

$$H_x^i = \frac{-1}{Z_1^a} \sin \frac{\pi x}{a} e^{-j\beta_1^a z}, \quad (4.91b)$$

where

$$\beta_n^a = \sqrt{k_0^2 - \left(\frac{n\pi}{a}\right)^2} \quad (4.92)$$

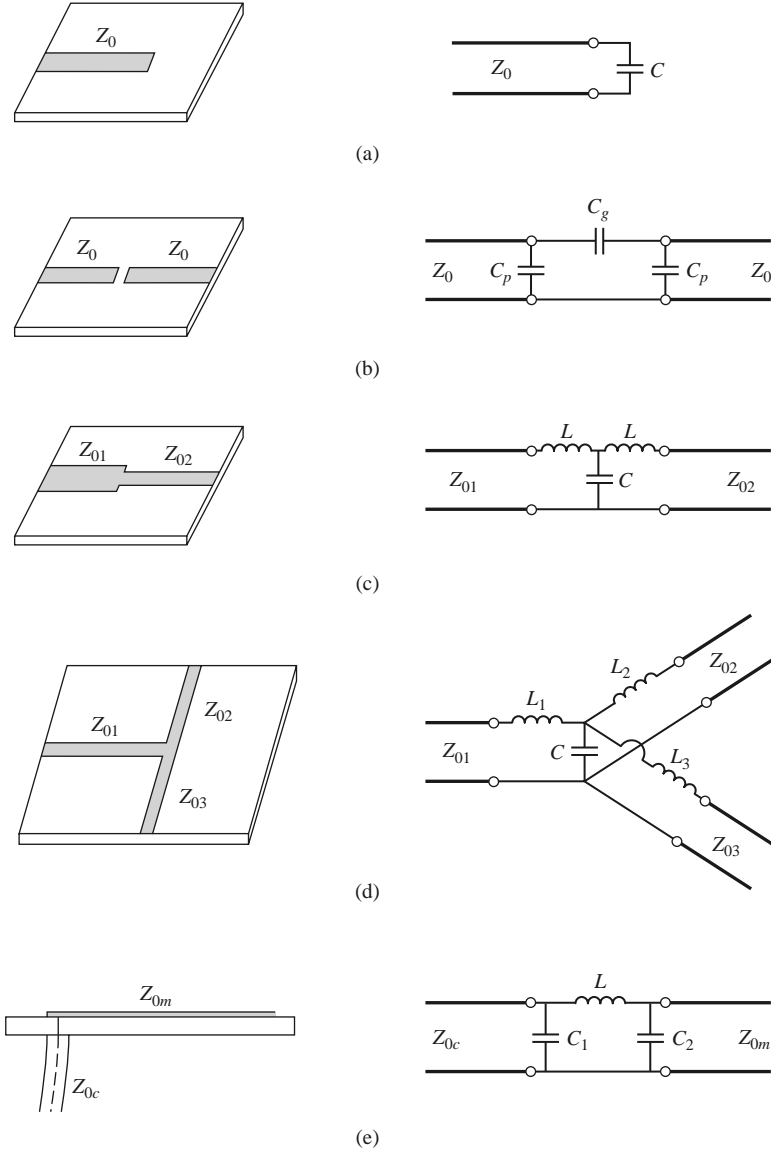


FIGURE 4.23 Some common microstrip discontinuities. (a) Open-ended microstrip. (b) Gap in microstrip. (c) Change in width. (d) T-junction. (e) Coax-to-microstrip junction.

is the propagation constant of the TE_{n0} mode in guide 1 (of width a), and

$$Z_n^a = \frac{k_0 \eta_0}{\beta_n^a} \quad (4.93)$$

is the wave impedance of the TE_{n0} mode in guide 1. Because of the discontinuity at $z = 0$ there will be reflected and transmitted waves in both guides, consisting of infinite sets of TE_{n0} modes in guides 1 and 2. Only the TE_{10} mode will propagate in guide 1, but higher order modes are also important in this problem because they account for stored energy,

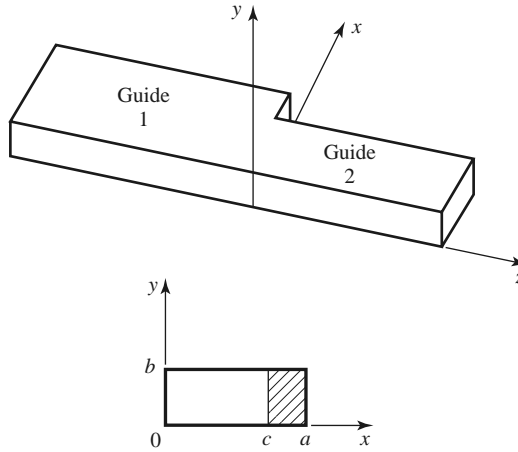


FIGURE 4.24 Geometry of an H -plane step (change in width) in a rectangular waveguide.

localized near $z = 0$. Because there is no y variation introduced by this discontinuity, TE_{nm} modes for $m \neq 0$ are not excited, nor are any TM modes. A more general discontinuity, however, may excite such modes.

The reflected modes in guide 1 may be written, for $z < 0$, as

$$E_y^r = \sum_{n=1}^{\infty} A_n \sin \frac{n\pi x}{a} e^{j\beta_n^a z}, \quad (4.94a)$$

$$H_x^r = \sum_{n=1}^{\infty} \frac{A_n}{Z_n^a} \sin \frac{n\pi x}{a} e^{j\beta_n^a z}, \quad (4.94b)$$

where A_n is the unknown amplitude coefficient of the reflected TE_{n0} mode in guide 1. The reflection coefficient of the incident TE_{10} mode is then A_1 . Similarly, the transmitted modes into guide 2 can be written, for $z > 0$, as

$$E_y^t = \sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{c} e^{-j\beta_n^c z}, \quad (4.95a)$$

$$H_x^t = - \sum_{n=1}^{\infty} \frac{B_n}{Z_n^c} \sin \frac{n\pi x}{c} e^{-j\beta_n^c z}, \quad (4.95b)$$

where the propagation constant in guide 2 is

$$\beta_n^c = \sqrt{k_0^2 - \left(\frac{n\pi}{c}\right)^2}, \quad (4.96)$$

and the wave impedance in guide 2 is

$$Z_n^c = \frac{k_0 \eta_0}{\beta_n^c}. \quad (4.97)$$

At $z = 0$, the transverse fields (E_y , H_x) must be continuous for $0 < x < c$; in addition, E_y must be zero for $c < x < a$ because of the step. Enforcing these boundary conditions leads to the following equations:

$$E_y = \sin \frac{\pi x}{a} + \sum_{n=1}^{\infty} A_n \sin \frac{n\pi x}{a} = \begin{cases} \sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{c} & \text{for } 0 < x < c, \\ 0 & \text{for } c < x < a, \end{cases} \quad (4.98a)$$

$$H_x = \frac{-1}{Z_1^a} \sin \frac{\pi x}{a} + \sum_{n=1}^{\infty} \frac{A_n}{Z_n^a} \sin \frac{n\pi x}{a} = - \sum_{n=1}^{\infty} \frac{B_n}{Z_n^c} \sin \frac{n\pi x}{c} \text{ for } 0 < x < c. \quad (4.98b)$$

Equations (4.98a) and (4.98b) constitute a doubly infinite set of linear equations for the modal coefficients A_n and B_n . We will first eliminate the B_n and then truncate the resulting equation to a finite number of terms and solve for the A_n .

Multiplying (4.98a) by $\sin(m\pi x/a)$, integrating from $x = 0$ to a , and using the orthogonality relations from Appendix D yields

$$\frac{a}{2} \delta_{m1} + \frac{a}{2} A_m = \sum_{n=1}^{\infty} B_n I_{mn} = \sum_{k=1}^{\infty} B_k I_{mk}, \quad (4.99)$$

where

$$I_{mn} = \int_{x=0}^c \sin \frac{m\pi x}{a} \sin \frac{n\pi x}{c} dx \quad (4.100)$$

is an integral that can be easily evaluated, and

$$\delta_{mn} = \begin{cases} 1 & \text{if } m = n \\ 0 & \text{if } m \neq n \end{cases} \quad (4.101)$$

is the Kronecker delta symbol. Now solve (4.98b) for B_k by multiplying (4.98b) by $\sin(k\pi x/c)$ and integrating from $x = 0$ to c . After using orthogonality relations, we obtain

$$\frac{-1}{Z_1^a} I_{k1} + \sum_{n=1}^{\infty} \frac{A_n}{Z_n^a} I_{kn} = \frac{-c B_k}{2 Z_k^c}. \quad (4.102)$$

Substituting B_k from (4.102) into (4.99) gives an infinite set of linear equations for the A_n , where $m = 1, 2, \dots$,

$$\frac{a}{2} A_m + \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \frac{2 Z_k^c I_{mk} I_{kn} A_n}{c Z_n^a} = \sum_{k=1}^{\infty} \frac{2 Z_k^c I_{mk} I_{k1}}{c Z_1^a} - \frac{a}{2} \delta_{m1}. \quad (4.103)$$

For numerical calculation we can truncate these summations to N terms, which will result in N linear equations for the first N coefficients, A_n . For example, let $N = 1$. Then (4.103) reduces to

$$\frac{a}{2} A_1 + \frac{2 Z_1^c I_{11}^2}{c Z_1^a} A_1 = \frac{2 Z_1^c I_{11}^2}{c Z_1^a} - \frac{a}{2}. \quad (4.104)$$

Solving for A_1 (the reflection coefficient of the incident TE₁₀ mode) gives

$$A_1 = \frac{Z_\ell - Z_1^a}{Z_\ell + Z_1^a} \text{ for } N = 1, \quad (4.105)$$

where $Z_\ell = 4Z_1^c I_{11}^2 / ac$, which looks like an effective load impedance to guide 1. Accuracy is improved by using larger values of N and leads to a set of equations that can be written in matrix form as

$$[Q][A] = [P], \quad (4.106)$$

where $[Q]$ is a square $N \times N$ matrix of coefficients,

$$Q_{mn} = \frac{a}{2} \delta_{mn} + \sum_{k=1}^N \frac{2Z_k^c I_{mk} I_{kn}}{cZ_n^a}, \quad (4.107)$$

$[P]$ is an $N \times 1$ column vector of coefficients given by

$$P_m = \sum_{k=1}^N \frac{2Z_k^c I_{mk} I_{k1}}{cZ_1^a} - \frac{a}{2} \delta_{m1}, \quad (4.108)$$

and $[A]$ is an $N \times 1$ column vector of the coefficients A_n . After the A_n are found, the B_n can be calculated from (4.102), if desired. Equations (4.106)–(4.108) lend themselves well to computer implementation, and Figure 4.25 shows the results of such a calculation for various matrix sizes.

If the width c of guide 2 is such that all modes are cut off (evanescent), then no real power can be transmitted into guide 2, and all the incident power is reflected back into guide 1. The evanescent fields on both sides of the discontinuity store reactive power, however, which implies that the step discontinuity and guide 2 beyond the discontinuity look like a reactance (in this case an inductive reactance) to an incident TE_{10} mode in guide 1. Thus the equivalent circuit of the H -plane step looks like a shunt inductor at the $z = 0$ plane of guide 1, as shown in Figure 4.22e. The equivalent reactance can be found from the reflection coefficient A_1 [after solving (4.106)] as

$$X = -jZ_1^a \frac{1 + A_1}{1 - A_1}. \quad (4.109)$$

Figure 4.25 shows the normalized equivalent inductance versus the ratio of the guide widths c/a for a free-space wavelength $\lambda = 1.4a$ and for $N = 1, 2$, and 10 equations. The

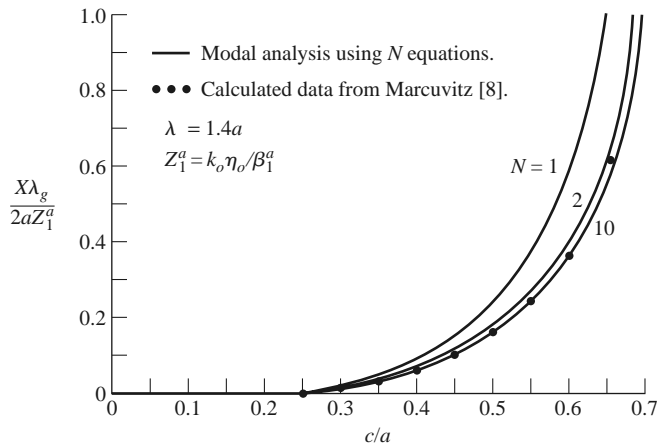


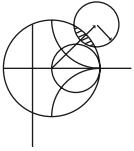
FIGURE 4.25 Equivalent inductance of an H -plane asymmetric step.

modal analysis results are compared to data from reference [8]. Note that the solution converges very quickly (because of the fast exponential decay of the higher order evanescent modes), and that the result using just two modes is very close to the data of reference [8].

The fact that the H -plane step appears inductive is a result of the actual value of the reflection coefficient, A_1 , but we can verify the inductive nature of the discontinuity by computing the complex power flow into the evanescent modes on either side of the discontinuity. For example, the complex power flow into guide 2 can be found as

$$\begin{aligned}
 P &= \int_{x=0}^c \int_{y=0}^b \bar{\mathbf{E}} \times \bar{\mathbf{H}}^* \Big|_{z=0^+} \cdot \hat{\mathbf{z}} dx dy \\
 &= -b \int_{x=0}^c E_y H_x^* dx \\
 &= -b \int_{x=0}^c \left[\sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{c} \right] \left[- \sum_{m=1}^{\infty} \frac{B_m^*}{Z_m^{c*}} \sin \frac{m\pi x}{c} \right] dx \\
 &= \frac{bc}{2} \sum_{n=1}^{\infty} \frac{|B_n|^2}{Z_n^{c*}} \\
 &= \frac{jbc}{2k_0\eta_0} \sum_{n=1}^{\infty} |B_n|^2 |\beta_n^c|,
 \end{aligned} \tag{4.110}$$

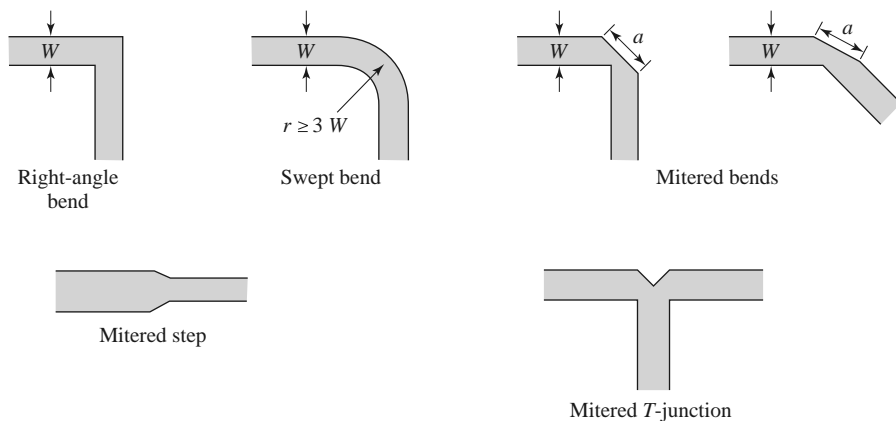
where the orthogonality property of the sine functions was used, as well as (4.95)–(4.97). Equation (4.110) shows that the complex power flow into guide 2 is positive imaginary, implying stored magnetic energy and an inductive reactance. A similar result can be derived for the evanescent modes in guide 1; this is left as a problem.



POINT OF INTEREST: Microstrip Discontinuity Compensation

Because a microstrip circuit is easy to fabricate and allows the convenient integration of passive and active components, many types of microwave circuits and subsystems are made in microstrip form. One problem with microstrip circuits (and other planar circuits) is that the inevitable discontinuities at bends, step changes in widths, and junctions can cause degradation in circuit performance. This is because such discontinuities introduce parasitic reactances that can lead to phase and amplitude errors, input and output mismatch, and possibly spurious coupling or radiation. One approach for eliminating such effects is to construct an equivalent circuit for the discontinuity (perhaps by measurement), including it in the design of the circuit, and compensating for its effect by adjusting other circuit parameters (such as line lengths and characteristic impedances, or tuning stubs). Another approach is to minimize the effect of a discontinuity by compensating the discontinuity directly, often by chamfering or mitering the conductor.

Consider the case of a bend in a microstrip line. The straightforward right-angle bend shown below has a parasitic discontinuity capacitance caused by the increased conductor area at the corner of the bend. This effect could be eliminated by making a smooth, “swept” bend with a radius $r \geq 3W$, but this takes up more space. Alternatively, the right-angle bend can be compensated by mitering the corner, which has the effect of reducing the excess capacitance at the bend. As shown later, this technique can be applied to bends of arbitrary angle. The optimum value of the miter length, a , depends on the characteristic impedance and the bend angle, but a value of $a = 1.8W$ is often used in practice. The technique of mitering can also be used to compensate step and T-junction discontinuities, as shown on the next page.



Reference: T. C. Edwards, *Foundations for Microwave Circuit Design*, John Wiley & Sons, New York, 1981.

4.7

EXCITATION OF WAVEGUIDES—ELECTRIC AND MAGNETIC CURRENTS

So far we have considered the propagation, reflection, and transmission of guided waves in the absence of sources, but obviously the waveguide or transmission line must be coupled to a generator or some other source of power. For TEM or quasi-TEM lines, there is usually only one propagating mode that can be excited by a given source, although there may be reactance (stored energy) associated with a given feed. In the waveguide case, it may be possible for several propagating modes to be excited, along with evanescent modes that store energy. In this section we will develop a formalism for determining the excitation of a given waveguide mode due to an arbitrary electric or magnetic current source. This theory can then be used to find the excitation and input impedance of probe and loop feeds and, in the next section, to determine the excitation of waveguides by apertures.

Current Sheets That Excite Only One Waveguide Mode

Consider an infinitely long rectangular waveguide with a transverse sheet of electric surface current density at $z = 0$, as shown in Figure 4.26. First assume that this current has \hat{x} and \hat{y} components given as

$$\bar{J}_s^{\text{TE}}(x, y) = -\hat{x} \frac{2A_{mn}^+ n\pi}{b} \cos \frac{m\pi x}{a} \sin \frac{n\pi y}{b} + \hat{y} \frac{2A_{mn}^+ m\pi}{a} \sin \frac{m\pi x}{a} \cos \frac{n\pi y}{b}. \quad (4.111)$$

We will show that such a current excites a single TE_{mn} waveguide mode traveling away from the current source in both the $+z$ and $-z$ directions.

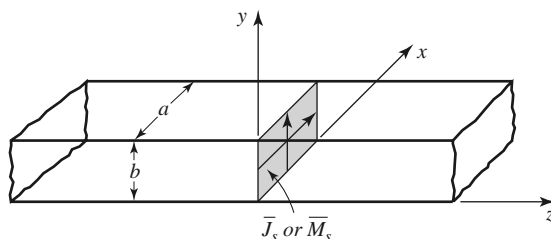


FIGURE 4.26 An infinitely long rectangular waveguide with surface current densities at $z = 0$.

From Table 3.2, the transverse fields for positive and negative traveling TE_{mn} waveguide modes can be written as

$$E_x^\pm = Z_{TE} \left(\frac{n\pi}{b} \right) A_{mn}^\pm \cos \frac{m\pi x}{a} \sin \frac{n\pi y}{b} e^{\mp j\beta z}, \quad (4.112a)$$

$$E_y^\pm = -Z_{TE} \left(\frac{m\pi}{a} \right) A_{mn}^\pm \sin \frac{m\pi x}{a} \cos \frac{n\pi y}{b} e^{\mp j\beta z}, \quad (4.112b)$$

$$H_x^\pm = \pm \left(\frac{m\pi}{a} \right) A_{mn}^\pm \sin \frac{m\pi x}{a} \cos \frac{n\pi y}{b} e^{\mp j\beta z}, \quad (4.112c)$$

$$H_y^\pm = \pm \left(\frac{n\pi}{b} \right) A_{mn}^\pm \cos \frac{m\pi x}{a} \sin \frac{n\pi y}{b} e^{\mp j\beta z}, \quad (4.112d)$$

where the \pm notation refers to waves traveling in the $+z$ direction or $-z$ direction with amplitude coefficients A_{mn}^+ and A_{mn}^- , respectively.

From (1.36) and (1.37), the following boundary conditions must be satisfied at $z = 0$:

$$(\bar{E}^+ - \bar{E}^-) \times \hat{z} = 0, \quad (4.113a)$$

$$\hat{z} \times (\bar{H}^+ - \bar{H}^-) = \bar{J}_s. \quad (4.113b)$$

Equation (4.112a) states that the transverse components of the electric field must be continuous at $z = 0$, which when applied to (4.112a) and (4.112b), gives

$$A_{mn}^+ = A_{mn}^-. \quad (4.114)$$

Equation (4.113b) states that the discontinuity in the transverse magnetic field is equal to the electric surface current density. Thus, the surface current density at $z = 0$ must be

$$\begin{aligned} \bar{J}_s &= \hat{y}(H_x^+ - H_x^-) - \hat{x}(H_y^+ - H_y^-) \\ &= -\hat{x} \frac{2A_{mn}^+ n\pi}{b} \cos \frac{m\pi x}{a} \sin \frac{n\pi y}{b} + \hat{y} \frac{2A_{mn}^+ m\pi}{a} \sin \frac{m\pi x}{a} \cos \frac{n\pi y}{b}, \end{aligned} \quad (4.115)$$

where (4.114) was used. This current is seen to be the same as the current of (4.111), which shows, by the uniqueness theorem, that such a current will excite only the TE_{mn} mode propagating in each direction, since Maxwell's equations and all boundary conditions are satisfied.

The analogous electric current that excites only the TM_{mn} mode can be shown to be

$$\bar{J}_s^{\text{TM}}(x, y) = \hat{x} \frac{2B_{mn}^+ m\pi}{a} \cos \frac{m\pi x}{a} \sin \frac{n\pi y}{b} + \hat{y} \frac{2B_{mn}^+ n\pi}{b} \sin \frac{m\pi x}{a} \cos \frac{n\pi y}{b}. \quad (4.116)$$

It is left as a problem to verify that this current excites TM_{mn} modes that satisfy the appropriate boundary conditions.

Similar results can be derived for magnetic surface current sheets. From (1.36) and (1.37) the appropriate boundary conditions are

$$(\bar{E}^+ - \bar{E}^-) \times \hat{z} = \bar{M}_s, \quad (4.117a)$$

$$\hat{z} \times (\bar{H}^+ - \bar{H}^-) = 0. \quad (4.117b)$$

For a magnetic current sheet at $z = 0$, the TE_{mn} waveguide mode fields of (4.112) must now have continuous H_x and H_y field components, due to (4.117b). This results in the condition that

$$A_{mn}^+ = -A_{mn}^-. \quad (4.118)$$

Then applying (4.117a) gives the source current as

$$\bar{M}_s^{\text{TE}} = \frac{-\hat{x}2Z_{\text{TE}}A_{mn}^+m\pi}{a} \sin \frac{m\pi x}{a} \cos \frac{n\pi y}{b} - \hat{y} \frac{2Z_{\text{TE}}A_{mn}^+n\pi}{b} \cos \frac{m\pi x}{a} \sin \frac{n\pi y}{b}. \quad (4.119)$$

The corresponding magnetic surface current that excites only the TM_{mn} mode can be shown to be

$$\bar{M}_s^{\text{TM}} = \frac{-\hat{x}2B_{mn}^+n\pi}{b} \sin \frac{m\pi x}{a} \cos \frac{n\pi y}{b} + \frac{\hat{y}2B_{mn}^+m\pi}{a} \cos \frac{m\pi x}{a} \sin \frac{n\pi y}{b}. \quad (4.120)$$

These results show that a single waveguide mode can be selectively excited, to the exclusion of all other modes, by either an electric or magnetic current sheet of the appropriate form. In practice, however, such currents are difficult to generate and are usually only approximated with one or two probes or loops. In this case many modes may be excited, but usually most of these modes are evanescent.

Mode Excitation from an Arbitrary Electric or Magnetic Current Source

We now consider the excitation of waveguide modes by an arbitrary electric or magnetic current source [4]. With reference to Figure 4.27, first consider an electric current source \bar{J} located between two transverse planes at z_1 and z_2 , which generates the fields \bar{E}^+ , \bar{H}^+ traveling in the $+z$ direction, and the fields \bar{E}^- , \bar{H}^- traveling in the $-z$ direction. These fields can be expressed in terms of the waveguide modes as follows:

$$\bar{E}^+ = \sum_n A_n^+ \bar{E}_n^+ = \sum_n A_n^+ (\bar{e}_n + \hat{z}e_{zn}) e^{-j\beta_n z}, \quad z > z_2, \quad (4.121a)$$

$$\bar{H}^+ = \sum_n A_n^+ \bar{H}_n^+ = \sum_n A_n^+ (\bar{h}_n + \hat{z}h_{zn}) e^{-j\beta_n z}, \quad z > z_2, \quad (4.121b)$$

$$\bar{E}^- = \sum_n A_n^- \bar{E}_n^- = \sum_n A_n^- (\bar{e}_n - \hat{z}e_{zn}) e^{j\beta_n z}, \quad z < z_1, \quad (4.121c)$$

$$\bar{H}^- = \sum_n A_n^- \bar{H}_n^- = \sum_n A_n^- (-\bar{h}_n + \hat{z}h_{zn}) e^{j\beta_n z}, \quad z < z_1, \quad (4.121d)$$

where the single index n is used to represent any possible TE or TM mode. For a given current \bar{J} , we can determine the unknown amplitude A_n^+ by using the Lorentz reciprocity theorem of (1.155) with $\bar{M}_1 = \bar{M}_2 = 0$ (since here we are only considering an electric current source),

$$\oint_S (\bar{E}_1 \times \bar{H}_2 - \bar{E}_2 \times \bar{H}_1) \cdot d\bar{s} = \int_V (\bar{E}_2 \cdot \bar{J}_1 - \bar{E}_1 \cdot \bar{J}_2) dv,$$

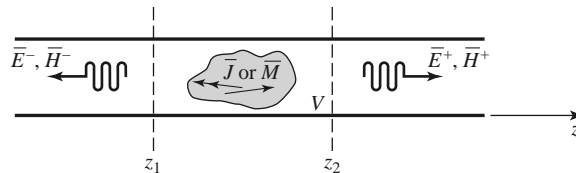


FIGURE 4.27 An arbitrary electric or magnetic current source in an infinitely long waveguide.

where S is a closed surface enclosing the volume V , and \bar{E}_i, \bar{H}_i are the fields due to the current source \bar{J}_i (for $i = 1$ or 2).

To apply the reciprocity theorem to the present problem we let the volume V be the region between the waveguide walls and the transverse cross-section planes at z_1 and z_2 . Then let $\bar{E}_1 = \bar{E}^\pm$ and $\bar{H}_1 = \bar{H}^\pm$, depending on whether $z \geq z_2$ or $z \leq z_1$, and let \bar{E}_2, \bar{H}_2 be the n th waveguide mode traveling in the negative z direction:

$$\begin{aligned}\bar{E}_2 &= \bar{E}_n^- = (\bar{e}_n - \hat{z}e_{zn})e^{j\beta_n z}, \\ \bar{H}_2 &= \bar{H}_n^- = (-\bar{h}_n + \hat{z}h_{zn})e^{j\beta_n z}.\end{aligned}$$

Substitution into the above form of the reciprocity theorem gives, with $\bar{J}_1 = \bar{J}$ and $\bar{J}_2 = 0$,

$$\oint_S (\bar{E}^\pm \times \bar{H}_n^- - \bar{E}_n^- \times \bar{H}^\pm) \cdot d\bar{s} = \int_V \bar{E}_n^- \cdot \bar{J} dv. \quad (4.122)$$

The portion of the surface integral over the waveguide walls vanishes because the tangential electric field is zero there; that is, $\bar{E} \times \bar{H} \cdot \hat{z} = \bar{H} \cdot (\hat{z} \times \bar{E}) = 0$ on the waveguide walls. This reduces the integration to the guide cross section, S_0 , at the planes z_1 and z_2 . In addition, the waveguide modes are orthogonal over the guide cross section:

$$\begin{aligned}\int_{S_0} \bar{E}_m^\pm \times \bar{H}_n^\pm \cdot d\bar{s} &= \int_{S_0} (\bar{e}_m \pm \hat{z}e_{zm}) \times (\pm \bar{h}_n + \hat{z}h_{zn}) \cdot \hat{z} ds \\ &= \pm \int_{S_0} \bar{e}_m \times \bar{h}_n \cdot \hat{z} ds = 0, \text{ for } m \neq n.\end{aligned} \quad (4.123)$$

Using (4.121) and (4.123) then reduces (4.122) to

$$\begin{aligned}A_n^+ \int_{z_2} (\bar{E}_n^+ \times \bar{H}_n^- - \bar{E}_n^- \times \bar{H}_n^+) \cdot d\bar{s} + A_n^- \int_{z_1} (\bar{E}_n^- \times \bar{H}_n^- - \bar{E}_n^- \times \bar{H}_n^-) \cdot d\bar{s} \\ = \int_V \bar{E}_n^- \cdot \bar{J} dv.\end{aligned}$$

Because the second integral vanishes, this further reduces to

$$\begin{aligned}A_n^+ \int_{z_2} [(\bar{e}_n + \hat{z}e_{zn}) \times (-\bar{h}_n + \hat{z}h_{zn}) - (\bar{e}_n - \hat{z}e_{zn}) \times (\bar{h}_n + \hat{z}h_{zn})] \cdot \hat{z} ds \\ = -2A_n^+ \int_{z_2} \bar{e}_n \times \bar{h}_n \cdot \hat{z} ds = \int_V \bar{E}_n^- \cdot \bar{J} dv,\end{aligned}$$

or

$$A_n^+ = \frac{-1}{P_n} \int_V \bar{E}_n^- \cdot \bar{J} dv = \frac{-1}{P_n} \int_V (\bar{e}_n - \hat{z}e_{zn}) \cdot \bar{J} e^{j\beta_n z} dv, \quad (4.124)$$

where

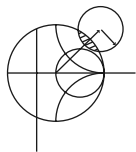
$$P_n = 2 \int_{S_0} \bar{e}_n \times \bar{h}_n \cdot \hat{z} ds \quad (4.125)$$

is a normalization constant proportional to the power flow of the n th mode.

By repeating the above procedure with $\bar{E}_2 = \bar{E}_n^+$ and $\bar{H}_2 = \bar{H}_n^+$, we can derive the amplitude of the negatively traveling waves as

$$A_n^- = \frac{-1}{P_n} \int_V \bar{E}_n^+ \cdot \bar{J} dv = \frac{-1}{P_n} \int_V (\bar{e}_n + \hat{z}e_{zn}) \cdot \bar{J} e^{-j\beta_n z} dv. \quad (4.126)$$

These results are quite general, being applicable to any type of waveguide (including planar lines such as stripline and microstrip), where modal fields can be defined. Example 4.8 applies this theory to the problem of a probe-fed rectangular waveguide.



EXAMPLE 4.8 PROBE-FED RECTANGULAR WAVEGUIDE

For the probe-fed rectangular waveguide shown in Figure 4.28, determine the amplitudes of the forward and backward traveling TE_{10} modes, and the input resistance seen by the probe. Assume that the TE_{10} mode is the only propagating mode.

Solution

If the current probe is assumed to have an infinitesimal diameter, the source volume current density \bar{J} can be written as

$$\bar{J}(x, y, z) = I_0 \delta\left(x - \frac{a}{2}\right) \delta(z) \hat{y} \text{ for } 0 \leq y \leq b.$$

From Chapter 3 the TE_{10} modal fields can be written as

$$\begin{aligned} \bar{e}_1 &= \hat{y} \sin \frac{\pi x}{a}, \\ \bar{h}_1 &= \frac{-\hat{x}}{Z_1} \sin \frac{\pi x}{a}, \end{aligned}$$

where $Z_1 = k_0 \eta_0 / \beta_1$ is the TE_{10} wave impedance. From (4.125) the normalization constant P_1 is

$$P_1 = \frac{2}{Z_1} \int_{x=0}^a \int_{y=0}^b \sin^2 \frac{\pi x}{a} dx dy = \frac{ab}{Z_1}.$$

Then from (4.124) the amplitude A_1^+ is

$$A_1^+ = \frac{-1}{P_1} \int_V \sin \frac{\pi x}{a} e^{j\beta_1 z} I_0 \delta\left(x - \frac{a}{2}\right) \delta(z) dx dy dz = \frac{-I_0 b}{P_1} = \frac{-Z_1 I_0}{a}.$$

Similarly,

$$A_1^- = \frac{-Z_1 I_0}{a}.$$

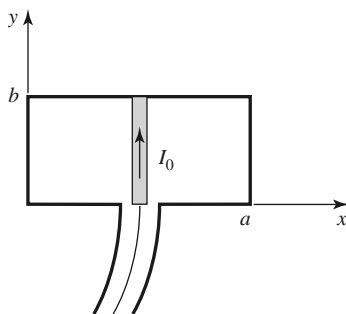


FIGURE 4.28 A uniform current probe in a rectangular waveguide.

If the TE_{10} mode is the only propagating mode in the waveguide, then this mode carries all of the average power, which can be calculated for real Z_1 as

$$\begin{aligned}
 P &= \frac{1}{2} \int_{S_0} \bar{E}^+ \times \bar{H}^{+*} \cdot d\bar{s} + \frac{1}{2} \int_{S_0} \bar{E}^- \times \bar{H}^{-*} \cdot d\bar{s} \\
 &= \int_{S_0} \bar{E}^+ \times \bar{H}^{+*} \cdot d\bar{s} \\
 &= \int_{x=0}^a \int_{y=0}^b \frac{|A_1^+|^2}{Z_1} \sin^2 \frac{\pi x}{a} dx dy \\
 &= \frac{ab|A_1^+|^2}{2Z_1}.
 \end{aligned}$$

If the input resistance seen looking into the probe is R_{in} , and the terminal current is I_0 , then $P = I_0^2 R_{in}/2$, so that the input resistance is

$$R_{in} = \frac{2P}{I_0^2} = \frac{ab|A_1^+|^2}{I_0^2 Z_1} = \frac{bZ_1}{a},$$

which is real for real Z_1 (corresponding to a propagating TE_{10} mode). ■

A similar derivation can be carried out for a magnetic current source \bar{M} (e.g., a small loop). This source will also generate positively and negatively traveling waves, which can be expressed as a superposition of waveguide modes, as in (4.121). For $\bar{J}_1 = \bar{J}_2 = 0$, the reciprocity theorem of (1.155) reduces to

$$\oint_S (\bar{E}_1 \times \bar{H}_2 - \bar{E}_2 \times \bar{H}_1) \cdot d\bar{s} = \int_V (\bar{H}_1 \cdot \bar{M}_2 - \bar{H}_2 \cdot \bar{M}_1) dv. \quad (4.127)$$

By following the same procedure as for the electric current case, we can derive the excitation coefficients of the n th waveguide mode as

$$A_n^+ = \frac{1}{P_n} \int_V \bar{H}_n^- \cdot \bar{M} dv = \frac{1}{P_n} \int_V (-\bar{h}_n + \hat{z}h_{zn}) \cdot \bar{M} e^{j\beta_n z} dv, \quad (4.128)$$

$$A_n^- = \frac{1}{P_n} \int_V \bar{H}_n^+ \cdot \bar{M} dv = \frac{1}{P_n} \int_V (\bar{h}_n + \hat{z}h_{zn}) \cdot \bar{M} e^{-j\beta_n z} dv, \quad (4.129)$$

where P_n is defined in (4.125).

4.8

EXCITATION OF WAVEGUIDES—APERTURE COUPLING

Besides the probe and loop feeds of the previous section, waveguides and other transmission lines can also be coupled through small apertures. One common application of such coupling is in directional couplers and power dividers, where power from one guide is coupled to another guide through small apertures in a common wall. Figure 4.29 shows a variety of waveguide and other transmission line configurations in which aperture coupling can be employed. We will first develop an intuitive explanation for the fact that a small aperture can be represented as an infinitesimal electric and/or magnetic dipole, then we will use the results of Section 4.7 to find the fields generated by these equivalent currents. Our analysis will be somewhat phenomenological [4, 10]; a more advanced theory of aperture coupling based on the equivalence theorem can be found in reference [11].

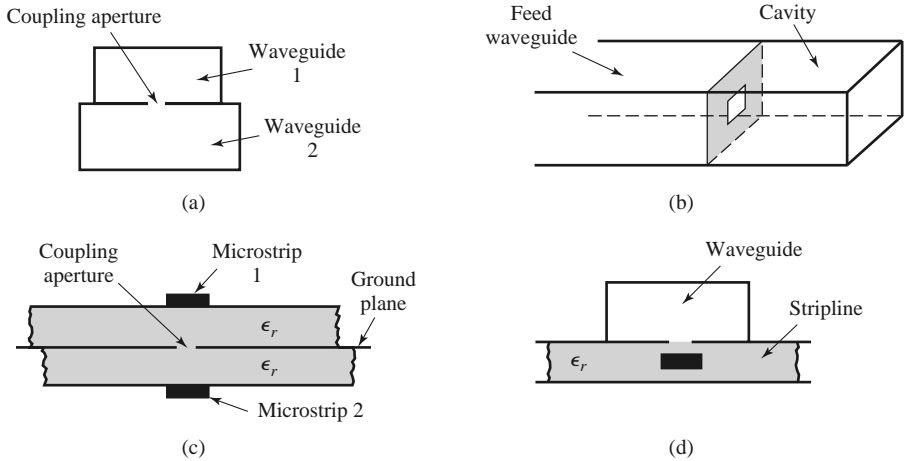


FIGURE 4.29 Various waveguide and other transmission line configurations using aperture coupling. (a) Coupling between two waveguides via an aperture in the common broad wall. (b) Coupling to a waveguide cavity via an aperture in a transverse wall. (c) Coupling between two microstrip lines via an aperture in the common ground plane. (d) Coupling from a waveguide to a stripline via an aperture.

Consider Figure 4.30a, which shows the normal electric field lines near a conducting wall (the tangential electric field is zero near the wall). If a small aperture is cut into the conductor, the electric field lines will fringe through and around the aperture as shown in Figure 4.30b. Now consider Figure 4.30c, which shows the fringing field lines around two infinitesimal electric polarization currents, \bar{P}_e , normal to a conducting wall (without

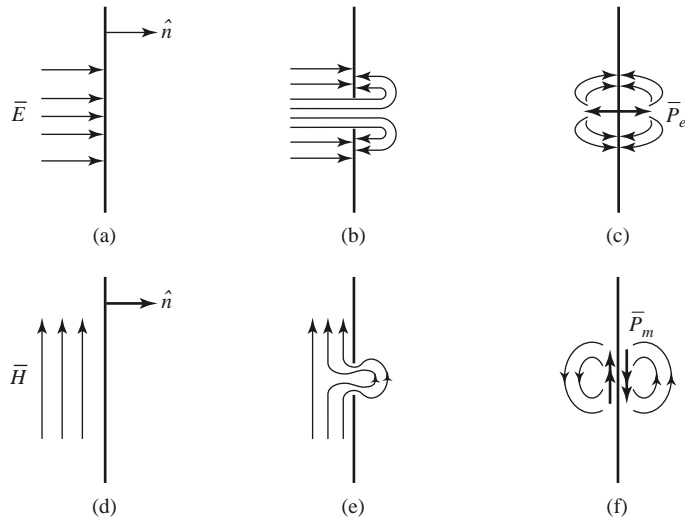


FIGURE 4.30 Illustrating the development of equivalent electric and magnetic polarization currents at an aperture in a conducting wall. (a) Normal electric field at a conducting wall. (b) Electric field lines around an aperture in a conducting wall. (c) Electric field lines around electric polarization currents normal to a conducting wall. (d) Magnetic field lines near a conducting wall. (e) Magnetic field lines near an aperture in a conducting wall. (f) Magnetic field lines near magnetic polarization currents parallel to a conducting wall.

an aperture). The similarity of the field lines of Figures 4.30c and 4.30b suggests that an aperture excited by a normal electric field can be represented by two oppositely directed infinitesimal electric polarization currents, \bar{P}_e , normal to the closed conducting wall. The strength of this polarization current is proportional to the normal electric field; thus,

$$\bar{P}_e = \epsilon_0 \alpha_e \hat{n} E_n \delta(x - x_0) \delta(y - y_0) \delta(z - z_0), \quad (4.130)$$

where the proportionality constant α_e is defined as the *electric polarizability* of the aperture, and (x_0, y_0, z_0) are the coordinates of the center of the aperture.

Similarly, Figure 4.30e shows the fringing of tangential magnetic field lines (the normal magnetic field is zero at the conductor) near a small aperture. Because these field lines are similar to those produced by two magnetic polarization currents located parallel to the conducting wall (as shown in Figure 4.30f), we can conclude that the aperture can be replaced by two oppositely directed infinitesimal polarization currents, \bar{P}_m , where

$$\bar{P}_m = -\alpha_m \bar{H}_t \delta(x - x_0) \delta(y - y_0) \delta(z - z_0). \quad (4.131)$$

In (4.131), α_m is defined as the *magnetic polarizability* of the aperture.

The electric and magnetic polarizabilities are constants that depend on the size and shape of the aperture and have been derived for a variety of simple shapes [3, 10, 11]. The polarizabilities for circular and rectangular apertures, which are probably the most commonly used shapes, are given in Table 4.3.

We now show that the electric and magnetic polarization currents, \bar{P}_e and \bar{P}_m , can be related to electric and magnetic current sources, \bar{J} and \bar{M} , respectively. From Maxwell's equations (1.27a) and (1.27b) we have

$$\nabla \times \bar{E} = -j\omega\mu\bar{H} - \bar{M}, \quad (4.132a)$$

$$\nabla \times \bar{H} = j\omega\epsilon\bar{E} + \bar{J}. \quad (4.132b)$$

Then using (1.15) and (1.23), which define \bar{P}_e and \bar{P}_m , we obtain

$$\nabla \times \bar{E} = -j\omega\mu_0\bar{H} - j\omega\mu_0\bar{P}_m - \bar{M}, \quad (4.133a)$$

$$\nabla \times \bar{H} = j\omega\epsilon_0\bar{E} + j\omega\bar{P}_e + \bar{J}. \quad (4.133b)$$

Thus, since \bar{M} has the same role in these equations as $j\omega\mu_0\bar{P}_m$, and \bar{J} has the same role as $j\omega\bar{P}_e$, we can define equivalent currents as

$$\bar{J} = j\omega\bar{P}_e, \quad (4.134a)$$

$$\bar{M} = j\omega\mu_0\bar{P}_m. \quad (4.134b)$$

These results allow us to use the formulas of (4.124), (4.126), (4.128), and (4.129) to compute the fields from these currents.

TABLE 4.3 Electric and Magnetic Polarizations

Aperture Shape	α_e	α_m
Round hole	$\frac{2r_0^3}{3}$	$\frac{4r_0^3}{3}$
Rectangular slot (\bar{H} across slot)	$\frac{\pi \ell d^2}{16}$	$\frac{\pi \ell d^2}{16}$

The above theory is approximate because of various assumptions involved in the evaluation of the polarizabilities, but generally it gives reasonable results for apertures that are small (where the term *small* implies small relative to an electrical wavelength), and not located too close to edges or corners of the guide. In addition, it is important to realize that the equivalent dipoles given by (4.130) and (4.131) radiate in the presence of the conducting wall to give the fields transmitted through the aperture. The fields on the input side of the conducting wall are also affected by the presence of the aperture, and this effect is accounted for by the equivalent dipoles on the incident side of the conductor (which are the negative of those on the output side). In this way, continuity of tangential fields is preserved across the aperture. In both cases, the presence of the (closed) conducting wall can be accounted for by using image theory to remove the wall and double the strength of the dipoles. These details will be clarified by applying this theory to apertures in transverse and broad walls of waveguides.

Coupling Through an Aperture in a Transverse Waveguide Wall

Consider a small circular aperture centered in the transverse wall of a waveguide, as shown in Figure 4.31a. Assume that only the TE₁₀ mode propagates in the guide, and is incident on the transverse wall from $z < 0$. Then, if the aperture is assumed to be closed, as in Figure 4.31b, the standing wave fields in the region $z < 0$ can be written as

$$E_y = A(e^{-j\beta z} - e^{j\beta z}) \sin \frac{\pi x}{a}, \quad (4.135a)$$

$$H_x = \frac{-A}{Z_{10}}(e^{-j\beta z} + e^{j\beta z}) \sin \frac{\pi x}{a}, \quad (4.135b)$$

where β and Z_{10} are the propagation constant and wave impedance of the TE₁₀ mode. From (4.130) and (4.131) we can determine the equivalent electric and magnetic polarization currents from the above fields as

$$\bar{P}_e = \hat{z}\epsilon_0\alpha_e E_z \delta\left(x - \frac{a}{2}\right) \delta\left(y - \frac{b}{2}\right) \delta(z) = 0, \quad (4.136a)$$

$$\begin{aligned} \bar{P}_m &= -\hat{x}\alpha_m H_x \delta\left(x - \frac{a}{2}\right) \delta\left(y - \frac{b}{2}\right) \delta(z) \\ &= \hat{x} \frac{2A\alpha_m}{Z_{10}} \delta\left(x - \frac{a}{2}\right) \delta\left(y - \frac{b}{2}\right) \delta(z), \end{aligned} \quad (4.136b)$$

since $E_z = 0$ for a TE mode. Now, by (4.134b), the magnetic polarization current \bar{P}_m is equivalent to the magnetic current density

$$\bar{M} = j\omega\mu_0\bar{P}_m = \hat{x} \frac{2j\omega\mu_0 A\alpha_m}{Z_{10}} \delta\left(x - \frac{a}{2}\right) \delta\left(y - \frac{b}{2}\right) \delta(z). \quad (4.137)$$

As shown in Figure 4.31d, the fields scattered by the aperture are considered as being produced by the equivalent currents \bar{P}_m and $-\bar{P}_m$ on either side of the closed wall. The presence of the conducting wall is easily accounted for using image theory, which has the effect of doubling the dipole strengths and removing the wall, as depicted in Figure 4.31e (for $z < 0$) and Figure 4.31f (for $z > 0$). Thus the coefficients of the transmitted and reflected waves caused by the equivalent aperture currents can be found by using (4.137) in (4.128) and (4.129) to give

$$A_{10}^+ = \frac{-1}{P_{10}} \int \bar{h}_{10} \cdot (2j\omega\mu_0\bar{P}_m) dv = \frac{4jA\omega\mu_0\alpha_m}{abZ_{10}} = \frac{4jA\beta\alpha_m}{ab}, \quad (4.138a)$$

$$A_{10}^- = \frac{-1}{P_{10}} \int \bar{h}_{10} \cdot (-2j\omega\mu_0\bar{P}_m) dv = \frac{4jA\omega\mu_0\alpha_m}{abZ_{10}} = \frac{4jA\beta\alpha_m}{ab}, \quad (4.138b)$$

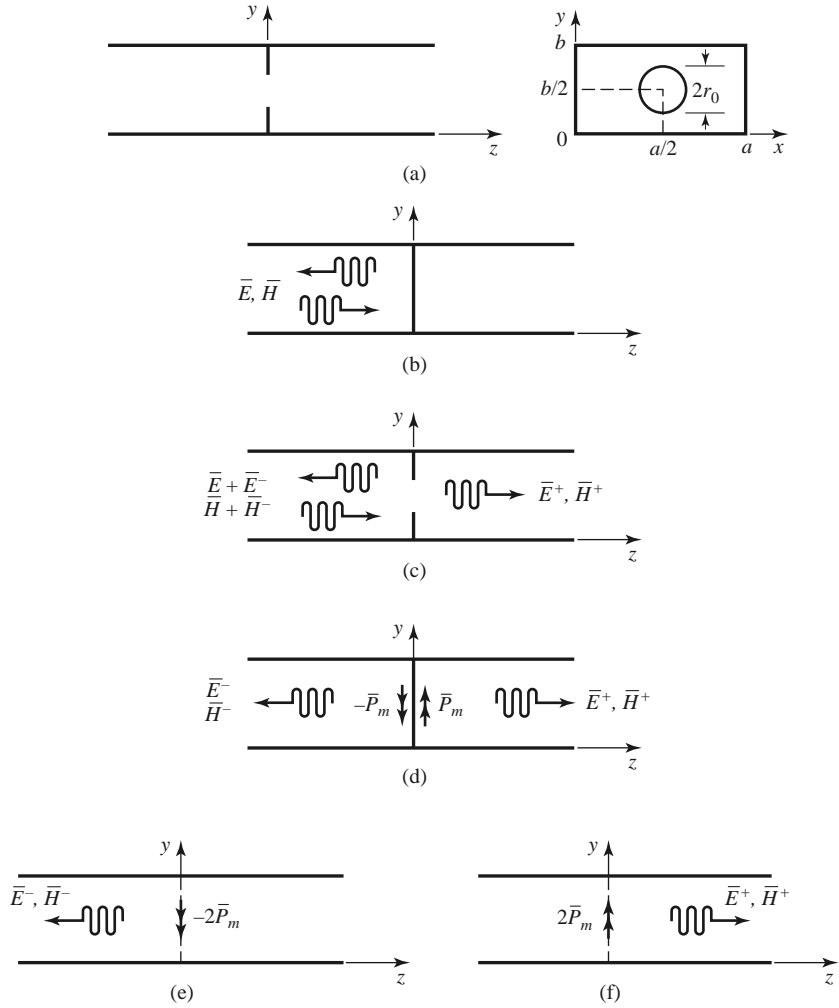


FIGURE 4.31 Applying small-hole coupling theory and image theory to the problem of an aperture in the transverse wall of a waveguide. (a) Geometry of a circular aperture in the transverse wall of a waveguide. (b) Fields with aperture closed. (c) Fields with aperture open. (d) Fields with aperture closed and replaced with equivalent dipoles. (e) Fields radiated by equivalent dipoles for $z < 0$; wall removed by image theory. (f) Fields radiated by equivalent dipoles for $z > 0$; wall removed by image theory.

since $\bar{h}_{10} = (-\hat{x}/Z_{10}) \sin(\pi x/a)$, and $P_{10} = ab/Z_{10}$. The magnetic polarizability α_m is given in Table 4.3. The complete fields can now be written as

$$E_y = [Ae^{-j\beta z} + (A_{10}^- - A)e^{j\beta z}] \sin \frac{\pi x}{a}, \quad \text{for } z < 0, \quad (4.139a)$$

$$H_x = \frac{1}{Z_{10}} [-Ae^{-j\beta z} + (A_{10}^- - A)e^{j\beta z}] \sin \frac{\pi x}{a}, \quad \text{for } z < 0, \quad (4.139b)$$

and

$$E_y = A_{10}^+ e^{-j\beta z} \sin \frac{\pi x}{a}, \quad \text{for } z > 0, \quad (4.140a)$$

$$H_x = \frac{-A_{10}^+}{Z_{10}} e^{-j\beta z} \sin \frac{\pi x}{a}, \quad \text{for } z > 0. \quad (4.140b)$$

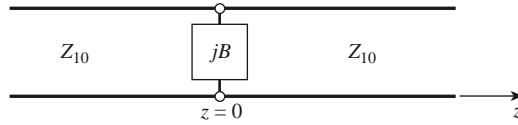


FIGURE 4.32 Equivalent circuit of the aperture in a transverse waveguide wall.

Then the reflection and transmission coefficients can be found as

$$\Gamma = \frac{A_{10}^- - A}{A} = \frac{4j\beta\alpha_m}{ab} - 1, \quad (4.141a)$$

$$T = \frac{A_{10}^+}{A} = \frac{4j\beta\alpha_m}{ab}, \quad (4.141b)$$

since $Z_{10} = k_0\eta_0/\beta$. Note that $|\Gamma| > 1$; this physically unrealizable result (for a passive network) is an artifact of the approximations used in the above theory. An equivalent circuit for this problem can be obtained by comparing the reflection coefficient of (4.141a) with that of the transmission line with a normalized shunt susceptance, jB , shown in Figure 4.32. The reflection coefficient seen looking into this line is

$$\Gamma = \frac{1 - y_{in}}{1 + y_{in}} = \frac{1 - (1 + jB)}{1 + (1 + jB)} = \frac{-jB}{2 + jB}.$$

If the shunt susceptance is very large (low impedance), Γ can be approximated as

$$\Gamma = \frac{-1}{1 + (2/jB)} \simeq -1 - j\frac{2}{B}.$$

Comparison with (4.141a) suggests that the aperture is equivalent to a normalized inductive susceptance,

$$B = \frac{-ab}{2\beta\alpha_m}.$$

Coupling Through an Aperture in the Broad Wall of a Waveguide

Another common configuration for aperture coupling is shown in Figure 4.33, where two parallel waveguides share a common broad wall and are coupled with a small centered aperture. We will assume a TE_{10} mode incident from $z < 0$ in the lower guide (guide 1),

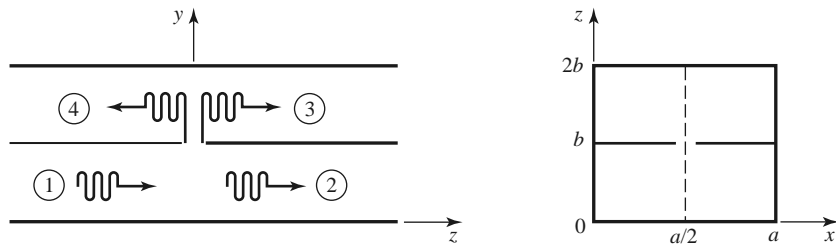


FIGURE 4.33 Two parallel waveguides coupled through an aperture in a common broad wall.

and compute the fields coupled to the upper guide. The incident fields can be written as

$$E_y = A \sin \frac{\pi x}{a} e^{-j\beta z}, \quad (4.142a)$$

$$H_x = \frac{-A}{Z_{10}} \sin \frac{\pi x}{a} e^{-j\beta z}. \quad (4.142b)$$

The excitation field at the center of the aperture at $(x = a/2, y = b, z = 0)$ is

$$E_y = A, \quad (4.143a)$$

$$H_x = \frac{-A}{Z_{10}}. \quad (4.143b)$$

(If the aperture were not centered at $x = a/2$, the H_z field would be nonzero and would have to be included.)

From (4.130), (4.131), and (4.134), the equivalent electric and magnetic dipoles for coupling to the fields in the upper guide are

$$J_y = j\omega\epsilon_0\alpha_e A \delta\left(x - \frac{a}{2}\right) \delta(y - b) \delta(z), \quad (4.144a)$$

$$M_x = \frac{j\omega\mu_0\alpha_m A}{Z_{10}} \delta\left(x - \frac{a}{2}\right) \delta(y - b) \delta(z). \quad (4.144b)$$

Note that in this case we have excited both an electric and a magnetic dipole. Let the fields in the upper guide be expressed as

$$E_y^- = A^- \sin \frac{\pi x}{a} e^{+j\beta z} \quad \text{for } z < 0, \quad (4.145a)$$

$$H_x^- = \frac{A^-}{Z_{10}} \sin \frac{\pi x}{a} e^{+j\beta z} \quad \text{for } z < 0, \quad (4.145b)$$

$$E_y^+ = A^+ \sin \frac{\pi x}{a} e^{-j\beta z} \quad \text{for } z > 0, \quad (4.146a)$$

$$H_x^+ = \frac{-A^+}{Z_{10}} \sin \frac{\pi x}{a} e^{-j\beta z} \quad \text{for } z > 0, \quad (4.146b)$$

where A^+ , A^- are the unknown amplitudes of the forward and backward traveling waves in the upper guide, respectively.

By superposition, the total fields in the upper guide due to the electric and magnetic currents of (4.144) can be found from (4.124) and (4.128) for the forward wave as

$$A^+ = \frac{-1}{P_{10}} \int_V (E_y^- J_y - H_x^- M_x) dv = \frac{-j\omega A}{P_{10}} \left(\epsilon_0 \alpha_e - \frac{\mu_0 \alpha_m}{Z_{10}^2} \right), \quad (4.147a)$$

and from (4.126) and (4.129) for the backward wave as

$$A^- = \frac{-1}{P_{10}} \int_V (E_y^+ J_y - H_x^+ M_x) dv = \frac{-j\omega A}{P_{10}} \left(\epsilon_0 \alpha_e + \frac{\mu_0 \alpha_m}{Z_{10}^2} \right), \quad (4.147b)$$

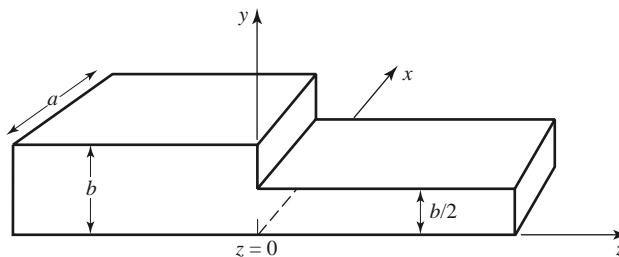
where $P_{10} = ab/Z_{10}$. Note that the electric dipole excites the same fields in both directions, but the magnetic dipole excites oppositely polarized fields in the forward and backward directions.

REFERENCES

- [1] S. Ramo, T. R. Whinnery, and T. van Duzer, *Fields and Waves in Communication Electronics*, John Wiley & Sons, New York, 1965.
- [2] A. A. Oliner, "Historical Perspectives on Microwave Field Theory," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-32, pp. 1022–1045, September 1984.
- [3] C. G. Montgomery, R. H. Dicke, and E. M. Purcell, eds., *Principles of Microwave Circuits*, MIT Radiation Laboratory Series, Vol. 8, McGraw-Hill, New York, 1948.
- [4] R. E. Collin, *Foundations for Microwave Engineering*, 2nd edition, McGraw-Hill, New York, 1992.
- [5] J. Rahola, "Power Waves and Conjugate Matching," *IEEE Transactions on Circuits and Systems*, vol. 55, pp. 92–96, January 2008.
- [6] J. S. Wright, O. P. Jain, W. J. Chudobiak, and V. Makios, "Equivalent Circuits of Microstrip Impedance Discontinuities and Launchers," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-22, pp. 48–52, January 1974.
- [7] G. F. Engen and C. A. Hoer, "Thru-Reflect-Line: An Improved Technique for Calibrating the Dual Six-Port Automatic Network Analyzer," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-27, pp. 987–998, December 1979.
- [8] N. Marcuvitz, ed., *Waveguide Handbook*, MIT Radiation Laboratory Series, Vol. 10, McGraw-Hill, New York, 1948.
- [9] K. C. Gupta, R. Garg, and I. J. Bahl, *Microstrip Lines and Slotlines*, Artech House, Dedham, Mass., 1979.
- [10] G. Matthaei, L. Young, and E. M. T. Jones, *Microwave Filters, Impedance-Matching Networks, and Coupling Structures*, Artech House, Dedham, Mass., 1980, Chapter 5.
- [11] R. E. Collin, *Field Theory of Guided Waves*, McGraw-Hill, New York, 1960.

PROBLEMS

- 4.1 Consider the reflection of a TE_{10} mode, incident from $z < 0$, at a step change in the height of a rectangular waveguide, as shown below. Show that if the method of Example 4.2 is used, the result $\Gamma = 0$ is obtained. Do you think this is the correct solution? Why? (This problem shows that the one-mode impedance viewpoint does not always provide a correct analysis.)

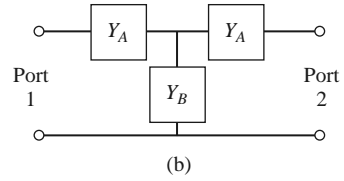
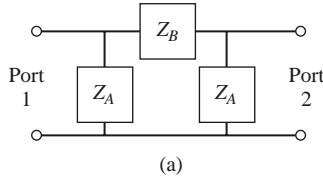


- 4.2 Consider a series RLC circuit with a current I . Calculate the power lost and the stored electric and magnetic energies, and show that the input impedance can be expressed as in (4.17).
- 4.3 Show that the input impedance Z of a parallel RLC circuit satisfies the condition that $Z(-\omega) = Z^*(\omega)$.
- 4.4 A two-port network is driven at both ports such that the port voltages and currents have the following values ($Z_0 = 50 \Omega$):

$$\begin{aligned} V_1 &= 10\angle 90^\circ, & I_1 &= 0.2\angle 90^\circ, \\ V_2 &= 8\angle 0^\circ, & I_2 &= 0.16\angle -90^\circ. \end{aligned}$$

Determine the input impedance seen at each port, and find the incident and reflected voltages at each port.

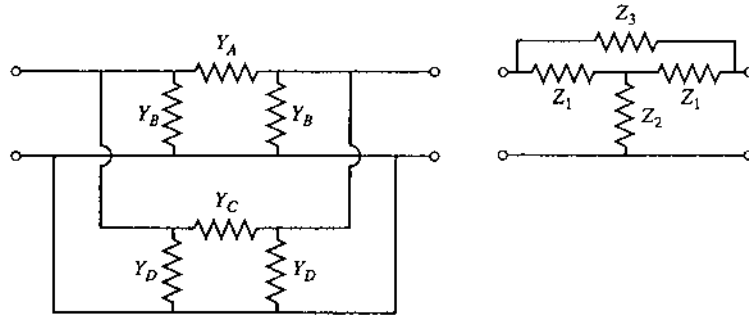
- 4.5 Show that the admittance matrix of a lossless N -port network has purely imaginary elements.
- 4.6 Does a nonreciprocal lossless network always have a purely imaginary impedance matrix?
- 4.7 Derive the $[Z]$ and $[Y]$ matrices for the two-port networks shown in the figure below.



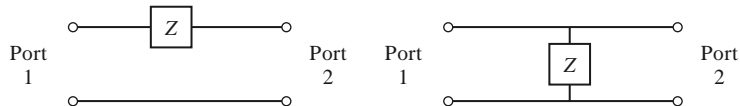
- 4.8 Consider a two-port network, and let $Z_{SC}^{(1)}$, $Z_{SC}^{(2)}$, $Z_{OC}^{(1)}$, and $Z_{OC}^{(2)}$ be the input impedance seen when port 2 is short-circuited, when port 1 is short-circuited, when port 2 is open-circuited, and when port 1 is open-circuited, respectively. Show that the impedance matrix elements are given by

$$Z_{11} = Z_{OC}^{(1)}, \quad Z_{22} = Z_{OC}^{(2)}, \quad Z_{12}^2 = Z_{21}^2 = (Z_{OC}^{(1)} - Z_{SC}^{(1)}) Z_{OC}^{(2)}.$$

- 4.9 Find the impedance parameters of a section of transmission line with length ℓ , characteristic impedance Z_0 , and propagation constant β .
- 4.10 Show that the admittance matrix of the two parallel-connected two-port π networks shown below can be found by adding the admittance matrices of the individual two-ports. Apply this result to find the admittance matrix of the bridged-T circuit shown. What is the corresponding result for the impedance matrix of two series-connected T-networks?



- 4.11 Find the scattering parameters for the series and shunt loads shown below. Show that $S_{12} = 1 - S_{11}$ for the series case, and that $S_{12} = 1 + S_{11}$ for the shunt case. Assume a characteristic impedance Z_0 .



- 4.12 Consider two two-port networks with individual scattering matrices $[S^A]$ and $[S^B]$. Show that the overall S_{21} parameter of the cascade of these networks is given by

$$S_{21} = \frac{S_{21}^A S_{21}^B}{1 - S_{22}^A S_{11}^B}.$$

- 4.13 Consider a lossless two-port network. (a) If the network is reciprocal, show that $|S_{21}|^2 = 1 - |S_{11}|^2$. (b) If the network is nonreciprocal, show that it is impossible to have unidirectional transmission, where $S_{12} = 0$ and $S_{21} \neq 0$.

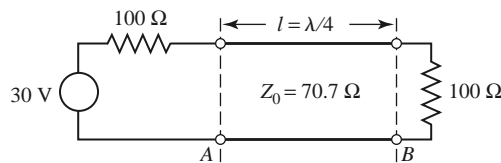
- 4.14** A four-port network has the scattering matrix shown as follows. (a) Is this network lossless? (b) Is this network reciprocal? (c) What is the return loss at port 1 when all other ports are terminated with matched loads? (d) What is the insertion loss and phase delay between ports 2 and 4 when all other ports are terminated with matched loads? (e) What is the reflection coefficient seen at port 1 if a short circuit is placed at the terminal plane of port 3 and all other ports are terminated with matched loads?

$$[S] = \begin{bmatrix} 0.178\angle 90^\circ & 0.6\angle 45^\circ & 0.4\angle 45^\circ & 0 \\ 0.6\angle 45^\circ & 0 & 0 & 0.3\angle -45^\circ \\ 0.4\angle 45^\circ & 0 & 0 & 0.5\angle -45^\circ \\ 0 & 0.3\angle -45^\circ & 0.5\angle -45^\circ & 0 \end{bmatrix}.$$

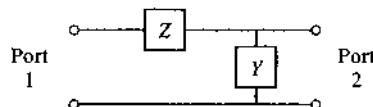
- 4.15** Show that it is impossible to construct a three-port network that is lossless, reciprocal, and matched at all ports. Is it possible to construct a nonreciprocal three-port network that is lossless and matched at all ports?
- 4.16** Prove the following *decoupling theorem*: For any lossless reciprocal three-port network, one port (say port 3) can be terminated in a reactance so that the other two ports (say ports 1 and 2) are decoupled (no power flow from port 1 to port 2, or from port 2 to port 1).
- 4.17** A certain three-port network is lossless and reciprocal, and has $S_{13} = S_{23}$ and $S_{11} = S_{22}$. Show that if port 2 is terminated with a matched load, then port 1 can be matched by placing an appropriate reactance at port 3.
- 4.18** A four-port network has the scattering matrix shown as follows. If ports 3 and 4 are connected with a lossless matched transmission line with an electrical length of 45° , find the resulting insertion loss and phase delay between ports 1 and 2.

$$[S] = \begin{bmatrix} 0.2\angle 50^\circ & 0 & 0 & 0.4\angle -45^\circ \\ 0 & 0.6\angle 45^\circ & 0.7\angle -45^\circ & 0 \\ 0 & 0.7\angle -45^\circ & 0.6\angle 45^\circ & 0 \\ 0.4\angle -45^\circ & 0 & 0 & 0.5\angle 45^\circ \end{bmatrix}.$$

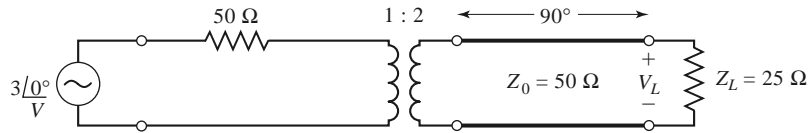
- 4.19** When normalized to a single characteristic impedance Z_0 , a certain two-port network has scattering parameters S_{ij} . Find the generalized scattering parameters, S_{ij}^p , in terms of the real reference impedances, R_{01} and R_{02} , at ports 1 and 2, respectively.
- 4.20** At reference plane A, for the circuit shown below, choose an appropriate reference impedance, find the power wave amplitudes, and compute the power delivered to the load. Repeat this procedure for reference plane B. Assume the transmission line is lossless.



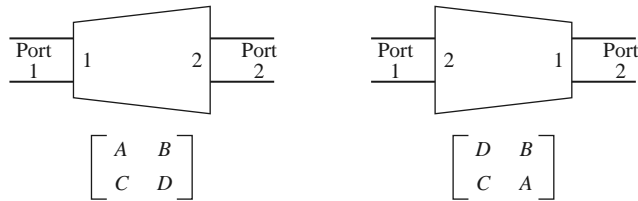
- 4.21** The $ABCD$ parameters of the first entry in Table 4.1 were derived in Example 4.6. Verify the $ABCD$ parameters for the second, third, and fourth entries.
- 4.22** Derive expressions that give the impedance parameters in terms of the $ABCD$ parameters.
- 4.23** Find the $ABCD$ matrix for the circuit shown below by direct calculation using the definition of the $ABCD$ matrix, and compare with the $ABCD$ matrix of the appropriate cascade of canonical circuits from Table 4.1.



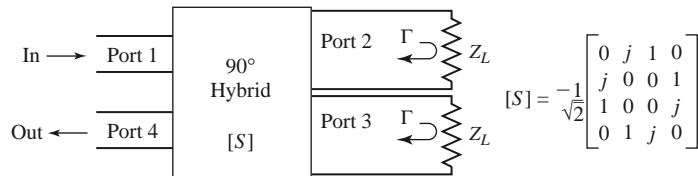
- 4.24 Use $ABCD$ matrices to find the voltage V_L across the load resistor in the circuit shown below.



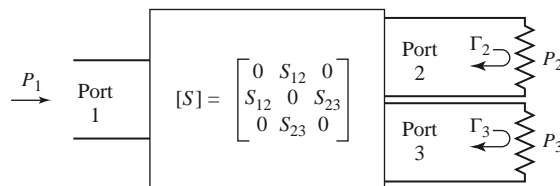
- 4.25 A reciprocal two-port network with its $ABCD$ matrix is shown below at left. Prove that the network with ports 1 and 2 in reversed positions has the $ABCD$ matrix shown below at right. Choose a simple asymmetrical network to demonstrate this result.



- 4.26 Derive the expressions for S parameters in terms of the $ABCD$ parameters, as given in Table 4.2.
- 4.27 As shown in the figure below, a variable attenuator can be implemented using a four-port 90° hybrid coupler by terminating ports 2 and 3 with equal but adjustable loads. (a) Using the given scattering matrix for the coupler, show that the transmission coefficient between the input (port 1) and the output (port 4) is given as $T = j\Gamma$, where Γ is the reflection coefficient of the mismatch at ports 2 and 3. Also show that the input port is matched for all values of Γ . (b) Plot the attenuation, in dB, from the input to the output as a function of Z_L/Z_0 , for $0 \leq Z_L/Z_0 \leq 10$ (let Z_L be real).



- 4.28 Use signal flow graphs to find the power ratios P_2/P_1 and P_3/P_1 for the mismatched three-port network shown in the accompanying figure.



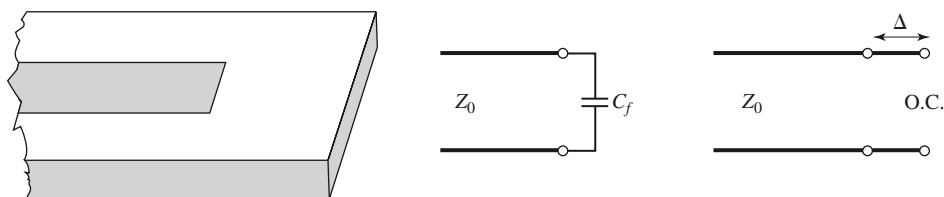
- 4.29 The $ABCD$ parameters are useful for treating cascades of two-port networks in terms of the total port voltages and currents, but it is also possible to use incident and reflected voltages to treat cascades. One way of doing this is with the *transfer*, or T -, *parameters*, defined as follows:

$$\begin{bmatrix} a_1 \\ b_1 \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} b_2 \\ a_2 \end{bmatrix},$$

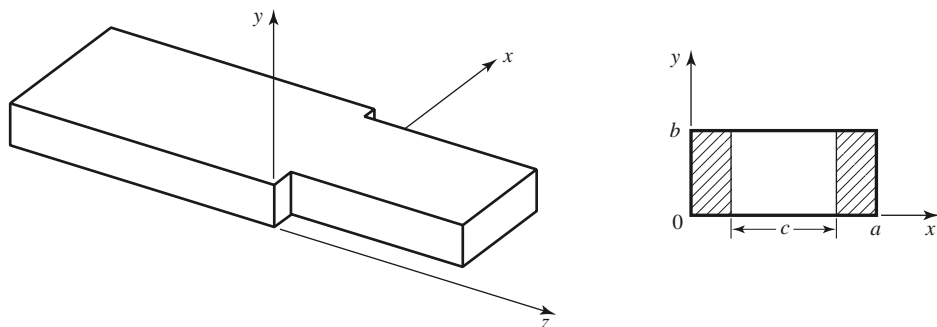
where a_1, b_1 and a_2, b_2 are the incident and reflected voltages at ports 1 and 2, respectively. Derive the T -parameters in terms of the scattering parameters of a two-port network. Show how the T -parameters can be used for a cascade of two two-port networks.

- 4.30** The end of an open-circuited microstrip line has fringing fields that can be modeled as a shunt capacitor, C_f , at the end of the line, as shown below. This capacitance can be replaced with an additional length, Δ , of microstrip line. Derive an expression for the length extension in terms of the fringing capacitance. Evaluate the length extension for a $50\ \Omega$ open-circuited microstrip line on a substrate with $d = 0.158\text{ cm}$ and $\epsilon_r = 2.2$ ($w = 0.487\text{ cm}$, $\epsilon_e = 1.894$), if the fringing capacitance is known to be $C_f = 0.075\text{ pF}$. Compare your result with the approximation given by Hammerstad and Bekkadal:

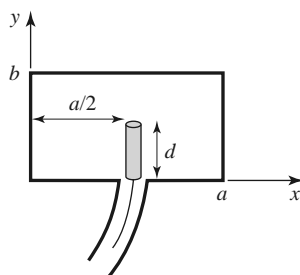
$$\Delta = 0.412d \left(\frac{\epsilon_e + 0.3}{\epsilon_e - 0.258} \right) \left(\frac{w + 0.262d}{w + 0.813d} \right).$$



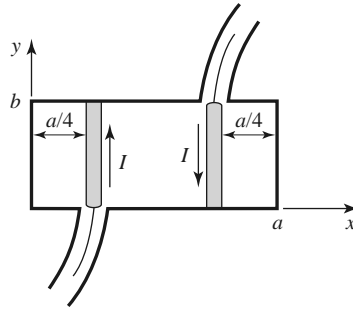
- 4.31** For the H -plane step analysis of Section 4.6, compute the complex power flow in the reflected modes in guide 1, and show that the reactive power is inductive.
- 4.32** Derive the modal analysis equations for the symmetric H -plane step shown below. (HINT: Because of symmetry, only the TE_{n0} modes for n odd will be excited.)



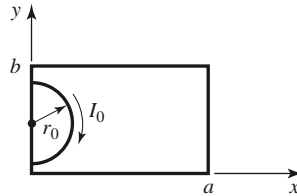
- 4.33** Find the transverse \vec{E} and \vec{H} fields excited by the current of (4.116) by postulating traveling TM_{mn} modes on either side of the source at $z = 0$ and applying the appropriate boundary conditions.
- 4.34** An infinitely long rectangular waveguide is fed with a probe of length d as shown below. The current on this probe can be approximated as $I(y) = I_0 \sin k(d - y)/\sin kd$. If the TE_{10} mode is the only propagating mode in the waveguide, compute the input resistance seen at the probe terminals.



- 4.35** Consider the infinitely long waveguide fed with two probes driven 180° out of phase, as shown below. What are the resulting excitation coefficients for the TE_{10} and TE_{20} modes? What other modes can be excited by this feeding arrangement?



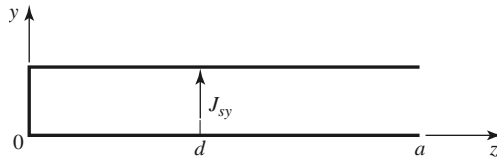
- 4.36** Consider a small current loop on the sidewall of a rectangular waveguide, as shown below. Find the TE_{10} fields excited by this loop if the loop is of radius r_0 .

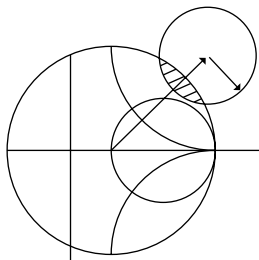


- 4.37** A rectangular waveguide is shorted at $z = 0$ and has an electric current sheet, J_{sy} , located at $z = d$, where

$$J_{sy} = \frac{2\pi A}{a} \sin \frac{\pi x}{a}$$

(see the accompanying figure). Find expressions for the fields generated by this current by assuming standing wave fields for $0 < z < d$, and traveling wave fields for $z > d$, and applying boundary conditions at $z = 0$ and $z = d$. Now solve the problem using image theory, by placing a current sheet $-J_{sy}$ at $z = -d$, and removing the shorting wall at $z = 0$. Use the results of Section 4.7 and superposition to find the fields radiated by these two currents, which should be the same as the first results for $z > 0$.





Impedance Matching and Tuning

This chapter marks a turning point, in that we now begin to apply the theory and techniques of previous chapters to practical problems in microwave engineering. We start with the topic of *impedance matching*, which is often an important part of a larger design process for a microwave component or system. The basic idea of impedance matching is illustrated in Figure 5.1, which shows an impedance matching network placed between a load impedance and a transmission line. The matching network is ideally lossless, to avoid unnecessary loss of power, and is usually designed so that the impedance seen looking into the matching network is Z_0 . Then reflections will be eliminated on the transmission line to the left of the matching network, although there will usually be multiple reflections between the matching network and the load. This procedure is sometimes referred to as *tuning*. Impedance matching or tuning is important for the following reasons:

- Maximum power is delivered when the load is matched to the line (assuming the generator is matched), and power loss in the feed line is minimized.
- Impedance matching sensitive receiver components (antenna, low-noise amplifier, etc.) may improve the signal-to-noise ratio of the system.
- Impedance matching in a power distribution network (such as an antenna array feed network) may reduce amplitude and phase errors.

As long as the load impedance, Z_L , has a positive real part, a matching network can always be found. Many choices are available, however, and we will discuss the design and performance of several types of practical matching networks. Factors that may be important in the selection of a particular matching network include the following:

- *Complexity*—As with most engineering solutions, the simplest design that satisfies the required specifications is generally preferable. A simpler matching network is usually cheaper, smaller, more reliable, and less lossy than a more complex design.
- *Bandwidth*—Any type of matching network can ideally give a perfect match (zero reflection) at a single frequency. In many applications, however, it is desirable to match a load over a band of frequencies. There are several ways of doing this, with, of course, a corresponding increase in complexity.



FIGURE 5.1 A lossless network matching an arbitrary load impedance to a transmission line.

- *Implementation*—Depending on the type of transmission line or waveguide being used, one type of matching network may be preferable to another. For example, tuning stubs are much easier to implement in waveguide than are multisection quarter-wave transformers.
- *Adjustability*—In some applications the matching network may require adjustment to match a variable load impedance. Some types of matching networks are more amenable than others in this regard.

5.1

MATCHING WITH LUMPED ELEMENTS (L NETWORKS)

Probably the simplest type of matching network is the L -section, which uses two reactive elements to match an arbitrary load impedance to a transmission line. There are two possible configurations for this network, as shown in Figure 5.2. If the normalized load impedance, $z_L = Z_L/Z_0$, is inside the $1 + jx$ circle on the Smith chart, then the circuit of Figure 5.2a should be used. If the normalized load impedance is outside the $1 + jx$ circle on the Smith chart, the circuit of Figure 5.2b should be used. The $1 + jx$ circle is the resistance circle on the impedance Smith chart for which $r = 1$.

In either of the configurations of Figure 5.2, the reactive elements may be either inductors or capacitors, depending on the load impedance. Thus, there are eight distinct possibilities for the matching circuit for various load impedances. If the frequency is low enough and/or the circuit size is small enough, actual lumped-element capacitors and inductors can be used. This may be feasible for frequencies up to about 1 GHz or so, although modern microwave integrated circuits may be small enough such that lumped elements can be used at higher frequencies as well. There is, however, a large range of frequencies and circuit sizes where lumped elements may not be realizable. This is a limitation of the L -section

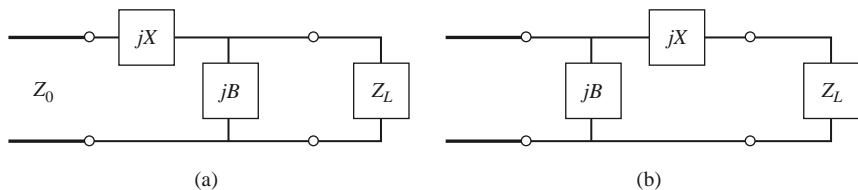


FIGURE 5.2 L -section matching networks. (a) Network for z_L inside the $1 + jx$ circle. (b) Network for z_L outside the $1 + jx$ circle.

matching technique. We will first derive analytic expressions for the matching network elements of the two cases in Figure 5.2, and then illustrate an alternative design procedure using the Smith chart.

Analytic Solutions

Although we will discuss a simple graphical solution using the Smith chart, it is also useful to have simple expressions for the L -section matching network components. These expressions can be used in a computer-aided design program for L -section matching, or when it is necessary to have more accuracy than the Smith chart can provide.

Consider first the circuit of Figure 5.2a, and let $Z_L = R_L + jX_L$. We stated that this circuit would be used when $z_L = Z_L/Z_0$ is inside the $1 + jx$ circle on the Smith chart, which implies that $R_L > Z_0$ for this case. The impedance seen looking into the matching network, followed by the load impedance, must be equal to Z_0 for an impedance-matched condition:

$$Z_0 = jX + \frac{1}{jB + 1/(R_L + jX_L)}. \quad (5.1)$$

Rearranging and separating into real and imaginary parts gives two equations for the two unknowns, X and B :

$$B(XR_L - X_L Z_0) = R_L - Z_0, \quad (5.2a)$$

$$X(1 - BX_L) = BZ_0 R_L - X_L. \quad (5.2b)$$

Solving (5.2a) for X and substituting into (5.2b) gives a quadratic equation for B . The solution is

$$B = \frac{X_L \pm \sqrt{R_L/Z_0} \sqrt{R_L^2 + X_L^2 - Z_0 R_L}}{R_L^2 + X_L^2}. \quad (5.3a)$$

Note that since $R_L > Z_0$, the argument of the second square root is always positive. Then the series reactance can be found as

$$X = \frac{1}{B} + \frac{X_L Z_0}{R_L} - \frac{Z_0}{BR_L}. \quad (5.3b)$$

Equation (5.3a) indicates that two solutions are possible for B and X . Both of these solutions are physically realizable since both positive and negative values of B and X are possible (positive X implies an inductor and negative X implies a capacitor, while positive B implies a capacitor and negative B implies an inductor). One solution, however, may result in significantly smaller values for the reactive components, or may be the preferred solution if the bandwidth of the match is better, or if the SWR on the line between the matching network and the load is smaller.

Next consider the circuit of Figure 5.2b. This circuit is used when z_L is outside the $1 + jx$ circle on the Smith chart, which implies that $R_L < Z_0$. The admittance seen looking into the matching network, followed by the load impedance, must be equal to $1/Z_0$ for an impedance-matched condition:

$$\frac{1}{Z_0} = jB + \frac{1}{R_L + j(X + X_L)}. \quad (5.4)$$

Rearranging and separating into real and imaginary parts gives two equations for the two unknowns, X and B :

$$BZ_0(X + X_L) = Z_0 - R_L, \quad (5.5a)$$

$$(X + X_L) = BZ_0R_L. \quad (5.5b)$$

Solving for X and B gives

$$X = \pm\sqrt{R_L(Z_0 - R_L)} - X_L, \quad (5.6a)$$

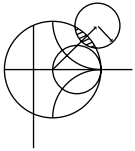
$$B = \pm \frac{\sqrt{(Z_0 - R_L)/R_L}}{Z_0}. \quad (5.6b)$$

Because $R_L < Z_0$, the arguments of the square roots are always positive. Again, note that two solutions are possible.

In order to match an arbitrary complex load to a line of characteristic impedance Z_0 , the real part of the input impedance to the matching network must be Z_0 , while the imaginary part must be zero. This implies that a general matching network must have at least two degrees of freedom; in the *L*-section matching circuit these two degrees of freedom are provided by the values of the two reactive components.

Smith Chart Solutions

Instead of the above formulas, the Smith chart can be used to quickly and accurately design *L*-section matching networks. The procedure is best illustrated by an example.



EXAMPLE 5.1 *L*-SECTION IMPEDANCE MATCHING

Design an *L*-section matching network to match a series *RC* load with an impedance $Z_L = 200 - j100 \, \Omega$ to a $100 \, \Omega$ line at a frequency of 500 MHz.

Solution

The normalized load impedance is $z_L = 2 - j1$, which is plotted on the Smith chart of Figure 5.3a. This point is inside the $1 + jx$ circle, so we use the matching circuit of Figure 5.2a. Because the first element from the load is a shunt susceptance, it makes sense to convert to admittance by drawing the SWR circle through the load, and a straight line from the load through the center of the chart, as shown in Figure 5.3a. After we add the shunt susceptance and convert back to impedance, we want to be on the $1 + jx$ circle so that we can add a series reactance to cancel jx and match the load. This means that the shunt susceptance must move us from y_L to the $1 + jx$ circle on the *admittance* Smith chart. Thus, we construct the rotated $1 + jx$ circle as shown in Figure 5.3a (center at $r = 0.333$). (A combined *ZY* chart may be convenient to use here, if it is not too confusing.) Then we see that adding a susceptance of $jb = j0.3$ will move us along a constant-conductance circle to $y = 0.4 + j0.5$ (this choice is the shortest distance from y_L to the shifted $1 + jx$ circle). Converting back to impedance leaves us at $z = 1 - j1.2$, indicating that a series reactance of $x = j1.2$ will bring us to the center of the chart. For comparison, the formulas (5.3a) and (5.3b) give the solution as $b = 0.29$, $x = 1.22$.

This matching circuit consists of a shunt capacitor and a series inductor, as shown in Figure 5.3b. For a matching frequency of 500 MHz, the capacitor has a value of

$$C = \frac{b}{2\pi f Z_0} = 0.92 \text{ pF},$$

and the inductor has a value of

$$L = \frac{x Z_0}{2\pi f} = 38.8 \text{ nH}.$$

It is also interesting to look at the second solution to this matching problem. If instead of adding a shunt susceptance of $b = 0.3$, we use a shunt susceptance of $b = -0.7$, we will move to a point on the lower half of the shifted $1 + jx$ circle, to $y = 0.4 - j0.5$. Then converting to impedance and adding a series reactance of $x = -1.2$ leads to a match as well. Formulas (5.3a) and (5.3b) give this solution as $b = -0.69$, $x = -1.22$. This matching circuit is also shown in Figure 5.3b, and is seen to have the positions of the inductor and capacitor reversed from the first matching network. At a frequency of $f = 500 \text{ MHz}$, the capacitor has a value of

$$C = \frac{-1}{2\pi f x Z_0} = 2.61 \text{ pF},$$

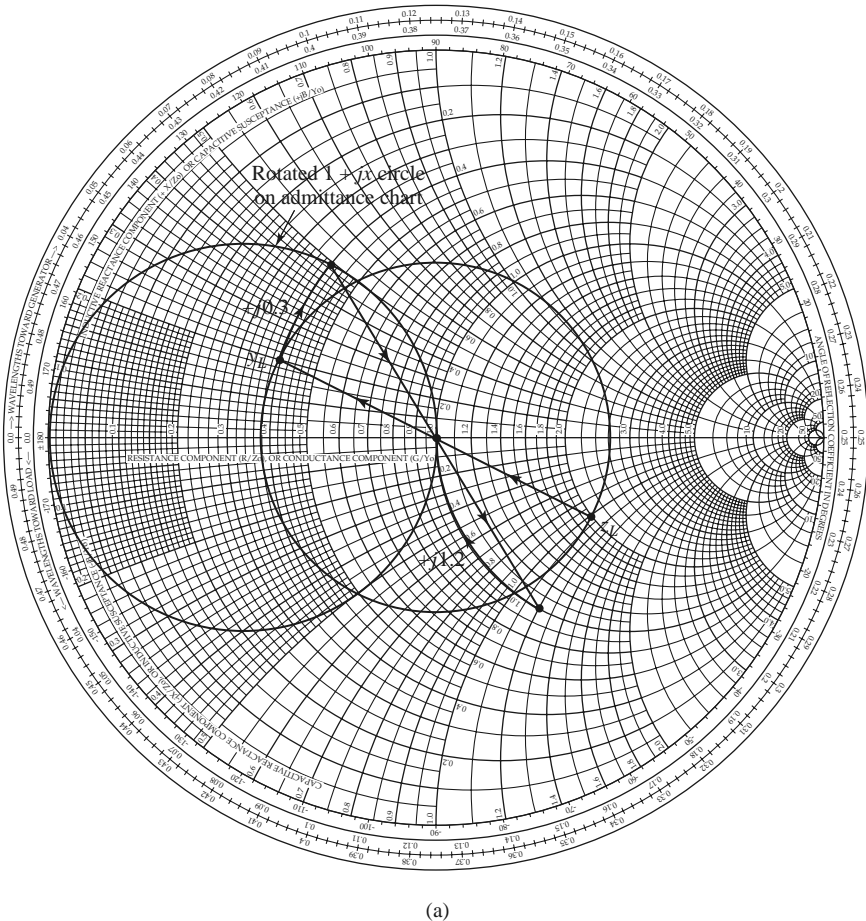


FIGURE 5.3 Solution to Example 5.1. (a) Smith chart for the L -section matching networks.

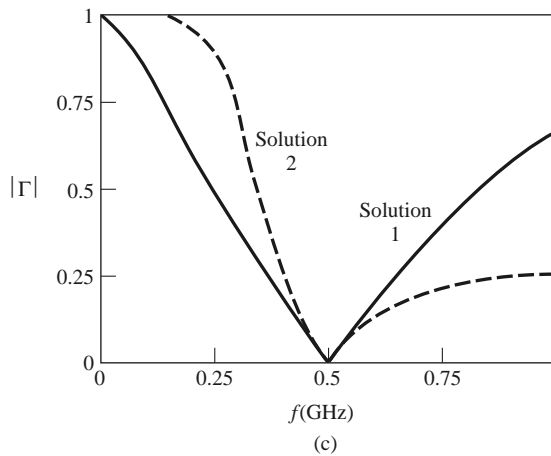
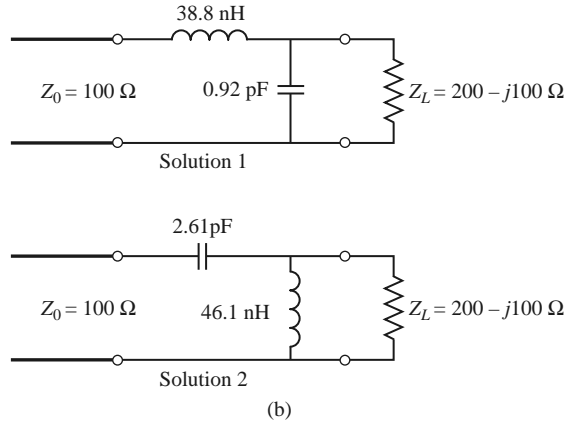
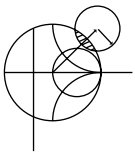


FIGURE 5.3 Continued. (b) The two possible L -section matching circuits. (c) Reflection coefficient magnitudes versus frequency for the matching circuits of (b).

while the inductor has a value of

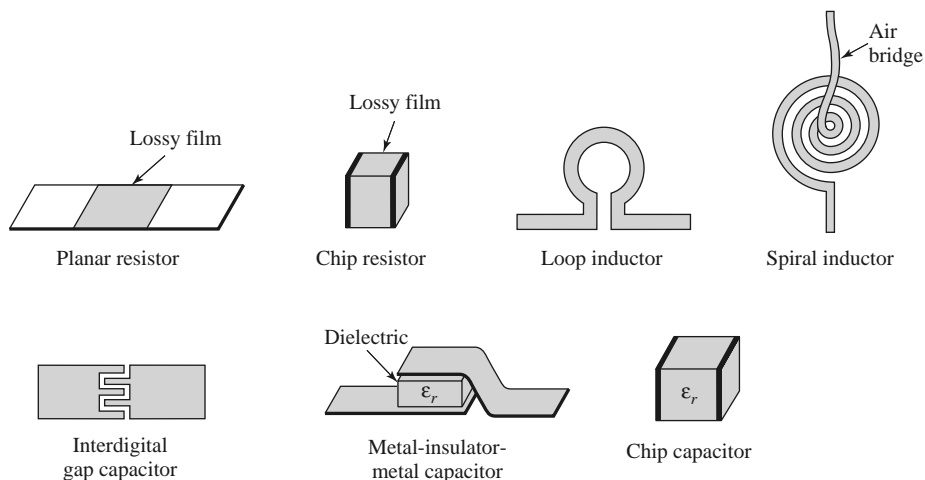
$$L = \frac{-Z_0}{2\pi f b} = 46.1 \text{ nH}.$$

Figure 5.3c shows the reflection coefficient magnitude versus frequency for these two matching networks, assuming that the load impedance of $Z_L = 200 - j100 \Omega$ at 500 MHz consists of a 200Ω resistor and a 3.18 pF capacitor in series. There is not a substantial difference in bandwidth for these two solutions. ■



POINT OF INTEREST: Lumped Elements for Microwave Integrated Circuits

Lumped R , L , and C elements can be practically realized at microwave frequencies if the length, ℓ , of the component is very small relative to the operating wavelength. Over a limited range of values, such components can be used in hybrid and monolithic microwave integrated circuits at frequencies up to 60 GHz, or higher, if the condition that $\ell < \lambda/10$ is satisfied. Usually, however, the characteristics of such an element are far from ideal, requiring that undesirable effects such as parasitic capacitance and/or inductance, spurious resonances, fringing fields, loss, and perturbations caused by a ground plane be incorporated in the design via a CAD model (see the Point of Interest concerning CAD).



Resistors are fabricated with thin films of lossy material such as nichrome, tantalum nitride, or doped semiconductor material. In monolithic circuits such films can be deposited or grown, whereas chip resistors made from a lossy film deposited on a ceramic chip can be bonded or soldered in a hybrid circuit. Low resistances are hard to obtain.

Small values of inductance can be realized with a short length or loop of transmission line, and larger values (up to about 10 nH) can be obtained with a spiral inductor, as shown in the following figures. Larger inductance values generally incur more loss and more shunt capacitance; this leads to a resonance that limits the maximum operating frequency.

Capacitors can be fabricated in several ways. A short transmission line stub can provide a shunt capacitance in the range of 0–0.1 pF. A single gap, or an interdigital set of gaps, in a transmission line can provide a series capacitance up to about 0.5 pF. Greater values (up to about 25 pF) can be obtained using a metal-insulator-metal sandwich in either monolithic or chip (hybrid) form.

5.2 SINGLE-STUB TUNING

Another popular matching technique uses a single open-circuited or short-circuited length of transmission line (a *stub*) connected either in parallel or in series with the transmission feed line at a certain distance from the load, as shown in Figure 5.4. Such a *single-stub tuning* circuit is often very convenient because the stub can be fabricated as part of the transmission line media of the circuit, and lumped elements are avoided. Shunt stubs are preferred for microstrip line or stripline, while series stubs are preferred for slotline or coplanar waveguide.

In single-stub tuning the two adjustable parameters are the distance, d , from the load to the stub position, and the value of susceptance or reactance provided by the stub. For the shunt-stub case, the basic idea is to select d so that the admittance, Y , seen looking into the line at distance d from the load is of the form $Y_0 + jB$. Then the stub susceptance is chosen as $-jB$, resulting in a matched condition. For the series-stub case, the distance d is selected so that the impedance, Z , seen looking into the line at a distance d from the load is of the form $Z_0 + jX$. Then the stub reactance is chosen as $-jX$, resulting in a matched condition.

As discussed in Chapter 2, the proper length of an open or shorted transmission line section can provide any desired value of reactance or susceptance. For a given susceptance or reactance, the difference in lengths of an open- or short-circuited stub is $\lambda/4$.

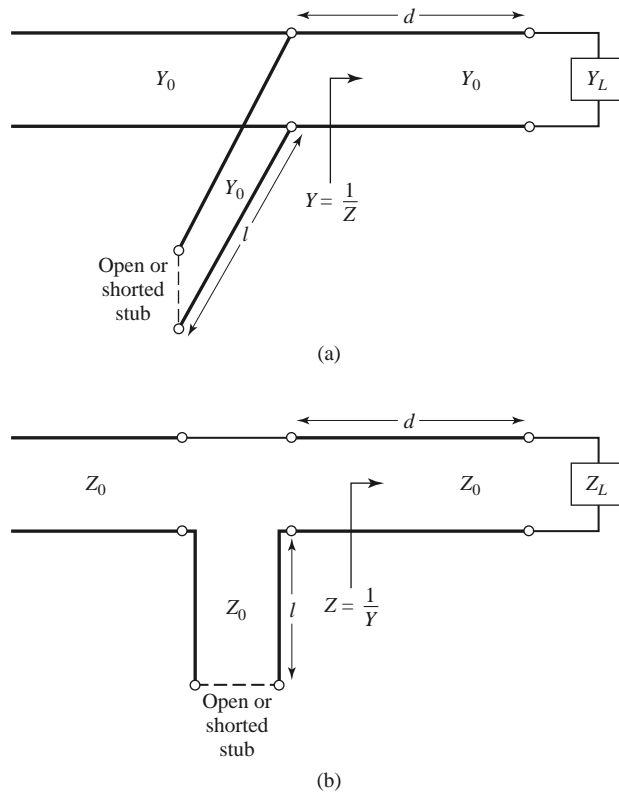


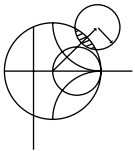
FIGURE 5.4 Single-stub tuning circuits. (a) Shunt stub. (b) Series stub.

For transmission line media such as microstrip or stripline, open-circuited stubs are easier to fabricate since a via hole through the substrate to the ground plane is not needed. For lines like coax or waveguide, however, short-circuited stubs are usually preferred because the cross-sectional area of such an open-circuited line may be large enough (electrically) to radiate, in which case the stub is no longer purely reactive.

We will discuss both Smith chart and analytic solutions for shunt- and series-stub tuning. The Smith chart solutions are fast, intuitive, and usually accurate enough in practice. The analytic expressions are more precise, and are useful for computer analysis.

Shunt Stubs

The single-stub shunt tuning circuit is shown in Figure 5.4a. We will first discuss an example illustrating the Smith chart solution and then derive formulas for d and ℓ .



EXAMPLE 5.2 SINGLE-STUB SHUNT TUNING

For a load impedance $Z_L = 60 - j80 \, \Omega$, design two single-stub (short circuit) shunt tuning networks to match this load to a $50 \, \Omega$ line. Assuming that the load is matched at 2 GHz and that the load consists of a resistor and capacitor in series, plot the reflection coefficient magnitude from 1 to 3 GHz for each solution.

Solution

The first step is to plot the normalized load impedance $z_L = 1.2 - j1.6$, construct the appropriate SWR circle, and convert to the load admittance, y_L , as shown on

the Smith chart in Figure 5.5a. For the remaining steps we consider the Smith chart as an admittance chart. Notice that the SWR circle intersects the $1 + jb$ circle at two points, denoted as y_1 and y_2 in Figure 5.5a. Thus the distance d from the load to the stub is given by either of these two intersections. Reading the WTG scale, we obtain

$$d_1 = 0.176 - 0.065 = 0.110\lambda,$$

$$d_2 = 0.325 - 0.065 = 0.260\lambda.$$

Actually, there is an infinite number of distances d around the SWR circle that intersect the $1 + jb$ circle. Usually it is desired to keep the matching stub as close as possible to the load to improve the bandwidth of the match and to reduce losses caused by a possibly large standing wave ratio on the line between the stub and the load.

At the two intersection points, the normalized admittances are

$$y_1 = 1.00 + j1.47,$$

$$y_2 = 1.00 - j1.47.$$

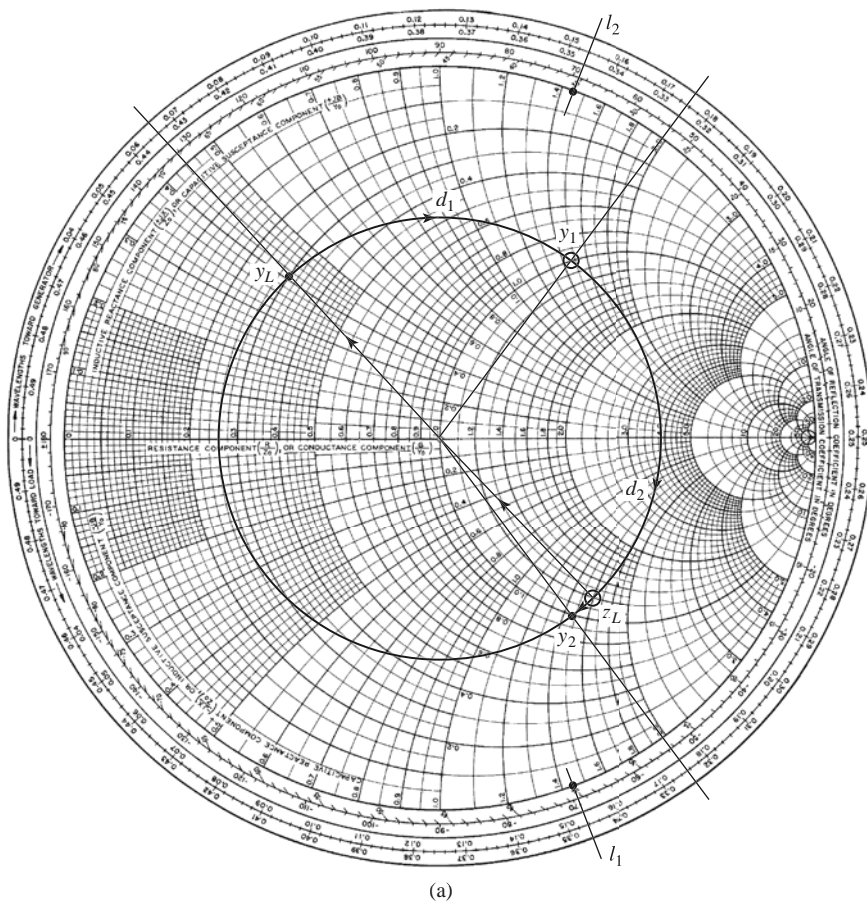


FIGURE 5.5 Solution to Example 5.2. (a) Smith chart for the shunt-stub tuners.

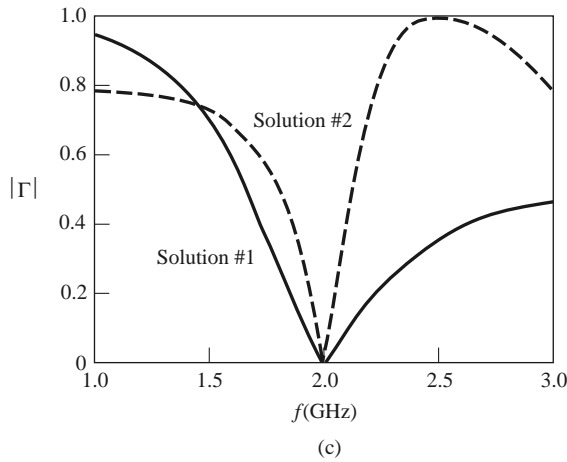
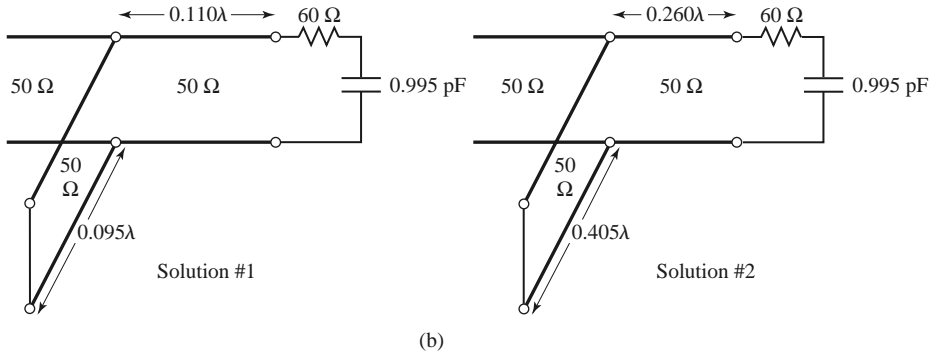


FIGURE 5.5 Continued. (b) The two shunt-stub tuning solutions. (c) Reflection coefficient magnitudes versus frequency for the tuning circuits of (b).

Thus, the first tuning solution requires a stub with a susceptance of $-j1.47$. The length of a short-circuited stub that gives this susceptance can be found on the Smith chart by starting at $y = \infty$ (the short circuit) and moving along the outer edge of the chart ($g = 0$) toward the generator to the $-j1.47$ point. The stub length is then

$$\ell_1 = 0.095\lambda.$$

Similarly, the required short-circuit stub length for the second solution is

$$\ell_2 = 0.405\lambda.$$

This completes the two tuner designs.

To analyze the frequency dependence of these two designs, we need to know the load impedance as a function of frequency. The series- RC load impedance is $Z_L = 60 - j80 \, \Omega$ at 2 GHz, so $R = 60 \, \Omega$ and $C = 0.995 \, \text{pF}$. The two tuning circuits are shown in Figure 5.5b. Figure 5.5c shows the calculated reflection coefficient magnitudes for these two solutions. Observe that solution 1 has a significantly better bandwidth than solution 2; this is because both d and ℓ are shorter for solution 1, which reduces the frequency variation of the match. ■

To derive formulas for d and ℓ , let the load impedance be written as $Z_L = 1/Y_L = R_L + jX_L$. Then the impedance Z down a length d of line from the load is

$$Z = Z_0 \frac{(R_L + jX_L) + jZ_0 t}{Z_0 + j(R_L + jX_L)t}, \quad (5.7)$$

where $t = \tan \beta d$. The admittance at this point is

$$Y = G + jB = \frac{1}{Z},$$

where

$$G = \frac{R_L(1 + t^2)}{R_L^2 + (X_L + Z_0 t)^2}, \quad (5.8a)$$

$$B = \frac{R_L^2 t - (Z_0 - X_L t)(X_L + Z_0 t)}{Z_0 [R_L^2 + (X_L + Z_0 t)^2]}. \quad (5.8b)$$

Now d (which implies t) is chosen so that $G = Y_0 = 1/Z_0$. From (5.8a), this results in a quadratic equation for t :

$$Z_0(R_L - Z_0)t^2 - 2X_L Z_0 t + (R_L Z_0 - R_L^2 - X_L^2) = 0.$$

Solving for t gives

$$t = \frac{X_L \pm \sqrt{R_L [(Z_0 - R_L)^2 + X_L^2]}/Z_0}{R_L - Z_0} \quad \text{for } R_L \neq Z_0. \quad (5.9)$$

If $R_L = Z_0$, then $t = -X_L/2Z_0$. Thus, the two principal solutions for d are

$$\frac{d}{\lambda} = \begin{cases} \frac{1}{2\pi} \tan^{-1} t & \text{for } t \geq 0 \\ \frac{1}{2\pi} (\pi + \tan^{-1} t) & \text{for } t < 0. \end{cases} \quad (5.10)$$

To find the required stub lengths, first use t in (5.8b) to find the stub susceptance, $B_s = -B$. Then, for an open-circuited stub,

$$\frac{\ell_o}{\lambda} = \frac{1}{2\pi} \tan^{-1} \left(\frac{B_s}{Y_0} \right) = \frac{-1}{2\pi} \tan^{-1} \left(\frac{B}{Y_0} \right), \quad (5.11a)$$

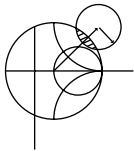
and for a short-circuited stub,

$$\frac{\ell_s}{\lambda} = \frac{-1}{2\pi} \tan^{-1} \left(\frac{Y_0}{B_s} \right) = \frac{1}{2\pi} \tan^{-1} \left(\frac{Y_0}{B} \right). \quad (5.11b)$$

If the length given by (5.11a) or (5.11b) is negative, $\lambda/2$ can be added to give a positive result.

Series Stubs

The series-stub tuning circuit is shown in Figure 5.4b. We will illustrate the Smith chart solution by an example, and then derive expressions for d and ℓ .



EXAMPLE 5.3 SINGLE-STUB SERIES TUNING

Match a load impedance of $Z_L = 100 + j80$ to a 50Ω line using a single series open-circuit stub. Assuming that the load is matched at 2 GHz and that the load

consists of a resistor and inductor in series, plot the reflection coefficient magnitude from 1 to 3 GHz.

Solution

First plot the normalized load impedance, $z_L = 2 + j1.6$, and draw the SWR circle. For the series-stub design the chart is an impedance chart. Note that the SWR circle intersects the $1 + jx$ circle at two points, denoted as z_1 and z_2 in Figure 5.6a. The shortest distance, d_1 , from the load to the stub is, from the WTG scale,

$$d_1 = 0.328 - 0.208 = 0.120\lambda,$$

and the second distance is

$$d_2 = (0.5 - 0.208) + 0.172 = 0.463\lambda.$$

As in the shunt-stub case, additional rotations around the SWR circle lead to additional solutions, but these are usually not of practical interest.

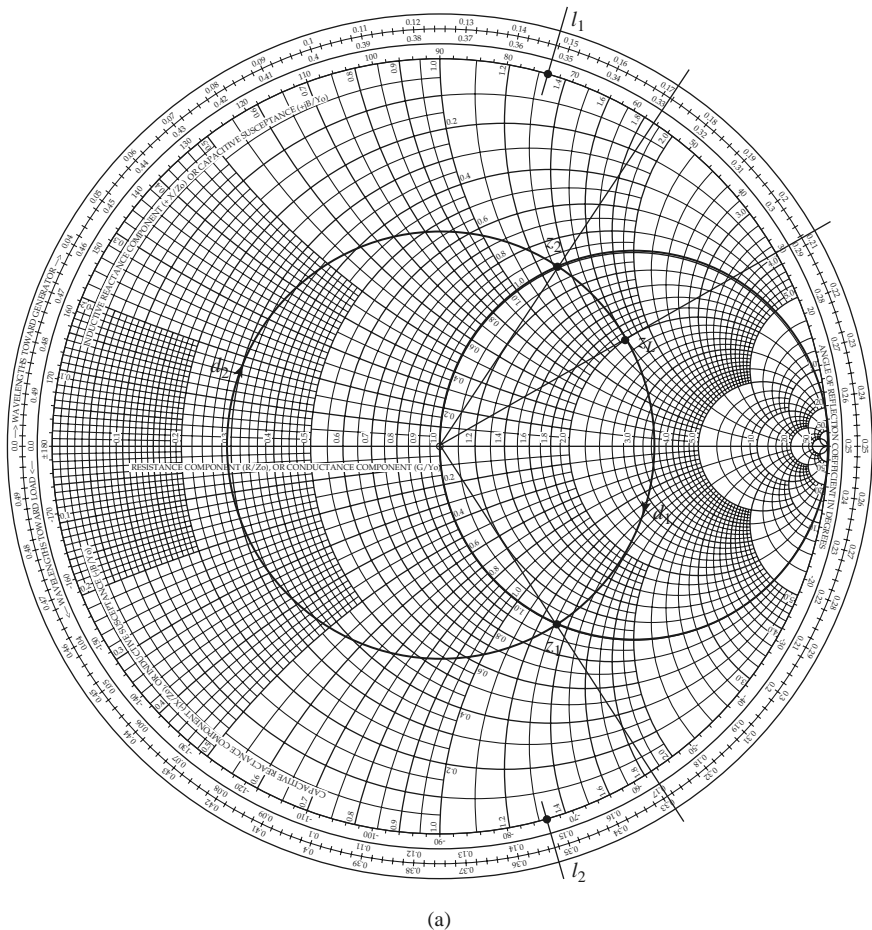


FIGURE 5.6 Solution to Example 5.3. (a) Smith chart for the series-stub tuners.

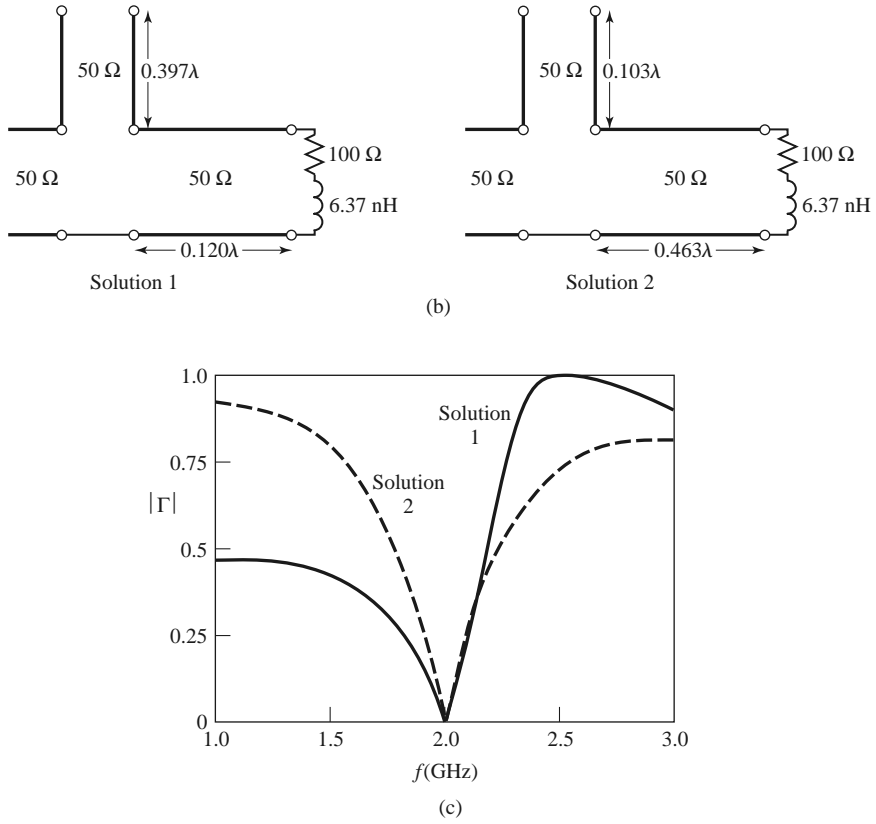


FIGURE 5.6 Continued. (b) The two series-stub tuning solutions. (c) Reflection coefficient magnitudes versus frequency for the tuning circuits of (b).

The normalized impedances at the two intersection points are

$$z_1 = 1 - j1.33,$$

$$z_2 = 1 + j1.33.$$

Thus, the first solution requires a stub with a reactance of $j1.33$. The length of an open-circuited stub that gives this reactance can be found on the Smith chart by starting at $z = \infty$ (open circuit), and moving along the outer edge of the chart ($r = 0$) toward the generator to the $j1.33$ point. This gives a stub length of

$$\ell_1 = 0.397\lambda.$$

Similarly, the required open-circuited stub length for the second solution is

$$\ell_2 = 0.103\lambda.$$

This completes the tuner designs.

If the load is a series resistor and inductor with $Z_L = 100 + j80 \Omega$ at 2 GHz, then $R = 100 \Omega$ and $L = 6.37 \text{ nH}$. The two matching circuits are shown in Figure 5.6b. Figure 5.6c shows the calculated reflection coefficient magnitudes versus frequency for the two solutions. ■

To derive formulas for d and ℓ for the series-stub tuner, let the load admittance be written as $Y_L = 1/Z_L = G_L + jB_L$. Then the admittance Y down a length d of line from the load is

$$Y = Y_0 \frac{(G_L + jB_L) + jtY_0}{Y_0 + jt(G_L + jB_L)}, \quad (5.12)$$

where $t = \tan \beta d$ and $Y_0 = 1/Z_0$. The impedance at this point is

$$Z = R + jX = \frac{1}{Y},$$

where

$$R = \frac{G_L(1 + t^2)}{G_L^2 + (B_L + Y_0 t)^2}, \quad (5.13a)$$

$$X = \frac{G_L^2 t - (Y_0 - tB_L)(B_L + tY_0)}{Y_0 [G_L^2 + (B_L + Y_0 t)^2]}. \quad (5.13b)$$

Now d (which implies t) is chosen so that $R = Z_0 = 1/Y_0$. From (5.13a), this results in a quadratic equation for t :

$$Y_0(G_L - Y_0)t^2 - 2B_L Y_0 t + (G_L Y_0 - G_L^2 - B_L^2) = 0.$$

Solving for t gives

$$t = \frac{B_L \pm \sqrt{G_L [(Y_0 - G_L)^2 + B_L^2] / Y_0}}{G_L - Y_0} \quad \text{for } G_L \neq Y_0. \quad (5.14)$$

If $G_L = Y_0$, then $t = -B_L/2Y_0$. Then the two principal solutions for d are

$$d/\lambda = \begin{cases} \frac{1}{2\pi} \tan^{-1} t & \text{for } t \geq 0 \\ \frac{1}{2\pi} (\pi + \tan^{-1} t) & \text{for } t < 0. \end{cases} \quad (5.15)$$

The required stub lengths are determined by first using t in (5.13b) to find the reactance X . This reactance is the negative of the necessary stub reactance, X_s . Thus, for a short-circuited stub,

$$\frac{\ell_s}{\lambda} = \frac{1}{2\pi} \tan^{-1} \left(\frac{X_s}{Z_0} \right) = \frac{-1}{2\pi} \tan^{-1} \left(\frac{X}{Z_0} \right), \quad (5.16a)$$

and for an open-circuited stub,

$$\frac{\ell_o}{\lambda} = \frac{-1}{2\pi} \tan^{-1} \left(\frac{Z_0}{X_s} \right) = \frac{1}{2\pi} \tan^{-1} \left(\frac{Z_0}{X} \right). \quad (5.16b)$$

If the length given by (5.16a) or (5.16b) is negative, $\lambda/2$ can be added to give a positive result.

5.3

DOUBLE-STUB TUNING

The single-stub tuner of the previous section is able to match any load impedance (having a positive real part) to a transmission line, but suffers from the disadvantage of requiring a variable length of line between the load and the stub. This may not be a problem for a fixed matching circuit, but would probably pose some difficulty if an adjustable tuner was

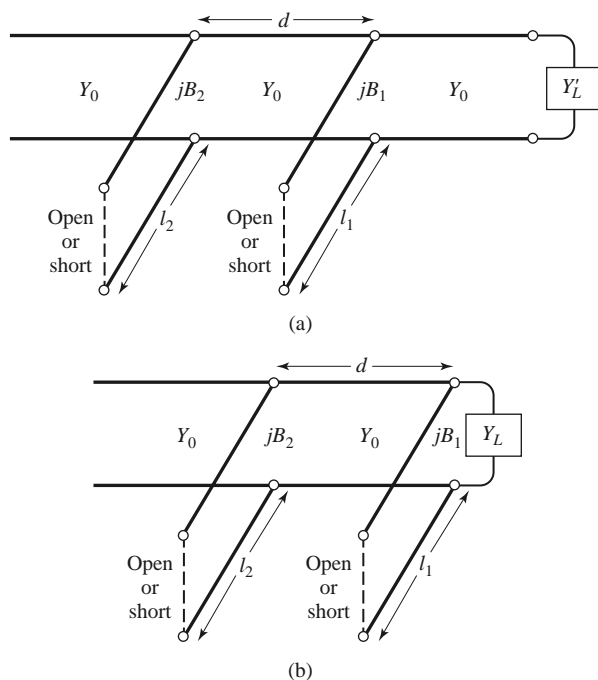


FIGURE 5.7 Double-stub tuning. (a) Original circuit with the load an arbitrary distance from the first stub. (b) Equivalent circuit with the load transformed to the first stub.

desired. In this case, the *double-stub tuner*, which uses two tuning stubs in fixed positions, can be used. Such tuners are often fabricated in coaxial line with adjustable stubs connected in shunt to the main coaxial line. We will see, however, that a double-stub tuner cannot match all load impedances.

The double-stub tuner circuit is shown in Figure 5.7a, where the load may be an arbitrary distance from the first stub. Although this is more representative of a practical situation, the circuit of Figure 5.7b, where the load Y_L' has been transformed back to the position of the first stub, is easier to deal with and does not lose any generality. The shunt stubs shown in Figure 5.7 can be conveniently implemented for some types of transmission lines, while series stubs are more appropriate for other types of lines. In either case, the stubs can be open-circuited or short-circuited.

Smith Chart Solution

The Smith chart of Figure 5.8 illustrates the basic operation of the double-stub tuner. As in the case of the single-stub tuner, two solutions are possible. The susceptance of the first stub, b_1 (or b_1' , for the second solution), moves the load admittance to y_1 (or y_1'). These points lie on the rotated $1 + jb$ circle; the amount of rotation is d wavelengths toward the load, where d is the electrical distance between the two stubs. Then transforming y_1 (or y_1') toward the generator through a length d of line leaves us at the point y_2 (or y_2'), which must be on the $1 + jb$ circle. The second stub then adds a susceptance b_2 (or b_2'), which brings us to the center of the chart and completes the match.

Notice from Figure 5.8 that if the load admittance, y_L , were inside the shaded region of the $g_0 + jb$ circle, no value of stub susceptance b_1 could ever bring the load point to intersect the rotated $1 + jb$ circle. This shaded region thus forms a forbidden range of load admittances that cannot be matched with this particular double-stub tuner. A simple way

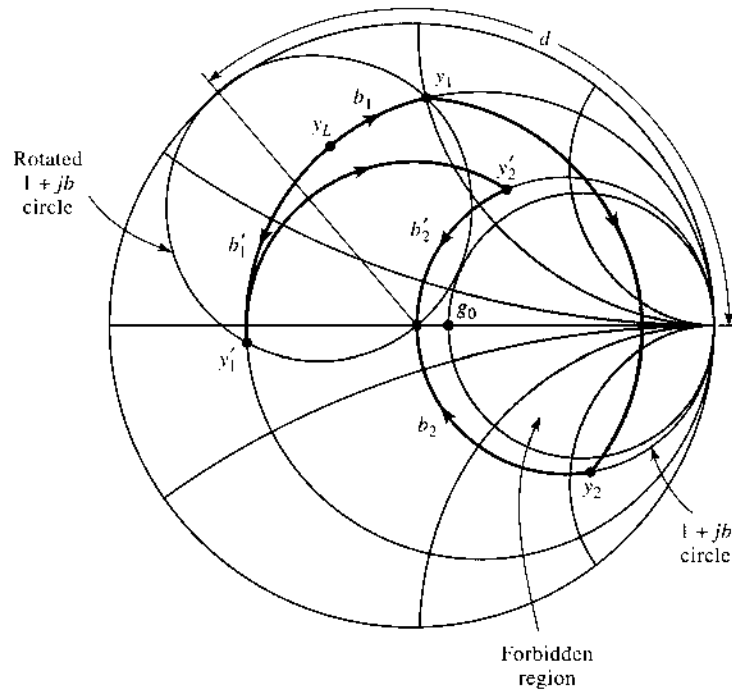
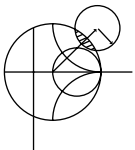


FIGURE 5.8 Smith chart diagram for the operation of a double-stub tuner.

of reducing the forbidden range is to reduce the distance d between the stubs. This has the effect of swinging the rotated $1 + jb$ circle back toward the $y = \infty$ point, but d must be kept large enough for the practical purpose of fabricating the two separate stubs. In addition, stub spacings near 0 or $\lambda/2$ lead to matching networks that are very frequency sensitive. In practice, stub spacings are usually chosen as $\lambda/8$ or $3\lambda/8$. If the length of line between the load and the first stub can be adjusted, then the load admittance y_L can always be moved out of the forbidden region.



EXAMPLE 5.4 DOUBLE-STUB TUNING

Design a double-stub shunt tuner to match a load impedance $Z_L = 60 - j80 \, \Omega$ to a $50 \, \Omega$ line. The stubs are to be open-circuited stubs and are spaced $\lambda/8$ apart. Assuming that this load consists of a series resistor and capacitor and that the match frequency is 2 GHz, plot the reflection coefficient magnitude versus frequency from 1 to 3 GHz.

Solution

The normalized load admittance is $y_L = 0.3 + j0.4$, which is plotted on the Smith chart of Figure 5.9a. Next we construct the rotated $1 + jb$ conductance circle by moving every point on the $g = 1$ circle $\lambda/8$ toward the load. We then find the susceptance of the first stub, which can be one of two possible values:

$$b_1 = 1.314 \quad \text{or} \quad b'_1 = -0.114.$$

We now transform through the $\lambda/8$ section of line by rotating along a constant-radius (SWR) circle $\lambda/8$ toward the generator. This brings the two solutions to the

following points:

$$y_2 = 1 - j3.38 \quad \text{or} \quad y'_2 = 1 + j1.38.$$

Then the susceptance of the second stub should be

$$b_2 = 3.38 \quad \text{or} \quad b'_2 = -1.38.$$

The lengths of the open-circuited stubs are then found as

$$\ell_1 = 0.146\lambda, \ell_2 = 0.204\lambda \quad \text{or} \quad \ell'_1 = 0.482\lambda, \ell'_2 = 0.350\lambda.$$

This completes both solutions for the double-stub tuner design.

At $f = 2$ GHz the resistor-capacitor load of $Z_L = 60 - j80 \, \Omega$ implies that $R = 60 \, \Omega$ and $C = 0.995$ pF. The two tuning circuits are then as shown in Figure 5.9b, and the reflection coefficient magnitudes are plotted versus frequency in Figure 5.9c. Note that the first solution has a much narrower bandwidth than the second (primed) solution due to the fact that both stubs for the first solution are somewhat longer (and closer to $\lambda/2$) than the stubs of the second solution. ■

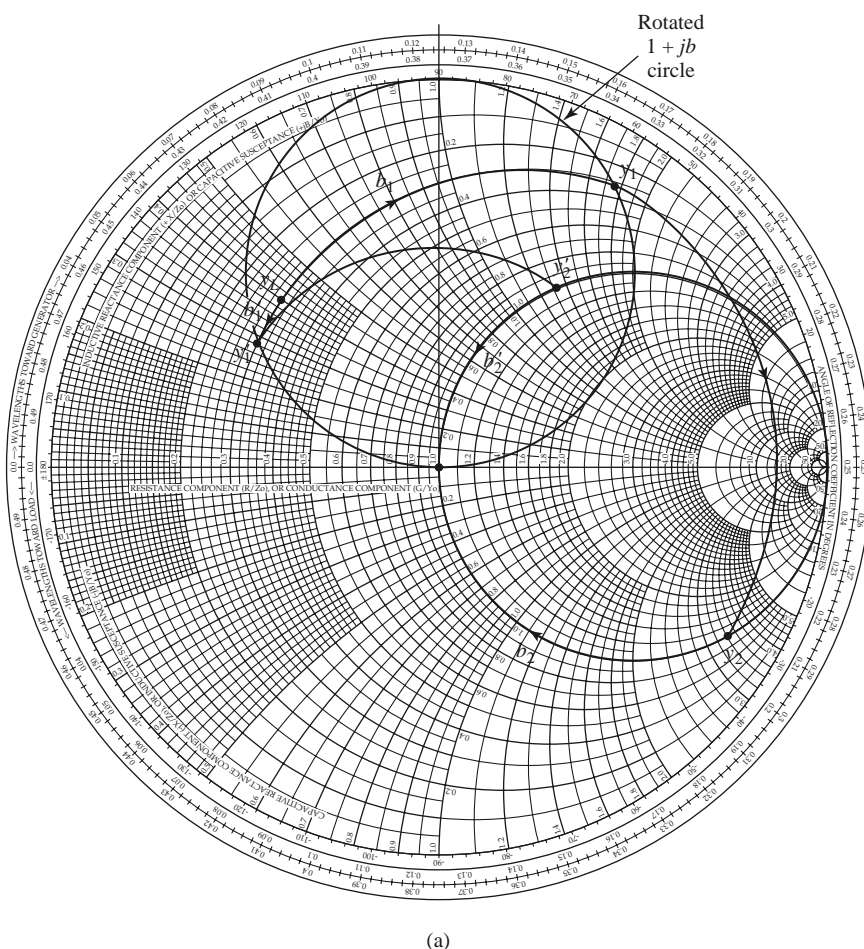


FIGURE 5.9 Solution to Example 5.4. (a) Smith chart for the double-stub tuners.

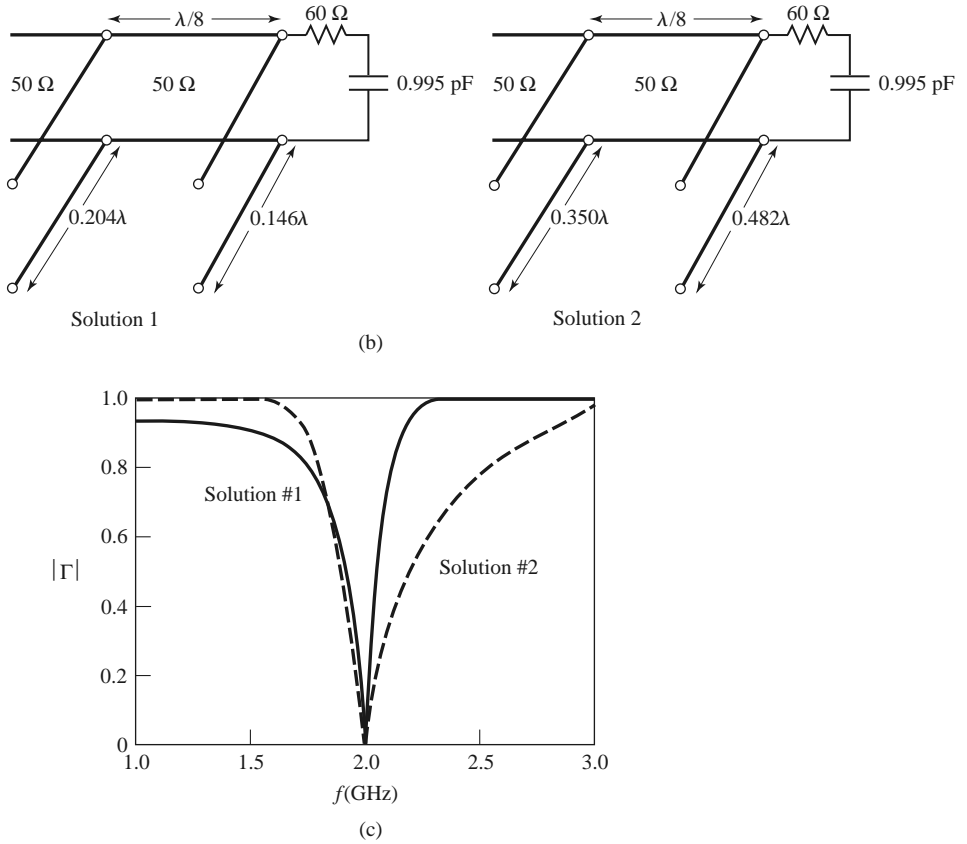


FIGURE 5.9 Continued. (b) The two double-stub tuning solutions. (c) Reflection coefficient magnitudes versus frequency for the tuning circuits of (b).

Analytic Solution

The admittance just to the left of the first stub in Figure 5.7b is

$$Y_1 = G_L + j(B_L + B_1), \quad (5.17)$$

where $Y_L = G_L + jB_L$ is the load admittance, and B_1 is the susceptance of the first stub. After transforming through a length d of transmission line, we find that the admittance just to the right of the second stub is

$$Y_2 = Y_0 \frac{G_L + j(B_L + B_1 + Y_0 t)}{Y_0 + j t(G_L + j B_L + j B_1)}, \quad (5.18)$$

where $t = \tan \beta d$ and $Y_0 = 1/Z_0$. At this point the real part of Y_2 must equal Y_0 , which leads to the equation

$$G_L^2 - G_L Y_0 \frac{1+t^2}{t^2} + \frac{(Y_0 - B_L t - B_1 t)^2}{t^2} = 0. \quad (5.19)$$

Solving for G_L gives

$$G_L = Y_0 \frac{1+t^2}{2t^2} \left[1 \pm \sqrt{1 - \frac{4t^2(Y_0 - B_L t - B_1 t)^2}{Y_0^2(1+t^2)^2}} \right]. \quad (5.20)$$

Because G_L is real, the quantity within the square root must be nonnegative, and so

$$0 \leq \frac{4t^2(Y_0 - B_L t - B_1 t)^2}{Y_0^2(1 + t^2)^2} \leq 1.$$

This implies that

$$0 \leq G_L \leq Y_0 \frac{1 + t^2}{t^2} = \frac{Y_0}{\sin^2 \beta d}, \quad (5.21)$$

which gives the range on G_L that can be matched for a given stub spacing d . After d has been set, the first stub susceptance can be determined from (5.19) as

$$B_1 = -B_L + \frac{Y_0 \pm \sqrt{(1 + t^2)G_L Y_0 - G_L^2 t^2}}{t}. \quad (5.22)$$

Then the second stub susceptance can be found from the negative of the imaginary part of (5.18) to be

$$B_2 = \frac{\pm Y_0 \sqrt{Y_0 G_L (1 + t^2) - G_L^2 t^2} + G_L Y_0}{G_L t}. \quad (5.23)$$

The upper and lower signs in (5.22) and (5.23) correspond to the same solutions. The open-circuited stub length is found as

$$\frac{\ell_o}{\lambda} = \frac{1}{2\pi} \tan^{-1} \left(\frac{B}{Y_0} \right), \quad (5.24a)$$

and the short-circuited stub length is found as

$$\frac{\ell_s}{\lambda} = \frac{-1}{2\pi} \tan^{-1} \left(\frac{Y_0}{B} \right), \quad (5.24b)$$

where $B = B_1$ or B_2 .

5.4 THE QUARTER-WAVE TRANSFORMER

As introduced in Section 2.5, the quarter-wave transformer is a simple and useful circuit for matching a real load impedance to a transmission line. An additional feature of the quarter-wave transformer is that it can be extended to multisection designs in a methodical manner to provide broader bandwidth. If only a narrow band impedance match is required, a single-section transformer may suffice. However, as we will see in the next few sections, multisection quarter-wave transformer designs can be synthesized to yield optimum matching characteristics over a desired frequency band. We will see in Chapter 8 that such networks are closely related to bandpass filters.

One drawback of the quarter-wave transformer is that it can only match a real load impedance. A complex load impedance can always be transformed into a real impedance, however, by using an appropriate length of transmission line between the load and the transformer, or an appropriate series or shunt reactive element. These techniques will usually alter the frequency dependence of the load, and this often has the effect of reducing the bandwidth of the match.

In Section 2.5 we analyzed the operation of a quarter-wave transformer from both an impedance viewpoint and a multiple reflection viewpoint. Here we will concentrate on the bandwidth performance of the transformer as a function of the load mismatch; this

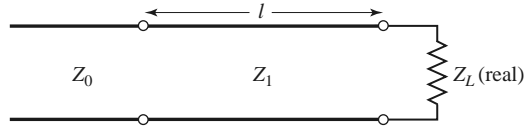


FIGURE 5.10 A single-section quarter-wave matching transformer. $\ell = \lambda_0/4$ at the design frequency f_0 .

discussion will also serve as a prelude to the more general case of multisection transformers in the sections to follow.

The single-section quarter-wave matching transformer circuit is shown in Figure 5.10, with the characteristic impedance of the matching section given as

$$Z_1 = \sqrt{Z_0 Z_L}. \quad (5.25)$$

At the design frequency, f_0 , the electrical length of the matching section is $\lambda_0/4$, but at other frequencies the length is different, so a perfect match is no longer obtained. We will derive an approximate expression for the resulting impedance mismatch versus frequency.

The input impedance seen looking into the matching section is

$$Z_{\text{in}} = Z_1 \frac{Z_L + j Z_1 t}{Z_1 + j Z_L t}, \quad (5.26)$$

where $t = \tan \beta \ell = \tan \theta$, and $\beta \ell = \theta = \pi/2$ at the design frequency f_0 . The resulting reflection coefficient is

$$\Gamma = \frac{Z_{\text{in}} - Z_0}{Z_{\text{in}} + Z_0} = \frac{Z_1(Z_L - Z_0) + jt(Z_1^2 - Z_0 Z_L)}{Z_1(Z_L + Z_0) + jt(Z_1^2 + Z_0 Z_L)}. \quad (5.27)$$

Because $Z_1^2 = Z_0 Z_L$, this reduces to

$$\Gamma = \frac{Z_L - Z_0}{Z_L + Z_0 + j 2t \sqrt{Z_0 Z_L}}. \quad (5.28)$$

The reflection coefficient magnitude is

$$\begin{aligned} |\Gamma| &= \frac{|Z_L - Z_0|}{[(Z_L + Z_0)^2 + 4t^2 Z_0 Z_L]^{1/2}} \\ &= \frac{1}{\{(Z_L + Z_0)^2/(Z_L - Z_0)^2 + [4t^2 Z_0 Z_L/(Z_L - Z_0)^2]\}^{1/2}} \\ &= \frac{1}{\{1 + [4Z_0 Z_L/(Z_L - Z_0)^2] + [4Z_0 Z_L t^2/(Z_L - Z_0)^2]\}^{1/2}} \\ &= \frac{1}{\{1 + [4Z_0 Z_L/(Z_L - Z_0)^2] \sec^2 \theta\}^{1/2}}, \end{aligned} \quad (5.29)$$

since $1 + t^2 = 1 + \tan^2 \theta = \sec^2 \theta$.

If we assume that the operating frequency is near the design frequency f_0 , then $\ell \simeq \lambda_0/4$ and $\theta \simeq \pi/2$. Then $\sec^2 \theta \gg 1$, and (5.29) simplifies to

$$|\Gamma| \simeq \frac{|Z_L - Z_0|}{2\sqrt{Z_0 Z_L}} |\cos \theta| \quad \text{for } \theta \text{ near } \pi/2. \quad (5.30)$$

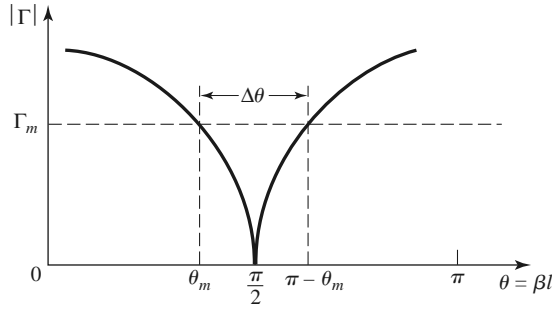


FIGURE 5.11 Approximate behavior of the reflection coefficient magnitude for a single-section quarter-wave transformer operating near its design frequency.

This result gives the approximate mismatch of the quarter-wave transformer near the design frequency, as sketched in Figure 5.11.

If we set a maximum value, Γ_m , for an acceptable reflection coefficient magnitude, then the bandwidth of the matching transformer can be defined as

$$\Delta\theta = 2\left(\frac{\pi}{2} - \theta_m\right), \quad (5.31)$$

since the response of (5.29) is symmetric about $\theta = \pi/2$, and $\Gamma = \Gamma_m$ at $\theta = \theta_m$ and at $\theta = \pi - \theta_m$. Equating Γ_m to the exact expression for the reflection coefficient magnitude in (5.29) allows us to solve for θ_m :

$$\frac{1}{\Gamma_m^2} = 1 + \left(\frac{2\sqrt{Z_0 Z_L}}{Z_L - Z_0} \sec \theta_m\right)^2,$$

or

$$\cos \theta_m = \frac{\Gamma_m}{\sqrt{1 - \Gamma_m^2}} \frac{2\sqrt{Z_0 Z_L}}{|Z_L - Z_0|}. \quad (5.32)$$

If we assume TEM lines, then

$$\theta = \beta\ell = \frac{2\pi f}{v_p} \frac{v_p}{4f_0} = \frac{\pi f}{2f_0},$$

and so the frequency of the lower band edge at $\theta = \theta_m$ is

$$f_m = \frac{2\theta_m f_0}{\pi},$$

and the fractional bandwidth is, using (5.32),

$$\begin{aligned} \frac{\Delta f}{f_0} &= \frac{2(f_0 - f_m)}{f_0} = 2 - \frac{2f_m}{f_0} = 2 - \frac{4\theta_m}{\pi} \\ &= 2 - \frac{4}{\pi} \cos^{-1} \left[\frac{\Gamma_m}{\sqrt{1 - \Gamma_m^2}} \frac{2\sqrt{Z_0 Z_L}}{|Z_L - Z_0|} \right]. \end{aligned} \quad (5.33)$$

Fractional bandwidth is usually expressed as a percentage, $100\Delta f/f_0\%$. Note that the bandwidth of the transformer increases as Z_L becomes closer to Z_0 (a less mismatched load).

The above results are strictly valid only for TEM lines. When non-TEM lines (such as waveguides) are used, the propagation constant is no longer a linear function of frequency, and the wave impedance will be frequency dependent. These factors serve to complicate

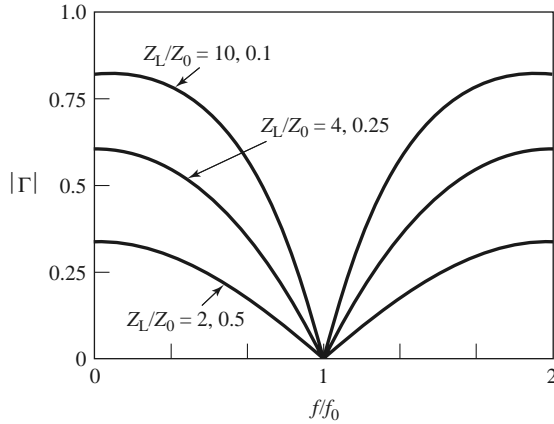
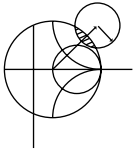


FIGURE 5.12 Reflection coefficient magnitude versus frequency for a single-section quarter-wave matching transformer with various load mismatches.

the general behavior of quarter-wave transformers for non-TEM lines, but in practice the bandwidth of the transformer is often small enough that these complications do not substantially affect the result. Another factor ignored in the above analysis is the effect of reactances associated with discontinuities when there is a step change in the dimensions of a transmission line. This can often be compensated by making a small adjustment in the length of the matching section.

Figure 5.12 shows a plot of the reflection coefficient magnitude versus normalized frequency for various mismatched loads. Note the trend of increased bandwidth for smaller load mismatches.



EXAMPLE 5.5 QUARTER-WAVE TRANSFORMER BANDWIDTH

Design a single-section quarter-wave matching transformer to match a $10\ \Omega$ load to a $50\ \Omega$ transmission line at $f_0 = 3\ \text{GHz}$. Determine the percent bandwidth for which the $\text{SWR} \leq 1.5$.

Solution

From (5.25), the characteristic impedance of the matching section is

$$Z_1 = \sqrt{Z_0 Z_L} = \sqrt{(50)(10)} = 22.36\ \Omega,$$

and the length of the matching section is $\lambda/4$ at $3\ \text{GHz}$ (the physical length depends on the dielectric constant of the line). An SWR of 1.5 corresponds to a reflection coefficient magnitude of

$$\Gamma_m = \frac{\text{SWR} - 1}{\text{SWR} + 1} = \frac{1.5 - 1}{1.5 + 1} = 0.2.$$

The fractional bandwidth is computed from (5.33) as

$$\begin{aligned} \frac{\Delta f}{f_0} &= 2 - \frac{4}{\pi} \cos^{-1} \left[\frac{\Gamma_m}{\sqrt{1 - \Gamma_m^2}} \frac{2\sqrt{Z_0 Z_L}}{|Z_L - Z_0|} \right] \\ &= 2 - \frac{4}{\pi} \cos^{-1} \left[\frac{0.2}{\sqrt{1 - (0.2)^2}} \frac{2\sqrt{(50)(10)}}{|10 - 50|} \right] \\ &= 0.29, \text{ or } 29\%. \end{aligned}$$

5.5 THE THEORY OF SMALL REFLECTIONS

The quarter-wave transformer provides a simple means of matching any real load impedance to any transmission line impedance. For applications requiring more bandwidth than a single quarter-wave section can provide, *multisection transformers* can be used. The design of such transformers is the subject of the next two sections, but prior to that material we need to derive some approximate results for the total reflection coefficient caused by the partial reflections from several small discontinuities. This topic is generally referred to as the *theory of small reflections* [1].

Single-Section Transformer

We will derive an approximate expression for the overall reflection coefficient, Γ , for the single-section matching transformer shown in Figure 5.13. The partial reflection and transmission coefficients are

$$\Gamma_1 = \frac{Z_2 - Z_1}{Z_2 + Z_1}, \quad (5.34)$$

$$\Gamma_2 = -\Gamma_1, \quad (5.35)$$

$$\Gamma_3 = \frac{Z_L - Z_2}{Z_L + Z_2}, \quad (5.36)$$

$$T_{21} = 1 + \Gamma_1 = \frac{2Z_2}{Z_1 + Z_2}, \quad (5.37)$$

$$T_{12} = 1 + \Gamma_2 = \frac{2Z_1}{Z_1 + Z_2}. \quad (5.38)$$

We can compute the total reflection, Γ , seen by the feed line using either the impedance method, or the multiple reflection method, as discussed in Section 2.5. For our present

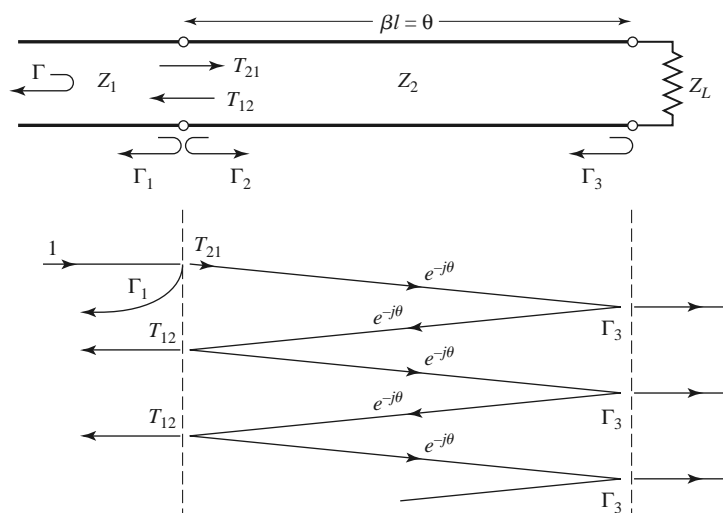


FIGURE 5.13 Partial reflections and transmissions on a single-section matching transformer.

purpose the latter technique is preferred, so we express the total reflection as an infinite sum of partial reflections and transmissions as follows:

$$\begin{aligned}\Gamma &= \Gamma_1 + T_{12}T_{21}\Gamma_3e^{-2j\theta} + T_{12}T_{21}\Gamma_3^2\Gamma_2e^{-4j\theta} + \dots \\ &= \Gamma_1 + T_{12}T_{21}\Gamma_3e^{-2j\theta} \sum_{n=0}^{\infty} \Gamma_2^n \Gamma_3^n e^{-2jn\theta}.\end{aligned}\quad (5.39)$$

The summation of the geometric series

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \quad \text{for } |x| < 1$$

allows us to express (5.39) in closed form as

$$\Gamma = \Gamma_1 + \frac{T_{12}T_{21}\Gamma_3e^{-2j\theta}}{1 - \Gamma_2\Gamma_3e^{-2j\theta}}. \quad (5.40)$$

From (5.35), (5.37), and (5.38), we use $\Gamma_2 = -\Gamma_1$, $T_{21} = 1 + \Gamma_1$, and $T_{12} = 1 - \Gamma_1$ in (5.40) to give

$$\Gamma = \frac{\Gamma_1 + \Gamma_3e^{-2j\theta}}{1 + \Gamma_1\Gamma_3e^{-2j\theta}}. \quad (5.41)$$

If the discontinuities between the impedances Z_1 , Z_2 and Z_2 , Z_L are small, then $|\Gamma_1\Gamma_3| \ll 1$, so we can approximate (5.41) as

$$\Gamma \simeq \Gamma_1 + \Gamma_3e^{-2j\theta}. \quad (5.42)$$

This result expresses the intuitive idea that the total reflection is dominated by the reflection from the initial discontinuity between Z_1 and Z_2 (Γ_1), and the first reflection from the discontinuity between Z_2 and Z_L ($\Gamma_3e^{-2j\theta}$). The $e^{-2j\theta}$ term accounts for the phase delay when the incident wave travels up and down the line. The accuracy of this approximation is illustrated in Problem 5.14.

Multisection Transformer

Now consider the multisection transformer shown in Figure 5.14, which consists of N equal-length (*commensurate*) sections of transmission lines. We will derive an approximate expression for the total reflection coefficient Γ .

Partial reflection coefficients can be defined at each junction, as follows:

$$\Gamma_0 = \frac{Z_1 - Z_0}{Z_1 + Z_0}, \quad (5.43a)$$

$$\Gamma_n = \frac{Z_{n+1} - Z_n}{Z_{n+1} + Z_n}, \quad (5.43b)$$

$$\Gamma_N = \frac{Z_L - Z_N}{Z_L + Z_N}. \quad (5.43c)$$

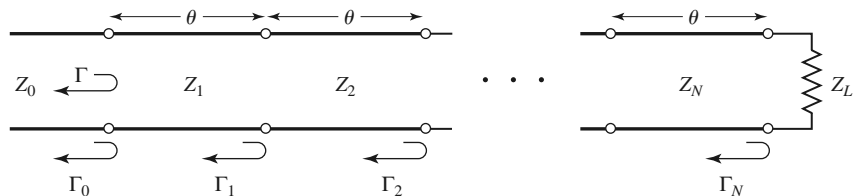


FIGURE 5.14 Partial reflection coefficients for a multisection matching transformer.

We also assume that all Z_n increase or decrease monotonically across the transformer and that Z_L is real. This implies that all Γ_n will be real and of the same sign ($\Gamma_n > 0$ if $Z_L > Z_0$; $\Gamma_n < 0$ if $Z_L < Z_0$). Using the results of the previous section allows us to approximate the overall reflection coefficient as

$$\Gamma(\theta) = \Gamma_0 + \Gamma_1 e^{-2j\theta} + \Gamma_2 e^{-4j\theta} + \cdots + \Gamma_N e^{-2jN\theta}. \quad (5.44)$$

Further assume that the transformer can be made symmetrical, so that $\Gamma_0 = \Gamma_N$, $\Gamma_1 = \Gamma_{N-1}$, $\Gamma_2 = \Gamma_{N-2}$, and so on. (Note that this does *not* imply that the Z_n are symmetrical.) Then (5.44) can be written as

$$\Gamma(\theta) = e^{-jN\theta} \left\{ \Gamma_0 [e^{jN\theta} + e^{-jN\theta}] + \Gamma_1 [e^{j(N-2)\theta} + e^{-j(N-2)\theta}] + \cdots \right\}. \quad (5.45)$$

If N is odd, the last term is $\Gamma_{(N-1)/2} (e^{j\theta} + e^{-j\theta})$, while if N is even, the last term is $\Gamma_{N/2}$. Equation (5.45) is seen to be of the form of a finite Fourier cosine series in θ , which can be written as

$$\begin{aligned} \Gamma(\theta) = 2e^{-jN\theta} \left[\Gamma_0 \cos N\theta + \Gamma_1 \cos(N-2)\theta + \cdots + \Gamma_n \cos(N-2n)\theta \right. \\ \left. + \cdots + \frac{1}{2} \Gamma_{N/2} \right] \quad \text{for } N \text{ even,} \end{aligned} \quad (5.46a)$$

$$\begin{aligned} \Gamma(\theta) = 2e^{-jN\theta} [\Gamma_0 \cos N\theta + \Gamma_1 \cos(N-2)\theta + \cdots + \Gamma_n \cos(N-2n)\theta \\ + \cdots + \Gamma_{(N-1)/2} \cos \theta] \quad \text{for } N \text{ odd.} \end{aligned} \quad (5.46b)$$

The importance of these results lies in the fact that we can synthesize any desired reflection coefficient response as a function of frequency (θ) by properly choosing the Γ_n and using enough sections (N). This should be clear from the realization that a Fourier series can approximate an arbitrary smooth function if enough terms are used. In the next two sections we will show how to use this theory to design multisection transformers for two of the most commonly used passband responses: the *binomial* (maximally flat) response, and the *Chebyshev* (equal-ripple) response.

5.6

BINOMIAL MULTISECTION MATCHING TRANSFORMERS

The passband response (the frequency band where a good impedance match is achieved) of a binomial matching transformer is optimum in the sense that, for a given number of sections, the response is as flat as possible near the design frequency. This type of response, which is also known as *maximally flat*, is determined for an N -section transformer by setting the first $N-1$ derivatives of $|\Gamma(\theta)|$ to zero at the center frequency, f_0 . Such a response can be obtained with a reflection coefficient of the following form:

$$\Gamma(\theta) = A(1 + e^{-2j\theta})^N. \quad (5.47)$$

Then the reflection coefficient magnitude is

$$\begin{aligned} |\Gamma(\theta)| &= |A| |e^{-j\theta}|^N |e^{j\theta} + e^{-j\theta}|^N \\ &= 2^N |A| |\cos \theta|^N \end{aligned} \quad (5.48)$$

Note that $|\Gamma(\theta)| = 0$ for $\theta = \pi/2$, and that $d^n |\Gamma(\theta)| / d\theta^n = 0$ at $\theta = \pi/2$ for $n = 1, 2, \dots, N-1$. ($\theta = \pi/2$ corresponds to the center frequency, f_0 , for which $\ell = \lambda/4$ and $\theta = \beta\ell = \pi/2$.)

We can determine the constant A by letting $f \rightarrow 0$. Then $\theta = \beta\ell = 0$, and (5.47) reduces to

$$\Gamma(0) = 2^N A = \frac{Z_L - Z_0}{Z_L + Z_0},$$

since for $f = 0$ all sections are of zero electrical length. The constant A can then be written as

$$A = 2^{-N} \frac{Z_L - Z_0}{Z_L + Z_0}. \quad (5.49)$$

Next we expand $\Gamma(\theta)$ in (5.47) according to the binomial expansion:

$$\Gamma(\theta) = A(1 + e^{-2j\theta})^N = A \sum_{n=0}^N C_n^N e^{-2jn\theta}, \quad (5.50)$$

where

$$C_n^N = \frac{N!}{(N-n)!n!} \quad (5.51)$$

are the binomial coefficients. Note that $C_n^N = C_{N-n}^N$, $C_0^N = 1$, and $C_1^N = N = C_{N-1}^N$. The key step is now to equate the desired passband response, given by (5.50), to the actual response as given (approximately) by (5.44):

$$\Gamma(\theta) = A \sum_{n=0}^N C_n^N e^{-2jn\theta} = \Gamma_0 + \Gamma_1 e^{-2j\theta} + \Gamma_2 e^{-4j\theta} + \cdots + \Gamma_N e^{-2jN\theta}.$$

This shows that the Γ_n must be chosen as

$$\Gamma_n = AC_n^N. \quad (5.52)$$

where A is given by (5.49) and C_n^N is a binomial coefficient.

At this point, the characteristic impedances, Z_n , can be found via (5.43), but a simpler solution can be obtained using the following approximation [1]. Because we assumed that the Γ_n are small, we can write

$$\Gamma_n = \frac{Z_{n+1} - Z_n}{Z_{n+1} + Z_n} \simeq \frac{1}{2} \ln \frac{Z_{n+1}}{Z_n},$$

since $\ln x \simeq 2(x - 1)/(x + 1)$ for x close to unity. Then, using (5.52) and (5.49) gives

$$\ln \frac{Z_{n+1}}{Z_n} \simeq 2\Gamma_n = 2AC_n^N = 2(2^{-N}) \frac{Z_L - Z_0}{Z_L + Z_0} C_n^N \simeq 2^{-N} C_n^N \ln \frac{Z_L}{Z_0}, \quad (5.53)$$

which can be used to find Z_{n+1} , starting with $n = 0$. This technique has the advantage of ensuring self-consistency, in that Z_{N+1} computed from (5.53) will be equal to Z_L , as it should.

Exact design results, including the effect of multiple reflections in each section, can be found by using the transmission line equations for each section and numerically solving for the characteristic impedances [2]. The results of such calculations are listed in Table 5.1, which gives the exact line impedances for $N = 2$ -, 3-, 4-, 5-, and 6-section

TABLE 5.1 Binomial Transformer Design

Z_L/Z_0	$N = 2$		$N = 3$		$N = 4$			
	Z_1/Z_0	Z_2/Z_0	Z_1/Z_0	Z_2/Z_0	Z_1/Z_0	Z_2/Z_0	Z_3/Z_0	Z_4/Z_0
1.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1.5	1.1067	1.3554	1.0520	1.2247	1.0257	1.1351	1.3215	1.4624
2.0	1.1892	1.6818	1.0907	1.4142	1.0444	1.2421	1.6102	1.9150
3.0	1.3161	2.2795	1.1479	1.7321	1.0718	1.4105	2.1269	2.7990
4.0	1.4142	2.8285	1.1907	2.0000	1.0919	1.5442	2.5903	3.6633
6.0	1.5651	3.8336	1.2544	2.4495	1.1215	1.7553	3.4182	5.3500
8.0	1.6818	4.7568	1.3022	2.8284	1.1436	1.9232	4.1597	6.9955
10.0	1.7783	5.6233	1.3409	3.1623	1.1613	2.0651	4.8424	8.6110

Z_L/Z_0	$N = 5$					$N = 6$				
	Z_1/Z_0	Z_2/Z_0	Z_3/Z_0	Z_4/Z_0	Z_5/Z_0	Z_1/Z_0	Z_2/Z_0	Z_3/Z_0	Z_4/Z_0	Z_5/Z_0
1.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1.5	1.0128	1.0790	1.2247	1.3902	1.4810	1.0064	1.0454	1.1496	1.3048	1.4349
2.0	1.0220	1.1391	1.4142	1.7558	1.9569	1.0110	1.0790	1.2693	1.5757	1.8536
3.0	1.0354	1.2300	1.7321	2.4390	2.8974	1.0176	1.1288	1.4599	2.0549	2.6577
4.0	1.0452	1.2995	2.0000	3.0781	3.8270	1.0225	1.1661	1.6129	2.4800	3.4302
6.0	1.0596	1.4055	2.4495	4.2689	5.6625	1.0296	1.2219	1.8573	3.2305	4.9104
8.0	1.0703	1.4870	2.8284	5.3800	7.4745	1.0349	1.2640	2.0539	3.8950	6.3291
10.0	1.0789	1.5541	3.1623	6.4346	9.2687	1.0392	1.2982	2.2215	4.5015	7.7030

binomial matching transformers for various ratios of load impedance, Z_L , to feed line impedance, Z_0 . The table gives results only for $Z_L/Z_0 > 1$; if $Z_L/Z_0 < 1$, the results for Z_0/Z_L should be used but with Z_1 starting at the load end. This is because the solution is symmetric about $Z_L/Z_0 = 1$; the same transformer that matches Z_L to Z_0 can be reversed and used to match Z_0 to Z_L . More extensive tables can be found in reference [2].

The bandwidth of the binomial transformer can be evaluated as follows. As in Section 5.4, let Γ_m be the maximum value of reflection coefficient that can be tolerated over the passband. Then from (5.48),

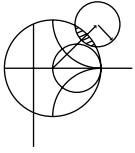
$$\Gamma_m = 2^N |A| \cos^N \theta_m,$$

where $\theta_m < \pi/2$ is the lower edge of the passband, as shown in Figure 5.11. Thus,

$$\theta_m = \cos^{-1} \left[\frac{1}{2} \left(\frac{\Gamma_m}{|A|} \right)^{1/N} \right], \quad (5.54)$$

and using (5.33) gives the fractional bandwidth as

$$\begin{aligned} \frac{\Delta f}{f_0} &= \frac{2(f_0 - f_m)}{f_0} = 2 - \frac{4\theta_m}{\pi} \\ &= 2 - \frac{4}{\pi} \cos^{-1} \left[\frac{1}{2} \left(\frac{\Gamma_m}{|A|} \right)^{1/N} \right]. \end{aligned} \quad (5.55)$$



EXAMPLE 5.6 BINOMIAL TRANSFORMER DESIGN

Design a three-section binomial transformer to match a $50 \, \Omega$ load to a $100 \, \Omega$ line and calculate the bandwidth for $\Gamma_m = 0.05$. Plot the reflection coefficient magnitude versus normalized frequency for the exact designs using 1, 2, 3, 4, and 5 sections.

Solution

For $N = 3$, $Z_L = 50 \, \Omega$, and $Z_0 = 100 \, \Omega$ we have, from (5.49) and (5.53),

$$A = 2^{-N} \frac{Z_L - Z_0}{Z_L + Z_0} \simeq \frac{1}{2^{N+1}} \ln \frac{Z_L}{Z_0} = -0.0433.$$

From (5.55) the bandwidth is

$$\begin{aligned} \frac{\Delta f}{f_0} &= 2 - \frac{4}{\pi} \cos^{-1} \left[\frac{1}{2} \left(\frac{\Gamma_m}{|A|} \right)^{1/N} \right] \\ &= 2 - \frac{4}{\pi} \cos^{-1} \left[\frac{1}{2} \left(\frac{0.05}{0.0433} \right)^{1/3} \right] = 0.70, \text{ or } 70\%. \end{aligned}$$

The necessary binomial coefficients are

$$C_0^3 = \frac{3!}{3!0!} = 1,$$

$$C_1^3 = \frac{3!}{2!1!} = 3,$$

$$C_2^3 = \frac{3!}{1!2!} = 3.$$

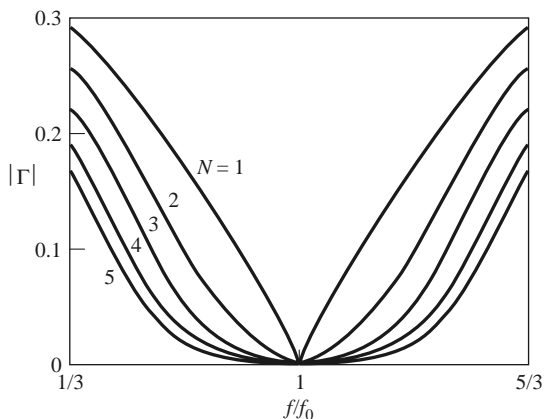


FIGURE 5.15 Reflection coefficient magnitude versus frequency for multisection binomial matching transformers of Example 5.6. $Z_L = 50 \Omega$ and $Z_0 = 100 \Omega$.

Using (5.53) gives the required characteristic impedances as

$$\begin{aligned}
 n = 0: \quad \ln Z_1 &= \ln Z_0 + 2^{-N} C_0^3 \ln \frac{Z_L}{Z_0} \\
 &= \ln 100 + 2^{-3}(1) \ln \frac{50}{100} = 4.518, \\
 Z_1 &= 91.7 \Omega; \\
 n = 1: \quad \ln Z_2 &= \ln Z_1 + 2^{-N} C_1^3 \ln \frac{Z_L}{Z_0} \\
 &= \ln 91.7 + 2^{-3}(3) \ln \frac{50}{100} = 4.26, \\
 Z_2 &= 70.7 \Omega; \\
 n = 2: \quad \ln Z_3 &= \ln Z_2 + 2^{-N} C_2^3 \ln \frac{Z_L}{Z_0} \\
 &= \ln 70.7 + 2^{-3}(3) \ln \frac{50}{100} = 4.00, \\
 Z_3 &= 54.5 \Omega.
 \end{aligned}$$

To use the data in Table 5.1 we reverse the source and load impedances and consider the problem of matching a 100Ω load to a 50Ω line. Then $Z_L/Z_0 = 2.0$, and we obtain the exact characteristic impedances as $Z_1 = 91.7 \Omega$, $Z_2 = 70.7 \Omega$, and $Z_3 = 54.5 \Omega$, which agree with the approximate results to three significant digits. Figure 5.15 shows the reflection coefficient magnitude versus frequency for exact designs using $N = 1, 2, 3, 4$, and 5 sections. Observe that greater bandwidth is obtained for transformers using more sections. ■

5.7

CHEBYSHEV MULTISECTION MATCHING TRANSFORMERS

In contrast with the binomial transformer, the multisection *Chebyshev matching transformer* optimizes bandwidth at the expense of passband ripple. Compromising on the flatness of the passband response leads to a bandwidth that is substantially better than that of the binomial transformer for a given number of sections. The Chebyshev transformer is

designed by equating $\Gamma(\theta)$ to a Chebyshev polynomial, which has the optimum characteristics needed for this type of transformer. We will first discuss the properties of Chebyshev polynomials and then derive a design procedure for Chebyshev matching transformers using the small-reflection theory of Section 5.5.

Chebyshev Polynomials

The n th-order Chebyshev polynomial is a polynomial of degree n , denoted by $T_n(x)$. The first four Chebyshev polynomials are

$$T_1(x) = x, \quad (5.56a)$$

$$T_2(x) = 2x^2 - 1, \quad (5.56b)$$

$$T_3(x) = 4x^3 - 3x, \quad (5.56c)$$

$$T_4(x) = 8x^4 - 8x^2 + 1. \quad (5.56d)$$

Higher order polynomials can be found using the following recurrence formula:

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x). \quad (5.57)$$

The first four Chebyshev polynomials are plotted in Figure 5.16, from which the following very useful properties of Chebyshev polynomials can be noted:

- For $-1 \leq x \leq 1$, $|T_n(x)| \leq 1$. In this range the Chebyshev polynomials oscillate between ± 1 . This is the *equal-ripple* property, and this region will be mapped to the passband of the matching transformer.
- For $|x| > 1$, $|T_n(x)| > 1$. This region will map to the frequency range outside the passband.
- For $|x| > 1$, the $|T_n(x)|$ increases faster with x as n increases.

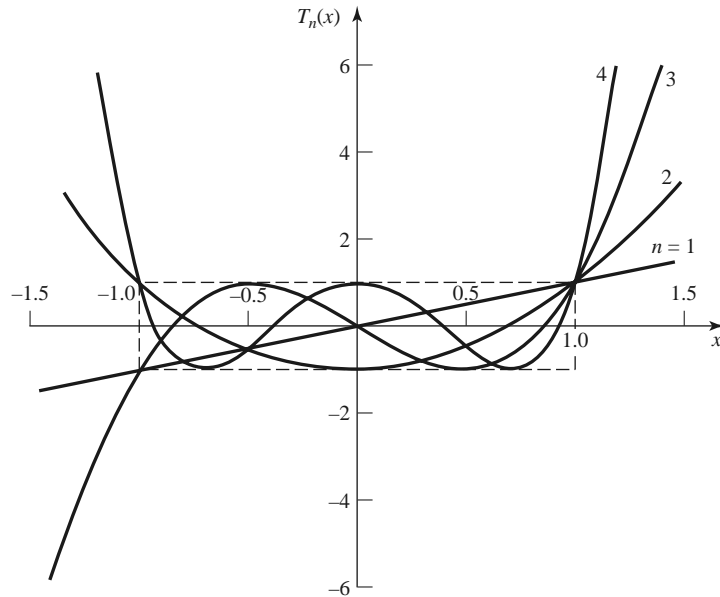


FIGURE 5.16 The first four Chebyshev polynomials, $T_n(x)$.

Now let $x = \cos \theta$ for $|x| < 1$. Then it can be shown that the Chebyshev polynomials can be expressed as

$$T_n(\cos \theta) = \cos n\theta,$$

or more generally as

$$T_n(x) = \cos(n \cos^{-1} x) \quad \text{for } |x| < 1, \quad (5.58a)$$

$$T_n(x) = \cosh(n \cosh^{-1} x) \quad \text{for } x > 1. \quad (5.58b)$$

We desire equal ripple for the passband response of the transformer, so it is necessary to map θ_m to $x = 1$ and $\pi - \theta_m$ to $x = -1$, where θ_m and $\pi - \theta_m$ are the lower and upper edges of the passband, respectively, as shown in Figure 5.11. This can be accomplished by replacing $\cos \theta$ in (5.58a) with $\cos \theta / \cos \theta_m$:

$$T_n\left(\frac{\cos \theta}{\cos \theta_m}\right) = T_n(\sec \theta_m \cos \theta) = \cos n \left[\cos^{-1} \left(\frac{\cos \theta}{\cos \theta_m} \right) \right]. \quad (5.59)$$

Then $|\sec \theta_m \cos \theta| \leq 1$ for $\theta_m < \theta < \pi - \theta_m$, so $|T_n(\sec \theta_m \cos \theta)| \leq 1$ over this same range.

Because $\cos^n \theta$ can be expanded into a sum of terms of the form $\cos(n - 2m)\theta$, the Chebyshev polynomials of (5.56) can be rewritten in the following useful form:

$$T_1(\sec \theta_m \cos \theta) = \sec \theta_m \cos \theta, \quad (5.60a)$$

$$T_2(\sec \theta_m \cos \theta) = \sec^2 \theta_m (1 + \cos 2\theta) - 1, \quad (5.60b)$$

$$T_3(\sec \theta_m \cos \theta) = \sec^3 \theta_m (\cos 3\theta + 3 \cos \theta) - 3 \sec \theta_m \cos \theta, \quad (5.60c)$$

$$T_4(\sec \theta_m \cos \theta) = \sec^4 \theta_m (\cos 4\theta + 4 \cos 2\theta + 3) - 4 \sec^2 \theta_m (\cos 2\theta + 1) + 1. \quad (5.60d)$$

These results can be used to design matching transformers with up to four sections, and will also be used in later chapters for the design of directional couplers and filters.

Design of Chebyshev Transformers

We can now synthesize a Chebyshev equal-ripple passband by making $\Gamma(\theta)$ proportional to $T_N(\sec \theta_m \cos \theta)$, where N is the number of sections in the transformer. Thus, using (5.46), we have

$$\begin{aligned} \Gamma(\theta) &= 2e^{-jN\theta} [\Gamma_0 \cos N\theta + \Gamma_1 \cos(N-2)\theta + \cdots + \Gamma_n \cos(N-2n)\theta + \cdots] \\ &= Ae^{-jN\theta} T_N(\sec \theta_m \cos \theta), \end{aligned} \quad (5.61)$$

where the last term in the series of (5.61) is $(1/2)\Gamma_{N/2}$ for N even and $\Gamma_{(N-1)/2} \cos \theta$ for N odd. As in the binomial transformer case, we can find the constant A by letting $\theta = 0$, corresponding to zero frequency. Thus,

$$\Gamma(0) = \frac{Z_L - Z_0}{Z_L + Z_0} = AT_N(\sec \theta_m),$$

so we have

$$A = \frac{Z_L - Z_0}{Z_L + Z_0} \frac{1}{T_N(\sec \theta_m)}. \quad (5.62)$$

If the maximum allowable reflection coefficient magnitude in the passband is Γ_m , then from (5.61) $\Gamma_m = |A|$ since the maximum value of $T_n(\sec \theta_m \cos \theta)$ in the passband is unity.

TABLE 5.2 Chebyshev Transformer Design

Z_L/Z_0	$N = 2$				$N = 3$					
	$\Gamma_m = 0.05$		$\Gamma_m = 0.20$		$\Gamma_m = 0.05$			$\Gamma_m = 0.20$		
	Z_1/Z_0	Z_2/Z_0	Z_1/Z_0	Z_2/Z_0	Z_1/Z_0	Z_2/Z_0	Z_3/Z_0	Z_1/Z_0	Z_2/Z_0	Z_3/Z_0
1.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1.5	1.1347	1.3219	1.2247	1.2247	1.1029	1.2247	1.3601	1.2247	1.2247	1.2247
2.0	1.2193	1.6402	1.3161	1.5197	1.1475	1.4142	1.7429	1.2855	1.4142	1.5558
3.0	1.3494	2.2232	1.4565	2.0598	1.2171	1.7321	2.4649	1.3743	1.7321	2.1829
4.0	1.4500	2.7585	1.5651	2.5558	1.2662	2.0000	3.1591	1.4333	2.0000	2.7908
6.0	1.6047	3.7389	1.7321	3.4641	1.3383	2.4495	4.4833	1.5193	2.4495	3.9492
8.0	1.7244	4.6393	1.8612	4.2983	1.3944	2.8284	5.7372	1.5766	2.8284	5.0742
10.0	1.8233	5.4845	1.9680	5.0813	1.4385	3.1623	6.9517	1.6415	3.1623	6.0920

$N = 4$									
Z_L/Z_0	$\Gamma_m = 0.05$				$\Gamma_m = 0.20$				
	Z_1/Z_0	Z_2/Z_0	Z_3/Z_0	Z_4/Z_0	Z_1/Z_0	Z_2/Z_0	Z_3/Z_0	Z_4/Z_0	
1.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
1.5	1.0892	1.1742	1.2775	1.3772	1.2247	1.2247	1.2247	1.2247	
2.0	1.1201	1.2979	1.5409	1.7855	1.2727	1.3634	1.4669	1.5715	
3.0	1.1586	1.4876	2.0167	2.5893	1.4879	1.5819	1.8965	2.0163	
4.0	1.1906	1.6414	2.4369	3.3597	1.3692	1.7490	2.2870	2.9214	
6.0	1.2290	1.8773	3.1961	4.8820	1.4415	2.0231	2.9657	4.1623	
8.0	1.2583	2.0657	3.8728	6.3578	1.4914	2.2428	3.5670	5.3641	
10.0	1.2832	2.2268	4.4907	7.7930	1.5163	2.4210	4.1305	6.5950	

Equating similar terms in $\cos n\theta$ gives the following results:

$$\begin{aligned}\cos 3\theta: \quad 2\Gamma_0 &= A \sec^3 \theta_m, \\ \Gamma_0 &= 0.0698; \\ \cos \theta: \quad 2\Gamma_1 &= 3A(\sec^3 \theta_m - \sec \theta_m), \\ \Gamma_1 &= 0.1037.\end{aligned}$$

From symmetry we also have that

$$\begin{aligned}\Gamma_3 &= \Gamma_0 = 0.0698, \\ \Gamma_2 &= \Gamma_1 = 0.1037.\end{aligned}$$

Then the characteristic impedances are:

$$\begin{aligned}n = 0: \quad \ln Z_1 &= \ln Z_0 + 2\Gamma_0 \\ &= \ln 50 + 2(0.0698) = 4.051 \\ Z_1 &= 57.5 \, \Omega; \\ n = 1: \quad \ln Z_2 &= \ln Z_1 + 2\Gamma_1 \\ &= \ln 57.5 + 2(0.1037) = 4.259 \\ Z_2 &= 70.7 \, \Omega; \\ n = 2: \quad \ln Z_3 &= \ln Z_2 + 2\Gamma_2 \\ &= \ln 70.7 + 2(0.1037) = 4.466 \\ Z_3 &= 87.0 \, \Omega.\end{aligned}$$

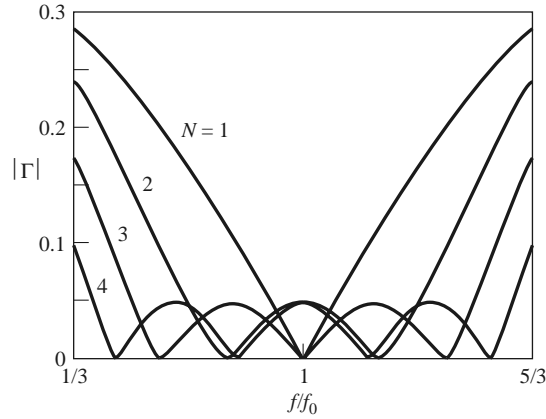


FIGURE 5.17 Reflection coefficient magnitude versus frequency for the multisecton matching transformers of Example 5.7.

These values can be compared to the exact values from Table 5.2 of $Z_1 = 57.37 \, \Omega$, $Z_2 = 70.71 \, \Omega$, and $Z_3 = 87.15 \, \Omega$. The bandwidth, from (5.64), is

$$\frac{\Delta f}{f_0} = 2 - \frac{4\theta_m}{\pi} = 2 - 4 \left(\frac{44.7^\circ}{180^\circ} \right) = 1.01,$$

or 101%. This is significantly greater than the bandwidth of the binomial transformer of Example 5.6 (70%), which involved the same impedance mismatch. The trade-off, of course, is a nonzero ripple in the passband of the Chebyshev transformer.

Figure 5.17 shows reflection coefficient magnitudes versus frequency for the exact designs from Table 5.2 for $N = 1, 2, 3$, and 4 sections. ■

5.8 TAPERED LINES

In the preceding sections we discussed how an arbitrary real load impedance could be matched to a line over a desired bandwidth by using multisecton matching transformers. As the number N of discrete transformer sections increases, the step changes in characteristic impedance between the sections become smaller, and the transformer geometry approaches a continuously tapered line. In practice, of course, a matching transformer must be of finite length—often no more than a few sections long. This suggests that, instead of discrete sections, the transformer can be continuously tapered, as shown in Figure 5.18a. Different passband characteristics can be obtained by using different types of taper.

In this section we will derive an approximate theory, again based on the theory of small reflections, to predict the reflection coefficient response as a function of the impedance taper versus position, $Z(z)$. We will apply these results to a few common types of impedance tapers.

Consider the continuously tapered line of Figure 5.18a as being made up of a number of incremental sections of length Δz , with an impedance change $\Delta Z(z)$ from one

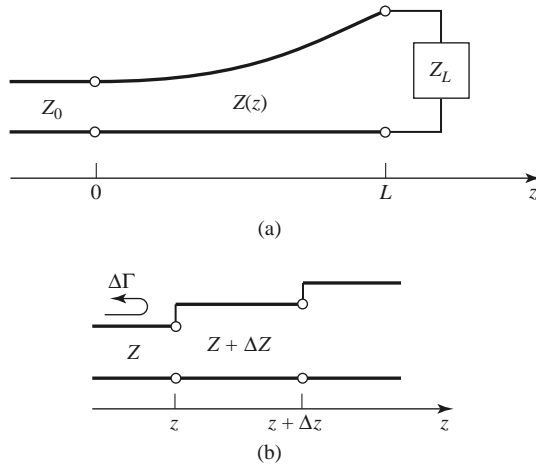


FIGURE 5.18 A tapered transmission line matching section and the model for an incremental length of tapered line. (a) The tapered transmission line matching section. (b) Model for an incremental step change in impedance of the tapered line.

section to the next, as shown in Figure 5.18b. The incremental reflection coefficient from the impedance step at z is given by

$$\Delta\Gamma = \frac{(Z + \Delta Z) - Z}{(Z + \Delta Z) + Z} \simeq \frac{\Delta Z}{2Z}. \quad (5.65)$$

In the limit as $\Delta z \rightarrow 0$ we have an exact differential:

$$d\Gamma = \frac{dZ}{2Z} = \frac{1}{2} \frac{d(\ln Z/Z_0)}{dz} dz, \quad (5.66)$$

since

$$\frac{d(\ln f(z))}{dz} = \frac{1}{f} \frac{df(z)}{dz}.$$

By using the theory of small reflections, we can find the total reflection coefficient at $z = 0$ by summing all the partial reflections with their appropriate phase shifts:

$$\Gamma(\theta) = \frac{1}{2} \int_{z=0}^L e^{-2j\beta z} \frac{d}{dz} \ln \left(\frac{Z}{Z_0} \right) dz, \quad (5.67)$$

where $\theta = 2\beta\ell$. If $Z(z)$ is known, $\Gamma(\theta)$ can be found as a function of frequency. Alternatively, if $\Gamma(\theta)$ is specified, then in principle $Z(z)$ can be found by inversion. This latter procedure is difficult, and is generally avoided in practice; the reader is referred to references [1] and [4] for further discussion of this topic. Here we will consider three special cases of $Z(z)$ impedance tapers, and evaluate the resulting responses.

Exponential Taper

Consider first an *exponential taper*, where

$$Z(z) = Z_0 e^{az} \quad \text{for } 0 < z < L, \quad (5.68)$$

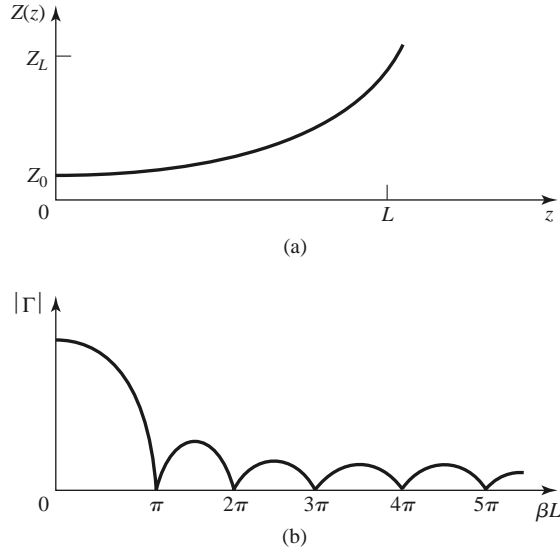


FIGURE 5.19 A matching section with an exponential impedance taper. (a) Variation of impedance. (b) Resulting reflection coefficient magnitude response.

as indicated in Figure 5.19a. At $z = 0$, $Z(0) = Z_0$, as desired. At $z = L$ we wish to have $Z(L) = Z_L = Z_0 e^{aL}$, which determines the constant a as

$$a = \frac{1}{L} \ln \left(\frac{Z_L}{Z_0} \right). \quad (5.69)$$

We find $\Gamma(\theta)$ by using (5.68) and (5.69) in (5.67):

$$\begin{aligned} \Gamma &= \frac{1}{2} \int_0^L e^{-2j\beta z} \frac{d}{dz} (\ln e^{az}) dz \\ &= \frac{\ln Z_L/Z_0}{2L} \int_0^L e^{-2j\beta z} dz \\ &= \frac{\ln Z_L/Z_0}{2} e^{-j\beta L} \frac{\sin \beta L}{\beta L}. \end{aligned} \quad (5.70)$$

Observe that this derivation assumes that β , the propagation constant of the tapered line, is not a function of z —an assumption generally valid only for TEM lines.

The magnitude of the reflection coefficient in (5.70) is sketched in Figure 5.19b; note that the peaks in $|\Gamma|$ decrease with increasing length, as one might expect, and that the length should be greater than $\lambda/2$ ($\beta L > \pi$) to minimize the mismatch at low frequencies.

Triangular Taper

Next consider a *triangular taper* for $d \ln (Z/Z_0) / dz$, that is,

$$Z(z) = \begin{cases} Z_0 e^{2(z/L)^2 \ln Z_L/Z_0} & \text{for } 0 \leq z \leq L/2 \\ Z_0 e^{(4z/L - 2z^2/L^2 - 1) \ln Z_L/Z_0} & \text{for } L/2 \leq z \leq L, \end{cases} \quad (5.71)$$

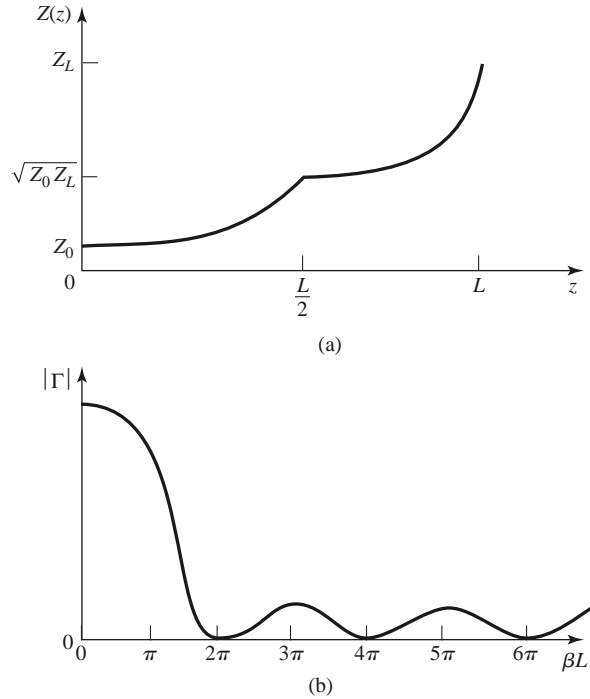


FIGURE 5.20 A matching section with a triangular taper for $d(\ln Z/Z_0)/dz$. (a) Variation of impedance. (b) Resulting reflection coefficient magnitude response.

so that the derivative is triangular in form:

$$\frac{d(\ln Z/Z_0)}{dz} = \begin{cases} 4z/L^2 \ln Z_L/Z_0 & \text{for } 0 \leq z \leq L/2 \\ (4/L - 4z/L^2) \ln Z_L/Z_0 & \text{for } L/2 \leq z \leq L. \end{cases} \quad (5.72)$$

$Z(z)$ is plotted in Figure 5.20a. Evaluating Γ from (5.67) gives

$$\Gamma(\theta) = \frac{1}{2} e^{-j\beta L} \ln \left(\frac{Z_L}{Z_0} \right) \left[\frac{\sin(\beta L/2)}{\beta L/2} \right]^2. \quad (5.73)$$

The magnitude of this result is sketched in Figure 5.20b. Note that, for $\beta L > 2\pi$, the peaks of the triangular taper are lower than the corresponding peaks of the exponential case. However, the first null for the triangular taper occurs at $\beta L = 2\pi$, whereas for the exponential taper it occurs at $\beta L = \pi$.

Klopfenstein Taper

Considering the fact that there is an infinite number of possibilities for choosing an impedance matching taper, it is logical to ask if there is a design that is “best.” For a given taper length (greater than some critical value), the *Klopfenstein impedance taper* [4, 5] has been shown to be optimum in the sense that the reflection coefficient is minimum over the passband. Alternatively, for a maximum reflection coefficient specification in the passband, the Klopfenstein taper yields the shortest matching section.

The Klopfenstein taper is derived from a stepped Chebyshev transformer as the number of sections increases to infinity, and is analogous to the Taylor distribution of antenna array theory. We will not present the details of this derivation, which can be found in

references [1] and [4]; only the necessary results for the design of Klopfenstein tapers are given in what follows.

The logarithm of the characteristic impedance variation for the Klopfenstein taper is given by

$$\ln Z(z) = \frac{1}{2} \ln Z_0 Z_L + \frac{\Gamma_0}{\cosh A} A^2 \phi(2z/L - 1, A) \quad \text{for } 0 \leq z \leq L, \quad (5.74)$$

where the function $\phi(x, A)$ is defined as

$$\phi(x, A) = -\phi(-x, A) = \int_0^x \frac{I_1(A\sqrt{1-y^2})}{A\sqrt{1-y^2}} dy \quad \text{for } |x| \leq 1, \quad (5.75)$$

where $I_1(x)$ is the modified Bessel function. The function of (5.75) has the following special values:

$$\begin{aligned} \phi(0, A) &= 0 \\ \phi(x, 0) &= \frac{x}{2} \\ \phi(1, A) &= \frac{\cosh A - 1}{A^2}, \end{aligned}$$

but otherwise (5.75) must be calculated numerically. A simple and efficient method for doing this is available [6].

The resulting reflection coefficient is given by

$$\Gamma(\theta) = \Gamma_0 e^{-j\beta L} \frac{\cos \sqrt{(\beta L)^2 - A^2}}{\cosh A} \quad \text{for } \beta L > A. \quad (5.76)$$

If $\beta L < A$, the $\cos \sqrt{(\beta L)^2 - A^2}$ term becomes $\cosh \sqrt{A^2 - (\beta L)^2}$.

In (5.74) and (5.76), Γ_0 is the reflection coefficient at zero frequency, given as

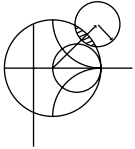
$$\Gamma_0 = \frac{Z_L - Z_0}{Z_L + Z_0} \simeq \frac{1}{2} \ln \left(\frac{Z_L}{Z_0} \right). \quad (5.77)$$

The passband is defined as $\beta L \geq A$, and so the maximum ripple in the passband is

$$\Gamma_m = \frac{\Gamma_0}{\cosh A} \quad (5.78)$$

because $\Gamma(\theta)$ oscillates between $\pm \Gamma_0 / \cosh A$ for $\beta L > A$.

It is interesting to note that the impedance taper of (5.74) has steps at $z = 0$ and L (the ends of the tapered section) and so does not smoothly join the source and load impedances. A typical Klopfenstein impedance taper and its response are given in the following example.



EXAMPLE 5.8 DESIGN OF TAPERED MATCHING SECTIONS

Design a triangular taper, an exponential taper, and a Klopfenstein taper (with $\Gamma_m = 0.02$) to match a 50Ω load to a 100Ω line. Plot the impedance variations and resulting reflection coefficient magnitudes versus βL .

Solution

Triangular taper: From (5.71) the impedance variation is

$$Z(z) = Z_0 \begin{cases} e^{2(z/L)^2 \ln Z_L/Z_0} & \text{for } 0 \leq z \leq L/2 \\ e^{(4z/L - 2z^2/L^2 - 1) \ln Z_L/Z_0} & \text{for } L/2 \leq z \leq L, \end{cases}$$

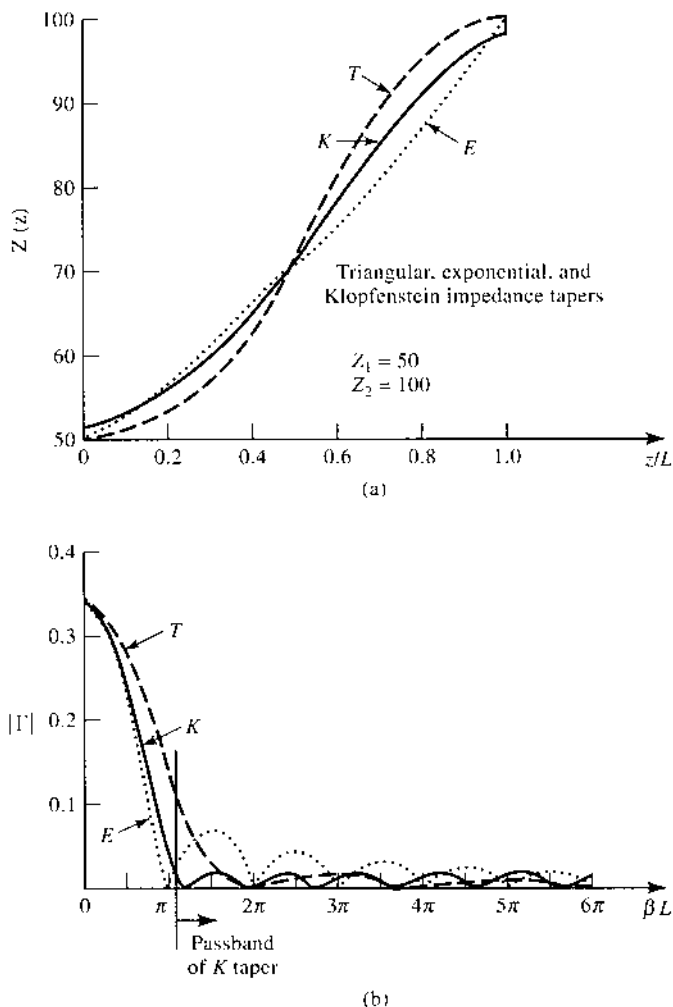


FIGURE 5.21 Solution to Example 5.8. (a) Impedance variations for the triangular, exponential, and Klopfenstein tapers. (b) Resulting reflection coefficient magnitude versus frequency for the tapers of (a).

with $Z_0 = 100 \Omega$ and $Z_L = 50 \Omega$. The resulting reflection coefficient response is given by (5.73):

$$|\Gamma(\theta)| = \frac{1}{2} \ln \left(\frac{Z_L}{Z_0} \right) \left[\frac{\sin(\beta L/2)}{\beta L/2} \right]^2.$$

Exponential taper: From (5.68) the impedance variation is

$$Z(z) = Z_0 e^{az} \quad \text{for } 0 < z < L,$$

with $a = (1/L) \ln Z_L/Z_0 = 0.693/L$. The reflection coefficient response is, from (5.70),

$$|\Gamma(\theta)| = \frac{1}{2} \ln \left(\frac{Z_L}{Z_0} \right) \frac{\sin \beta L}{\beta L}.$$

Klopfenstein taper: Using (5.77) gives Γ_0 as

$$\Gamma_0 = \frac{1}{2} \ln \left(\frac{Z_L}{Z_0} \right) = 0.346,$$

and (5.78) gives A as

$$A = \cosh^{-1} \left(\frac{\Gamma_0}{\Gamma_m} \right) = \cosh^{-1} \left(\frac{0.346}{0.02} \right) = 3.543.$$

The impedance taper must be numerically evaluated from (5.74). The reflection coefficient magnitude is given by (5.76):

$$|\Gamma(\theta)| = \Gamma_0 \frac{\cos \sqrt{(\beta L)^2 - A^2}}{\cosh A}.$$

The passband for the Klopfenstein taper is defined as $\beta L > A = 3.543 = 1.13\pi$.

Figure 5.21 shows the impedance variations (vs. z/L), and the resulting reflection coefficient magnitude (vs. βL) for the three types of tapers. The Klopfenstein taper gives the desired response of $|\Gamma| \leq \Gamma_m = 0.02$ for $\beta L \geq 1.13\pi$, which is smaller than the corresponding lengths of either the triangular or the exponential taper transformer. Also note that, like the stepped-Chebyshev matching transformer, the response of the Klopfenstein taper has equal-ripple lobes versus frequency in its passband. ■

5.9

THE BODE–FANO CRITERION

In this chapter we discussed several techniques for matching an arbitrary load at a single frequency, using lumped elements, tuning stubs, and single-section quarter-wave transformers. We presented multisection matching transformers and tapered lines as a means of obtaining broader bandwidths with various passband characteristics. We close our study of impedance matching with a somewhat qualitative discussion of the theoretical limits that constrain the performance of an impedance matching network.

We limit our discussion to the circuit of Figure 5.1, where a lossless network is used to match an arbitrary complex load, generally over a nonzero bandwidth. From a very general perspective, we might raise the following questions in regard to this problem:

- Can we achieve a perfect match (zero reflection) over a specified bandwidth?
- If not, how well can we do? What is the trade-off between Γ_m , the maximum allowable reflection in the passband, and the bandwidth?
- How complex must the matching network be for a given specification?

These questions can be answered by the Bode–Fano criterion [7, 8] which gives, for certain canonical types of load impedances, a theoretical limit on the minimum reflection coefficient magnitude that can be obtained with an arbitrary matching network. The Bode–Fano criterion thus represents an optimum result that can be ideally achieved, even though such a result may only be approximated in practice. Such optimal results are always important, however, because they specify an upper limit of performance, and so provide a benchmark against which a practical design can be compared.

Figure 5.22a shows a lossless network used to match a parallel RC load impedance. The Bode–Fano criterion states that

$$\int_0^\infty \ln \frac{1}{|\Gamma(\omega)|} d\omega \leq \frac{\pi}{RC}, \quad (5.79)$$

where $\Gamma(\omega)$ is the reflection coefficient seen looking into the arbitrary lossless matching network. The derivation of this result is beyond the scope of this text (the interested reader is referred to references [7] and [8]); our goal here is to discuss the implications of this result.

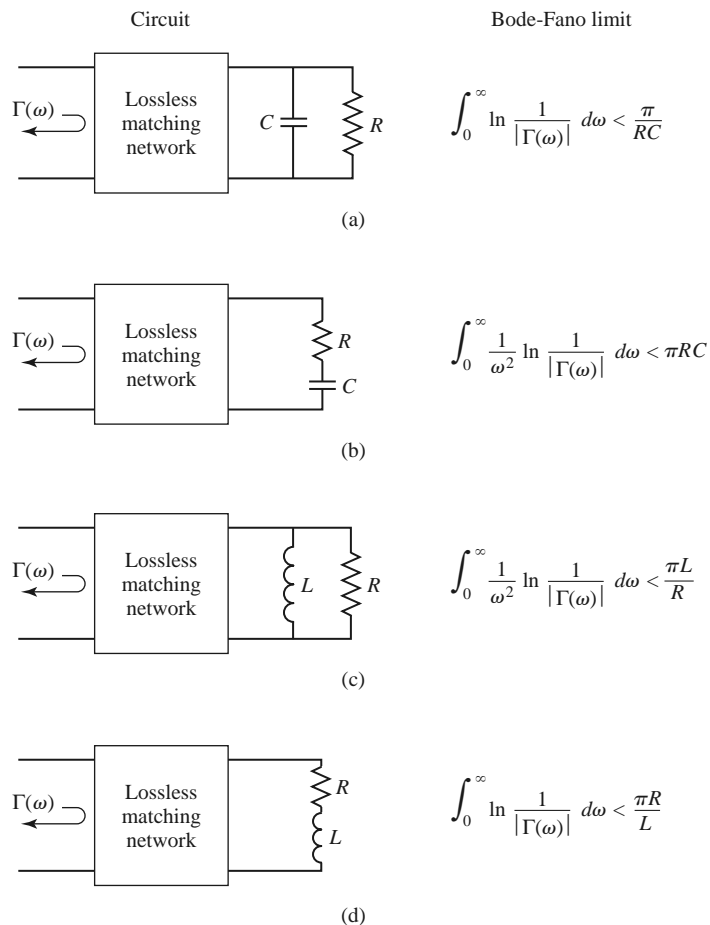


FIGURE 5.22 The Bode–Fano limits for RC and RL loads matched with passive and lossless networks (ω_0 is the center frequency of the matching bandwidth). (a) Parallel RC . (b) Series RC . (c) Parallel RL . (d) Series RL .

Assume that we desire to synthesize a matching network with a reflection coefficient response like that shown in Figure 5.23a. Applying (5.79) to this function gives

$$\int_0^\infty \ln \frac{1}{|\Gamma|} d\omega = \int_{\Delta\omega} \ln \frac{1}{\Gamma_m} d\omega = \Delta\omega \ln \frac{1}{\Gamma_m} \leq \frac{\pi}{RC}, \quad (5.80)$$

which leads to the following conclusions:

- For a given load (a fixed RC product), a broader bandwidth ($\Delta\omega$) can be achieved only at the expense of a higher reflection coefficient in the passband (Γ_m).
- The passband reflection coefficient, Γ_m , cannot be zero unless $\Delta\omega = 0$. Thus a perfect match can be achieved only at a finite number of discrete frequencies, as illustrated in Figure 5.23b.
- As R and/or C increases, the quality of the match ($\Delta\omega$ and/or $1/\Gamma_m$) must decrease. Thus, higher- Q circuits are intrinsically harder to match than are lower- Q circuits (we will discuss Q in Chapter 6).

Because $\ln(1/|\Gamma|)$ is proportional to the return loss (in dB) at the input of the matching network, (5.79) can be interpreted as requiring that the area between the return loss curve

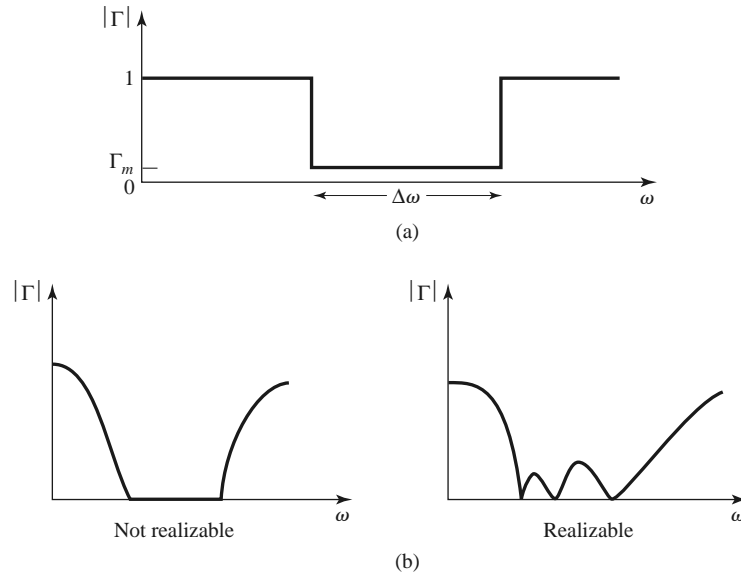


FIGURE 5.23 Illustrating the Bode-Fano criterion. (a) A possible reflection coefficient response. (b) Nonrealizable and realizable reflection coefficient responses.

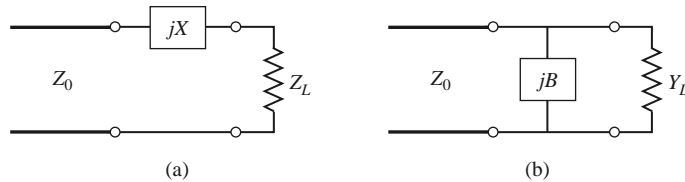
and the $|\Gamma| = 1$ ($RL = 0$ dB) axis must be less than or equal to a particular constant. Optimization then implies that the return loss curve be adjusted so that $|\Gamma| = \Gamma_m$ over the passband and $|\Gamma| = 1$ elsewhere, as in Figure 5.23a. In this way, no area under the return loss curve is wasted outside the passband, or lost in regions within the passband for which $|\Gamma| < \Gamma_m$. The square-shaped response of Figure 5.23a is therefore the optimum response, but cannot be realized in practice because it would require an infinite number of elements in the matching network. It can be approximated, however, with a reasonably small number of elements, as described in reference [8]. Finally, note that the Chebyshev matching transformer can be considered as a close approximation to the ideal passband of Figure 5.23a when the ripple of the Chebyshev response is made equal to Γ_m . Figure 5.22 lists the Bode-Fano limits for other types of RC and RL loads.

REFERENCES

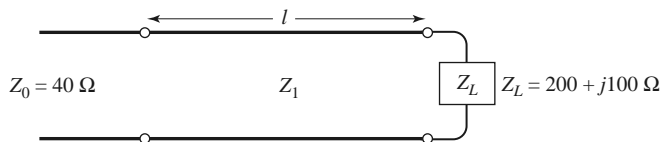
- [1] R. E. Collin, *Foundations for Microwave Engineering*, 2nd edition, McGraw-Hill, New York, 1992.
- [2] G. L. Matthaei, L. Young, and E. M. T. Jones, *Microwave Filters, Impedance-Matching Networks, and Coupling Structures*, Artech House Books, Dedham, Mass. 1980.
- [3] P. Bhartia and I. J. Bahl, *Millimeter Wave Engineering and Applications*, Wiley Interscience, New York, 1984.
- [4] R. E. Collin, "The Optimum Tapered Transmission Line Matching Section," *Proceedings of the IRE*, vol. 44, pp. 539–548, April 1956.
- [5] R. W. Klopfenstein, "A Transmission Line Taper of Improved Design," *Proceedings of the IRE*, vol. 44, pp. 31–15, January 1956.
- [6] M. A. Grossberg, "Extremely Rapid Computation of the Klopfenstein Impedance Taper," *Proceedings of the IEEE*, vol. 56, pp. 1629–1630, September 1968.
- [7] H. W. Bode, *Network Analysis and Feedback Amplifier Design*, Van Nostrand, New York, 1945.
- [8] R. M. Fano, "Theoretical Limitations on the Broad-Band Matching of Arbitrary Impedances," *Journal of the Franklin Institute*, vol. 249, pp. 57–83, January 1950, and pp. 139–154, February 1950.

PROBLEMS

- 5.1 Design two lossless L -section matching circuits to match each of the following loads to a $100\ \Omega$ generator at 3 GHz. (a) $Z_L = 150 - j200\ \Omega$ and (b) $Z_L = 20 - j90\ \Omega$.
- 5.2 We have seen that the matching of an arbitrary load impedance requires a network with at least two degrees of freedom. Determine the types of load impedances/admittances that can be matched with the two single-element networks shown below.

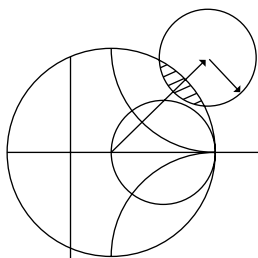


- 5.3 A load impedance $Z_L = 100 + j80\ \Omega$ is to be matched to a $75\ \Omega$ line using a single shunt-stub tuner. Find two designs using open-circuited stubs.
- 5.4 Repeat Problem 5.3 using short-circuited stubs.
- 5.5 A load impedance $Z_L = 90 + j60\ \Omega$ is to be matched to a $75\ \Omega$ line using a single series-stub tuner. Find two designs using open-circuited stubs.
- 5.6 Repeat Problem 5.5 using short-circuited stubs.
- 5.7 In the circuit shown below a load $Z_L = 200 + j100\ \Omega$ is to be matched to a $40\ \Omega$ line, using a length ℓ of lossless transmission line of characteristic impedance Z_1 . Find ℓ and Z_1 . Determine, in general, what type of load impedances can be matched using such a circuit.



- 5.8 An open-circuit tuning stub is to be made from a lossy transmission line with an attenuation constant α . What is the maximum value of normalized reactance that can be obtained with this stub? What is the maximum value of normalized reactance that can be obtained with a shorted stub of the same type of transmission line? Assume $\alpha\ell$ is small.
- 5.9 Design a double-stub tuner using open-circuited stubs with a $\lambda/8$ spacing to match a load admittance $Y_L = (0.4 + j1.2)Y_0$.
- 5.10 Repeat Problem 5.9 using a double-stub tuner with short-circuited stubs and a $3\lambda/8$ spacing.
- 5.11 Derive the design equations for a double-stub tuner using two series stubs spaced a distance d apart. Assume the load impedance is $Z_L = R_L + jX_L$.
- 5.12 Consider matching a load $Z_L = 200\ \Omega$ to a $100\ \Omega$ line, using single shunt-stub, single series stub, and double shunt-stub tuners, with short-circuited stubs. Which tuner will give the best bandwidth? Justify your answer by calculating the reflection coefficient for all six solutions at $1.1f_0$, where f_0 is the match frequency, or use CAD to plot the reflection coefficient versus frequency.
- 5.13 Design a single-section quarter-wave matching transformer to match a $350\ \Omega$ load to a $100\ \Omega$ line. What is the percent bandwidth of this transformer, for $\text{SWR} \leq 2$? If the design frequency is 4 GHz, sketch the layout of a microstrip circuit, including dimensions, to implement this matching transformer. Assume the substrate is 0.159 cm thick, with a relative permittivity of 2.2.
- 5.14 Consider the quarter-wave transformer of Figure 5.13 with $Z_1 = 100\ \Omega$, $Z_2 = 150\ \Omega$, and $Z_L = 225\ \Omega$. Evaluate the worst-case percent error in computing $|\Gamma|$ from the approximate expression (5.42), compared to the exact result.

- 5.15** A waveguide load with an equivalent TE_{10} wave impedance of $377\ \Omega$ must be matched to an air-filled X-band rectangular guide at 10 GHz. A quarter-wave matching transformer is to be used, and is to consist of a section of guide filled with dielectric. Find the required dielectric constant and physical length of the matching section. What restrictions on the load impedance apply to this technique?
- 5.16** A four-section binomial matching transformer is to be used to match a $12.5\ \Omega$ load to a $50\ \Omega$ line at a center frequency of 1 GHz. (a) Design the matching transformer, and compute the bandwidth for $\Gamma_m = 0.05$. Use CAD to plot the input reflection coefficient versus frequency. (b) Lay out the microstrip implementation of this circuit on an FR4 substrate having $\epsilon_r = 4.2$, $d = 0.158\text{ cm}$, and $\tan \delta = 0.02$, with copper conductors 0.5 mil thick. Use CAD to plot the insertion loss versus frequency.
- 5.17** Derive the exact characteristic impedance for a two-section binomial matching transformer for a normalized load impedance $Z_L/Z_0 = 1.5$. Check your results with Table 5.1.
- 5.18** Calculate and plot the percent bandwidth for an $N = 1$ -, 2 -, and 4 -section binomial matching transformer versus $Z_L/Z_0 = 1.5$ to 6 for $\Gamma_m = 0.2$.
- 5.19** Design a four-section Chebyshev matching transformer to match a $50\ \Omega$ line to a $30\ \Omega$ load. The maximum permissible SWR over the passband is 1.25. What is the resulting bandwidth? Use the approximate theory developed in the text, as opposed to the tables. Use CAD to plot the input SWR versus frequency.
- 5.20** Derive the exact characteristic impedances for a two-section Chebyshev matching transformer for a normalized load impedance $Z_L/Z_0 = 1.5$. Check your results with Table 5.2 for $\Gamma_m = 0.05$.
- 5.21** A load of $Z_L/Z_0 = 1.5$ is to be matched to a feed line using a multisection transformer, and it is desired to have a passband response with $|\Gamma(\theta)| = A(0.1 + \cos^2 \theta)$ for $0 \leq \theta \leq \pi$. Use the approximate theory for multisection transformers to design a two-section transformer.
- 5.22** A tapered matching section has $d \ln(Z/Z_0)/dz = A \sin \pi z/L$. Find the constant A so that $Z(0) = Z_0$ and $Z(L) = Z_L$. Compute Γ , and plot $|\Gamma|$ versus βL .
- 5.23** Design an exponentially tapered matching transformer to match a $100\ \Omega$ load to a $50\ \Omega$ line. Plot $|\Gamma|$ versus βL , and find the length of the matching section (at the center frequency) required to obtain $|\Gamma| \leq 0.05$ over a 100% bandwidth. How many sections would be required if a Chebyshev matching transformer were used to achieve the same specifications?
- 5.24** An ultra wideband (UWB) radio transmitter, operating from 3.1 to 10.6 GHz, drives a parallel RC load with $R = 75\ \Omega$ and $C = 0.6\text{ pF}$. What is the best return loss that can be obtained with an optimum matching network?
- 5.25** Consider a series RL load with $R = 80\ \Omega$ and $L = 5\text{ nH}$. Design a lumped-element L -section matching network to match this load to a $50\ \Omega$ line at 2 GHz. Plot $|\Gamma|$ versus frequency for this network to determine the bandwidth for which $|\Gamma| \leq \Gamma_m = 0.1$. Compare this with the maximum possible bandwidth for this load, as given by the Bode–Fano criterion. (Assume a square reflection coefficient response like that of Figure 5.23a.)



Microwave Resonators

Microwave resonators are used in a variety of applications, including filters, oscillators, frequency meters, and tuned amplifiers. Because the operation of microwave resonators is very similar to that of lumped-element resonators of circuit theory, we will begin by reviewing the basic characteristics of series and parallel *RLC* resonant circuits. We will then discuss various implementations of resonators at microwave frequencies using distributed elements such as transmission lines, rectangular and circular waveguides, and dielectric cavities. We will also discuss the excitation of resonators using apertures and current sheets.

6.1

SERIES AND PARALLEL RESONANT CIRCUITS

At frequencies near resonance, a microwave resonator can usually be modeled by either a series or parallel *RLC* lumped-element equivalent circuit, and so we will now review some of the basic properties of these circuits.

Series Resonant Circuit

A series *RLC* resonant circuit is shown in Figure 6.1a. The input impedance is

$$Z_{\text{in}} = R + j\omega L - j\frac{1}{\omega C}, \quad (6.1)$$

and the complex power delivered to the resonator is

$$\begin{aligned} P_{\text{in}} &= \frac{1}{2}VI^* = \frac{1}{2}Z_{\text{in}}|I|^2 = \frac{1}{2}Z_{\text{in}}\left|\frac{V}{Z_{\text{in}}}\right|^2 \\ &= \frac{1}{2}|I|^2\left(R + j\omega L - j\frac{1}{\omega C}\right). \end{aligned} \quad (6.2)$$

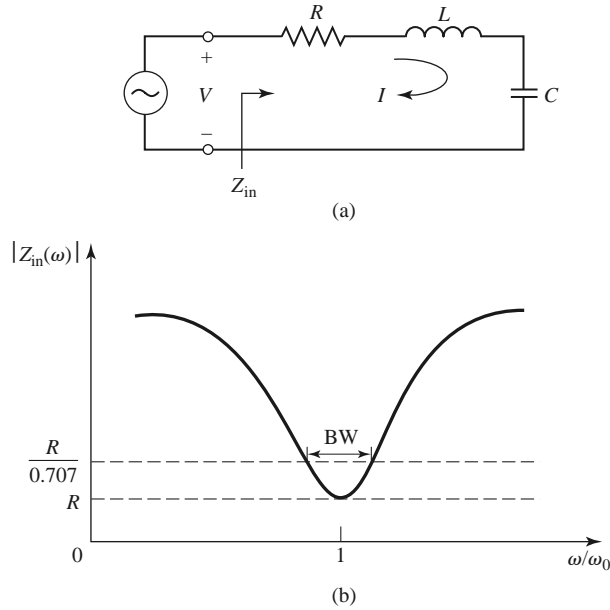


FIGURE 6.1 A series RLC resonator and its response. (a) A series RLC resonator circuit. (b) Input impedance magnitude versus frequency.

The power dissipated by the resistor R is

$$P_{\text{loss}} = \frac{1}{2} |I|^2 R, \quad (6.3a)$$

the average magnetic energy stored in the inductor L is

$$W_m = \frac{1}{4} |I|^2 L, \quad (6.3b)$$

and the average electric energy stored in the capacitor C is

$$W_e = \frac{1}{4} |V_c|^2 C = \frac{1}{4} |I|^2 \frac{1}{\omega^2 C}, \quad (6.3c)$$

where V_c is the voltage across the capacitor. Then the complex power of (6.2) can be rewritten as

$$P_{\text{in}} = P_{\text{loss}} + 2j\omega(W_m - W_e), \quad (6.4)$$

and the input impedance of (6.1) can be rewritten as

$$Z_{\text{in}} = \frac{2P_{\text{in}}}{|I|^2} = \frac{P_{\text{loss}} + 2j\omega(W_m - W_e)}{\frac{1}{2} |I|^2}. \quad (6.5)$$

Resonance occurs when the average stored magnetic and electric energies are equal, or $W_m = W_e$. Then from (6.5) and (6.3a), the input impedance at resonance is

$$Z_{\text{in}} = \frac{P_{\text{loss}}}{\frac{1}{2} |I|^2} = R,$$

which is purely real. From (6.3b,c), $W_m = W_e$ implies that the resonant frequency, ω_0 , can be defined as

$$\omega_0 = \frac{1}{\sqrt{LC}}. \quad (6.6)$$

Another important parameter of a resonant circuit is its Q , or *quality factor*, which is defined as

$$\begin{aligned} Q &= \omega \frac{\text{average energy stored}}{\text{energy loss/second}} \\ &= \omega \frac{W_m + W_e}{P_{\text{loss}}}. \end{aligned} \quad (6.7)$$

Thus Q is a measure of the loss of a resonant circuit—lower loss implies a higher Q . Resonator losses may be due to conductor loss, dielectric loss, or radiation loss, and are represented by the resistance, R , of the equivalent circuit. An external connecting network may introduce additional loss. Each of these loss mechanisms will have the effect of lowering the Q . The Q of the resonator itself, disregarding external loading effects, is called the *unloaded* Q , denoted as Q_0 .

For the series resonant circuit of Figure 6.1a, the unloaded Q can be evaluated from (6.7), using (6.3) and the fact that $W_m = W_e$ at resonance, to give

$$Q_0 = \omega_0 \frac{2W_m}{P_{\text{loss}}} = \frac{\omega_0 L}{R} = \frac{1}{\omega_0 RC}, \quad (6.8)$$

which shows that Q increases as R decreases.

Next, consider the behavior of the input impedance of this resonator near its resonant frequency [1]. Let $\omega = \omega_0 + \Delta\omega$, where $\Delta\omega$ is small. The input impedance can then be rewritten from (6.1) as

$$\begin{aligned} Z_{\text{in}} &= R + j\omega L \left(1 - \frac{1}{\omega^2 LC} \right) \\ &= R + j\omega L \left(\frac{\omega^2 - \omega_0^2}{\omega^2} \right), \end{aligned}$$

since $\omega_0^2 = 1/LC$. Now $\omega^2 - \omega_0^2 = (\omega - \omega_0)(\omega + \omega_0) = \Delta\omega(2\omega - \Delta\omega) \simeq 2\omega\Delta\omega$ for small $\Delta\omega$. Thus,

$$\begin{aligned} Z_{\text{in}} &\simeq R + j2L\Delta\omega \\ &\simeq R + j \frac{2RQ_0\Delta\omega}{\omega_0}. \end{aligned} \quad (6.9)$$

This form will be useful for identifying equivalent circuits with distributed element resonators.

Alternatively, a resonator with loss can be modeled as a lossless resonator whose resonant frequency, ω_0 , has been replaced with a complex effective resonant frequency:

$$\omega_0 \leftarrow \omega_0 \left(1 + \frac{j}{2Q_0} \right). \quad (6.10)$$

This can be seen by considering the input impedance of a series resonator with no loss, as given by (6.9) with $R = 0$:

$$Z_{\text{in}} = j2L(\omega - \omega_0).$$

Then substituting the complex frequency of (6.10) for ω_0 gives

$$\begin{aligned} Z_{\text{in}} &= j2L \left(\omega - \omega_0 - j \frac{\omega_0}{2Q_0} \right) \\ &= \frac{\omega_0 L}{Q_0} + j2L(\omega - \omega_0) = R + j2L\Delta\omega, \end{aligned}$$

which is identical to (6.9). This is a useful procedure because for most practical resonators the loss is very small, so the Q can be found using the perturbation method, beginning with the solution for the lossless case. Then the effect of loss can be added to the input impedance by replacing ω_0 with the complex resonant frequency given in (6.10).

Finally, consider the *half-power fractional bandwidth* of the resonator. Figure 6.1b shows the variation of the magnitude of the input impedance versus frequency. When the frequency is such that $|Z_{\text{in}}|^2 = 2R^2$, then by (6.2) the average (real) power delivered to the circuit is one-half that delivered at resonance. If BW is the fractional bandwidth, then $\Delta\omega/\omega_0 = \text{BW}/2$ at the upper band edge. Using (6.9) gives

$$|R + jRQ_0(\text{BW})|^2 = 2R^2,$$

or

$$\text{BW} = \frac{1}{Q_0}. \quad (6.11)$$

Parallel Resonant Circuit

The parallel RLC resonant circuit, shown in Figure 6.2a, is the dual of the series RLC circuit. The input impedance is

$$Z_{\text{in}} = \left(\frac{1}{R} + \frac{1}{j\omega L} + j\omega C \right)^{-1}, \quad (6.12)$$

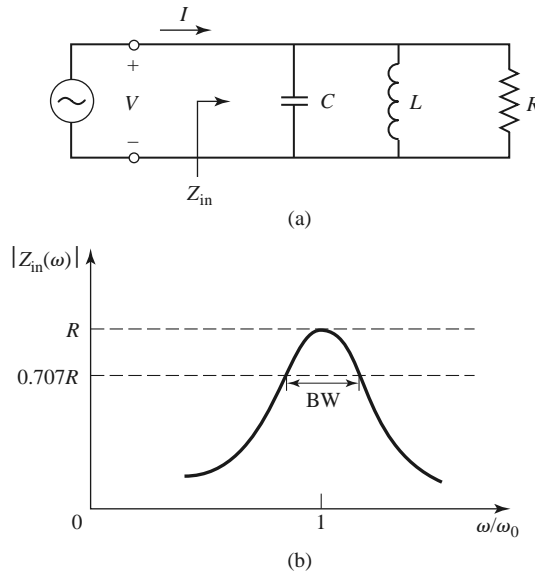


FIGURE 6.2 A parallel RLC resonator and its response. (a) A parallel RLC circuit. (b) Input impedance magnitude versus frequency.

and the complex power delivered to the resonator is

$$\begin{aligned} P_{\text{in}} &= \frac{1}{2} VI^* = \frac{1}{2} Z_{\text{in}} |I|^2 = \frac{1}{2} |V|^2 \frac{1}{Z_{\text{in}}^*} \\ &= \frac{1}{2} |V|^2 \left(\frac{1}{R} + \frac{j}{\omega L} - j\omega C \right). \end{aligned} \quad (6.13)$$

The power dissipated by the resistor, R , is

$$P_{\text{loss}} = \frac{1}{2} \frac{|V|^2}{R}, \quad (6.14a)$$

the average electric energy stored in the capacitor, C , is

$$W_e = \frac{1}{4} |V|^2 C, \quad (6.14b)$$

and the average magnetic energy stored in the inductor, L , is

$$W_m = \frac{1}{4} |I_L|^2 L = \frac{1}{4} |V|^2 \frac{1}{\omega^2 L}, \quad (6.14c)$$

where I_L is the current through the inductor. Then the complex power of (6.13) can be rewritten as

$$P_{\text{in}} = P_{\text{loss}} + 2j\omega(W_m - W_e), \quad (6.15)$$

which is identical to (6.4). Similarly, the input impedance can be expressed as

$$Z_{\text{in}} = \frac{2P_{\text{in}}}{|I|^2} = \frac{P_{\text{loss}} + 2j\omega(W_m - W_e)}{\frac{1}{2}|I|^2}, \quad (6.16)$$

which is identical to (6.5).

As in the series case, resonance occurs when $W_m = W_e$. Then from (6.16) and (6.14a) the input impedance at resonance is

$$Z_{\text{in}} = \frac{P_{\text{loss}}}{\frac{1}{2}|I|^2} = R,$$

which is a purely real impedance. From (6.14b) and (6.14c), $W_m = W_e$ implies that the resonant frequency, ω_0 , can be defined as

$$\omega_0 = \frac{1}{\sqrt{LC}}, \quad (6.17)$$

which is identical to the series resonant circuit case. Resonance in the case of a parallel RLC circuit is sometimes referred to as an *antiresonance*.

From the definition of (6.7), and the results in (6.14), the unloaded Q of the parallel resonant circuit can be expressed as

$$Q_0 = \omega_0 \frac{2W_m}{P_{\text{loss}}} = \frac{R}{\omega_0 L} = \omega_0 RC, \quad (6.18)$$

since $W_m = W_e$ at resonance. This result shows that the Q of the parallel resonant circuit increases as R increases.

Near resonance, the input impedance of (6.12) can be simplified using the series expansion result that

$$\frac{1}{1+x} \simeq 1 - x + \dots$$

Again letting $\omega = \omega_0 + \Delta\omega$, where $\Delta\omega$ is small, allows (6.12) to be rewritten as [1]

$$\begin{aligned} Z_{\text{in}} &\simeq \left(\frac{1}{R} + \frac{1 - \Delta\omega/\omega_0}{j\omega_0 L} + j\omega_0 C + j\Delta\omega C \right)^{-1} \\ &\simeq \left(\frac{1}{R} + j \frac{\Delta\omega}{\omega_0^2 L} + j\Delta\omega C \right)^{-1} \\ &\simeq \left(\frac{1}{R} + 2j\Delta\omega C \right)^{-1} \\ &\simeq \frac{R}{1 + 2j\Delta\omega RC} = \frac{R}{1 + 2jQ_0\Delta\omega/\omega_0}, \end{aligned} \quad (6.19)$$

since $\omega_0^2 = 1/LC$. When $R = \infty$ (6.19) reduces to

$$Z_{\text{in}} = \frac{1}{j2C(\omega - \omega_0)}.$$

As in the series resonator case, the effect of loss can be accounted for by replacing ω_0 in this expression with a complex effective resonant frequency:

$$\omega_0 \leftarrow \omega_0 \left(1 + \frac{j}{2Q_0} \right). \quad (6.20)$$

Figure 6.2b shows the behavior of the magnitude of the input impedance versus frequency. The half-power bandwidth edges occur at frequencies ($\Delta\omega/\omega_0 = BW/2$) such that

$$|Z_{\text{in}}|^2 = \frac{R^2}{2},$$

which, from (6.19), implies that

$$BW = \frac{1}{Q_0}, \quad (6.21)$$

as in the series resonance case.

Loaded and Unloaded Q

The unloaded Q , Q_0 , defined in the preceding sections is a characteristic of the resonator itself, in the absence of any loading effects caused by external circuitry. In practice, however, a resonator is invariably coupled to other circuitry, which will have the effect of lowering the overall, or *loaded* Q , Q_L , of the circuit. Figure 6.3 depicts a resonator coupled to an

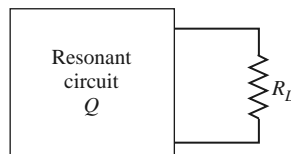


FIGURE 6.3 A resonant circuit connected to an external load, R_L .

TABLE 6.1 Summary of Results for Series and Parallel Resonators

Quantity	Series Resonator	Parallel Resonator
Input impedance/admittance	$Z_{\text{in}} = R + j\omega L - j\frac{1}{\omega C}$ $\simeq R + j\frac{2RQ_0\Delta\omega}{\omega_0}$	$Y_{\text{in}} = \frac{1}{R} + j\omega C - j\frac{1}{\omega L}$ $\simeq \frac{1}{R} + j\frac{2Q_0\Delta\omega}{R\omega_0}$
Power loss	$P_{\text{loss}} = \frac{1}{2} I ^2 R$	$P_{\text{loss}} = \frac{1}{2}\frac{ V ^2}{R}$
Stored magnetic energy	$W_m = \frac{1}{4} I ^2 L$	$W_m = \frac{1}{4} V ^2 \frac{1}{\omega^2 L}$
Stored electric energy	$W_e = \frac{1}{4} I ^2 \frac{1}{\omega^2 C}$	$W_e = \frac{1}{4} V ^2 C$
Resonant frequency	$\omega_0 = \frac{1}{\sqrt{LC}}$	$\omega_0 = \frac{1}{\sqrt{LC}}$
Unloaded Q	$Q_0 = \frac{\omega_0 L}{R} = \frac{1}{\omega_0 RC}$	$Q_0 = \omega_0 RC = \frac{R}{\omega_0 L}$
External Q	$Q_e = \frac{\omega_0 L}{R_L}$	$Q_e = \frac{R_L}{\omega_0 L}$

external load resistor, R_L . If the resonator is a series RLC circuit, the load resistor R_L adds in series with R , so the effective resistance in (6.8) is $R + R_L$. If the resonator is a parallel RLC circuit, the load resistor R_L combines in parallel with R , so the effective resistance in (6.18) is $RR_L/(R + R_L)$. If we define an *external* Q , Q_e , as

$$Q_e = \begin{cases} \frac{\omega_0 L}{R_L} & \text{for series circuits} \\ \frac{R_L}{\omega_0 L} & \text{for parallel circuits,} \end{cases} \quad (6.22)$$

then the loaded Q can be expressed as

$$\frac{1}{Q_L} = \frac{1}{Q_e} + \frac{1}{Q_0}. \quad (6.23)$$

Table 6.1 summarizes the above results for series and parallel resonant circuits.

6.2 TRANSMISSION LINE RESONATORS

As we have seen, ideal lumped circuit elements are often unattainable at microwave frequencies, so distributed elements are frequently used. In this section we will study the use of transmission line sections with various lengths and terminations (usually open- or short-circuited) to form resonators. Because we are interested in the Q of these resonators, we must consider transmission lines with losses.

Short-Circuited $\lambda/2$ Line

A length of lossy transmission line, short circuited at one end, is shown in Figure 6.4. The line has a characteristic impedance, Z_0 , propagation constant, β , and attenuation

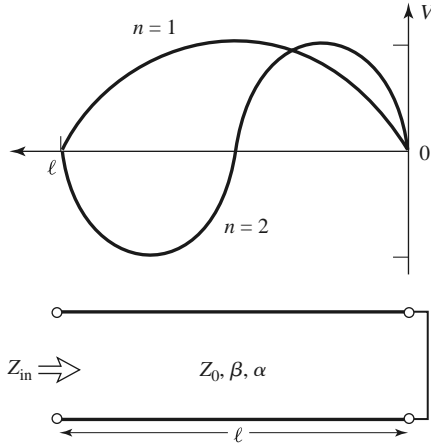


FIGURE 6.4 A short-circuited length of lossy transmission line, and the voltage distributions for $n = 1$ ($\ell = \lambda/2$) and $n = 2$ ($\ell = \lambda$) resonators.

constant, α . At the resonant frequency $\omega = \omega_0$, the length of the line is $\ell = \lambda/2$. From (2.91), the input impedance is

$$Z_{\text{in}} = Z_0 \tanh(\alpha + j\beta)\ell.$$

Using an identity for the hyperbolic tangent gives

$$Z_{\text{in}} = Z_0 \frac{\tanh \alpha \ell + j \tan \beta \ell}{1 + j \tan \beta \ell \tanh \alpha \ell}. \quad (6.24)$$

Observe that $Z_{\text{in}} = jZ_0 \tan \beta \ell$ if $\alpha = 0$ (a lossless line).

In practice it is usually desirable to use a low-loss transmission line, so we assume that $\alpha \ell \ll 1$, and then $\tanh \alpha \ell \simeq \alpha \ell$. Again let $\omega = \omega_0 + \Delta\omega$, where $\Delta\omega$ is small. Then, assuming a TEM line, we have

$$\beta \ell = \frac{\omega \ell}{v_p} = \frac{\omega_0 \ell}{v_p} + \frac{\Delta\omega \ell}{v_p},$$

where v_p is the phase velocity of the transmission line. Because $\ell = \lambda/2 = \pi v_p/\omega_0$ for $\omega = \omega_0$, we have

$$\beta \ell = \pi + \frac{\Delta\omega \pi}{\omega_0},$$

and then

$$\tan \beta \ell = \tan \left(\pi + \frac{\Delta\omega \pi}{\omega_0} \right) = \tan \frac{\Delta\omega \pi}{\omega_0} \simeq \frac{\Delta\omega \pi}{\omega_0}.$$

Using these results in (6.24) gives

$$Z_{\text{in}} \simeq Z_0 \frac{\alpha \ell + j(\Delta\omega \pi/\omega_0)}{1 + j(\Delta\omega \pi/\omega_0)\alpha \ell} \simeq Z_0 \left(\alpha \ell + j \frac{\Delta\omega \pi}{\omega_0} \right), \quad (6.25)$$

since $\Delta\omega \alpha \ell/\omega_0 \ll 1$.

Equation (6.25) is of the form

$$Z_{\text{in}} = R + 2jL\Delta\omega,$$

which is the input impedance of a series RLC resonant circuit, as given by (6.9). We can identify the resistance of the equivalent circuit as

$$R = Z_0 \alpha \ell, \quad (6.26a)$$

and the inductance of the equivalent circuit as

$$L = \frac{Z_0 \pi}{2\omega_0}. \quad (6.26b)$$

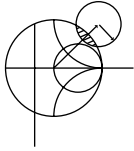
The capacitance of the equivalent circuit can be found from (6.6) as

$$C = \frac{1}{\omega_0^2 L}. \quad (6.26c)$$

The resonator of Figure 6.4 thus resonates for $\Delta\omega = 0$ ($\ell = \lambda/2$), and its input impedance at resonance is $Z_{in} = R = Z_0 \alpha \ell$. Resonance also occurs for $\ell = n\lambda/2$, $n = 1, 2, 3, \dots$. The voltage distributions for the $n = 1$ and $n = 2$ resonant modes are shown in Figure 6.4. The unloaded Q of this resonator can be found from (6.8) and (6.26) as

$$Q_0 = \frac{\omega_0 L}{R} = \frac{\pi}{2\alpha \ell} = \frac{\beta}{2\alpha}, \quad (6.27)$$

since $\beta \ell = \pi$ at the first resonance. This result shows that the Q decreases as the attenuation of the line increases, as expected.



EXAMPLE 6.1 Q OF HALF-WAVE COAXIAL LINE RESONATORS

A $\lambda/2$ resonator is made from a piece of copper coaxial line having an inner conductor radius of 1 mm and an outer conductor radius of 4 mm. If the resonant frequency is 5 GHz, compare the unloaded Q of an air-filled coaxial line resonator to that of a Teflon-filled coaxial line resonator.

Solution

We first compute the attenuation of the coaxial line, using the results of Examples 2.6 or 2.7. From Appendix F, the conductivity of copper is $\sigma = 5.813 \times 10^7$ S/m. The surface resistivity at 5 GHz is

$$R_s = \sqrt{\frac{\omega \mu_0}{2\sigma}} = 1.84 \times 10^{-2} \Omega.$$

The attenuation due to conductor loss for the air-filled line is

$$\begin{aligned} \alpha_c &= \frac{R_s}{2\eta \ln b/a} \left(\frac{1}{a} + \frac{1}{b} \right) \\ &= \frac{1.84 \times 10^{-2}}{2(377) \ln (0.004/0.001)} \left(\frac{1}{0.001} + \frac{1}{0.004} \right) = 0.022 \text{ Np/m.} \end{aligned}$$