Chapter 7

Transferred Electron Devices (TEDs)

7-0 INTRODUCTION

The application of two-terminal semiconductor devices at microwave frequencies has been increased usage during the past decades. The CW, average, and peak power outputs of these devices at higher microwave frequencies are much larger than those obtainable with the best power transistor. The common characteristic of all active two-terminal solid-state devices is their negative resistance. The real part of their impedance is negative over a range of frequencies. In a positive resistance the current through the resistance and the voltage across it are in phase. The voltage drop across a positive resistance is positive and a power of (I^2R) is dissipated in the resistance. In a negative resistance, however, the current and voltage are out of phase by 180°. The voltage drop across a negative resistance is negative, and a power of $(-I^2R)$ is generated by the power supply associated with the negative resistance. In other words, positive resistances absorb power (passive devices), whereas negative resistances generate power (active devices). In this chapter the transferred electron devices (TEDs) are analyzed.

The differences between microwave transistors and transferred electron devices (TEDs) are fundamental. Transistors operate with either junctions or gates, but TEDs are bulk devices having no junctions or gates. The majority of transistors are fabricated from elemental semiconductors, such as silicon or germanium, whereas TEDs are fabricated from compound semiconductors, such as gallium arsenide (GaAs), indium phosphide (InP), or cadmium telluride (CdTe). Transistors operate with "warm" electrons whose energy is not much greater than the thermal energy (0.026 eV at room temperature) of electrons in the semiconductor, whereas TEDs operate with "hot" electrons whose energy is very much greater than the thermal energy. Because of these fundamental differences, the theory and technology of transistors cannot be applied to TEDs.

7-1 GUNN-EFFECT DIODES—GaAs DIODE

Gunn-effect diodes are named after J. B. Gunn, who in 1963 discovered a periodic fluctuations of current passing through the *n*-type gallium arsenide (GaAs) specimen when the applied voltage exceeded a certain critical value. Two years later, in 1965, B. C. DeLoach, R. C. Johnston, and B. G. Cohen discovered the impact ionization avalanche transit-time (IMPATT) mechanism in silicon, which employs the avalanching and transit-time properties of the diode to generate microwave frequencies. In later years the limited space-charge-accumulation diode (LSA diode) and the indium phosphide diode (InP diode) were also successfully developed. These are bulk devices in the sense that microwave amplification and oscillation are derived from the bulk negative-resistance property of uniform semiconductors rather than from the junction negative-resistance property between two different semiconductors, as in the tunnel diode.

7-1-1 Background

After inventing the transistor, Shockley suggested in 1954 that two-terminal negative-resistance devices using semiconductors may have advantages over transistors at high frequencies [1]. In 1961 Ridley and Watkins described a new method for obtaining negative differential mobility in semiconductors [2]. The principle involved is to heat carriers in a light-mass, high-mobility subband with an electric field so that the carriers can transfer to a heavy-mass, low-mobility, higher-energy subband when they have a high enough temperature. Ridley and Watkins also mentioned that Ge–Si alloys and some III–V compounds may have suitable subband structures in the conduction bands. Their theory for achieving negative differential mobility in bulk semiconductors by transferring electrons from high-mobility energy bands to low-mobility energy bands was taken a step further by Hilsum in 1962 [3]. Hilsum carefully calculated the transferred electron effect in several III–V compounds and was the first to use the terms transferred electron amplifiers (TEAs) and oscillators (TEOs). He predicted accurately that a TEA bar of semi-insulating GaAs would be operated at 373°K at a field of 3200 V/cm. Hilsum's attempts to verify his theory experimentally failed because the GaAs diode available to him at that time was not of sufficiently high quality.

It was not until 1963 that J. B. Gunn of IBM discovered the so-called Gunn effect from thin disks of *n*-type GaAs and *n*-type InP specimens while studying the noise properties of semiconductors [4]. Gunn did not connect—and even immediately rejected—his discoveries with the theories of Ridley, Watkins, and Hilsum. In 1963 Ridley predicted [5] that the field domain is continually moving down through the crystal, disappearing at the anode and then reappearing at a favored nucleating center, and starting the whole cycle once more. Finally, Kroemer stated [6] that the origin of the negative differential mobility is Ridley–Watkins–Hilsum's mechanism of electron transfer into the satellite valleys that occur in the conduction bands of both the n-type GaAs and the n-type InP and that the properties of the Gunn effect are the current oscillations caused by the periodic nucleation and disappearance of traveling space-charge instability domains. Thus the correlation of theoretical predictions and experimental discoveries completed the theory of transferred electron devices (TEDs).

7-1-2 Gunn Effect

A schematic diagram of a uniform n-type GaAs diode with ohmic contacts at the end surfaces is shown in Fig. 7-1-1. J. B. Gunn observed the Gunn effect in the n-type GaAs bulk diode in 1963, an effect best explained by Gunn himself, who published several papers about his observations [7 to 9]. He stated in his first paper [7] that

> Above some critical voltage, corresponding to an electric field of 2000–4000 volts/cm, the current in every specimen became a fluctuating function of time. In the GaAs specimens, this fluctuation took the form of a periodic oscillation superimposed upon the pulse current. . . . The frequency of oscillation was determined mainly by the specimen, and not by the external circuit. . . . The period of oscillation was usually inversely proportional to the specimen length and closely equal to the transit time of electrons between the electrodes, calculated from their estimated velocity of slightly over 10^7 cm/s. . . . The peak pulse microwave power delivered by the GaAs specimens to a matched load was measured. Value as high as 0.5 W at 1 Gc/s, and 0.15 W at 3 Gc/s, were found, corresponding to 1–2% of the pulse input power.*

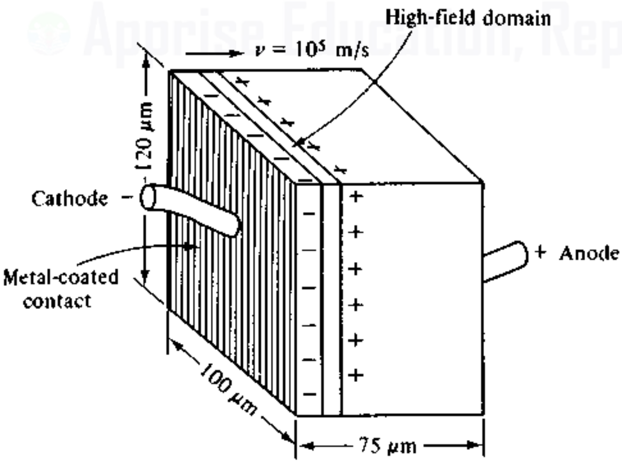

Figure 7-1-1 Schematic diagram for n-type GaAs diode.

*After J. B. Gunn [7]; reproduced by permission of IBM, Inc.

From Gunn's observation the carrier drift velocity is linearly increased from zero to a maximum when the electric field is varied from zero to a threshold value. When the electric field is beyond the threshold value of 3000 V/cm for the n-type GaAs, the drift velocity is decreased and the diode exhibits negative resistance. This situation is shown in Fig. 7-1-2.

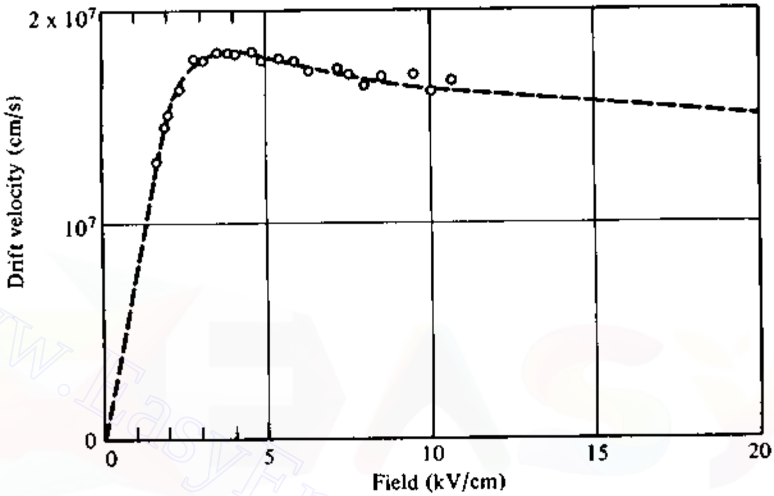

Figure 7-1-2 Drift velocity of electrons in n-type GaAs versus electric field. (*After J. B. Gunn [8]; reprinted by permission of IBM, Inc.*)

The current fluctuations are shown in Fig. 7-1-3. The current waveform was produced by applying a voltage pulse of 16-V amplitude and 10-ns duration to a specimen of n-type GaAs 2.5×10^{-3} cm in length. The oscillation frequency was 4.5 GHz. The lower trace had 2 ns/cm in the horizontal axis and 0.23 A/cm in the vertical axis. The upper trace was the expanded view of the lower trace. Gunn found that the period of these oscillations was equal to the transit time of the electrons through the specimen calculated from the threshold current.

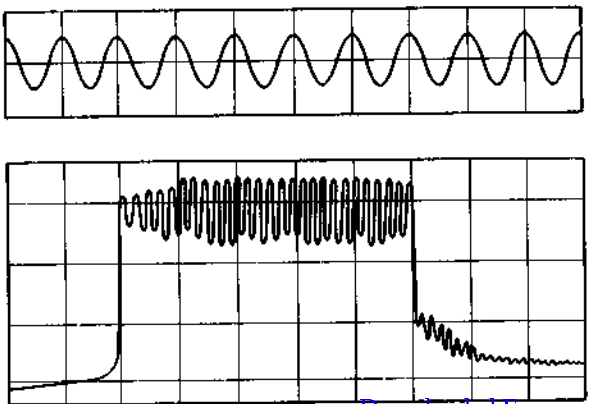

Figure 7-1-3 Current waveform of n-type GaAs reported by Gunn. (*After J. B. Gunn [8]; reprinted by permission of IBM, Inc.*)

Gunn also discovered that the threshold electric field E_{th} varied with the length and type of material. He developed an elaborate capacitive probe for plotting the electric field distribution within a specimen of n-type GaAs of length $L = 210\ \mu m$ and cross-sectional area $3.5 \times 10^{-3}\ cm^2$ with a low-field resistance of 16 Ω. Current instabilities occurred at specimen voltages above 59 V, which means that the threshold field is

$$E_{th} = \frac{V}{L} = \frac{59}{210 \times 10^{-6} \times 10^2} = 2810 \text{ volts/cm} \tag{7-1-1}$$

7-2 RIDLEY–WATKINS–HILSUM (RWH) THEORY

Many explanations have been offered for the Gunn effect. In 1964 Kroemer [6] suggested that Gunn's observations were in complete agreement with the Ridley–Watkins–Hilsum (RWH) theory.

7-2-1 Differential Negative Resistance

The fundamental concept of the Ridley–Watkins–Hilsum (RWH) theory is the differential negative resistance developed in a bulk solid-state III-V compound when either a voltage (or electric field) or a current is applied to the terminals of the sample. There are two modes of negative-resistance devices: voltage-controlled and current-controlled modes as shown in Fig. 7-2-1(a) and (b), respectively [5].

Figure 7-2-1 Diagram of negative resistance. (*From B. K. Ridley [5]; reprinted by permission of the Institute of Physics.*)

In the voltage-controlled mode the current density can be multivalued, whereas in the current-controlled mode the voltage can be multivalued. The major effect of the appearance of a differential negative-resistance region in the current-density-field curve is to render the sample electrically unstable. As a result, the initially homogeneous sample becomes electrically heterogeneous in an attempt to reach stability. In the voltage-controlled negative-resistance mode high-field domains are formed, separating two low-field regions. The interfaces separating low- and high-field domains lie along equipotentials; thus they are in planes perpendicular to the current direction as shown in Fig. 7-2-2(a). In the current-controlled negative-resistance mode splitting the sample results in high-current filaments running along the field direction as shown in Fig. 7-2-2(b).

Chapter 0

Introduction

The central theme of this book concerns the basic principles and applications of microwave devices and circuits. Microwave techniques have been increasingly adopted in such diverse applications as radio astronomy, long-distance communications, space navigation, radar systems, medical equipment, and missile electronic systems. As a result of the accelerating rate of growth of microwave technology in research and industry, students who are preparing themselves for, and electronics engineers who are working in, the microwave area are faced with the need to understand the theoretical and experimental design and analysis of microwave devices and circuits.

0-1 MICROWAVE FREQUENCIES

The term *microwave frequencies* is generally used for those wavelengths measured in centimeters, roughly from 30 cm to 1 mm (1 to 300 GHz). However, *microwave* really indicates the wavelengths in the micron ranges. This means microwave frequencies are up to infrared and visible-light regions. In this revision, microwave frequencies refer to those from 1 GHz up to 10^6 GHz. The microwave band designation that derived from World War II radar security considerations has never been officially sanctioned by any industrial, professional, or government organization. In August 1969 the United States Department of Defense, Office of Joint Chiefs of Staff, by message to all services, directed the use of a new frequency band breakdown as shown in Table 0-1. On May 24, 1970, the Department of Defense adopted another band designation for microwave frequencies as listed in Table 0-2. The Institute of Electrical and Electronics Engineers (IEEE) recommended new microwave band designations as shown in Table 0-3 for comparison.

TABLE 0-1 U.S. MILITARY MICROWAVE BANDS

Designation	Frequency range in gigahertz
P band	0.225– 0.390
L band	0.390– 1.550
S band	1.550– 3.900
C band	3.900– 6.200
X band	6.200– 10.900
K band	10.900– 36.000
Q band	36.000– 46.000
V band	46.000– 56.000
W band	56.000–100.000

TABLE 0-2 U.S. NEW MILITARY MICROWAVE BANDS

Designation	Frequency range in gigahertz	Designation	Frequency range in gigahertz
A band	0.100–0.250	H band	6.000– 8.000
B band	0.250–0.500	I band	8.000– 10.000
C band	0.500–1.000	J band	10.000– 20.000
D band	1.000–2.000	K band	20.000– 40.000
E band	2.000–3.000	L band	40.000– 60.000
F band	3.000–4.000	M band	60.000–100.000
G band	4.000–6.000		

TABLE 0-3 IEEE MICROWAVE FREQUENCY BANDS

Designation	Frequency range in gigahertz
HF	0.003– 0.030
VHF	0.030– 0.300
UHF	0.300– 1.000
L band	1.000– 2.000
S band	2.000– 4.000
C band	4.000– 8.000
X band	8.000– 12.000
Ku band	12.000– 18.000
K band	18.000– 27.000
Ka band	27.000– 40.000
Millimeter	40.000–300.000
Submillimeter	>300.000

0-2 MICROWAVE DEVICES

In the late 1930s it became evident that as the wavelength approached the physical dimensions of the vacuum tubes, the electron transit angle, interelectrode capacitance, and lead inductance appeared to limit the operation of vacuum tubes in microwave frequencies. In 1935 A. A. Heil and O. Heil suggested that microwave voltages be generated by using transit-time effects together with lumped tuned circuits. In 1939 W. C. Hahn and G. F. Metcalf proposed a theory of velocity modulation for microwave tubes. Four months later R. H. Varian and S. F. Varian described a two-cavity klystron amplifier and oscillator by using velocity modulation. In 1944 R. Kompfner invented the helix-type traveling-wave tube (TWT). Ever since then the concept of microwave tubes has deviated from that of conventional vacuum tubes as a result of the application of new principles in the amplification and generation of microwave energy.

Historically microwave generation and amplification were accomplished by means of velocity-modulation theory. In the past two decades, however, microwave solid-state devices—such as tunnel diodes, Gunn diodes, transferred electron devices (TEDs), and avalanche transit-time devices have been developed to perform these functions. The conception and subsequent development of TEDs and avalanche transit-time devices were among the outstanding technical achievements. B. K. Ridley and T. B. Watkins in 1961 and C. Hilsum in 1962 independently predicted that the transferred electron effect would occur in GaAs (gallium arsenide). In 1963 J. B. Gunn reported his "Gunn effect." The common characteristic of all microwave solid-state devices is the negative resistance that can be used for microwave oscillation and amplification. The progress of TEDs and avalanche transit-time devices has been so swift that today they are firmly established as one of the most important classes of microwave solid-state devices.

0-3 MICROWAVE SYSTEMS

A microwave system normally consists of a transmitter subsystem, including a microwave oscillator, waveguides, and a transmitting antenna, and a receiver subsystem that includes a receiving antenna, transmission line or waveguide, a microwave amplifier, and a receiver. Figure 0-1 shows a typical microwave system.

In order to design a microwave system and conduct a proper test of it, an adequate knowledge of the components involved is essential. Besides microwave devices, the text therefore describes microwave components, such as resonators, cavities, microstrip lines, hybrids, and microwave integrated circuits.

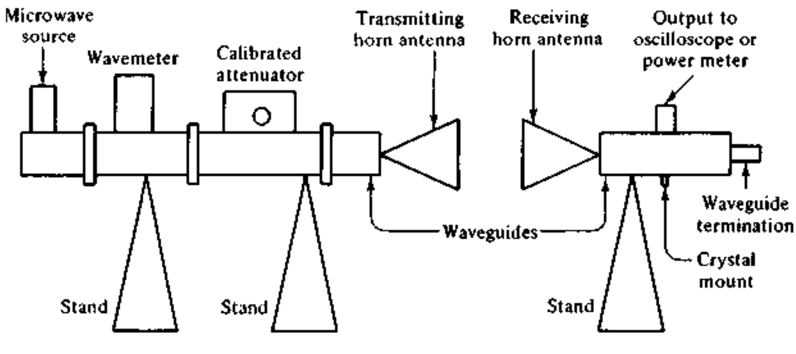

Figure 0-1 Microwave system.

Chapter 3

Microwave Transmission Lines

3-0 INTRODUCTION

Conventional two-conductor transmission lines are commonly used for transmitting microwave energy. If a line is properly matched to its characteristic impedance at each terminal, its efficiency can reach a maximum.

In ordinary circuit theory it is assumed that all impedance elements are lumped constants. This is not true for a long transmission line over a wide range of frequencies. Frequencies of operation are so high that inductances of short lengths of conductors and capacitances between short conductors and their surroundings cannot be neglected. These inductances and capacitances are distributed along the length of a conductor, and their effects combine at each point of the conductor. Since the wavelength is short in comparison to the physical length of the line, distributed parameters cannot be represented accurately by means of a lumped-parameter equivalent circuit. Thus microwave transmission lines can be analyzed in terms of voltage, current, and impedance only by the distributed-circuit theory. If the spacing between the lines is smaller than the wavelength of the transmitted signal, the transmission line must be analyzed as a waveguide.

3-1 TRANSMISSION-LINE EQUATIONS AND SOLUTIONS

3-1-1 Transmission-Line Equations

A transmission line can be analyzed either by the solution of Maxwell's field equations or by the methods of distributed-circuit theory. The solution of Maxwell's equations involves three space variables in addition to the time variable. The distributed-circuit method, however, involves only one space variable in addition to the time variable. In this section the latter method is used to analyze a transmission line in terms of the voltage, current, impedance, and power along the line.

Based on uniformly distributed-circuit theory, the schematic circuit of a conventional two-conductor transmission line with constant parameters R, L, G, and C is shown in Fig. 3-1-1. The parameters are expressed in their respective names per unit length, and the wave propagation is assumed in the positive z direction.

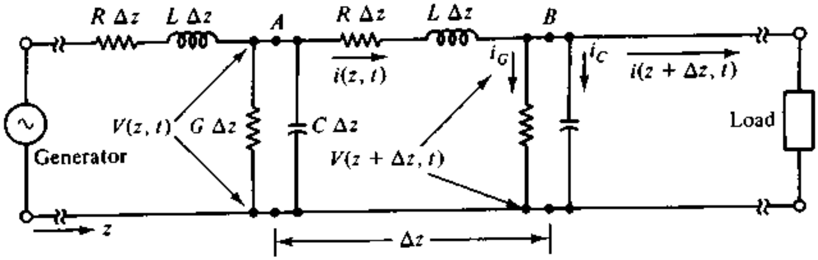

Figure 3-1-1 Elementary section of a transmission line.

By Kirchhoff's voltage law, the summation of the voltage drops around the central loop is given by

$$v(z, t) = i(z, t)R\,\Delta z + L\,\Delta z\frac{\partial i(z, t)}{\partial t} + v(z, t) + \frac{\partial v(z, t)}{dz}\Delta z \qquad \text{(3-1-1)}$$

Rearranging this equation, dividing it by Δz, and then omitting the argument (z, t), which is understood, we obtain

$$-\frac{\partial v}{\partial z} = Ri + L\frac{\partial i}{\partial t} \qquad \text{(3-1-2)}$$

Using Kirchhoff's current law, the summation of the currents at point B in Fig. 3-1-1 can be expressed as

$$\begin{aligned} i(z, t) &= v(z + \Delta z, t)G\,\Delta z + C\,\Delta z\,\frac{\partial v(z + \Delta z, t)}{\partial t} + i(z + \Delta z, t) \\ &= \left[v(z, t) + \frac{\partial v(z, t)}{\partial z}\Delta z\right]G\,\Delta z \qquad \text{(3-1-3)} \\ &\quad + C\,\Delta z\,\frac{\partial}{\partial t}\left[v(z, t) + \frac{\partial v(z, t)}{\partial z}\Delta z\right] + i(z, t) + \frac{\partial i(z, t)}{\partial z}\Delta z \end{aligned}$$

By rearranging the preceding equation, dividing it by Δz, omitting (z, t), and assuming Δz equal to zero, we have

$$-\frac{\partial i}{\partial z} = Gv + C\frac{\partial v}{\partial t} \qquad \text{(3-1-4)}$$

Then by differentiating Eq. (3-1-2) with respect to z and Eq. (3-1-4) with respect to t and combining the results, the final transmission-line equation in voltage form is found to be

$$\frac{\partial^2 v}{\partial z^2} = RGv + (RC + LG)\frac{\partial v}{\partial t} + LC\frac{\partial^2 v}{\partial t^2} \tag{3-1-5}$$

Also, by differentiating Eq. (3-1-2) with respect to t and Eq. (3-1-4) with respect to z and combining the results, the final transmission-line equation in current form is

$$\frac{\partial^2 i}{\partial z^2} = RGi + (RC + LG)\frac{\partial i}{\partial t} + LC\frac{\partial^2 i}{\partial t^2} \tag{3-1-6}$$

All these transmission-line equations are applicable to the general transient solution. The voltage and current on the line are the functions of both position z and time t. The instantaneous line voltage and current can be expressed as

$$v(z, t) = \text{Re}\ \mathbf{V}(z)e^{j\omega t} \tag{3-1-7}$$

$$i(z, t) = \text{Re}\ \mathbf{I}(z)e^{j\omega t} \tag{3-1-8}$$

where Re stands for "real part of." The factors $\mathbf{V}(z)$ aand $\mathbf{I}(z)$ are complex quantities of the sinusoidal functions of position z on the line and are known as *phasors*. The phasors give the magnitudes and phases of the sinusoidal function at each position of z, and they can be expressed as

$$\mathbf{V}(z) = \mathbf{V}_+ e^{-\gamma z} + \mathbf{V}_- e^{\gamma z} \tag{3-1-9}$$

$$\mathbf{I}(z) = \mathbf{I}_+ e^{-\gamma z} + \mathbf{I}_- e^{\gamma z} \tag{3-1-10}$$

$$\gamma = \alpha + j\beta \qquad \text{(propagation constant)} \tag{3-1-11}$$

where $\mathbf{V}_+$ and $\mathbf{I}_+$ indicate complex amplitudes in the positive z direction, $\mathbf{V}_-$ and $\mathbf{I}_-$ signify complex amplitudes in the negative z direction, α is the attenuation constant in nepers per unit length, and β is the phase constant in radians per unit length.

If we substitute $j\omega$ for $\partial/\partial t$ in Eqs. (3-1-2), (3-1-4), (3-1-5), and (3-1-6) and divide each equation by $e^{j\omega t}$, the transmission-line equations in phasor form of the frequency domain become

$$\frac{d\mathbf{V}}{dz} = -\mathbf{ZI} \tag{3-1-12}$$

$$\frac{d\mathbf{I}}{dz} = -\mathbf{YV} \tag{3-1-13}$$

$$\frac{d^2\mathbf{V}}{dz^2} = \gamma^2\mathbf{V} \tag{3-1-14}$$

$$\frac{d^2\mathbf{I}}{dz^2} = \gamma^2\mathbf{I} \tag{3-1-15}$$

in which the following substitutions have been made:

$$\mathbf{Z} = R + j\omega L \qquad \text{(ohms per unit length)} \tag{3-1-16}$$

$$\mathbf{Y} = G + j\omega C \qquad \text{(mhos per unit length)} \tag{3-1-17}$$

$$\gamma = \sqrt{\mathbf{Z}\mathbf{Y}} = \alpha + j\beta \qquad \text{(propagation constant)} \tag{3-1-18}$$

For a lossless line, $R = G = 0$, and the transmission-line equations are expressed as

$$\frac{d\mathbf{V}}{dz} = -j\omega L\mathbf{I} \tag{3-1-19}$$

$$\frac{d\mathbf{I}}{dz} = -j\omega C\mathbf{V} \tag{3-1-20}$$

$$\frac{d^2\mathbf{V}}{dz^2} = -\omega^2 LC\mathbf{V} \tag{3-1-21}$$

$$\frac{d^2\mathbf{I}}{dz^2} = -\omega^2 LC\mathbf{I} \tag{3-1-22}$$

It is interesting to note that Eqs. (3-1-14) and (3-1-15) for a transmission line are similar to equations of the electric and magnetic waves, respectively. The only difference is that the transmission-line equations are one-dimensional.

3-1-2 Solutions of Transmission-Line Equations

The one possible solution for Eq. (3-1-14) is

$$\mathbf{V} = \mathbf{V}_+e^{-\gamma z} + \mathbf{V}_-e^{\gamma z} = \mathbf{V}_+e^{-\alpha z}e^{-j\beta z} + \mathbf{V}_-e^{\alpha z}e^{j\beta z} \tag{3-1-23}$$

The factors $\mathbf{V}_+$ and $\mathbf{V}_-$ represents complex quantities. The term involving $e^{-j\beta z}$ shows a wave traveling in the positive z direction, and the term with the factor $e^{j\beta z}$ is a wave going in the negative z direction. The quantity βz is called the *electrical length of the line* and is measured in radians.

Similarly, the one possible solution for Eq. (3-1-15) is

$$\mathbf{I} = \mathbf{Y}_0(\mathbf{V}_+e^{-\gamma z} - \mathbf{V}_-e^{\gamma z}) = \mathbf{Y}_0(\mathbf{V}_+e^{-\alpha z}e^{-j\beta z} - \mathbf{V}_-e^{\alpha z}e^{j\beta z}) \tag{3-1-24}$$

In Eq. (3-1-24) the characteristic impedance of the line is defined as

$$\mathbf{Z}_0 = \frac{1}{\mathbf{Y}_0} \equiv \sqrt{\frac{\mathbf{Z}}{\mathbf{Y}}} = \sqrt{\frac{R + j\omega L}{G + j\omega C}} = R_0 \pm jX_0 \tag{3-1-25}$$

The magnitude of both voltage and current waves on the line is shown in Fig. 3-1-2.

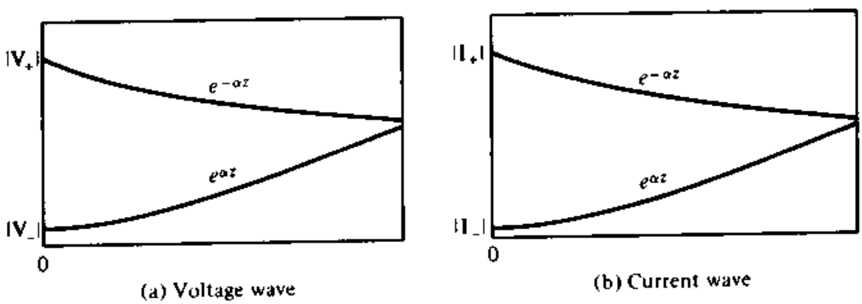

Figure 3-1-2 Magnitude of voltage and current traveling waves.

At microwave frequencies it can be seen that

$$R \ll \omega L \quad \text{and} \quad G \ll \omega C \tag{3-1-26}$$

By using the binomial expansion, the propagation constant can be expressed as

$$\begin{aligned}
\gamma &= \sqrt{(R + j\omega L)(G + j\omega C)} \\
&= \sqrt{(j\omega)^2 LC}\sqrt{\left(1 + \frac{R}{j\omega L}\right)\left(1 + \frac{G}{j\omega C}\right)} \\
&\approx j\omega\sqrt{LC}\left[\left(1 + \frac{1}{2}\frac{R}{j\omega L}\right)\left(1 + \frac{1}{2}\frac{G}{j\omega C}\right)\right] \\
&\approx j\omega\sqrt{LC}\left[1 + \frac{1}{2}\left(\frac{R}{j\omega L} + \frac{G}{j\omega C}\right)\right] \\
&= \frac{1}{2}\left(R\sqrt{\frac{C}{L}} + G\sqrt{\frac{L}{C}}\right) + j\omega\sqrt{LC}
\end{aligned} \tag{3-1-27}$$

Therefore the attenuation and phase constants are, respectively, given by

$$\alpha = \frac{1}{2}\left(R\sqrt{\frac{C}{L}} + G\sqrt{\frac{L}{C}}\right) \tag{3-1-28}$$

$$\beta = \omega\sqrt{LC} \tag{3-1-29}$$

Similarly, the characteristic impedance is found to be

$$\begin{aligned}
\mathbf{Z}_0 &= \sqrt{\frac{R + j\omega L}{G + j\omega C}} \\
&= \sqrt{\frac{L}{C}}\left(1 + \frac{R}{j\omega L}\right)^{1/2}\left(1 + \frac{G}{j\omega C}\right)^{-1/2} \\
&\approx \sqrt{\frac{L}{C}}\left(1 + \frac{1}{2}\frac{R}{j\omega L}\right)\left(1 - \frac{1}{2}\frac{G}{j\omega C}\right) \\
&\approx \sqrt{\frac{L}{C}}\left[1 + \frac{1}{2}\left(\frac{R}{j\omega L} - \frac{G}{j\omega C}\right)\right] \\
&\approx \sqrt{\frac{L}{C}}
\end{aligned} \tag{3-1-30}$$

From Eq. (3-1-29) the phase velocity is

$$v_p = \frac{\omega}{\beta} = \frac{1}{\sqrt{LC}} \tag{3-1-31}$$

The product of LC is independent of the size and separation of the conductors and depends only on the permeability μ and permittivity of ϵ of the insulating medium. If a lossless transmission line used for microwave frequencies has an air dielectric and contains no ferromagnetic materials, free-space parameters can be assumed.

Thus the numerical value of $1/\sqrt{LC}$ for air-insulated conductors is approximately equal to the velocity of light in vacuum. That is,

$$v_p = \frac{1}{\sqrt{LC}} = \frac{1}{\sqrt{\mu_0 \epsilon_0}} = c = 3 \times 10^8 \text{ m/s} \tag{3-1-32}$$

When the dielectric of a lossy microwave transmission line is not air, the phase velocity is smaller than the velocity of light in vacuum and is given by

$$v_\epsilon = \frac{1}{\sqrt{\mu\epsilon}} = \frac{c}{\sqrt{\mu_r \epsilon_r}} \tag{3-1-33}$$

In general, the relative phase velocity factory can be defined as

$$\text{Velocity factor} = \frac{\text{actual phase velocity}}{\text{velocity of light in vacuum}}$$

$$v_r = \frac{v_\epsilon}{c} = \frac{1}{\sqrt{\mu_r \epsilon_r}} \tag{3-1-34}$$

A low-loss transmission line filled only with dielectric medium, such as a coaxial line with solid dielectric between conductors, has a velocity factor on the order of about 0.65.

Example 3-1-1: Line Characteristic Impedance and Propagation Constant

A transmission line has the following parameters:

$$R = 2\ \Omega/\text{m} \qquad G = 0.5 \text{ mmho/m} \qquad f = 1 \text{ GHz}$$

$$L = 8 \text{ nH/m} \qquad C = 0.23 \text{ pF}$$

Calculate: **(a)** the characteristic impedance; **(b)** the propagation constant.

Solution

a. From Eq. (3-1-25) the line characteristic impedance is

$$\mathbf{Z}_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}} = \sqrt{\frac{2 + j2\pi \times 10^9 \times 8 \times 10^{-9}}{0.5 \times 10^{-3} + j2\pi \times 10^9 \times 0.23 \times 10^{-12}}}$$

$$= \sqrt{\frac{50.31\underline{/87.72^\circ}}{15.29 \times 10^{-4}\underline{/70.91^\circ}}} = 181.39\underline{/8.40^\circ} = 179.44 + j26.50$$

b. From Eq. (3-1-18) the propagation constant is

$$\gamma = \sqrt{(R + j\omega L)(G + j\omega C)} = \sqrt{(50.31\underline{/87.72^\circ})(15.29 \times 10^{-4}\underline{/70.91^\circ})}$$

$$= \sqrt{769.24 \times 10^{-4}\underline{/158.63^\circ}}$$

$$= 0.2774\underline{/79.31^\circ} = 0.051 + j0.273$$

3-2 REFLECTION COEFFICIENT AND TRANSMISSION COEFFICIENT

3-2-1 Reflection Coefficient

In the analysis of the solutions of transmission-line equations in Section 3-1, the traveling wave along the line contains two components: one traveling in the positive z direction and the other traveling the negative z direction. If the load impedance is equal to the line characteristic impedance, however, the reflected traveling wave does not exist.

Figure 3-2-1 shows a transmission line terminated in an impedance $\mathbf{Z}_\ell$. It is usually more convenient to start solving the transmission-line problem from the receiving rather than the sending end, since the voltage-to-current relationship at the load point is fixed by the load impedance. The incident voltage and current waves traveling along the transmission line are given by

$$\mathbf{V} = \mathbf{V}_+ e^{-\gamma z} + \mathbf{V}_- e^{+\gamma z} \tag{3-2-1}$$

$$\mathbf{I} = \mathbf{I}_+ e^{-\gamma z} + \mathbf{I}_- e^{+\gamma z} \tag{3-2-2}$$

in which the current wave can be expressed in terms of the voltage by

$$\mathbf{I} = \frac{\mathbf{V}_+}{\mathbf{Z}_0} e^{-\gamma z} - \frac{\mathbf{V}^-}{\mathbf{Z}_0} e^{\gamma z} \tag{3-2-3}$$

If the line has a length of ℓ, the voltage and current at the receiving end become

$$\mathbf{V}_\ell = \mathbf{V}_+ e^{-\gamma \ell} + \mathbf{V}_- e^{\gamma \ell} \tag{3-2-4}$$

$$\mathbf{I}_\ell = \frac{1}{\mathbf{Z}_0} (\mathbf{V}_+ e^{-\gamma \ell} - \mathbf{V}_- e^{\gamma \ell}) \tag{3-2-5}$$

The ratio of the voltage to the current at the receiving end is the load impedance. That is,

$$\mathbf{Z}_\ell = \frac{\mathbf{V}_\ell}{\mathbf{I}_\ell} = \mathbf{Z}_0 \frac{\mathbf{V}_+ e^{-\gamma \ell} + \mathbf{V}_- e^{\gamma \ell}}{\mathbf{V}_+ e^{-\gamma \ell} - \mathbf{V}_- e^{\gamma \ell}} \tag{3-2-6}$$

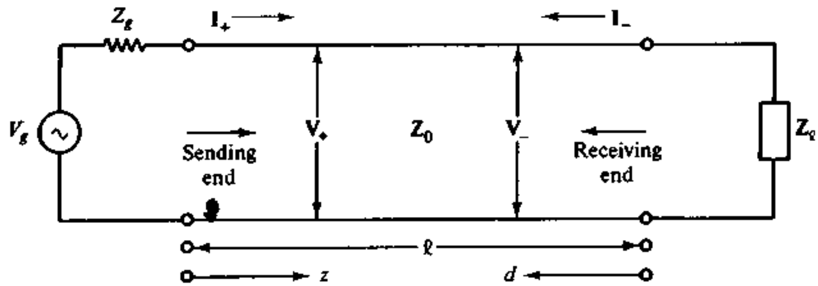

Figure 3-2-1 Transmission line terminated in a load impedance.

The reflection coefficient, which is designated by Γ (gamma), is defined as

$$\text{Reflection coefficient} \equiv \frac{\text{reflected voltage or current}}{\text{incident voltage or current}}$$

$$\Gamma \equiv \frac{\mathbf{V}_{\text{ref}}}{\mathbf{V}_{\text{inc}}} = \frac{-\mathbf{I}_{\text{ref}}}{\mathbf{I}_{\text{inc}}} \tag{3-2-7}$$

If Eq. (3-2-6) is solved for the ratio of the reflected voltage at the receiving end, which is $\mathbf{V}_- e^{\gamma\ell}$, to the incident voltage at the receiving end, which is $\mathbf{V}_+ e^{\gamma\ell}$, the result is the reflection coefficient at the receiving end:

$$\Gamma_\ell = \frac{\mathbf{V}_- e^{\gamma\ell}}{\mathbf{V}_+ e^{-\gamma\ell}} = \frac{\mathbf{Z}_\ell - \mathbf{Z}_0}{\mathbf{Z}_\ell + \mathbf{Z}_0} \tag{3-2-8}$$

If the load impedance and/or the characteristic impedance are complex quantities, as is usually the case, the reflection coefficient is generally a complex quantity that can be expressed as

$$\Gamma_\ell = |\Gamma_\ell| e^{j\theta_\ell} \tag{3-2-9}$$

where $|\Gamma_\ell|$ is the magnitude and never greater than unity—that is, $|\Gamma_\ell| \le 1$. Note that θ_ℓ is the phase angle between the incident and reflected voltages at the receiving end. It is usually called the phase angle of the reflection coefficient.

The general solution of the reflection coefficient at any point on the line, then, corresponds to the incident and reflected waves at that point, each attenuated in the direction of its own progress along the line. The generalized reflection coefficient is defined as

$$\Gamma \equiv \frac{\mathbf{V}_- e^{\gamma z}}{\mathbf{V}_+ e^{-\gamma z}} \tag{3-2-10}$$

From Fig. 3-2-1 let $z = \ell - d$. Then the reflection coefficient at some point located a distance d from the receiving end is

$$\Gamma_d = \frac{\mathbf{V}_- e^{\gamma(\ell - d)}}{\mathbf{V}_+ e^{-\gamma(\ell - d)}} = \frac{\mathbf{V}_- e^{\gamma\ell}}{\mathbf{V}_+ e^{-\gamma\ell}} e^{-2\gamma d} = \Gamma_\ell e^{-2\gamma d} \tag{3-2-11}$$

Next, the reflection coefficient at that point can be expressed in terms of the reflection coefficient at the receiving end as

$$\Gamma_d = \Gamma_\ell e^{-2\alpha d} e^{-j2\beta d} = |\Gamma_\ell| e^{-2\alpha d} e^{j(\theta_\ell - 2\beta d)} \tag{3-2-12}$$

This is a very useful equation for determining the reflection coefficient at any point along the line. For a lossy line, both the magnitude and phase of the reflection coefficient are changing in an inward-spiral way as shown in Fig. 3-2-2. For a lossless line, $\alpha = 0$, the magnitude of the reflection coefficient remains constant, and only the phase of Γ is changing circularly toward the generator with an angle of $-2\beta d$ as shown in Fig. 3-2-3.

It is evident that Γ_ℓ will be zero and there will be no reflection from the receiving end when the terminating impedance is equal to the characteristic impedance

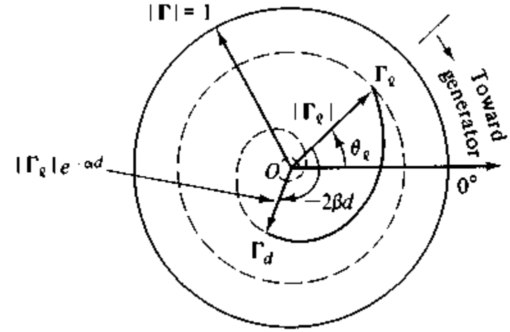

Figure 3-2-2 Reflection coefficient for lossy line.

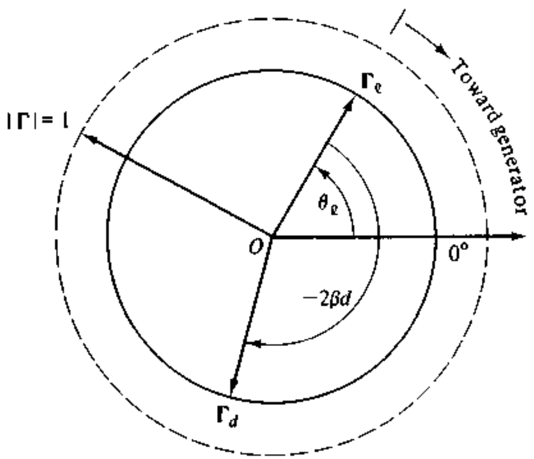

Figure 3-2-3 Reflection coefficient for lossless line.

of the line. Thus a terminating impedance that differs from the characteristic impedance will create a reflected wave traveling toward the source from the termination. The reflection, upon reaching the sending end, will itself be reflected if the source impedance is different from the line characteristic impedance at the sending end.

3-2-2 Transmission Coefficient

A transmission line terminated in its characteristic impedance $\mathbf{Z}_0$ is called a *properly terminated* line. Otherwise it is called an *improperly terminated* line. As described earlier, there is a reflection coefficient $\mathbf{\Gamma}$ at any point along an improperly terminated line. According to the principle of conservation of energy, the incident power minus the reflected power must be equal to the power transmitted to the load. This can be expressed as

$$1 - \mathbf{\Gamma}_\ell^2 = \frac{\mathbf{Z}_0}{\mathbf{Z}_\ell}\mathbf{T}^2 \qquad (3\text{-}2\text{-}13)$$

Equation (3-2-13) will be verified later. The letter **T** represents the transmission coefficient, which is defined as

$$\mathbf{T} = \frac{\text{transmitted voltage or current}}{\text{incident voltage or current}} = \frac{\mathbf{V}_{tr}}{\mathbf{V}_{inc}} = \frac{\mathbf{I}_{tr}}{\mathbf{I}_{inc}} \tag{3-2-14}$$

Figure 3-2-4 shows the transmission of power along a transmission line where P_{inc} is the incident power, P_{ref} the reflected power, and P_{tr} the transmitted power.

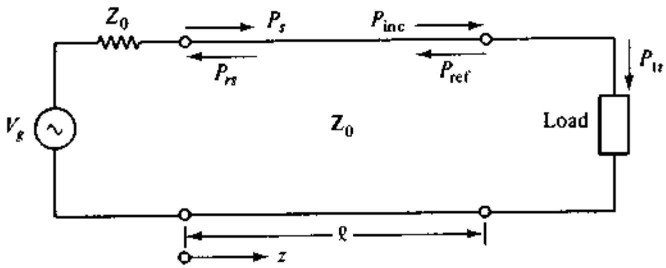

Figure 3-2-4 Power transmission on a line.

Let the traveling waves at the receiving end be

$$\mathbf{V}_+ e^{-\gamma\ell} + \mathbf{V}_- e^{\gamma\ell} = \mathbf{V}_{tr} e^{-\gamma\ell} \tag{3-2-15}$$

$$\frac{\mathbf{V}_+}{\mathbf{Z}_0} e^{-\gamma\ell} - \frac{\mathbf{V}_-}{\mathbf{Z}_0} e^{\gamma\ell} = \frac{\mathbf{V}_{tr}}{\mathbf{Z}_\ell} e^{-\gamma\ell} \tag{3-2-16}$$

Multiplication of Eq. (3-2-16) by $\mathbf{Z}_\ell$ and substitution of the result in Eq. (3-2-15) yield

$$\mathbf{\Gamma}_\ell = \frac{\mathbf{V}_- e^{\gamma\ell}}{\mathbf{V}_+ e^{-\gamma\ell}} = \frac{\mathbf{Z}_\ell - \mathbf{Z}_0}{\mathbf{Z}_\ell + \mathbf{Z}_0} \tag{3-2-17}$$

which, in turn, on substitution back into Eq. (3-2-15), results in

$$\mathbf{T} = \frac{\mathbf{V}_{tr}}{\mathbf{V}_+} = \frac{2\mathbf{Z}_\ell}{\mathbf{Z}_\ell + \mathbf{Z}_0} \tag{3-2-18}$$

The power carried by the two waves in the side of the incident and reflected waves is

$$P_{inr} = P_{inc} - P_{ref} = \frac{(\mathbf{V}_+ e^{-\alpha\ell})^2}{2\mathbf{Z}_0} - \frac{(\mathbf{V}_- e^{\alpha\ell})^2}{2\mathbf{Z}_0} \tag{3-2-19}$$

The power carried to the load by the transmitted waves is

$$P_{tr} = \frac{(\mathbf{V}_{tr} e^{-\alpha\ell})^2}{2\mathbf{Z}_\ell} \tag{3-2-20}$$

By setting $P_{inr} = P_{tr}$ and using Eqs. (3-2-17) and (3-2-18), we have

$$\mathbf{T}^2 = \frac{\mathbf{Z}_\ell}{\mathbf{Z}_0}(1 - \mathbf{\Gamma}_\ell^2) \tag{3-2-21}$$

This relation verifies the previous statement that the transmitted power is equal to the difference of the incident power and reflected power.

Example 3-2-1: Reflection Coefficient and Transmission Coefficient

A certain transmission line has a characteristic impedance of $75 + j0.01\ \Omega$ and is terminated in a load impedance of $70 + j50\ \Omega$. Compute **(a)** the reflection coefficient; **(b)** the transmission coefficient. Verify: **(c)** the relationship shown in Eq. (3-2-21); **(d)** the transmission coefficient equals the algebraic sum of 1 plus the reflection coefficient as shown in Eq. (2-3-18).

Solution

a. From Eq. (3-2-17) the reflection coefficient is

$$\mathbf{\Gamma} = \frac{\mathbf{Z}_\ell - \mathbf{Z}_0}{\mathbf{Z}_\ell + \mathbf{Z}_0} = \frac{70 + j50 - (75 + j0.01)}{70 + j50 + (75 + j0.01)}$$

$$= \frac{50.24\underline{/95.71^\circ}}{153.38\underline{/19.03^\circ}} = 0.33\underline{/76.68^\circ} = 0.08 + j0.32$$

b. From Eq. (3-2-18) the transmission coefficient is

$$\mathbf{T} = \frac{2\mathbf{Z}_\ell}{\mathbf{Z}_\ell + \mathbf{Z}_0} = \frac{2(70 + j50)}{70 + j50 + (75 + j0.01)}$$

$$= \frac{172.05\underline{/35.54^\circ}}{153.38\underline{/19.03^\circ}} = 1.12\underline{/16.51^\circ} = 1.08 + j0.32$$

c.

$$\mathbf{T}^2 = (1.12\underline{/16.51^\circ})^2 = 1.25\underline{/33.02^\circ}$$

$$\frac{\mathbf{Z}_\ell}{\mathbf{Z}_0}(1 - \mathbf{\Gamma}^2) = \frac{70 + j50}{75 + j0.01}[1 - (0.33)\underline{/76.68^\circ})^2]$$

$$= \frac{86\underline{/35.54^\circ}}{75\underline{/0^\circ}} \times 1.10\underline{/-2.6^\circ} = 1.25\underline{/33^\circ}$$

Thus Eq. (3-2-21) is verified.

d. From Eq. (2-3-18) we obtain

$$\mathbf{T} = 1.08 + j0.32 = 1 + 0.08 + j0.32 = 1 + \mathbf{\Gamma}$$

3-3 STANDING WAVE AND STANDING-WAVE RATIO

3-3-1 Standing Wave

The general solutions of the transmission-line equation consist of two waves traveling in opposite directions with unequal amplitude as shown in Eqs. (3-1-23) and (3-1-24). Equation (3-1-23) can be written

$$\begin{aligned}\mathbf{V} &= \mathbf{V}_+ e^{-\alpha z} e^{-j\beta z} + \mathbf{V}_- e^{\alpha z} e^{j\beta z} \\ &= \mathbf{V}_+ e^{-\alpha z}[\cos(\beta z) - j\sin(\beta z)] + \mathbf{V}_- e^{\alpha z}[\cos(\beta z) + j\sin(\beta z)] \qquad (3\text{-}3\text{-}1) \\ &= (\mathbf{V}_+ e^{-\alpha z} + \mathbf{V}_- e^{\alpha z})\cos(\beta z) - j(\mathbf{V}_+ e^{-\alpha z} - \mathbf{V}_- e^{\alpha z})\sin(\beta z)\end{aligned}$$

With no loss in generality it can be assumed that $\mathbf{V}_+e^{-\alpha z}$ and $\mathbf{V}_-e^{\alpha z}$ are real. Then the voltage-wave equation can be expressed as

$$\mathbf{V}_s = \mathbf{V}_0 e^{-j\phi} \tag{3-3-2}$$

This is called the *equation of the voltage standing wave*, where

$$\mathbf{V}_0 = [(\mathbf{V}_+e^{-\alpha z} + \mathbf{V}_-e^{\alpha z})^2 \cos^2(\beta z) + (\mathbf{V}_+e^{-\alpha z} - \mathbf{V}_-e^{-\alpha z})^2 \sin^2(\beta z)]^{1/2} \tag{3-3-3}$$

which is called the *standing-wave pattern* of the voltage wave or the amplitude of the standing wave, and

$$\phi = \arctan\left(\frac{\mathbf{V}_+e^{-\alpha z} - \mathbf{V}_-e^{\alpha z}}{\mathbf{V}_+e^{-\alpha z} + \mathbf{V}_-e^{\alpha z}} \tan(\beta z)\right) \tag{3-3-4}$$

which is called the *phase pattern of the standing wave*. The maximum and minimum values of Eq. (3-3-3) can be found as usual by differentiating the equation with respect to βz and equating the result to zero. By doing so and substituting the proper values of βz in the equation, we find that

1. The maximum amplitude is

$$\mathbf{V}_{max} = \mathbf{V}_+e^{-\alpha z} + \mathbf{V}_-e^{\alpha z} = \mathbf{V}_+e^{-\alpha z}(1 + |\Gamma|) \tag{3-3-5}$$

   and this occurs at $\beta z = n\pi$, where $n = 0, \pm 1, \pm 2, \ldots$.
2. The minimum amplitude is

$$\mathbf{V}_{min} = \mathbf{V}_+e^{-\alpha z} - V_-e^{\alpha z} = \mathbf{V}_+e^{-\alpha z}(1 - |\Gamma|) \tag{3-3-6}$$

   and this occurs at $\beta z = (2n - 1)\pi/2$, where $n = 0, \pm 1, \pm 2, \ldots$.
3. The distance between any two successive maxima or minima is one-half wavelength, since

$$\beta z = n\pi \qquad z = \frac{n\pi}{\beta} = \frac{n\pi}{2\pi/\lambda} = n\frac{\lambda}{2} \qquad (n = 0, \pm 1, \pm 2, \ldots)$$

Then

$$z_1 = \frac{\lambda}{2} \tag{3-3-7}$$

It is evident that there are no zeros in the minimum. Similarly,

$$\mathbf{I}_{max} = \mathbf{I}_+e^{-\alpha z} + \mathbf{I}_-e^{\alpha z} = \mathbf{I}_+e^{-\alpha z}(1 + |\Gamma|) \tag{3-3-8}$$

$$\mathbf{I}_{min} = \mathbf{I}_+e^{-\alpha z} - \mathbf{I}_-e^{\alpha z} = \mathbf{I}_+e^{-\alpha z}(1 - |\Gamma|) \tag{3-3-9}$$

The standing-wave patterns of two oppositely traveling waves with unequal amplitude in lossy or lossless line are shown in Figs. 3-3-1 and 3-3-2.

A further study of Eq. (3-3-3) reveals that

1. When $\mathbf{V}_+ \neq 0$ and $\mathbf{V}_- = 0$, the standing-wave pattern becomes

$$\mathbf{V}_0 = \mathbf{V}_+e^{-\alpha z} \tag{3-3-10}$$

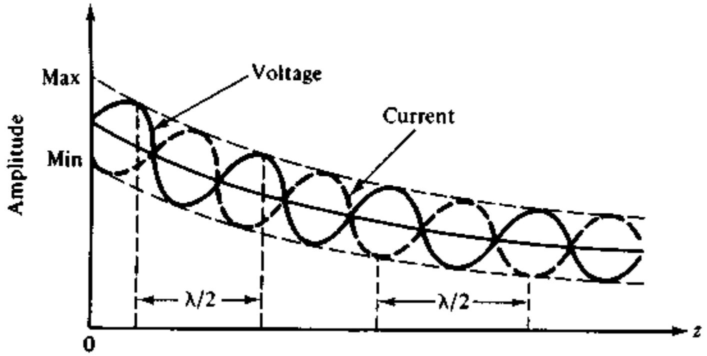

Figure 3-3-1 Standing-wave pattern in a lossy line.

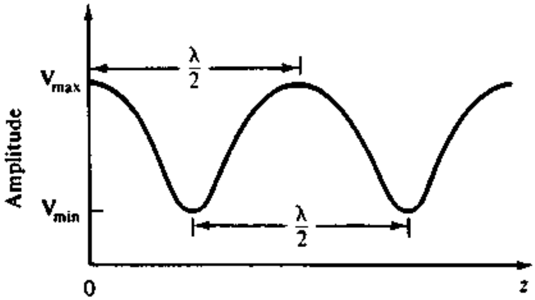

Figure 3-3-2 Voltage standing-wave pattern in a lossless line.

2. When $\mathbf{V}_+ = 0$ and $\mathbf{V}_- \neq 0$, the standing-wave pattern becomes

$$\mathbf{V}_0 = \mathbf{V}_- e^{\alpha z} \tag{3-3-11}$$

3. When the positive wave and the negative wave have equal amplitudes (that is, $|\mathbf{V}_+ e^{-\alpha z}| = |\mathbf{V}_- e^{\alpha z}|$) or the magnitude of the reflection coefficient is unity, the standing-wave pattern with a zero phase is given by

$$\mathbf{V}_s = 2\mathbf{V}_+ e^{-\alpha z} \cos (\beta z) \tag{3-3-12}$$

which is called a *pure standing wave*.

Similarly, the equation of a pure standing wave for the current is

$$\mathbf{I}_s = -j2\mathbf{Y}_0 \mathbf{V}_+ e^{-\alpha z} \sin (\beta z) \tag{3-3-13}$$

Equations (3-3-12) and (3-3-13) show that the voltage and current standing waves are 90° out of phase along the line. The points of zero current are called the *current nodes*. The voltage nodes and current nodes are interlaced a quarter wavelength apart.

The voltage and current may be expressed as real functions of time and space:

$$v_s = (z, t) = \text{Re}[\mathbf{V}_s(z)e^{j\omega t}] = 2\mathbf{V}_+ e^{-\alpha z} \cos (\beta z) \cos (\omega t) \tag{3-3-14}$$

$$i_s = (z, t) = \text{Re}[\mathbf{I}_s(z)e^{j\omega t}] = 2\mathbf{Y}_0 \mathbf{V}_+ e^{-\alpha z} \sin (\beta z) \sin (\omega t) \tag{3-3-15}$$

The amplitudes of Eqs. (3-3-14) and (3-3-15) vary sinusoidally with time; the voltage is a maximum at the instant when the current is zero and vice versa. Figure 3-3-3 shows the pure-standing-wave patterns of the phasor of Eqs. (3-3-12) and (3-3-13) for an open-terminal line.

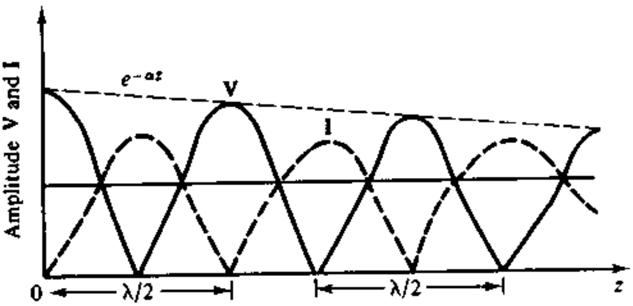

Figure 3-3-3 Pure standing waves of voltage and current.

3-3-2 Standing-Wave Ratio

Standing waves result from the simultaneous presence of waves traveling in opposite directions on a transmission line. The ratio of the maximum of the standing-wave pattern to the minimum is defined as the standing-wave ratio, designated by ρ. That is,

$$\text{Standing-wave ratio} \equiv \frac{\text{maximum voltage or current}}{\text{minimum voltage or current}}$$

$$\rho \equiv \frac{|\mathbf{V}_{\max}|}{|\mathbf{V}_{\min}|} = \frac{|\mathbf{I}_{\max}|}{|\mathbf{I}_{\min}|} \tag{3-3-16}$$

The standing-wave ratio results from the fact that the two traveling-wave components of Eq. (3-3-1) add in phase at some points and subtract at other points. The distance between two successive maxima or minima is $\lambda/2$. The standing-wave ratio of a pure traveling wave is unity and that of a pure standing wave is infinite. It should be noted that since the standing-wave ratios of voltage and current are identical, no distinctions are made between VSWR and ISWR.

When the standing-wave ratio is unity, there is no reflected wave and the line is called a *flat line*. The standing-wave ratio cannot be defined on a lossy line because the standing-wave pattern changes markedly from one position to another. On a low-loss line the ratio remains fairly constant, and it may be defined for some region. For a lossless line, the ratio stays the same throughout the line.

Since the reflected wave is defined as the product of an incident wave and its reflection coefficient, the standing-wave ratio ρ is related to the reflection coefficient $\mathbf{\Gamma}$ by

$$\rho = \frac{1 + |\mathbf{\Gamma}|}{1 - |\mathbf{\Gamma}|} \tag{3-3-17}$$

and vice versa

$$|\Gamma| = \frac{\rho - 1}{\rho + 1} \tag{3-3-18}$$

This relation is very useful for determining the reflection coefficient from the standing-wave ratio, which is usually found from the Smith chart. The curve in Fig. 3-3-4 shows the relationship between reflection coefficient $|\Gamma|$ and standing-wave ratio ρ.

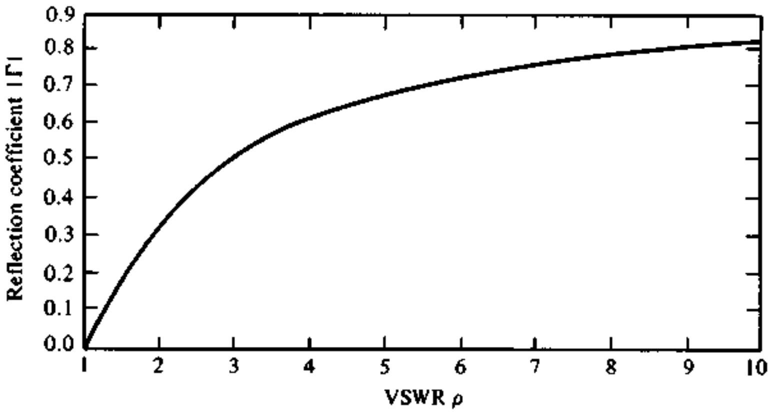

Figure 3-3-4 SWR versus reflection coefficient.

As a result of Eq. (3-3-17), since $|\Gamma| \leq 1$, the standing-wave ratio is a positive real number and never less than unity, $\rho \geq 1$. From Eq. (3-3-18) the magnitude of the reflection coefficient is never greater than unity.

Example 3-3-1: Standing-Wave Ratio

A transmission line has a characteristic impedance of $50 + j0.01\ \Omega$ and is terminated in a load impedance of $73 - j42.5\ \Omega$. Calculate: **(a)** the reflection coefficient; **(b)** the standing-wave ratio.

Solution

a. From Eq. (3-2-8) the reflection coefficient is

$$\mathbf{\Gamma} = \frac{\mathbf{Z}_\ell - \mathbf{Z}_0}{\mathbf{Z}_\ell + \mathbf{Z}_0} = \frac{73 - j42.5 - (50 + j0.01)}{73 - j42.5 + (50 + j0.01)} = 0.377\underline{/-42.7^\circ}$$

b. From Eq. (3-3-17) the standing-wave ratio is

$$\rho = \frac{1 + |\Gamma|}{1 - |\Gamma|} = \frac{1 + 0.377}{1 - 0.377} = 2.21$$

3-5 SMITH CHART

Many of the computations required to solve transmission-line problems involve the use of rather complicated equations. The solution of such problems is tedious and difficult because the accurate manipulation of numerous equations is necessary. To simplify their solution, we need a graphic method of arriving at a quick answer.

A number of impedance charts have been designed to facilitate the graphic solution of transmission-line problems. Basically all the charts are derived from the fundamental relationships expressed in the transmission equations. The most popular chart is that developed by Phillip H. Smith [1]. The purpose of this section is to present the graphic solutions of transmission-line problems by using the Smith chart.

The Smith chart consists of a plot of the normalized impedance or admittance with the angle and magnitude of a generalized complex reflection coefficient in a unity circle. The chart is applicable to the analysis of a lossless line as well as a lossy line. By simple rotation of the chart, the effect of the position on the line can be determined. To see how a Smith chart works, consider the equation of reflection coefficient at the load for a transmission line as shown in Eq. (3-2-8):

$$\boldsymbol{\Gamma}_\ell = \frac{\mathbf{Z}_\ell - \mathbf{Z}_0}{\mathbf{Z}_\ell + \mathbf{Z}_0} = |\boldsymbol{\Gamma}_\ell| e^{j\theta_\ell} = \Gamma_\gamma + j\Gamma_i \tag{3-5-1}$$

Since $|\Gamma_\ell| \le 1$, the value of Γ_ℓ must lie on or within the unity circle with a radius of 1. The reflection coefficient at any other location along a line as shown in Eq. (3-2-12) is

$$\boldsymbol{\Gamma}_d = \boldsymbol{\Gamma}_\ell e^{-2\alpha d} e^{-j2\beta d} = |\boldsymbol{\Gamma}_\ell| e^{-2\alpha d} e^{j(\theta_\ell - 2\beta d)} \tag{3-5-2}$$

which is also on or within the unity circle. Figure 3-5-1 shows circles for a constant reflection coefficient $\boldsymbol{\Gamma}$ and constant electrical-length radials βd.

From Eqs. (3-4-29) and (3-4-44) the normalized impedance along a line is given by

$$\mathbf{z} = \frac{\mathbf{Z}}{\mathbf{Z}_0} = \frac{1 + \boldsymbol{\Gamma}_\ell e^{-2\gamma d}}{1 - \boldsymbol{\Gamma}_\ell e^{-2\gamma d}} \tag{3-5-3}$$

With no loss in generality, it is assumed that $d = 0$; then

$$\mathbf{z} = \frac{1 + \boldsymbol{\Gamma}_\ell}{1 - \boldsymbol{\Gamma}_\ell} = \frac{\mathbf{Z}_\ell}{\mathbf{Z}_0} = \frac{R_\ell + jX_\ell}{\mathbf{Z}_0} = r + jx \tag{3-5-4}$$

and

$$\boldsymbol{\Gamma}_\ell = \frac{\mathbf{z} - 1}{\mathbf{z} + 1} = \Gamma_r + j\Gamma_i \tag{3-5-5}$$

Substitution of Eq. (3-5-5) into Eq. (3-5-4) yields

$$r = \frac{1 - \Gamma_r^2 - \Gamma_i^2}{(1 - \Gamma_r)^2 + \Gamma_i^2} \tag{3-5-6}$$

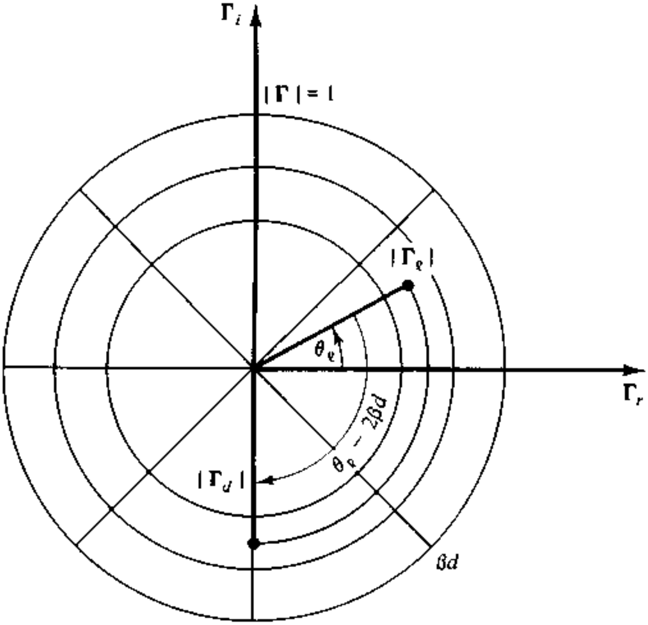

Figure 3-5-1 Constant Γ circles and electrical-length radials βd.

and

$$x = \frac{2\Gamma_i}{(1 - \Gamma_r)^2 + \Gamma_i^2} \tag{3-5-7}$$

Equations (3-5-6) and (3-5-7) can be rearranged as

$$\left(\Gamma_r - \frac{r}{1 + r}\right)^2 + \Gamma_i^2 = \left(\frac{1}{1 + r}\right)^2 \tag{3-5-8}$$

and

$$(\Gamma_r - 1)^2 + \left(\Gamma_i - \frac{1}{x}\right)^2 = \left(\frac{1}{x}\right)^2 \tag{3-5-9}$$

Equation (3-5-8) represents a family of circles in which each circle has a constant resistance r. The radius of any circle is $1/(1 + r)$, and the center of any circle is $r/(1 + r)$ along the real axis in the unity circle, where r varies from zero to infinity. All constant resistance circles are plotted in Fig. 3-5-2 according to Eq. (3-5-8).

Equation (3-5-9) also describes a family of circles, but each of these circles specifies a constant reactance x. The radius of any circle is $(1/x)$, and the center of any circle is at

$$\Gamma_r = 1 \qquad \Gamma_i = \frac{1}{x} \qquad (\text{where } -\infty \le x \le \infty)$$

All constant reactance circles are plotted in Fig. 3-5-3 according to Eq. (3-5-9).

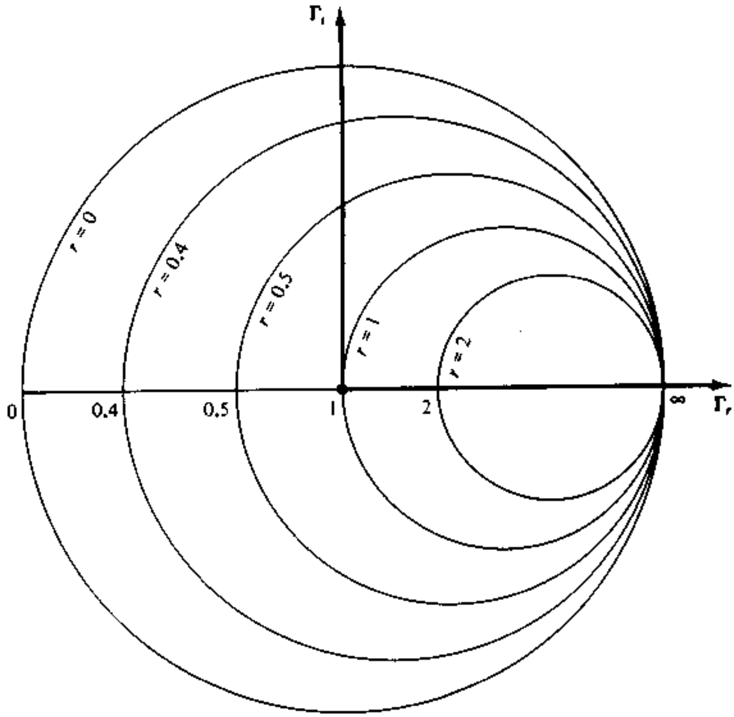

Figure 3-5-2 Constant resistance r circles.

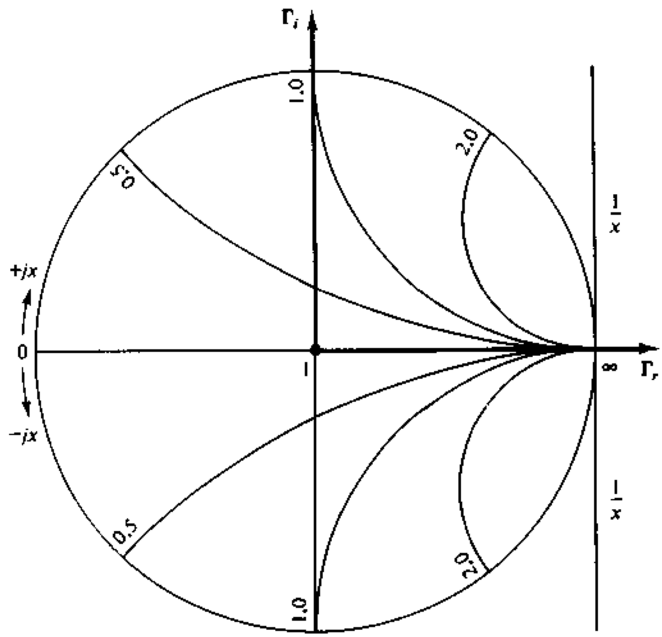

Figure 3-5-3 Constant reactance x circles.

There are relative distance scales in wavelength along the circumference of the Smith chart. Also, there is a phase scale specifying the angle of the reflection coefficient. When a normalized impedance **z** is located on the chart, the normalized impedance of any other location along the line can be found by use of Eq. (3-5-3):

$$\mathbf{z} = \frac{1 + \Gamma_\ell e^{-2\gamma d}}{1 - \Gamma_\ell e^{-2\gamma d}} \tag{3-5-10}$$

where

$$\Gamma_\ell e^{-2\gamma d} = |\Gamma_\ell| e^{-2\alpha d} e^{j(\theta_\ell - 2\beta d)} \tag{3-5-11}$$

The Smith chart may also be used for normalized admittance. This is evident since

$$\mathbf{Y}_0 = \frac{1}{\mathbf{Z}_0} = G_0 + jB_0 \quad \text{and} \quad \mathbf{Y} = \frac{1}{\mathbf{Z}} = G + jB \tag{3-5-12}$$

Then the normalized admittance is

$$\mathbf{y} = \frac{\mathbf{Y}}{\mathbf{Y}_0} = \frac{\mathbf{Z}_0}{\mathbf{Z}} = \frac{1}{\mathbf{z}} = g + jb \tag{3-5-13}$$

Figure 3-5-4 shows a Smith chart which superimposes Figs. 3-5-2 and 3-5-3 into one chart. The characteristics of the Smith chart are summarized as follows:

1. The constant r and constant x loci form two families of orthogonal circles in the chart.
2. The constant r and constant x circles all pass through the point ($\Gamma_r = 1$, $\Gamma_i = 0$).
3. The upper half of the diagram represents $+jx$.
4. The lower half of the diagram represents $-jx$.
5. For admittance the constant r circles become constant g circles, and the constant x circles become constant susceptance b circles.
6. The distance around the Smith chart once is one-half wavelength ($\lambda/2$).
7. At a point of $z_{min} = 1/\rho$, there is a V_{min} on the line.
8. At a point of $z_{max} = \rho$, there is a V_{max} on the line.
9. The horizontal radius to the right of the chart center corresponds to V_{max}, I_{min}, z_{max}, and ρ (SWR).
10. The horizontal radius to the left of the chart center corresponds to V_{min}, I_{max}, z_{min}, and $1/\rho$.
11. Since the normalized admittance **y** is a reciprocal of the normalized impedance **z**, the corresponding quantities in the admittance chart are 180° out of phase with those in the impedance chart.
12. The normalized impedance or admittance is repeated for every half wavelength of distance.
13. The distances are given in wavelengths toward the generator and also toward the load.

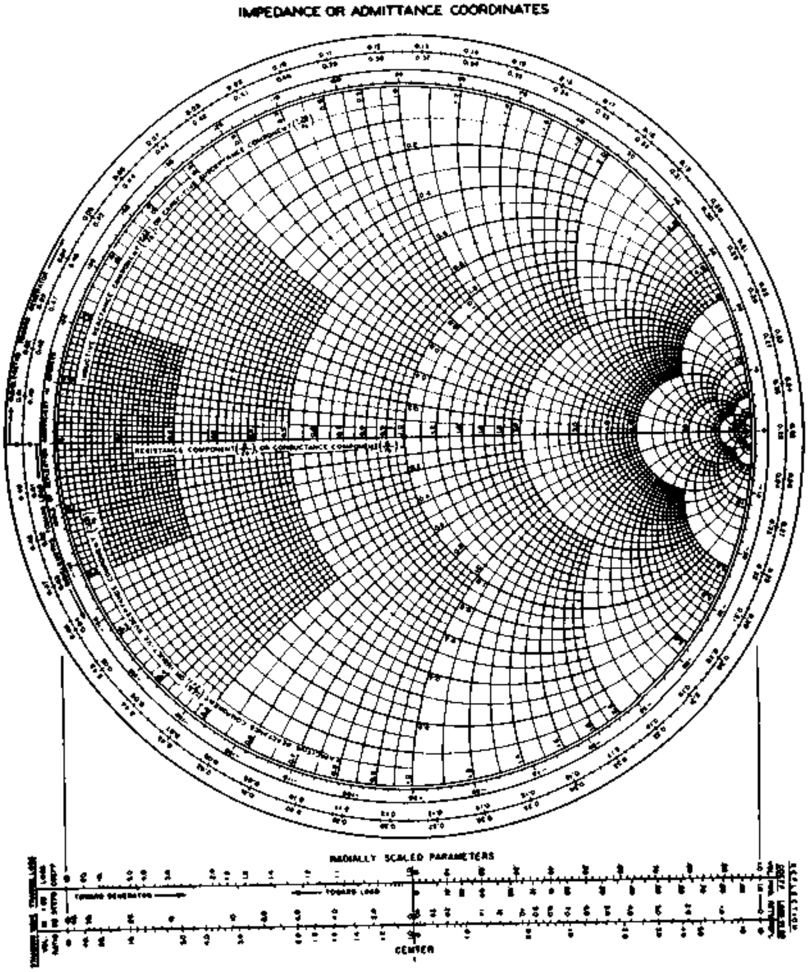

Figure 3-5-4 Smith chart.

The magnitude of the reflection coefficient is related to the standing-wave ratio by the following expression:

$$|\Gamma| = \frac{\rho - 1}{\rho + 1} \tag{3-5-14}$$

A Smith chart or slotted line can be used to measure a standing-wave pattern directly and then the magnitudes of the reflection coefficient, reflected power, transmitted power, and the load impedance can be calculated from it. The use of the Smith chart is illustrated in the following examples.

Example 3-5-1: Location Determination of Voltage Maximum and Minimum from Load

Given the normalized load impedance $z_\ell = 1 + j1$ and the operating wavelength

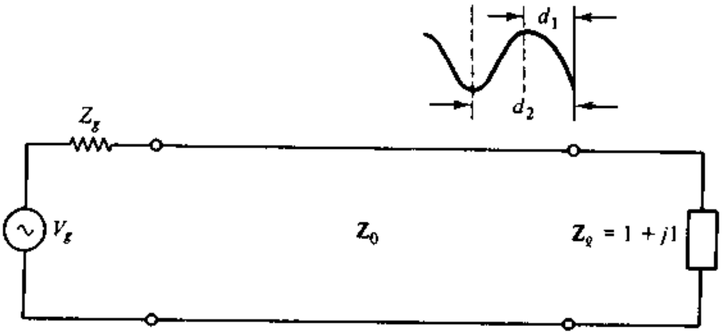

Figure 3-5-5 Diagram for Example 3-5-1.

$\lambda = 5$ cm, determine the first V_{max}, first V_{min} from the load, and the VSWR ρ as shown in Fig. 3-5-5.

Solution

1. Enter $z_\ell = 1 + j1$ on the chart as shown in Fig. 3-5-6.
2. Read 0.162λ on the distance scale by drawing a dashed-straight line from the center of the chart through the load point and intersecting the distance scale.
3. Move a distance from the point at 0.162λ toward the generator and first stop at the voltage maximum on the right-hand real axis at 0.25λ. Then

$$d_1(V_{max}) = (0.25 - 0.162)\lambda = (0.088)(5) = 0.44 \text{ cm}$$

4. Similarly, move a distance from the point of 0.162λ toward the generator and first stop at the voltage minimum on the left-hand real axis at 0.5λ. Then

$$d_2(V_{min}) = (0.5 - 0.162)\lambda = (0.338)(5) = 1.69 \text{ cm}$$

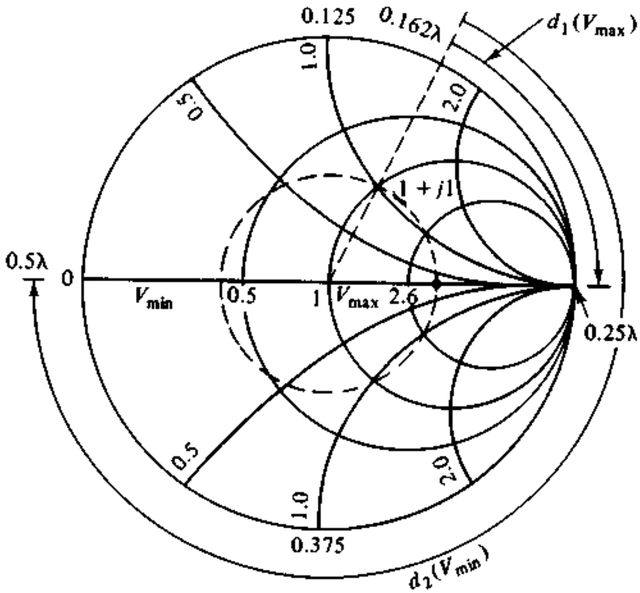

Figure 3-5-6 Graphic solution for Example 3-5-1.

5. Make a standing-wave circle with the center at (1, 0) and pass the circle through the point of $1 + j1$. The location intersected by the circle at the right portion of the real axis indicates the SWR. This is $\rho = 2.6$.

Example 3-5-2: Impedance Determination with Short-Circuit Minima Shift

The location of a minimum instead of a maximum is usually specified because it can be determined more accurately. Suppose that the characteristic impedance of the line R_0 is 50 Ω, and the SWR $\rho = 2$ when the line is loaded. When the load is shorted, the minima shift 0.15λ toward the load. Determine the load impedance. Figure 3-5-7 shows the diagram for the example.

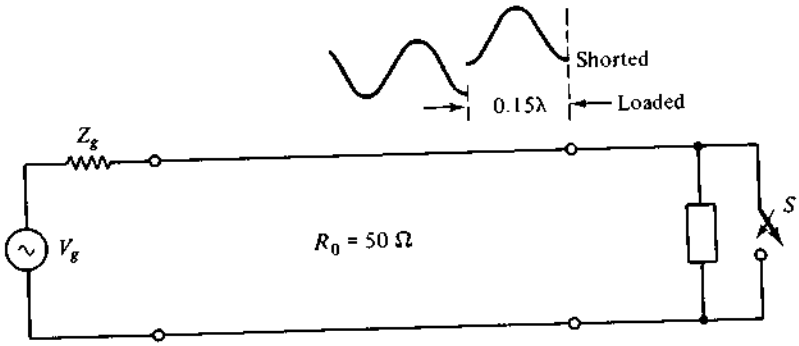

Figure 3-5-7 Diagram for Example 3-5-2.

Solution

1. When the line is shorted, the first voltage minimum occurs at the place of the load as shown in Fig. 3-5-8.

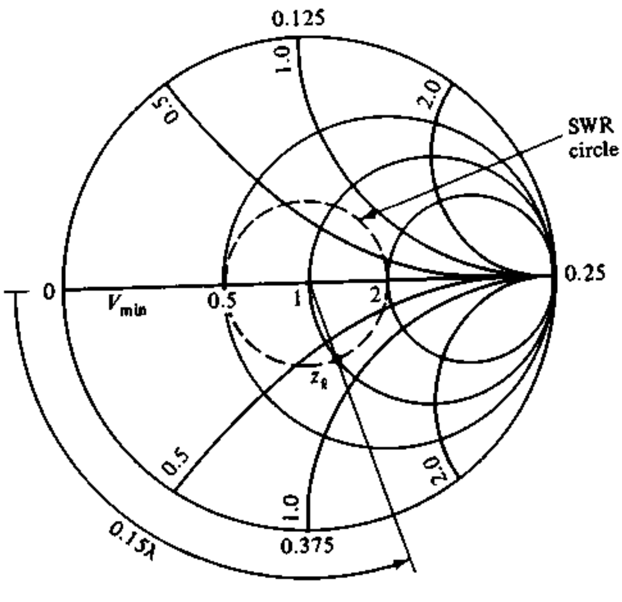

Figure 3-5-8 Graphic solution for Example 3-5-2.

2. When the line is loaded, the first voltage minimum shifts 0.15λ from the load. The distance between two successive minima is one-half wavelength.
3. Plot a SWR circle for $\rho = 2$.
4. Move a distance of 0.15λ from the minimum point along the distance scale toward the load and stop at 0.15λ.
5. Draw a line from this point to the center of the chart.
6. The intersection between the line and the SWR circle is

$$z_\ell = 1 - j0.65$$

7. The load impedance is

$$\mathbf{Z}_\ell = (1 - j0.65)(50) = 50 - j32.5\ \Omega$$

3-6 IMPEDANCE MATCHING

Impedance matching is very desirable with radio frequency (RF) transmission lines. Standing waves lead to increased losses and frequently cause the transmitter to malfunction. A line terminated in its characteristic impedance has a standing-wave ratio of unity and transmits a given power without reflection. Also, transmission efficiency is optimum where there is no reflected power. A "flat" line is nonresonant; that is, its input impedance always remains at the same value $\mathbf{Z}_0$ when the frequency changes.

Matching a transmission line has a special meaning, one differing from that used in circuit theory to indicate equal impedance seen looking both directions from a given terminal pair for maximum power transfer. In circuit theory, maximum power transfer requires the load impedance to be equal to the complex conjugate of the generator. This condition is sometimes referred to as a *conjugate match*. In transmission-line problems *matching* means simply terminating the line in its characteristic impedance.

A common application of RF transmission lines is the one in which there is a feeder connection between a transmitter and an antenna. Usually the input impedance to the antenna itself is not equal to the characteristic impedance of the line. Furthermore, the output impedance of the transmitter may not be equal to the $\mathbf{Z}_0$ of the line. Matching devices are necessary to flatten the line. A complete matched transmission-line system is shown in Fig. 3-6-1.

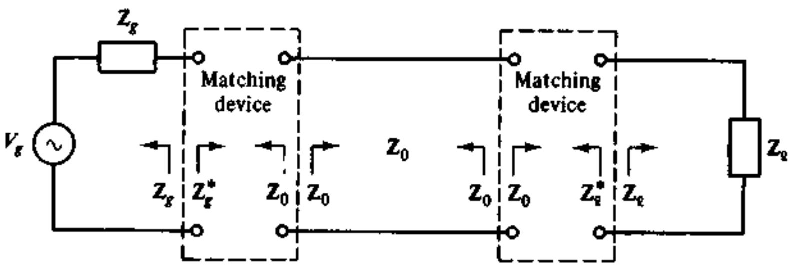

Figure 3-6-1 Matched transmission-line system.

For a low-loss or lossless transmission line at radio frequency, the characteristic impedance $\mathbf{Z}_0$ of the line is resistive. At every point the impedances looking in opposite directions are conjugate. If $\mathbf{Z}_0$ is real, it is its own conjugate. Matching can be tried first on the load side to flatten the line; then adjustment may be made on the transmitter side to provide maximum power transfer. At audio frequencies an iron-cored transformer is almost universally used as an impedance-matching device. Occasionally an iron-cored transformer is also used at radio frequencies. In a practical transmission-line system, the transmitter is ordinarily matched to the coaxial cable for maximum power transfer. Because of the variable loads, however, an impedance-matching technique is often required at the load side.

Since the matching problems involve parallel connections on the transmission line, it is necessary to work out the problems with admittances rather than impedances. The Smith chart itself can be used as a computer to convert the normalized impedance to admittance by a rotation of 180°, as described earlier.

3-6-1 Single-Stub Matching

Although single-lumped inductors or capacitors can match the transmission line, it is more common to use the susceptive properties of short-circuited sections of transmission lines. Short-circuited sections are preferable to open-circuited ones because a good short circuit is easier to obtain than a good open circuit.

For a lossless line with $\mathbf{Y}_g = \mathbf{Y}_0$, maximum power transfer requires $\mathbf{Y}_{11} = \mathbf{Y}_0$, where $\mathbf{Y}_{11}$ is the total admittance of the line and stub looking to the right at point 1-1 (see Fig. 3-6-2). The stub must be located at that point on the line where the real part of the admittance, looking toward the load, is $\mathbf{Y}_0$. In a normalized unit y_{11} must be in the form

$$y_{11} = y_d \pm y_s = 1 \tag{3-6-1}$$

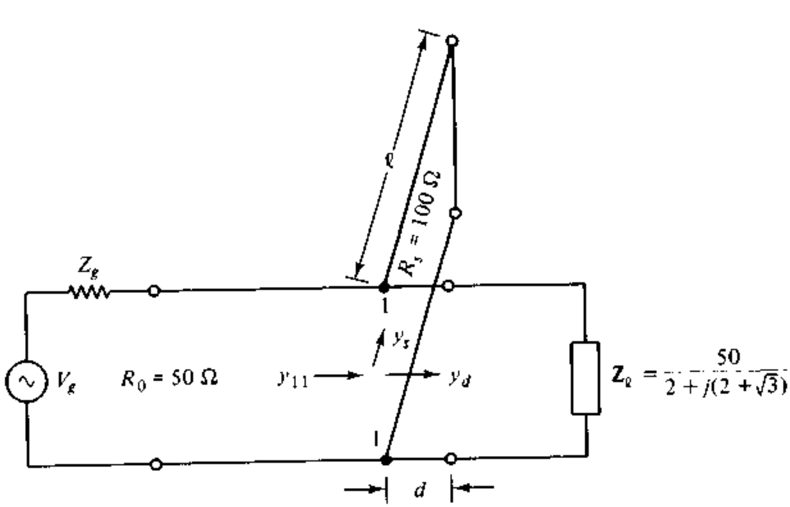

Figure 3-6-2 Single-stub matching for Example 3-6-1.

if the stub has the same characteristic impedance as that of the line. Otherwise

$$\mathbf{Y}_{11} = \mathbf{Y}_d \pm \mathbf{Y}_s = \mathbf{Y}_0 \tag{3-6-2}$$

The stub length is then adjusted so that its susceptance just cancels out the susceptance of the line at the junction.

Example 3-6-1: Single-Stub Matching

A lossless line of characteristic impedance $R_0 = 50\ \Omega$ is to be matched to a load $\mathbf{Z}_\ell = 50/[2 + j(2 + \sqrt{3})]\ \Omega$ by means of a lossless short-circuited stub. The characteristic impedance of the stub is 100 Ω. Find the stub position (closest to the load) and length so that a match is obtained.

Solution

1. Compute the normalized load admittance and enter it on the Smith chart (see Fig. 3-6-3).

$$y_\ell = \frac{1}{z_\ell} = \frac{R_0}{\mathbf{Z}_\ell} = 2 + j(2 + \sqrt{3}) = 2 + j3.732$$

2. Draw a SWR circle through the point of y_ℓ so that the circle intersects the unity circle at the point y_d.

$$y_d = 1 - j2.6$$

Note that there are an infinite number of y_d. Take the one that allows the stub to be attached as closely as possible to the load.

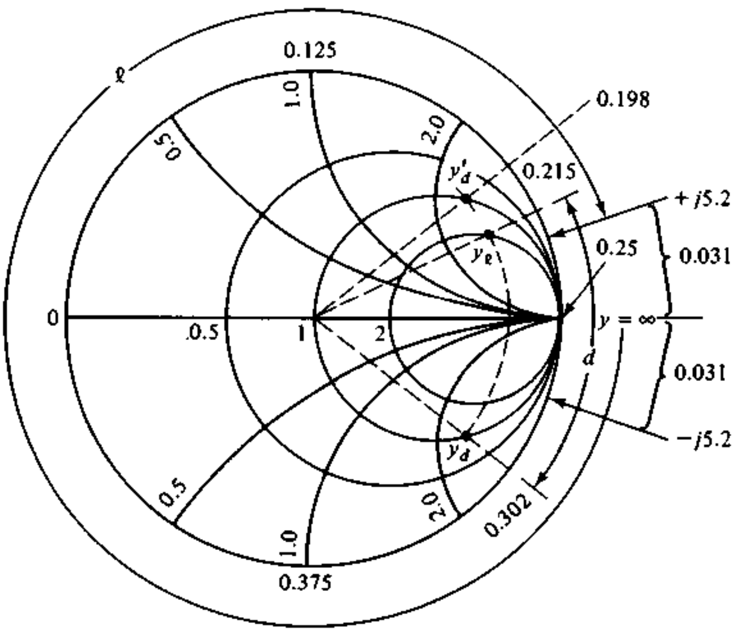

Figure 3-6-3 Graphic solution for Example 3-6-1.

3. Since the characteristic impedance of the stub is different from that of the line, the condition for impedance matching at the junction requires

$$\mathbf{Y}_{11} = \mathbf{Y}_d + \mathbf{Y}_s$$

where $\mathbf{Y}_s$ is the susceptance that the stub will contribute.

It is clear that the stub and the portion of the line from the load to the junction are in parallel, as seen by the main line extending to the generator. The admittances must be converted to normalized values for matching on the Smith chart. Then Eq. (3-6-2) becomes

$$y_{11}\mathbf{Y}_0 = y_d\mathbf{Y}_0 + y_s\mathbf{Y}_{0s}$$

$$y_s = (y_{11} - y_d)\left(\frac{\mathbf{Y}_0}{\mathbf{Y}_{0s}}\right) = [1 - (1 - j2.6)]\frac{100}{50} = +j5.20$$

4. The distance between the load and the stub position can be calculated from the distance scale as

$$d = (0.302 - 0.215)\lambda = 0.087\lambda$$

5. Since the stub contributes a susceptance of $+j5.20$, enter $+j5.20$ on the chart and determine the required distance ℓ from the short-circuited end ($z = 0$, $y = \infty$), which corresponds to the right side of the real axis on the chart, by transversing the chart toward the generator until the point of $+j5.20$ is reached. Then

$$\ell = (0.50 - 0.031)\lambda = 0.469\lambda$$

When a line is matched at the junction, there will be no standing wave in the line from the stub to the generator.

6. If an inductive stub is required,

$$y'_d = 1 + j2.6$$

the susceptance of the stub will be

$$y'_s = -j5.2$$

7. The position of the stub from the load is

$$d' = [0.50 - (0.215 - 0.198)]\lambda = 0.483\lambda$$

and the length of the short-circuited stub is

$$\ell' = 0.031\lambda$$

3-6-2 Double-Stub Matching

Since single-stub matching is sometimes impractical because the stub cannot be placed physically in the ideal location, double-stub matching is needed. Double-stub devices consist of two short-circuited stubs connected in parallel with a fixed length between them. The length of the fixed section is usually one-eighth, three-eighths, or five-eighths of a wavelength. The stub that is nearest the load is used to adjust the susceptance and is located at a fixed wavelength from the constant conductance

Question Bank

1. Write a short note on TEDs
2. Describe the mechanism of Gunn effect.
3. Describe the drift velocity and current fluctuation in n-type Ga-As.
4. Write a note on microwave frequencies
5. Write a short note on microwave devices and microwave systems
6. Derive the general transmission line equation to find voltage and current on the line in terms of position 'z' and time 't '
7. From the solutions of transmission line equations derive the equation for characteristic impedance and prorogation constant
8. A transmission line has the following parameters: R=2Ω/m, G=0.5m mho/m, f=1GHz. L=8nH/m, C=0.23pF/m. Calculate i) Characteristic Impedance, ii) Propagation Constant
9. Define reflection coefficient. Derive the equation for reflection coefficient at load end at a distance 'd' from the load
10. Define transmission coefficient. Derive the equation for transmission coefficient and also deduce its relation with reflection coefficient.
11. A certain transmission line has a characteristic impedance of 75+ j0.01Ω and is terminated with load impedance of 70+j50 Ω. Compute i) Reflection coefficient ii) Transmission coefficient iii) prove Relationship between reflection coefficient and transmission coefficient.
12. What is standing wave? Write necessary equation to describe the standing wave
13. Write a note on standing wave ratio
14. A transmission line has a characteristic impedance of 50+j0.01Ω and terminated in a load impedance of 73-j42.5Ω. Calculate i) Reflection coefficient ii) SWR