Principles of RF and Microwave Measurements
(Lecture Notes for ECEN 4634)

by
Zoya Popović and Edward F. Kuester

Electromagnetics Laboratory
Department of Electrical and Computer Engineering
Campus Box 425
University of Colorado
Boulder, Colorado 80309-0425

Contents

Lectures

Labs

Chapter 1

Introduction and Review

1.1 What Are Microwaves?

The word *microwaves* refers to AC signals that have frequencies between 0.3 and 300 GHz. To find the wavelength of a microwave signal, it is convenient to use the following expression:

$$\lambda_{(\text{in cm})} = \frac{30}{f_{(\text{in GHz})}}.$$

According to this formula, signals above 30 GHz have wavelengths on the order of millimeters, and are called *millimeter waves*. The microwave frequency region is divided into bands, as shown in Table 1.1. It should be noted that band designations have never been precisely standardized, and can vary from one source to the next.

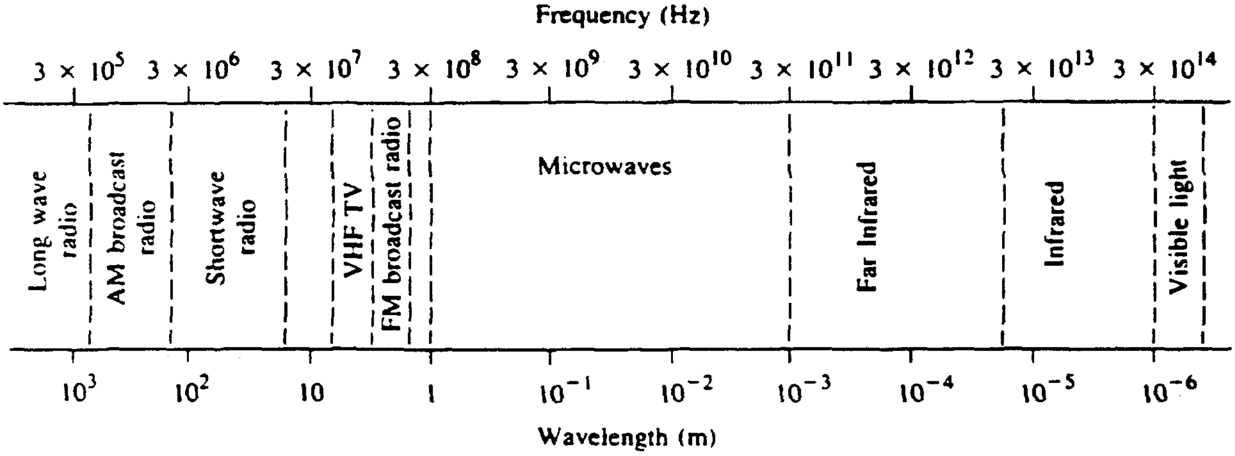
Microwave networks are harder to analyze than their low-frequency counterparts. The reason is that a microwave circuit is about the same size as the wavelength, so phase variation along a part of a circuit cannot be ignored as is the case at lower frequencies. In other words, Kirchhoff's laws do not apply, since they assume that the circuit is much smaller than a wavelength. On the other hand, in optics, everything is many thousands of wavelengths large, and rays and geometrical optics approximations can be used. The microwave region is the trickiest one to deal with mathematically. Rigorous analysis uses electromagnetic field theory, starting from Maxwell's equations, and is very complicated in most practical cases. Fortunately, we do not need all the information that such a field theory analysis gives us, and in many cases *transmission-line theory* is applicable.

Where are microwaves used?

1. ANTENNAS — the gain of an antenna is proportional to its size measured in wavelengths. (When the dimensions of an object are measured in wavelengths, we call this the *electrical size* of the object.) This means that, for a given gain (focusing capability), microwave antennas are small compared to lower frequency antennas.

2. COMMUNICATION — at higher frequencies, there is more freactional bandwidth available. For example, an analog TV channel takes up 6 MHz. At 600 MHz, a 1% bandwidth can accommodate only one TV channel, while at 60 GHz a 1% bandwidth covers 100 TV channels. (A 10-MHz digital TV channel requires even more bandwidth.)

3. SATELLITES — microwave signals travel by line-of-sight and are not bent by the ionosphere, unlike lower frequency signals. This makes communication links via satellite possible. Millimeter-wave frequencies, however, can be highly attenuated by the atmosphere, which makes them suitable for applications such as communications between satellites in which case interference with ground



Typical Frequencies	
AM broadcast band	535–1605 kHz
Shortwave radio	3–30 MHz
FM broadcast band	88–108 MHz
VHF TV (2–4)	54–72 MHz
VHF TV (5–6)	76–88 MHz
UHF TV (7–13)	174–216 MHz
UHF TV (14–83)	470–890 MHz
Microwave ovens	2.45 GHz

Approximate Band Designations	
L-band	1–2 GHz
S-band	2–4 GHz
C-band	4–8 GHz
X-band	8–12 GHz
Ku-band	12–18 GHz
K-band	18–26 GHz
Ka-band	26–40 GHz
U-band	40–60 GHz

Table 1.1: RF and microwave frequency bands.

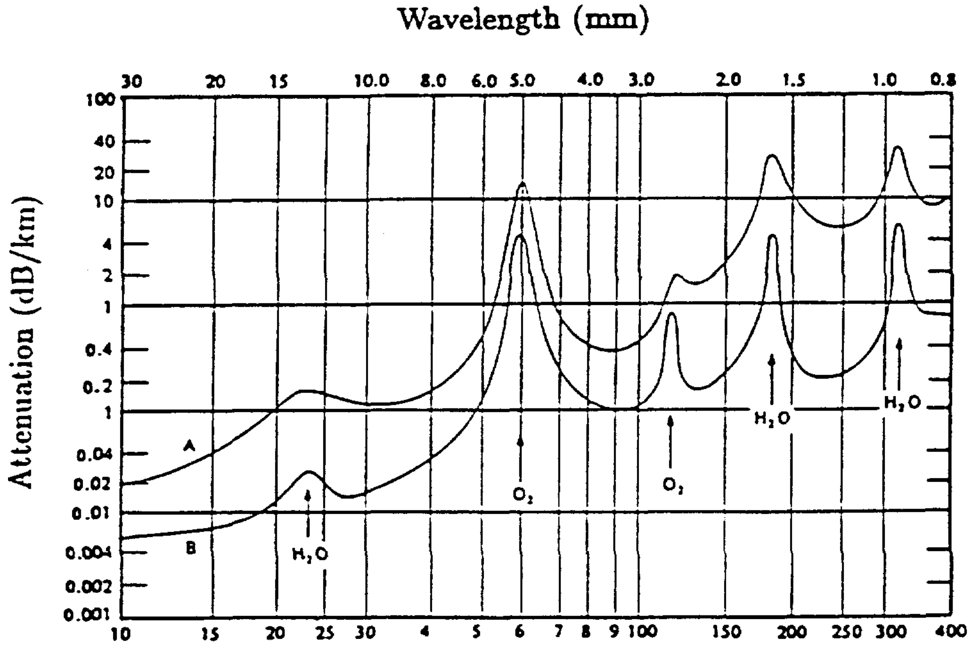


Figure 1.1: Attenuation of the atmosphere at sea level and 4 km altitude at microwave and millimeter-wave frequencies.

transmitters is avoided. Fig. 1.1 shows attenuation of a wave as it passes through the atmosphere as a function of frequency for two different altitudes.

4. RADAR — a radar's target effective reflection area is proportional to its size measured in wavelengths, and this, together with antenna size, makes microwaves the preferred radar frequency band. In other words, the resolution of a radar is much larger at higher frequencies for the same antenna size. Radar are used for target tracking, velocity determination and remote sensing (mapping of geography and weather), to name a few.

5. OTHER — molecular, atomic and nuclear resonances of a large number of materials occur at microwave frequencies, creating such applications as remote sensing, radio-astronomy, medical diagnostics and, of course, cooking (most microwave ovens work at 2.45 GHz). A large number of high-power microwave industrial heating applications also exist in the 900-MHz and 2.45-GHz heating designated bands. In the medical field, microwave hyperthermia has been proven to make radiation treatment of cancer more effective.

1.2 History

The history of microwaves started with Maxwell's theory in the nineteenth century. Maxwell mathematically showed that electromagnetic wave propagation exists, and that light is an electromagnetic wave. Not many people understood Maxwell's theory at the time. Two people, however, did: Heinrich Hertz, who verified the theory with a series of ingenious experiments about twenty years later, and Oliver Heaviside, who developed a mathematical language for Maxwell's theory that most engineers could understand and use.

Heaviside introduced vector notation and provided foundations for guided-wave and transmission-line theory. He was a telegrapher in his youth, and understood transmission lines very well.

Hertz was the first true microwave engineer. Between 1887 and 1891 he performed a series of experiments at wavelengths between 6 cm and 6 m. His most important experiment was probably the following. He used a high voltage spark (rich in high harmonics) to excite a half-wave dipole antenna at about 60 MHz. This was his transmitter. The receiver was an adjustable loop of wire with another spark gap. When he adjusted the resonance of the receiving antenna to that of the transmitting one, Hertz was able to show propagation of waves for the first time. Hertz demonstrated first reflector antennas, finite velocity of wave propagation in coaxial transmission lines, standing waves, and a number of microwave and RF techniques. Unfortunately, he died at an early age of 36 (from a tooth infection). He was a professor at Karlsruhe University in Germanny, and his original lab apparatus is kept operational at Bonn University, Germany.

The next important discovery for the development of microwaves were metal waveguides, discovered independently by Southworth at AT&T and Barrow at MIT. Southworth made his invention in 1932, but could not talk about it, because of company policies, until a meeting in 1936. Barrow was at the same time working on antennas, and came to a conclusion that a hollow metal tube could guide electromagnetic waves. His first experiments in 1935 were not successful, because he did not understand cutoff in waveguides, and tried to guide a 50 cm wave through a 4.5 cm tube (which is well below cutoff at λ=50 cm). He understood his mistake soon, though, and repeated his experiment with a tube 18 inches in diameter. Before the Second World War, a high power microwave source was invented—the magnetron. This triggered development of radar (Radio Detection And Ranging), which was under way simultaneously in Great Britain, the United States and Germany, but the first radar was built in Britain and played an important role in the victory of the Allies. In the United States, the microwave field prospered at that time at the MIT Radiation Labs. Most of the work used waveguides and coaxial lines as the transmission medium. A waveguide can handle high power levels, but is narrow band, whereas coax is broadband, but limited in power and achievable circuit complexity. In the early 50's, planar transmission lines, such as strip line and microstrip, were developed. Microstrip lines are currently used

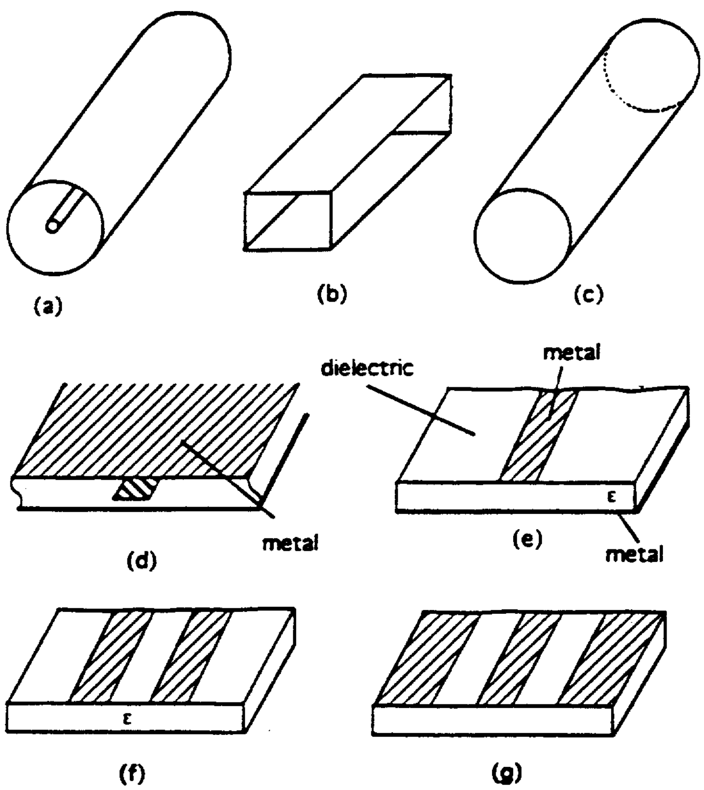


Figure 1.2: (a) coaxial line, (b) rectangular waveguide (c) cylindrical waveguide, (d) stripline (e) microstrip, (f) coplanar strips (CPS) and (g) coplanar waveguide (CPW).

for many microwave applications, since they are planar, low cost, compact and allow a large variety of circuits on a relatively small area. Other planar transmission lines on dielectric substrates, such as coplanar waveguide, are still a research topic today. Fig. 1.2 shows the most frequently used transmission media.

The development of active microwave devices started before the Second World war with the invention of the magnetron tube. Soon after that, in 1937, the klystron tube was invented. These tubes can work as both amplifiers and oscillators. Another important tube is the TWT (Travelling Wave Tube) invented in the 50's. All the tube sources are bulky and require large additional equipment such as power supplies and refrigerators. There was a clear need for smaller and cheaper active devices, which came with the development of semiconductor devices. The device most often used today at microwave frequencies is the GaAs MESFET, first made by Carver Mead at CalTech in 1965. People today mostly talk about MMIC's (Monolithic Microwave Integrated Circuits). This means that planar transmission lines and active devices are made simultaneously on one semiconductor substrate, typically GaAs. This is still a field of active research, especially at millimeter-wave frequencies and for more complex circuits.

1.3 Transmission Lines — Review

1.3.1 Transmission Lines in the Time Domain

A coaxial cable and a two-wire line consist of two wires. The current and voltage on these wires do not satisfy Kirchhoff's laws as you studied them in circuits classes and as you will see in Lab #1 – the voltage at the two coax ends is not necessarily the same, even if we assume the wires to be perfect conductors. Let us look at a very short piece Δz of cable, Fig. 1.3(a). There is a capacitance between the two wires, and since current flows through them, there is an associated magnetic field and an inductance. This is represented with a shunt capacitor C_Δ and a series inductor L_Δ which now represent a circuit equivalent of a short piece of cable. Any longer piece of cable of length z can be represented as a cascade of many short pieces, Fig. 1.3(b). By looking at the three sections in the middle of the cable, $(n-1)$, n and $(n+1)$, we can see that the voltage drop across the n-th inductor and the current through the n-th capacitor are:

$$L_\Delta \frac{\mathrm{d}i_n}{\mathrm{d}t} = v_n - v_{n+1} \qquad \text{and} \qquad C_\Delta \frac{\mathrm{d}v_n}{\mathrm{d}t} = i_{n-1} - i_n \tag{1.1}$$

When drawing the circuit in Fig. 1.3(b), we implicitly assumed that all of the short sections Δz have the same inductance and capacitance and that the total capacitance and inductance of the cable, L and C, is equal to the sum of all series inductors and shunt capacitors. Although you can measure the capacitance and inductance of different cable lengths and convince yourself, this is not obvious. If the capacitance and inductance of the cable are indeed proportional to its length, we can write:

$$L = \frac{L_\Delta}{\Delta z} \qquad \text{and} \qquad C = \frac{C_\Delta}{\Delta z} \tag{1.2}$$

L and C are called the *distributed* inductance and capacitance, and their units are Henry/meter and Farad/meter. From now on, when we are dealing with transmission lines, we will assume that L and C are quantities given per unit length. The meaning of distributed circuit elements is that they are not physically connected between two ends of the cable, but rather that they "accumulate" along a cable length. Now we can rewrite (1.2) as follows:

$$L \frac{\mathrm{d}i_n}{\mathrm{d}t} = \frac{v_n - v_{n+1}}{\Delta z} \qquad \text{and} \qquad C \frac{\mathrm{d}v_n}{\mathrm{d}t} = \frac{i_{n-1} - i_n}{\Delta z} \tag{1.3}$$

As Δz shrinks and approaches zero, the quotients become derivatives with respect to the distance z. Keep in mind that the current and voltage change along the cable, but they also change in time, since

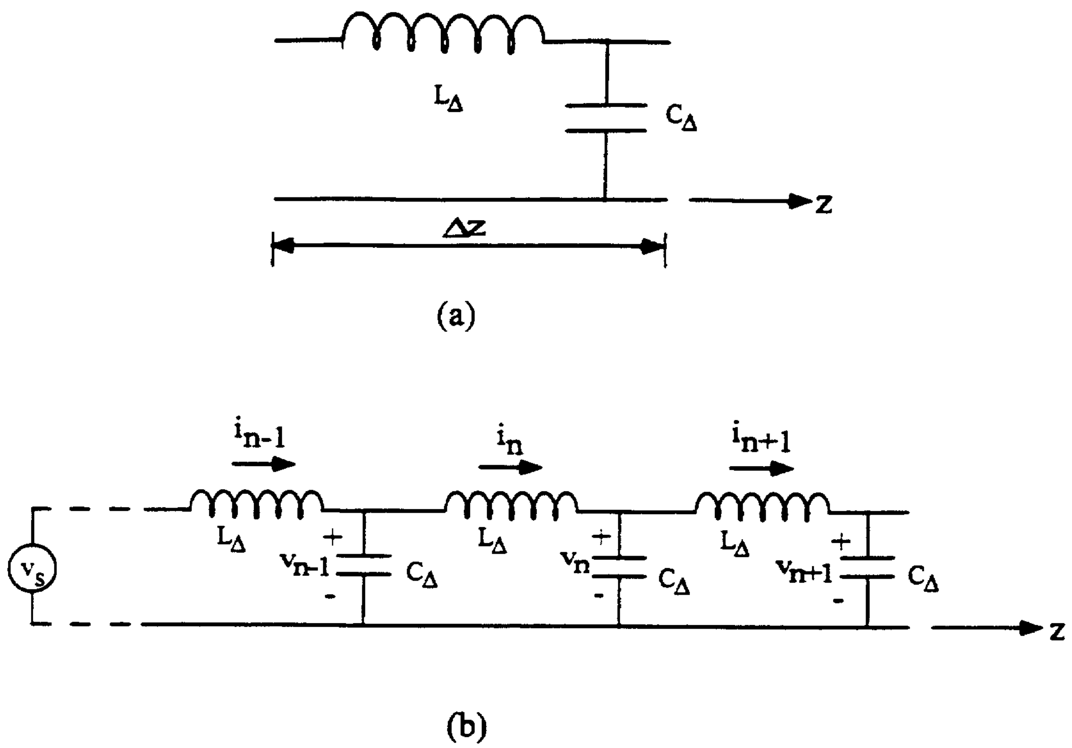


Figure 1.3: (a) A very short piece of lossless cable (Δz) can be represented as a lumped circuit consisting of a series inductor and a shunt (parallel) capacitor. (b) A longer piece of cable can be represented as many cascaded short sections Δz.

a function generator is at the beginning of the cable. We can now write (1.3) as:

$$\frac{\partial v(z,t)}{\partial z} = -L\frac{\partial i(z,t)}{\partial t} \qquad \text{and} \qquad \frac{\partial i(z,t)}{\partial z} = -C\frac{\partial v(z,t)}{\partial t} \tag{1.4}$$

Eqns. (1.4) are called the *telegrapher's equations* or the *transmission-line equations*. In your review homework, you will do some manipulation with these equations to eliminate the current, and get:

$$\frac{\partial^2 v}{\partial t^2} - \frac{1}{LC}\frac{\partial^2 v}{\partial z^2} = 0 \tag{1.5}$$

This is the *wave equation* in one dimension (z) and it describes the voltages (and currents) along a cable. The same type of equation can be used to describe the electric and magnetic fields in a radio wave or optical ray, sound waves in acoustics, and heat transfer in thermodynamics. This equation can also be derived in three dimensions from Maxwell's equations for electric and magnetic fields instead of voltages and currents.

In order to solve the second-order partial differential equation (1.5), let us first rewrite it somewhat. The quantity $1/LC$ has the dimensions of a velocity squared. If we define the velocity

$$c_0 = \frac{1}{\sqrt{LC}}, \tag{1.6}$$

then (1.5) takes the familiar form:

$$\frac{\partial^2 v}{\partial t^2} - c_0^2\frac{\partial^2 v}{\partial z^2} = 0 \tag{1.7}$$

The fields of a plane electromagnetic wave travelling through a homogeneous dielectric medium of relative permittivity ϵ_r obey (1.7), with c_0 replaced by the velocity of light in that medium: $c/\sqrt{\epsilon_r}$, where $c \simeq 3 \times 10^8$ m/s is the velocity of light in vacuum. The actual velocity c_0 of waves on a transmission line can be expressed in terms of an *effective permittivity* ϵ_e:

$$c_0 = \frac{c}{\sqrt{\epsilon_e}} \tag{1.8}$$

so that c_0 is the same as the velocity of a plane wave in a uniform dielectric of relative permittivity ϵ_e.

Now let us try a solution to (1.5) of the form $v(z,t) = f(z - at)$, where f is some arbitrary function and a is a quantity which has units of velocity. Substituting this into (1.5) and carrying out the differentiations using the chain rule, the following is obtained:

$$\frac{\partial^2 v}{\partial t^2} - c_0^2\frac{\partial^2 v}{\partial z^2} = a^2 f'' - c_0^2 f'' = 0 \tag{1.9}$$

Either f must be trivial (a constant or linear function), or we must have $a = \pm c_0$. Here the + sign corresponds to a *forward voltage wave* $f(z - c_0 t)$, and the − sign to a *backward traveling wave* $f(z + c_0 t)$. For example, let f be a rectangular pulse which starts at $t = 0$, Fig. 1.4(a). At a later time t_1 the pulse has moved to the right (in the $+z$ direction) by $c_0 t$. This is a forward wave. We will often denote the forward wave $v_+(z,t)$ as an *incident wave*, and the backward wave $v_-(z,t)$ as a *reflected wave*.

How fast are the voltage waves along a typical cable? For typical coaxial cables you use in the lab, the capacitance per unit length is about 1 pF/cm and the inductance is about 2.5 nH/cm, so the velocity is about 2/3 of the speed of light in air. (What is ϵ_e in this case? In which cases would the velocity be equal to the speed of light in vacuum?)

Similar expressions to the ones for the voltage wave can be written for the current wave along the cable. From (1.4), we find

$$v' = aLi' = \pm c_0 Li', \tag{1.10}$$

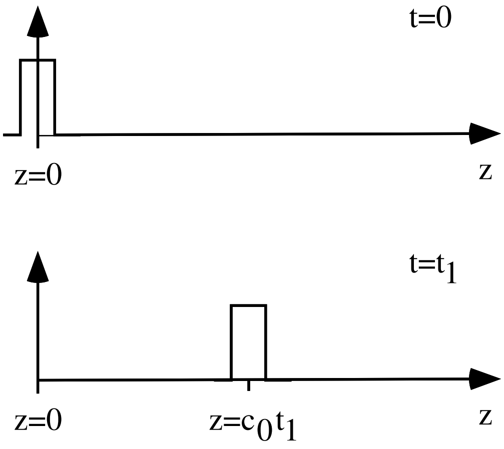


Figure 1.4: A voltage wave along a cable changes in time.

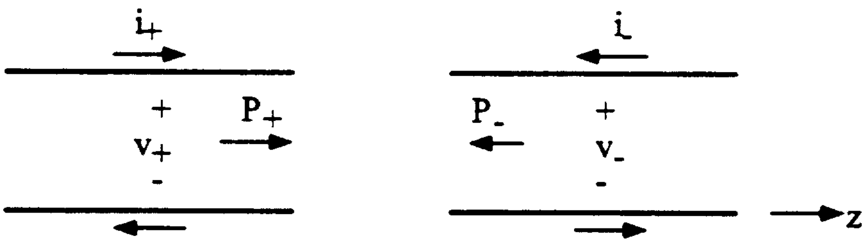


Figure 1.5: Forward and backward voltage and current waves in a transmission line.

which means that, assuming the DC voltages and currents are zero, the ratio of the voltage and current along the line is constant and equal to

$$\frac{v}{i} = \pm c_0 L = \pm\sqrt{\frac{L}{C}} = \pm Z_0 \tag{1.11}$$

Z_0 is called the *characteristic impedance* of the transmission line. The plus and minus signs apply to the forward and backward waves. If we assume that the voltage between the lines has the same sign for the two waves, then the current in a forward wave flows in the opposite direction to the current in the backward wave, as shown in Fig. 1.5. Physically, this is because a forward wave carries power to the right, while the backward wave carries power to the left. We will examine the flow of power more closely in the next section.

1.3.2 Power Flow and Decibels

The power flow in a forward wave on a transmission line is a positive number since it is produced by a source (function generator) connected at the left end of the line and it is delivering power to the line in the $+z$ direction:

$$P_+ = v_+ i_+ = v_+(v_+/Z_0) = v_+^2/Z_0 \tag{1.12}$$

In a backward wave, the power flow is in the opposite direction; it is flowing towards our source, appears to the source as another generator, and therefore it is negative:

$$P_- = v_- i_- = v_-(-v_-/Z_0) = -v_-^2/Z_0 \tag{1.13}$$

The unit for power is a Watt (W), but engineers often use a relative unit – the decibel (dB), to express the ratio of two powers in a convenient way:

$$P_{\text{dB}} = 10 \log \left(\frac{P_1}{P_2} \right)$$

and we say that P_1 is P_{dB} decibels above P_2. A single power level can be expressed in the form

$$P_{\text{dB}} = 10 \log \left(\frac{P}{P_{\text{ref}}} \right), \tag{1.14}$$

where P is the power we are measuring or calculating, and P_{ref} is some given reference power level. At microwave frequencies, very often this reference power level is 1 mW, and in this case the unit is called a dBm:

$$P_{\text{dBm}} = 10 \log \left(\frac{P}{1\text{mW}} \right) \tag{1.15}$$

A positive number of dB's corresponds to a ratio greater than 1 (gain), while a negative number of dB's represents a ratio less than 1 (loss). Decibels are convenient for two reasons: (1) they are easier to write (for example, the range between +63 dB to -153 dB corresponds to $2 \cdot 10^6$ to $0.5 \cdot 10^{-15}$), and (2) adding dB's corresponds to multiplication, which is useful whenever there are several stages cascaded in some system (for example, in a multistage amplifier, the gain can be found just by adding individual gains in dB).

Since power is proportional to the square of voltage in a circuit or on a transmission line, we can also use voltage ratios to obtain decibel levels for gain or loss, *provided the same impedance or characteristic impedance is used for both voltages being compared.* The same is, incidentally, true of currents. Thus, a voltage V_1 is said to be

$$P_{\text{dB}} = 20 \log \left(\frac{V_1}{V_2} \right)$$

decibels above the voltage V_2 on a transmission line of identical characteristic impedance. It is important to keep in mind that the decibel is always a measure of a *power* ratio; to say that voltages are so many dB apart is somewhat an abuse of the terminology, although a common one. In an obvious way, we can also define dB relative to some given voltage reference level, such as the dBμV: the number of decibels above 1 microvolt. Although it is done much less frequently, we can express ratios of currents in dB as well.

1.3.3 Time-Harmonic Steady State

In this course, we will deal mostly with sinusoidally time-varying voltages and currents and linear materials, which means that it is appropriate to use phasor (complex) notation. Both voltage and current in the time domain can be obtained from an assumed exponential time variation, for example:

$$v(z,t) = \text{Re}\left[\sqrt{2} V(z) e^{j\omega t}\right], \tag{1.16}$$

The presence of the factor $\sqrt{2}$ in (1.16) (and in a similar equation for current) indicates that the phasors V and I are *RMS* (root-mean-square) rather than peak values. This means that when we calculate the time-average power associated with a time-harmonic voltage $v(t)$ (as in (1.12) above, for example), we will have

$$\left[v^2(t)\right]_{\text{av}} = |V|^2 \tag{1.17}$$

without the factor of 1/2 which appears when using peak voltages (see (3.3)). Standard laboratory equipment normally displays RMS values for measured voltages and currents. Thus, unless otherwise stated, all phasor voltages and currents will hereinafter be understood to refer to RMS quantities.

Rather than deal with time derivatives explicitly, we can now leave the exponentials out and write equations involving the phasors V and I directly. The derivative with respect to time then becomes just a multiplication with $j\omega$, and we can write the transmission-line equations as:

$$V' = -j\omega LI, \qquad I' = -j\omega CV \tag{1.18}$$

The prime denotes a derivative with respect to z. Again, the current can be eliminated, to find

$$V'' = -\omega^2 LCV \tag{1.19}$$

As before, the two solutions to this second order differential equation are waves propagating in the $+z$ and $-z$ directions. In the case of sinusoidal voltages and currents, they are of the form

$$V_{\pm} \propto e^{\mp j\beta z}.$$

The subscript for the wave travelling in the $+z$ direction (forward) is $_+$, and for the one travelling in the $-z$ direction (backward) it is $_-$. Notice that the *forward-travelling* wave has a *minus* sign in the exponential. This means that the phase lags as you move along the z-direction.

The quantity $\beta = \omega\sqrt{LC}$ is called the *phase constant*, because it determines the phase of the voltage at a distance z from the beginning of the line (z=0). The phase constant is related to the wavelength on the transmission line, and actually, $\beta = \omega/c_0 = 2\pi f/c_0 = 2\pi/\lambda_g$, where

$$\lambda_g = \frac{2\pi}{\beta} = \frac{c_0}{f}$$

is the definition of the so-called *guided wavelength* of the transmission line. It is related to the wavelength $\lambda_0 = c/f$ of a plane wave in free space by $\lambda_g = \lambda_0/\sqrt{\epsilon_e}$. Often, when there is no risk of confusion, the guided wavelength will be denoted simply by λ. Finally, the phase constant is related to the wavenumber $k_0 = \omega/c$ of a plane wave in free space by

$$\beta = k_0\sqrt{\epsilon_e}$$

In order to find the corresponding currents (forward and backward currents), we substitute the expression for the voltage back into (1.18), and obtain

$$I_+ = V_+/Z_0 \quad \text{and} \quad I_- = -V_-/Z_0, \tag{1.20}$$

where $Z_0 = \sqrt{L/C}$ is the characteristic impedance. Notice again that the current and the voltage have the same sign for a forward travelling wave, but opposite signs for a backward travelling wave. Since the power is the product of the voltage and current, this means that the power flow with respect to the $+z$ reference direction is positive for a forward wave, and negative for a backward wave. As in the previous time-domain formulation, at any point along the line, or for every z, the total *complex* voltage and current are equal to the sum of the forward and backward voltage and current *at that point*:

$$\begin{aligned} V(z) &= V_+(z) + V_-(z) \\ I(z) &= I_+(z) + I_-(z) \end{aligned} \tag{1.21}$$

So far we have considered only transmission lines themselves, and we have not said what is connected at the ends of the lines. In reality, what we are really interested in is what happens when a load or a generator is connected at some point on the line, or more than one section of line is interconnected. The simplest case of a load impedance terminating a section of line is shown in Fig. 1.6. A forward, or incident, wave travels from the generator, reaches the load, and some of the power is delivered to the

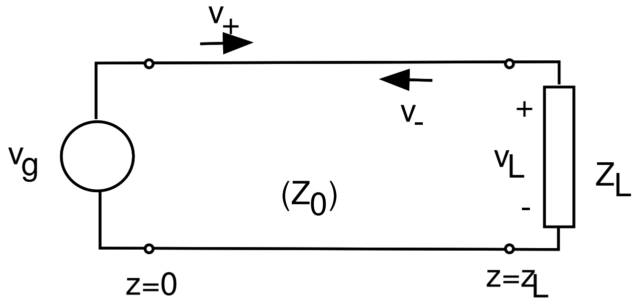


Figure 1.6: A transmission line of characteristic impedance Z_0 terminated in a load Z_L at a distance z_L from the generator.

load, while some of it can be reflected into a backward wave. Since the line is linear and L and C are constants, the reflected wave is proportional to the incident wave. The complex ratio of the two voltage waves at any point along the line is called the (voltage) *reflection coefficient*:

$$\rho(z) = \frac{V_-(z)}{V_+(z)} \tag{1.22}$$

Notice that the current reflection coefficient (defined as the ratio of backward to forward *current* waves) is the negative of the voltage reflection coefficient. The voltage that appears across the load, V_L is also proportional to the incident voltage, and the *load transfer coefficient* is a complex number defined as

$$\tau = \frac{V_L}{V_+(z_L)} \tag{1.23}$$

If the load impedance Z_L is replaced by a matched section of transmission line whose characteristic impedance is equal to Z_L, the same reflected wave exists on the first line, and the total voltage at the connection point is also equal to the forward (transmitted) voltage wave traveling on the second line. In that case, we call τ the (voltage) *transmission coefficient*.

The total voltage (or current) at any point along a linear transmission line are the sums of the incident and reflected voltages (or currents) at that point:

$$v(z,t) = v_+(z,t) + v_-(z,t) \qquad \text{and} \qquad i(z,t) = i_+(z,t) + i_-(z,t) \tag{1.24}$$

To find the *impedance* at any point along the line (remember, everything is a function of position), we divide the total voltage by the total current, just as in circuit theory:

$$Z(z) = \frac{v(z)}{i(z)} = \frac{v_+(z)}{i_+(z)} \frac{1 + v_-(z)/v_+(z)}{1 + i_-(z)/i_+(z)} = Z_0 \frac{1+\rho(z)}{1-\rho(z)} \tag{1.25}$$

So, the impedance along the line, $Z(z)$, and the reflection coefficient, $\rho(z)$, are related by

$$\frac{Z(z)}{Z_0} = \frac{1+\rho(z)}{1-\rho(z)} \qquad \text{and} \qquad \rho(z) = \frac{Z(z) - Z_0}{Z(z) + Z_0} \tag{1.26}$$

An important thing to remember is that the characteristic impedance Z_0 depends on the way the cable is made (its dimensions, shape in transverse plane and materials). On the other hand, the impedance

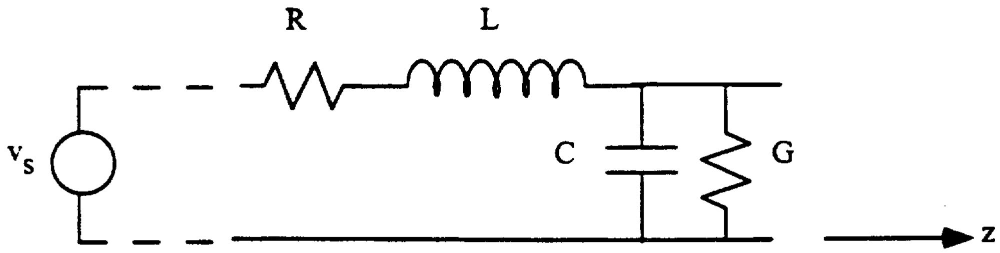


Figure 1.7: Schematic of a short section of transmission line with distributed losses included.

along the cable is a function of position and termination as well as being proportional to Z_0, and is the ratio of the *total* voltage and current.

Now we can also write the expression for the transmission coefficient. At the load, the voltage is the sum of the incident and reflected voltage: $v_L(z_L) = v_+(z_L) + v_-(z_L)$. Dividing by $v_+(z_L)$, and from (1.23) and (1.25), we obtain

$$\tau = 1 + \rho = \frac{2Z_L}{Z_L + Z_0} \tag{1.27}$$

Please keep in mind that (1.27) is *not* universally true, but was specifically derived for the situation shown in Fig. 1.6. Other circuits, particularly ones which contain series circuit elements in between the transmission line and the load, have a different relation between τ and ρ.

In RF and microwave engineering, usually all impedances are normalized to the characteristic impedance, and this is usually $50\,\Omega$. Normalized quantities are usually written in lower-case letters, so we can write

$$z = r + jx = \frac{1+\rho}{1-\rho} \qquad \text{and} \qquad \rho = \frac{z-1}{z+1} \tag{1.28}$$

Let us look at a few simple and extreme examples of terminations (loads): a short circuit, open circuit and a load $Z_L = Z_0$. At an open end of a cable, there is no current flowing between the two conductors, and the reflected voltage has to be equal to the incident voltage, so the reflection coefficient is equal to 1. You will see in the lab what effect this has on the reflected and transmitted voltage waves. On the other hand, at a short circuited cable end, there is not voltage drop between the two conductors of the cable. Since the total voltage at the end of the cable has to be zero, this means that the reflected voltage is the negative of the incident voltage, and the reflection coefficient is -1. For the case of a load impedance exactly equal to the cable characteristic impedance, from (1.26) we see that the reflection coefficient is zero and there is no reflected wave.

After going through some simple circuit theory, we have managed to show that voltage and current waves travel along transmission lines such as a coaxial cable. We have found that the capacitance and inductance of the cable determine its characteristic impedance, as well as the velocity of waves traveling along it. These inductances and capacitances can be found for any cable from Gauss' and Ampère's law.

1.4 Losses in Transmission Lines

The distributed circuit in Fig. 1.4 represents a perfect transmission line with no losses. A real transmission line has losses in the conductor, as well as in the dielectric between the conductors. These losses are represented in Fig. 1.7. as a distributed series resistance per unit length R in Ω/m, and a shunt conductance G in S/m, respectively. For such a transmission line, the characteristic impedance is given by:

$$Z_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}}, \tag{1.29}$$

and the propagation constant is now a complex number with both a real and imaginary part:

$$\gamma = \alpha + j\beta = \sqrt{(R + j\omega L)(G + j\omega C)} \tag{1.30}$$

What this means is that instead of $e^{-j\beta z}$ in the expressions for voltages and currents, we now have $e^{-(\alpha+j\beta)z} = e^{-\alpha z}e^{-j\beta z}$. You can see from this expression that in addition to traveling in the z-direction, the amplitude of the voltage and current waves also falls off in the direction of propagation. This is called *attenuation* and is a characteristic of every real transmission line. The phase of the wave is determined by β (phase constant), and its attenuation by α, which is called the *attenuation constant*. Since

$$\begin{aligned}\gamma &= \sqrt{j\omega L\, j\omega C\left(1 + \frac{R}{j\omega L}\right)\left(1 + \frac{G}{j\omega C}\right)} \\ &= j\omega\sqrt{LC}\sqrt{1 - j\left(\frac{R}{\omega L} + \frac{G}{\omega C}\right) - \frac{RG}{\omega^2 LC}}\end{aligned} \tag{1.31}$$

In the case where the cable is a good one with low losses, $R << \omega L$ and $G << \omega C$:

$$\gamma \approx j\omega\sqrt{LC}\sqrt{1 - j\left(\frac{R}{\omega L} + \frac{G}{\omega C}\right)} \approx j\omega\sqrt{LC}\left[1 - \frac{j}{2}\left(\frac{R}{\omega L} + \frac{G}{\omega C}\right)\right] \tag{1.32}$$

So, we find that

$$\alpha \approx \frac{1}{2}\left(R\sqrt{\frac{C}{L}} + G\sqrt{\frac{L}{C}}\right) \quad , \quad \beta \simeq \omega\sqrt{LC} \tag{1.33}$$

We usually assume that the in power that is transmitted along the transmission line varies with distance from the generators as $P(z) = P(0)e^{-2\alpha z}$.

The unit for α is Nepers per meter (after the Latin version of the name of John Napier, who invented the logarithm), but the attenuation is also often expressed in decibels per length of cable. The relationship between Nepers/m and dB/m is:

$$\alpha_{\mathrm{dB/m}} = 8.685\,\alpha_{\mathrm{Nepers/m}} \tag{1.34}$$

1.5 Loaded Transmission Lines

Consider a transmission-line of length l with a load Z_L, Fig. 1.8, and assume that the load has a reflection coefficient ρ and is located at $z = 0$. Moving back along the line, we can write the voltage and current at some distance z as

$$\begin{aligned}V(z) &= V_+(z) + V_-(z) = V_+(z)(1 + \rho(z)) = V_+(0)e^{-j\beta z}(1 + \rho(0)e^{2j\beta z}) \\ I(z) &= I_+(z) + I_-(z) = I_+(z)(1 - \rho(z)) = I_+(0)e^{-j\beta z}(1 - \rho(0)e^{2j\beta z})\end{aligned} \tag{1.35}$$

The impedance at point z is

$$Z(z) = \frac{V(z)}{I(z)} = Z_0\frac{1 + \rho(0)\,e^{j2\beta z}}{1 - \rho(0)\,e^{j2\beta z}} \tag{1.36}$$

We also know that

$$\rho(0) = \frac{Z_L - Z_0}{Z_L + Z_0}, \tag{1.37}$$

and by remembering Euler's formula ($e^{j\alpha} = \cos\alpha + j\sin\alpha$), we can write the impedance as a function of position along the line in the following form:

$$Z(z = -\ell) = Z_0\frac{Z_L + jZ_0\,\tan\beta\ell}{Z_0 + jZ_L\,\tan\beta\ell} \tag{1.38}$$

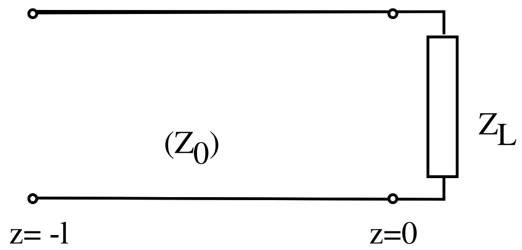


Figure 1.8: Reflection coefficient at different points along a transmission-line terminated with a load at $z = 0$.

Often, the "electrical length" $\beta\ell$ of a section of transmission line is identified and given the notation θ. Sometimes the electrical position βz along the line is also denoted by θ. The meaning should be clear from the context in which it is used.

The reflection coefficient at the load $\rho(0)$ is the ratio of the backward wave, $V_{-}(0)$ and the forward wave, $V_{+}(0)$. Moving down the line does not change the magnitudes of these two waves (for a lossless line), so the magnitude of the reflection coefficient does not change. However, the phase does change, since the backward wave lags as we move back down the line, and the forward wave leads. The phase of the reflection coefficient is the difference between the two phases, so it lags by $2\beta l$ as one moves away from the load:

$$|\rho(-\ell)| = |\rho(0)| \qquad \text{and} \qquad \angle\,\rho(-\ell) = \angle\,\rho(0) - 2\beta\ell \tag{1.39}$$

The most interesting case is when the length of the transmission-line is a quarter-wavelength. The reflection coefficient becomes

$$\rho(\ell = \frac{\lambda}{4}) = -\rho(0), \tag{1.40}$$

and the sign of the reflection coefficient changes. This means that the impedance converts to an admittance of the same value. In terms of un-normalized impedances, we get, from (1.38) when $\ell = \lambda/4$ ($\beta\ell = \pi/2$):

$$Z\left(\frac{\lambda}{4}\right) = \frac{Z_0^2}{Z_L} \tag{1.41}$$

You can see from this expression that the load impedance is transformed from a value Z_L to a value Z_0^2/Z_L. Quarter-wave long transmission-line sections often play the same role at microwave frequencies that impedance transformers play at lower frequencies. This is especially used for matching resistive loads. For example, if we want to match a 100-Ω load to a 50-Ω transmission-line, we could use a quarter-wavelength section of a line with a characteristic impedance of $50\sqrt{2} = 70.7\,\Omega$.

However, unlike in a low-frequency transformer, there is phase lag in the section of the transmission-line, and also the transformer effect works only in a narrow range of frequencies. In a transformer design, though, one can make this phase lag useful, and also one can add more quarter-wavelength sections to improve the bandwidth. The same ideas are used in optics to make anti-reflection coatings for lenses.

Another interesting case is the half-wavelength long transmission-line. The reflection coefficient in this case is unchanged: $\rho(\lambda/2) = \rho(0)$. This also means that the impedance is unchanged, or that adding a half-wavelength section has no effect. You can also see that from (1.38) when $\ell = \lambda/2$ ($\beta\ell = \pi$). This is used for making radomes that protect radars from mechanical damage. For example, most large airplanes have a radome on the nose with a radar hidden behind it, so that it does not get blown away. The radome is made of a piece of material that is half a wavelength thick at the operating frequency. This windowing effect is frequency dependent, just like the quarter-wave transformer is.

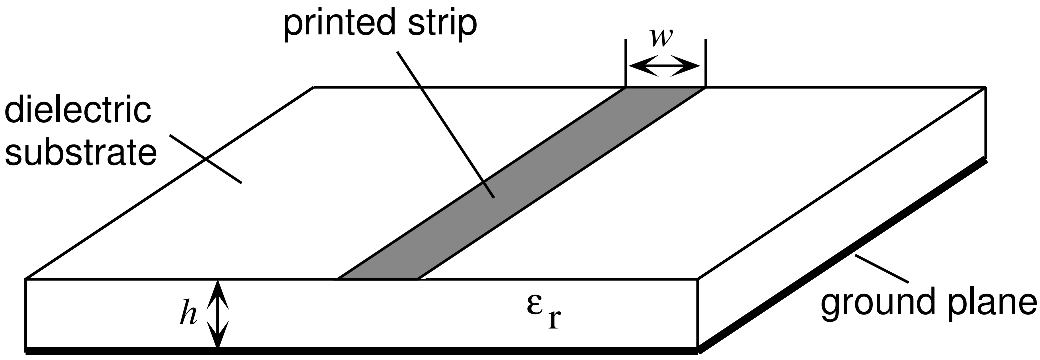


Figure 1.9: A microstrip transmission line printed on a grounded dielectric substrate.

1.6 Microstrip Circuits

In addition to traditional coaxial and waveguide components, many other types of transmission line and waveguide can be used at microwave frequencies. One of the most often used structures today is the microstrip, shown in Fig. 1.9. The wave is guided between the bottom ground plane and the top metal strip. Some of the fields spill over from the dielectric into air. Usually the mode guided in a microstrip circuit is called a quasi-TEM mode, because it almost looks like a TEM mode, but the fields do have a small component in the propagation direction. In reality, the guided mode is a hybrid TE and TM mode, and the analysis is quite complicated. If you imagine that there is no dielectric, just air, the mode could be TEM. The presence of the dielectric complicates things, but people have been able to use quasi-static analysis to obtain formulas for the impedances and propagation constants in microstrip lines. In these formulas, the so called *effective dielectric constant* is used. This is just some kind of average between the permittivity of air and the dielectric that gives you a rough picture of the portion of field that remains in the dielectric. It is usually found from:

$$\epsilon_e = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \frac{1}{\sqrt{1 + 12\ h/w}} \tag{1.42}$$

You can find formulas for microstrip line impedance and propagation constants in almost any microwave book today. We will just list them here, so you have them handy:

$$\begin{aligned} \beta &= k_0 \sqrt{\epsilon_e} \\ Z_0 &= \left\{ \begin{array}{ll} \frac{60}{\sqrt{\epsilon_e}} \ln\left(\frac{8h}{w} + \frac{w}{4h}\right), & \frac{w}{h} \leq 1 \\ \frac{120\pi}{\sqrt{\epsilon_e}\left[\frac{w}{h} + 1.393 + 0.667\ \ln\left(\frac{w}{h} + 1.444\right)\right]}, & \frac{w}{h} \geq 1 \end{array} \right. \end{aligned} \tag{1.43}$$

For a given characteristic impedance Z_0, and permittivity ϵ_r, the $\frac{w}{h}$ ratio can be found as

$$\frac{w}{h} = \left\{ \begin{array}{ll} \frac{8e^A}{e^{2A} - 2} & \frac{w}{h} < 2 \\ \frac{2}{\pi}\left[B - 1 - \ln(2B - 1) + \frac{\epsilon_r - 1}{2\epsilon_r}\left\{\ln(B - 1) + 0.39 - \frac{0.61}{\epsilon_r}\right\}\right] & \frac{w}{h} > 2 \end{array} \right\} \tag{1.44}$$

where

$$\begin{aligned} A &= \frac{Z_0}{60}\sqrt{\frac{\epsilon_r + 1}{2}} + \frac{\epsilon_r - 1}{\epsilon_r + 1}\left(0.23 + \frac{0.11}{\epsilon_r}\right) \\ B &= \frac{377\pi}{2 Z_0 \sqrt{\epsilon_r}} \end{aligned} \tag{1.45}$$

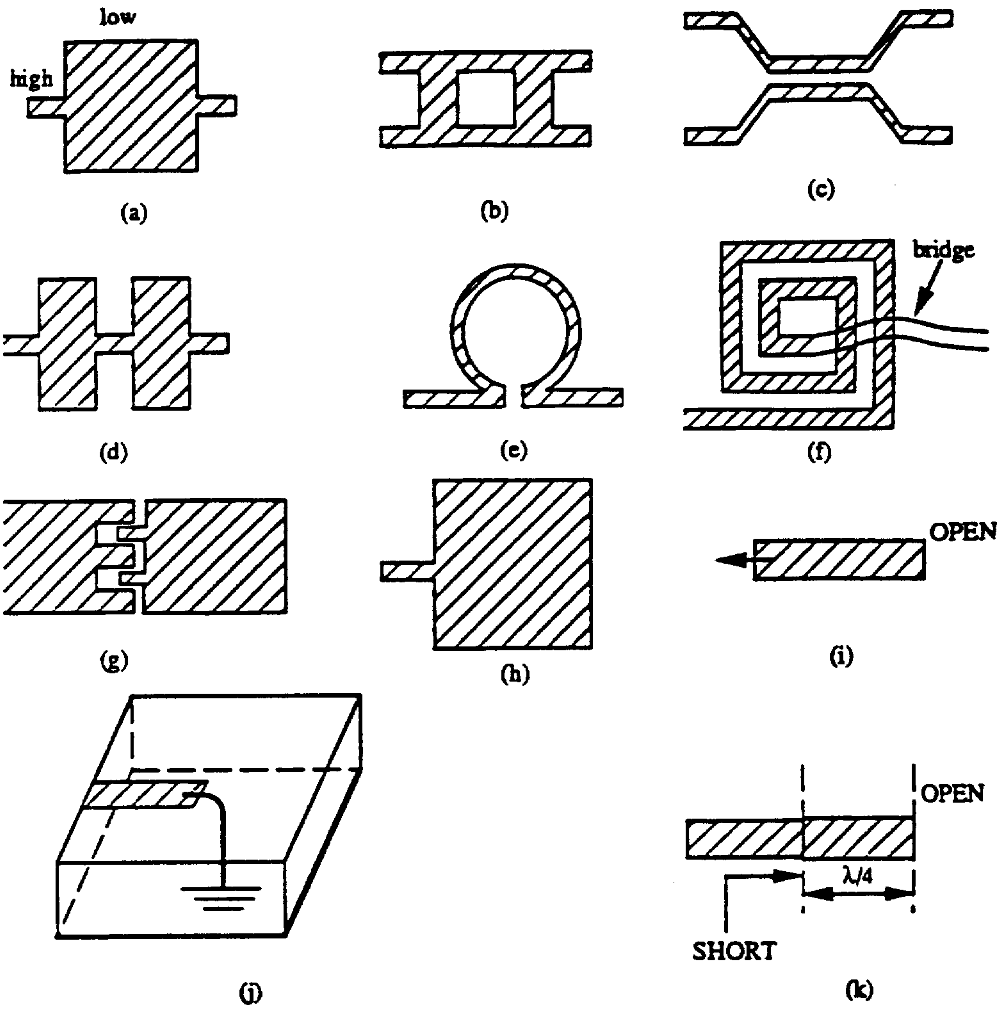


Figure 1.10: A microstrip impedance step (a), directional couplers (b and c), filter (d), inductors (e and f), capacitor (g), patch antenna (h), open circuit (i) and two ways of making a short circuit (j and k).

Here are some rules of thumb to remember:

- the higher the dielectric constant, the thinner the line, keeping the thickness of the dielectric and the impedance of the line constant.
- the thinner the dielectric, the thinner the line is, keeping the dielectric constant and impedance of the line constant.
- the higher the dielectric constant, the smaller the circuit is (why?).
- the wider the line, the lower the impedance.

Microstrip circuits became popular because they are planar (flat), small, easy and fast to make, and cheap. However, they cannot handle very high power levels and they are more lossy than coax or waveguide. At higher frequencies (above 20 GHz), the dielectric losses limit the performance. As an illustration, Fig. 1.10 shows a few geometries of typical microstrip circuits. We will cover a few microstrip circuits in the lab, including at least one active microstrip circuit such as an oscillator or amplifier.

1.7 Waveguides

Let us recall a few basic facts about rectangular waveguides. Closed metallic waveguides support electromagnetic fields known as *modes*, which have the form

$$\vec{E}(x,y,z) = e^{-\gamma z}\vec{\mathcal{E}}(x,y); \quad \vec{H}(x,y,z) = e^{-\gamma z}\vec{\mathcal{H}}(x,y) \tag{1.46}$$

where

$$\gamma = \alpha + j\beta \tag{1.47}$$

is the complex propagation constant of the mode, and $\vec{\mathcal{E}}$, $\vec{\mathcal{H}}$ are mode field patterns determined by frequency of operation, dielectric and magnetic properties μ and ϵ of the uniform material filling the waveguide, and waveguide cross section geometry. Unlike for a transmission line, the propagation constant of a waveguide may be real or imaginary, even though no losses are present. Thus we have

$$\gamma = \alpha = 2\pi f_c\sqrt{\mu\epsilon}\sqrt{1-\frac{f^2}{f_c^2}}; \qquad f < f_c \tag{1.48}$$

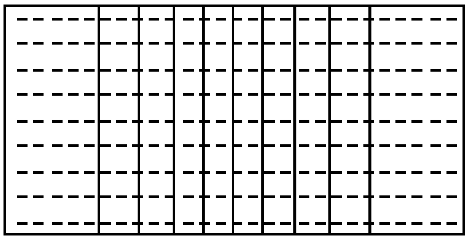
where f_c is the so-called *cutoff frequency* of the waveguide mode, while

$$\gamma = j\beta = j2\pi f\sqrt{\mu\epsilon}\sqrt{1-\frac{f_c^2}{f^2}}; \qquad f > f_c \tag{1.49}$$

Thus, if the operating frequency is below that of the mode with the lowest value of f_c for the waveguide, no propagation takes place; all modes decay exponentially with distance. If the operating frequency is greater than the cutoff frequency of the mode with lowest f_c (the so-called fundamental mode of the waveguide) but is smaller than those of all the other modes, then only this fundamental mode will propagate energy along the waveguide. Ordinarily, this is the desired way to use a waveguide, because otherwise interference between several propagating modes could take place.

Above the cutoff frequency of a waveguide mode, it behaves in a very similar manner to a transmission line. The electric field transverse to the direction of propagation (z) behaves like the voltage, and the transverse magnetic field like the current. We speak of the *guide wavelength* λ_g, defined by:

$$\lambda_g = \frac{2\pi}{\beta} = \frac{1}{f\sqrt{\mu\epsilon}\sqrt{1-\frac{f_c^2}{f^2}}} \tag{1.50}$$



(1) TE_{10}

Figure 1.11: Electric (—) and magnetic (- - -) transverse field distributions of the TE_{10} mode of a metallic rectangular waveguide.

EIA Waveguide Designation	a, in.	b, in.	f_c, GHz, for TE_{10} mode	Single-mode frequency range, GHz	Microwave band
WR-284	2.84	1.34	2.08	2.60-3.95	S
WR-187	1.872	0.872	3.16	3.95-5.85	C/S
WR-137	1.372	0.622	4.29	5.85-8.20	C/X
WR-112	1.122	0.497	5.26	7.05-10.00	C/X
WR-90	0.900	0.400	6.56	8.20-12.40	X
WR-62	0.622	0.311	9.49	12.40-18.00	Ku
WR-42	0.420	0.170	14.1	18.00-26.50	K
WR-28	0.280	0.140	21.1	26.50-40.00	Ka
WR-15	0.148	0.074	39.9	50.00-75.00	W
WR-12	0.122	0.061	48.4	60.00-90.00	W

Table 1.2: Standard air-filled metallic rectangular waveguide sizes and their parameters.

The guide wavelength is used in the same way as the ordinary wavelength of a transmission line. Both forward and backward propagating modes exist in a waveguide, resulting in reflection and transmission just like in a transmission line. The bottom line is that a waveguide mode can be treated for almost all purposes as a transmission line, except for the fact that genuine voltage and current cannot be identified for the waveguide.

If the dimensions of an air-filled metallic rectangular waveguide are $a \times b$, with $a > b$, then the fundamental mode is the TE_{10} mode, whose cutoff frequency is

$$f_c = \frac{c}{2a} \tag{1.51}$$

and whose electric field spans the shorter dimension of the waveguide as shown in Fig. 1.11. The electric field is largest in the center of the waveguide cross section, going to zero at the side walls as

$$\sin\left(\frac{\pi x}{a}\right) \tag{1.52}$$

Rectangular waveguide comes in a variety of standard sizes, according to the frequency band over which one desires single-mode operation. Some of the most popular sizes are listed in Table 1.2, together with the inner dimensions, the fundamental mode cutoff frequency, the single-mode operating band, and the common (though usually imprecise) microwave letter-band designation. Note how the waveguide is not used for frequencies immediately above the cutoff frequency of the fundamental mode. This is because the waveguide can be too lossy and too dispersive (it distorts pulses) if the operating frequency is too close to f_c.

1.8 Practice questions

1. How big are the wavelengths of a 1 GHz, 3 GHz, 10 GHz and 30 GHz wave?

2. What is an electrical size?

3. Write down Maxwell's equations (in any form)?

4. What does the cutoff frequency of a waveguide depend on? What is the formula for the cutoff frequency of the dominant mode in a rectangular waveguide?

5. What is a distributed impedance?

6. Can you make a transmission line out of a single conductor? Why? (What kind of a wave does a transmission line support?)

7. Why does it make sense for the voltage and current to have the same sign for a forward propagating wave on a transmission line, and opposite signs for a backward propagating wave?

8. What is the difference between the characteristic impedance of a transmission line and the impedance along the line?

9. What kind of a wave do you have to have if the propagation constant is complex?

10. How do the reflection coefficient and impedance of a transmission line terminated in some load change along the line (remember, these are complex numbers)?

11. List the formulas and definitions for (1) the characteristic impedance of a line (what does it depend on?), (2) the phase velocity, (3) the attenuation constant, (4) the propagation constant, (5) the voltage at any point on a line, (6) the current at any point on the line, (7) the reflection coefficient, (8) the transmission coefficient, and (9) the impedance at any point along a transmission line (what does it depend on?).

12. Derive the loss coefficient α in a coaxial cable assuming the loss is small.

13. Assume a coaxial cable has only resistive loss (the dielectric is perfect). In that case, what kind of coaxial cable would you fabricate to minimize the attenuation coefficient? The conclusion might seem surprising. Try to explain your conclusion based on electromagnetic field principles.

14. What is a microstrip? Sketch the electric and magnetic field lines in a microstrip line.

15. How do the thickness and permittivity of the substrate affect the way a 50-Ω microstrip line looks?

16. Why, based on qualitative physical arguments, does a thin microstrip line have a high impedance, and a wide one low impedance?

17. Why, based on qualitative physical arguments, is a series gap in a microstrip line capacitive, and a loop (such as the one in Fig. 1.10(e) inductive?

18. What are the main differences, from the electrical point of view, between a physical short, Fig. 1.10(j) and a short made using a $\lambda/4$ long section of open microstrip line, Fig. 1.10(k)?

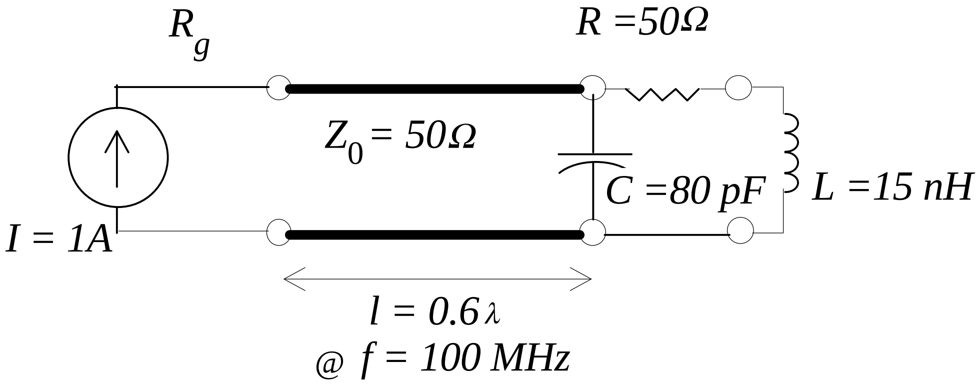


Figure 1.12: Transmission line terminated in lumped element load.

1.9 Homework Problems

1. Use Ansoft Designer (or SPICE) to simulate a circuit consisting of a resistor, a capacitor and an inductor at the end of a length of lossless transmission line as shown in Fig. 1.12. [SPICE instructions: The input end is driven by a current generator of 1 A, so that the voltage appearing at the input terminals of the transmission line is numerically equal to the input impedance of the loaded transmission line. Carry out the AC simulation over the frequency range of 10 MHz to 1000 MHz, using at least 101 frequency points. Plot the real and imaginary parts of the input impedance of the loaded line as seen by the current generator over this frequency range. Comment on the positions and sizes of the peaks you see in this plot, and on the low and high frequency limits of this plot.]

2. Four *lumped* elements are inserted into a transmission-line section, one at a time, as shown in Fig. 1.13. Find an expression for the reflection coefficient of each lumped element for a wave incident from the left. Assume the line is terminated to the right so that there is no reflection off the end of the line. Find simplified expressions that apply when R, C, L, and G are small.

3. A short circuit is connected to a 50 Ω transmission line at $z = 0$. Make a plot of the impedance (real and imaginary parts), normalized voltage amplitude and normalized current amplitude along the line up to $z = -3\lambda/2$ for each case. "Normalized" means you can divide by any constant to get a maximum of 1.

4. Repeat problem 3, except use an open circuit termination for the transmission line.

5. Repeat problem 3, except use a load impedance of $Z_L = 100\ \Omega$ connected at $z = 0$.

6. Repeat problem 3, except use a load impedance of $Z_L = j50\ \Omega$ connected at $z = 0$.

7. Use Ansoft Designer or SPICE to analyze the behavior of an unbalanced artificial transmission line having twelve sections, with $L_\Delta = 1\,\text{mH}$ and $C_\Delta = 10\,\text{nF}$. Use a frequency of $f = 10\,\text{kHz}$, a voltage generator of strength 1 V, and load impedances of $Z_L = 316\ \Omega$, $Z_L = 0\ \Omega$, and $Z_L =$ a 50 Ω resistor in series with a 1μF capacitor. Plot the magnitude of the voltage vs. the section number along the line for these three loads. Comment on the results in each case.

Note: If you are using PSPICE from OrCAD (formerly Microsim), run an AC simulation at the single frequency, then run Probe to display the desired voltages. Select each trace label by holding down SHIFT while clicking with the mouse (the label turns red), then EDIT/COPY to copy the data to the clipboard. This data can then be pasted into, e. g., EXCEL for editing and plotting.

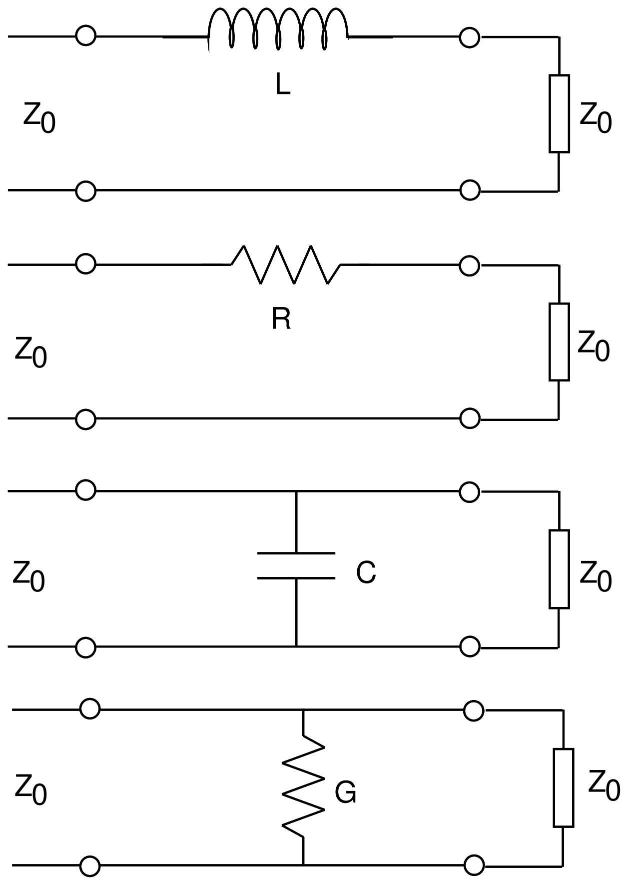


Figure 1.13: Lumped elements in a transmission line.

Next, replace the artificial transmission line with an actual lossless transmission line section with $Z_0 = 316\ \Omega$ that is 0.379λ long at $f = 10\,\text{kHz}$, and terminated with these same loads. Use Designer or SPICE to calculate the magnitude of the load voltage for each case, and compare with the results for the artificial line.

8. Use Designer or SPICE to analyze the behavior of the voltage and current waves along a long 50 Ω transmission line terminated by the various loads indicated below. In order to be able to "measure" the voltage and current at various points along the line, connect 9 sections of line whose electrical length is 20° each in cascade. Use a frequency of $f = 500\,\text{MHz}$ and a voltage generator of strength 1 V connected at the left end of the transmission line. At the right end (load end), connect load impedances of $Z_L = 50\ \Omega$, $Z_L = 100\ \Omega$, and $Z_L = 20\ \Omega$. Plot the magnitudes of the voltage and current vs. position along the line for each case. [See the note in the previous problem for hints on extracting data from SPICE analyses.]

9. Why don't we worry about impedance matching in the coaxial cables we use at lower frequencies (those used to connect audio equipment, for example)? Or should we? Let's examine this issue here.

Suppose an electrically "short" ($\beta l \ll 1$) length of lossless 50 Ω transmission line is connected to a load impedance of 50 kΩ (this would be typical of the input impedance to an audio amplifier). Consider an audio frequency of $f = 10\,\mathrm{kHz}$.

(a) If the dielectric filling the transmission line has a relative permittivity of 2.25, calculate β at this frequency.

(b) If the length of the line is $l = 2$ m, what is the impedance seen at the input end of the transmission line?

(c) Repeat part (b) if the length of the line is 10 m. Verify that this line still qualifies as electrically short.

(d) Does the input end of the line present a constant load impedance of 50 kΩ to a generator connected to that end as frequency is varied through the audio range (20 Hz to 20 kHz)? Explain what happens to a complicated audio signal passing through a line in this way, and why.

10. Repeat problem 9, using a load impedance of 4 Ω, as is comparable to typical values for loudspeakers. How long a transmission line is needed to observe a 1% change in the impedance seen at the input end of a 50 Ω line connected to such a load at $f = 20$ kHz?

11. A lab receiver measures an RF signal to have a strength of 35 dBμV. What is the signal level in volts? In watts? In dBm? Assume the instrument has an input impedance of 50 Ω.

12. A coaxial line attenuator is labeled as having a power attenuation of 6 dB. If an incident wave of 5 V is applied to one end, and if the reflection coefficient of the attenuator is zero, what is the amplitude of the voltage wave which emerges from the other end of the attenuator? What are the values of the incident and transmitted powers, if the characteristic impedance of the transmission lines is 75 Ω?

Chapter 2

Scattering Parameters and Review of Smith Chart

2.1 Scattering Parameters (S-Parameters)

In a coax, voltages and currents do make sense, but, for example, in a waveguide, they do not. We mentioned also that it is hard to measure voltages and currents in a transmission line, since any probe presents some load impedance, which changes what we are measuring. The standard quantities used at microwave frequencies to characterize microwave circuits are *wave variables* and *scattering parameters*. Usually, in a microwave circuit, we talk about *ports*, which are not simple wires, but transmission lines connected to a circuit. These transmission lines support waves travelling into and out of the circuit. This is shown in Fig. 2.1. A microwave circuit, in general, can have many ports. You can think of this in the following way: we send a wave into an unknown N-port circuit, and by measuring the reflected wave at the same port (like an echo) and the transmitted wave at some other port, we can find out what the microwave circuit is. The problem is the following. Let us say you look at an N-port, you send a wave through port 1, and you look at what gets reflected at port 1 and transmitted, say, at port 3. The result of your measurement will depend on what loads were connected to all of the other ports during the measurement. The convention is to terminate all other ports with the characteristic impedances of the transmission lines connected to the ports, so that there is no reflection from these ports. In other words, you can think of an S-parameter as a generalized reflection or transmission coefficient when all other ports of a multi-port circuit are matched.

Let us look at the generalized N-port microwave circuit in Fig. 2.1. The ports are indexed by the subscript i that goes from 1 to N. The normalized voltage waves a_i and b_i are defined as

$$a_i = \frac{V_i{}^+}{\sqrt{Z_{0i}}} \qquad b_i = \frac{V_i{}^-}{\sqrt{Z_{0i}}} \tag{2.1}$$

where $V_i^{\pm}$ are *RMS* voltages and Z_{0i} is a real normalizing impedance, usually chosen to be the characteristic impedance of the transmission line connected to port i. The a_i's and b_i's are complex numbers and are often called *wave amplitudes*. The waves going into the circuit are called incident, and the ones coming out are called scattered.

The magnitudes of a_i and b_i are related to power in the following way. The *total* currents and voltages expressed in terms of a_i and b_i are

$$V_i = (a_i + b_i)\sqrt{Z_{0i}} \qquad , \qquad I_i = \frac{a_i - b_i}{\sqrt{Z_{0i}}}. \tag{2.2}$$

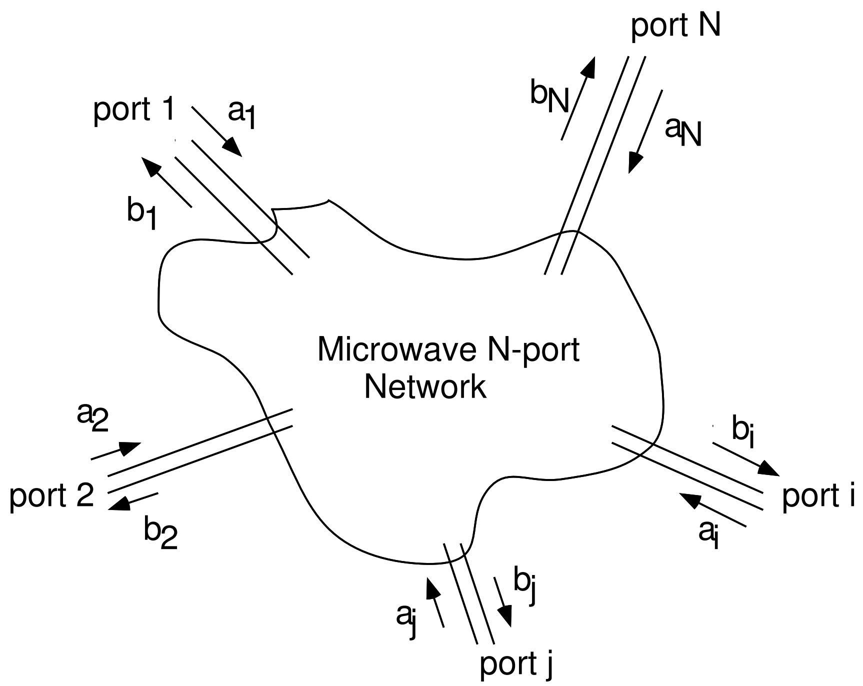


Figure 2.1: An N-port used for defining S-parameters.

while conversely,

$$a_i = \frac{1}{2}\left(\frac{V_i}{\sqrt{Z_{0i}}} + I_i\sqrt{Z_{0i}}\right) \quad , \quad b_i = \frac{1}{2}\left(\frac{V_i}{\sqrt{Z_{0i}}} - I_i\sqrt{Z_{0i}}\right) \tag{2.3}$$

Since we are using RMS quantities, the power going *into* port i is equal to

$$P_i = Re\left\{V_i\ I_i{}^*\right\} = Re\left\{(a_i + b_i)(a_i - b_i)^*\right\} = |a_i|^2 - |b_i|^2\,, \tag{2.4}$$

where the asterisk denotes the complex conjugate of a complex number. This formula means that we can interpret the *total* power going *into* port i as the incident power $|a_i|^2$ minus the scattered power $|b_i|^2$. This formula can be extended to calculate the power flowing into the entire circuit:

$$P_{1\mathrm{N}} = \sum_i P_i = \mathbf{a}^\dagger\mathbf{a} - \mathbf{b}^\dagger\mathbf{b}, \tag{2.5}$$

where $\mathbf{a}^\dagger$ is the Hermitian conjugate, that is the complex conjugate of $\mathbf{a}$ transposed. Here $\mathbf{a}$ is a column vector of order N consisting of all the a_i's. Usually this is defined as an input vector, and the vector $\mathbf{b}$ is defined as the output vector of a microwave network, and they are related by

$$\mathbf{b} = \mathbf{S}\mathbf{a}, \tag{2.6}$$

where $\mathbf{S}$ is called the *scattering matrix*.

In principle, we can measure the coefficients of the scattering matrix by terminating all the ports with their normalizing impedance, and driving port j with an incident wave a_j. All the other a_k waves

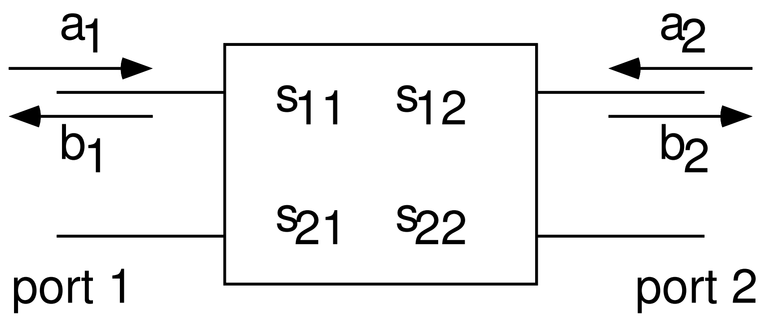


Figure 2.2: A two-port microwave network. The waves a_1 and a_2 are the input waves, the waves b_1 and b_2 are the output waves, and $\mathbf{S}$ is the scattering matrix for this network. One example of a two port network is just a section of transmission line.

will be zero, since the other terminations are matched and have no reflection. The scattering coefficients S_{ij} are then

$$S_{ij} = \left.\frac{b_i}{a_j}\right|_{a_k=0 \text{ for } k\neq j}. \tag{2.7}$$

As an example, for a typical two-port network as shown in Fig. 2.2, the scattering matrix is a 2×2 matrix, the scattering coefficient S_{11} is the reflection coefficient at port 1 with port 2 terminated in a matched load, and the scattering coefficient S_{21} is the transmission coefficient from port 1 to port 2. Mathematically, we have

$$b_1 = S_{11}a_1 + S_{12}a_2 \qquad , \qquad b_2 = S_{21}a_1 + S_{22}a_2 \tag{2.8}$$

where, following (2.7), we have

$$S_{11} = \frac{b_1}{a_1} \quad \text{when } a_2 = 0. \tag{2.9}$$

and so forth.

When modelling S-parameters of a network in a circuit analysis program such as SPICE which does not natively include the capability of handling incident and reflected waves, the following trick is often useful. Consider the circuit of Figure 2.3(a), where a load is connected at the end of a transmission line of characteristic impedance Z_0 on which an incident voltage wave v_+ is present. We can use the Thévenin equivalence theorem to replace the transmission line by an equivalent generator and equivalent Thévenin impedance. The generator voltage is found by open-circuiting the ends, and is equal to $v_{Th} = 2v_+(t)$ because the reflection coefficient of an open circuit is $+1$. The short-circuit current is $v_+(t)/Z_0$ because the *current* reflection coefficient of a short is -1 (the negative of the voltage reflection coefficient). Therefore, the Thévenin equivalent circuit of a lossless transmission line terminated in some load is shown in Fig. 2.3(b) and is given by

$$Z_{Th} = \frac{2v_+}{2v_+/Z_0} = Z_0, \quad \text{and} \quad V_{Th} = 2v_+. \tag{2.10}$$

This works in the frequency domain, and also in the time domain if the characteristic impedance Z_0 is equal to a frequency-independent resistance, as will be the case for a lossless line.

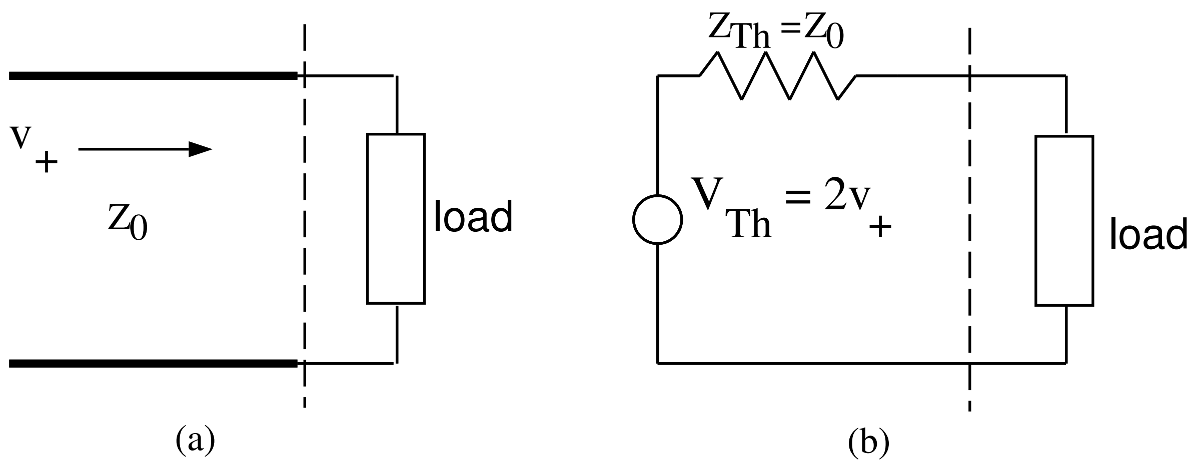


Figure 2.3: (a) Transmission line with an incident wave v_+, terminated by a load. (b) Equivalent circuit with the transmission line replaced by its Thévenin equivalent circuit.

2.2 Reciprocal and Lossless Networks

In general, a scattering matrix has many parameters that need to be determined for a specific network. For example, a 4-port network has a 4×4 scattering matrix, and in this case the network is determined by 32 real numbers (each scattering parameter is complex). Fortunately, in many cases it is possible to reduce the number of unknown coefficients knowing some of the properties of the network. One important property is reciprocity. A network is reciprocal if the power transfer and the phase do not change when the input and output are interchanged. This means that for reciprocal networks, the scattering matrices are symmetrical. In order for a network to be reciprocal, it has to be linear, time invariant, made of reciprocal materials, and there cannot be any dependent voltage or current sources in the network. For example, a transistor amplifier is not reciprocal because of the dependent current source, and you know from your circuits classes that an amplifier usually does not work well backwards. A nonreciprocal device used commonly in microwave engineering is an isolator, which contains a nonreciprocal material called a ferrite. In this case there is a static magnetic field that gives a preferred direction to the device. Isolators typically have a low loss in one direction, about 1 dB, and a very high loss in the other direction, usually about 20 dB or more. Isolators are often used to protect a transmitter, just like the one you will be using at the output of the sweepers in at least one of your lab experiments. For example, if you have a radar that is producing a megawatt (MW), in case of an open circuited output, you do not want the power to reflect back into the transmitter.

Reciprocal circuits have a symmetrical scattering matrix, which means that $S_{ij} = S_{ji}$. For example, the scattering matrix of a reciprocal two-port looks like

$$\mathbf{S} = \begin{bmatrix} S_{11} & S_{21} \\ S_{21} & S_{22} \end{bmatrix}. \tag{2.11}$$

Another simplification that can be made in a scattering matrix is when the network is lossless, which means it absorbs no power. This means that the scattered power is equal to the incident power, or mathematically

$$\mathbf{b}^\dagger \mathbf{b} = \mathbf{a}^\dagger \mathbf{a}. \tag{2.12}$$

It turns out that this is equivalent to saying that the scattering matrix is unitary, which is written as

$$\mathbf{S}^\dagger \mathbf{S} = \mathbf{I}, \tag{2.13}$$

where $\mathbf{I}$ is the identity matrix. This means that the dot product of any column of $\mathbf{S}$ with the complex conjugate of the corresponding row (or of the same column if the network is reciprocal) gives unity, and the dot product of a column with the complex conjugate of a different column gives a zero. The columns of the scattering matrix form an orthonormal set, which cuts down the number of independent S-parameters by a factor of two.

Example

As a simple example, consider the junction of two transmission lines, one (at port 1) with a characteristic impedance of Z_{01}, the other (at port 2) with a characteristic impedance of Z_{02}. Nothing else appears in the circuit. We know that if port 2 is connected to a matched load, the voltage reflection coefficient of a wave incident at port 1 is

$$\rho_1 = \frac{Z_{02} - Z_{01}}{Z_{02} + Z_{01}}$$

and the transmission coefficient is $\tau_1 = 1 + \rho_1$. now we note that

$$a_1 = \frac{V_1^+}{\sqrt{Z_{01}}}; \quad b_1 = \frac{V_1^-}{\sqrt{Z_{01}}}$$

so that

$$S_{11} = \frac{b_1}{a_1} = \frac{V_1^-}{V_1^+} = \rho_1$$

On the other hand,

$$b_2 = \frac{V_2^-}{\sqrt{Z_{02}}}$$

so that S_{21} is not quite so simple:

$$S_{21} = \frac{b_2}{a_1} = \tau_1 \frac{\sqrt{Z_{01}}}{\sqrt{Z_{02}}} = \frac{2\sqrt{Z_{01}Z_{02}}}{Z_{02} + Z_{01}}$$

If we next consider a wave incident at port 2 with a matched load connected to port 1, we get in a similar way:

$$S_{22} = \frac{Z_{01} - Z_{02}}{Z_{02} + Z_{01}}; \quad S_{12} = \frac{2\sqrt{Z_{01}Z_{02}}}{Z_{02} + Z_{01}}$$

Note that $S_{12} = S_{21}$, as we should expect for this reciprocal network.

So far, we have only talked about scattering parameters of two port networks. There are many applications when one might want to use a network with more ports: for example, if there is a need to split the power in one transmission line or waveguide into several others, or combine the power from several lines into one.

Let us say we wish to make a network that will have three ports and will be used as a two-way power splitter (or combiner). We would like this network to be lossless and matched at all ports. Such a circuit is also reciprocal if it is passive and contains no material anisotropy. From the reciprocity condition, we know that the scattering matrix has to be symmetrical, and from the matched condition we know that all three reflection coefficients at the three ports are zero, so the scattering matrix of such a device looks like:

$$\mathbf{S} = \begin{bmatrix} 0 & S_{12} & S_{13} \\ S_{12} & 0 & S_{23} \\ S_{13} & S_{23} & 0 \end{bmatrix}. \tag{2.14}$$

Now we can use the lossless condition, which tells us that the matrix is unitary. This gives us the following equations:

$$\begin{aligned} |S_{12}|^2 + |S_{13}|^2 &= 1 \qquad & {S_{13}}^* S_{23} = 0 \\ |S_{12}|^2 + |S_{23}|^2 &= 1 \qquad & {S_{23}}^* S_{12} = 0 \\ |S_{13}|^2 + |S_{23}|^2 &= 1 \qquad & {S_{12}}^* S_{13} = 0. \end{aligned} \tag{2.15}$$

The last three equations imply that at least two of the three parameters have to be equal to zero. This contradicts the first three equations. The conclusion is that a lossless, reciprocal and matched three-port is not possible.

If one of the three conditions is dropped, a feasible device can be made. For example, if we assume the device is not reciprocal, but is matched and lossless, a device called a *circulator* results. It has the property that power coming into port 1 will go out of port 2, and none will go out of port 3, power going into port 2 will only go out port 3, and power going into port 3 will only go out port 1. It is now obvious where the name comes from. This device is widely used in microwave engineering, and is physically realized by using ferrite materials, which give it a preferred direction by producing a static magnetic field in only one direction. If a three port is reciprocal and lossless, but not necessarily matched at all ports, we get a power divider, and finally, if we relax the lossless condition, we get a resistive power divider. Such resistive power dividers can be made such that $S_{23} = S_{32} = 0$ so that the two output ports are isolated. They can also be designed to have more than 2 outputs. All of these components are often used in microwave circuits.

Four-port microwave devices are also used very often, especially ones called *directional couplers*, and in waveguide systems the magic T (which is another special kind of *hybrid network*) is also used quite often. In connection with Lab 4 (multiport networks), we will study some three-ports and four-ports in more detail.

2.3 Return Loss and Standing Wave Ratio

When the load is not matched to the characteristic impedance of the line, not all of the available power from the generator is delivered to the load. This "loss" is called the *return loss* and is defined as

$$RL = -20\log|\rho|\,\mathrm{dB} \tag{2.16}$$

When a load is not matched to the line, the total voltage on the line is the sum of the incident and reflected voltages:

$$|V(z)| = |V_{0+}|\left|1 + \rho e^{-j2\beta\ell}\right| = |V_{0+}|\left|1 + |\rho| e^{j(\theta - 2\beta\ell)}\right|, \tag{2.17}$$

where $\ell = -z$ is the positive distance measured from the load at $z = 0$, and θ is the phase of the reflection coefficient at the load. The previous equation shows that the magnitude of the voltage oscillates between a maximum value of

$$V_{\max} = |V_{0+}|\ \ (1 + |\rho|), \tag{2.18}$$

corresponding to a phase term $e^{j(\theta - 2\beta l)} = 1$, and a minimum value

$$V_{\min} = |V_{0+}|\ \ (1 - |\rho|), \tag{2.19}$$

corresponding to the phase term equal to -1.

The distance between two successive voltage maxima or minima is $l = 2\pi/2\beta = \lambda/2$, and the distance between a maximum and its nearest minimum is $\lambda/4$. When ρ increases, the ratio of V_{max} to V_{min} increases, so we can measure the mismatch of a load through the *standing wave ratio* (SWR) defined as

$$SWR = \frac{V_{max}}{V_{min}} = \frac{1 + |\rho|}{1 - |\rho|}. \tag{2.20}$$

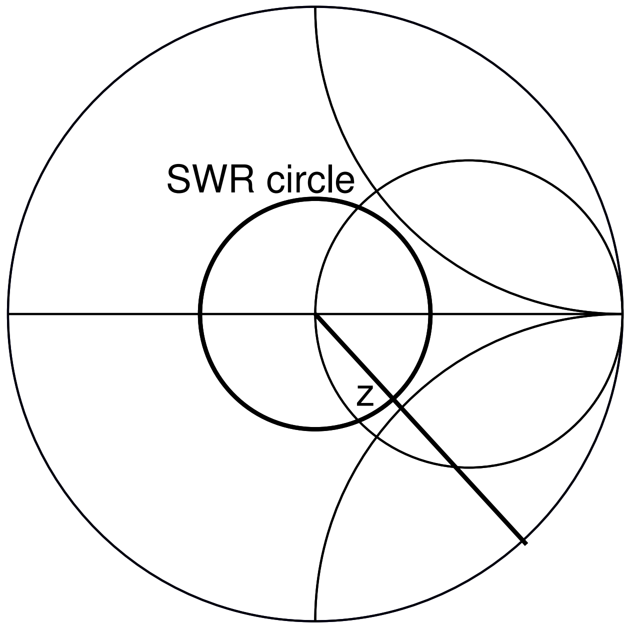


Figure 2.4: Solution to Example **??**.

Sometimes, it is also called the *voltage standing wave ratio* (VSWR). The SWR is a real number between 1 and ∞, where SWR=1 corresponds to a matched load.

2.3.1 Example

A $50\,\Omega$ line is terminated in a load impedance of $Z = 80 - j40\,\Omega$. Find the return loss in dB, SWR on the line, and the reflection coefficient at the load.

Since all impedances on a Smith chart are normalized, we first need to find the normalized load impedance:

$$z = \frac{Z}{Z_0} = 1.6 - j0.8,$$

which we plot on the Smith chart. By using a compass, we can then see that this corresponds to a reflection coefficient magnitude of $|\rho| = 0.36$. By using the same compass reading, we can find from the SWR and RL scales given below the Smith chart that SWR=2.2, and the return loss is $RL = 8.7\,\text{dB}$. The angle of the reflection coefficient, from the chart, is $-36°$. If we draw a circle through the load impedance point (centered at the origin where $\rho = 0$), the SWR can be read from the intersection with the horizontal axis for $r > 1$. Such a circle is called a SWR circle, since all points on it have the same SWR.

2.4 The Smith Chart

The Smith chart is a useful graphical way of solving transmission-line problems. It was developed by Smith, an engineer from Bell Labs, and is shown in Fig. 2.5. Do not get intimidated by the way it looks: there are a few simple rules for using it. The starting point is to realize that the reflection coefficient, which is a complex number in general, can be represented in polar form, $\rho = |\rho|e^{j\theta}$. The

Smith chart plots reflection coefficients in polar form. It also gives information about the impedances and admittances, since they are related to the reflection coefficient through a bilinear transformation. You can look up the way it is derived in any microwave book, for example in *Microwave Engineering* by Pozar. Here are a few simple rules to follow:

1. The Smith chart is a normalized graph: the largest circle represents a reflection coefficient equal to $|\rho| = 1$.

2. Circles concentric to the origin are circles of constant reflection coefficient.

3. Circles of constant resistance are not concentric. The centers of these circles lie on the horizontal diameter of the chart. Point L corresponds to a resistance of zero, which is a short circuit, and the resistance increases as you move to the right. The right end of the horizontal axis corresponds to an open circuit ($r \to \infty$). The resistance is also normalized, usually to the characteristic impedance of the transmission line: $r = R/Z_0$. Circle 2 in Fig. 2.6 shows the $r = 1$ circle. All points on this circle correspond to a resistance equal to the characteristic impedance of the line (usually $50\,\Omega$.)

4. Circles of constant (again, normalized) reactance $x = X/Z_0$ have centers on the vertical axis on the right side of the chart. The top part of the chart corresponds to inductances, and the bottom to capacitances. For example, all points on circle 3 in Fig. 2.6 have a normalized inductive reactance of $x = 1$. This means that, for a $50\,\Omega$ line at, say, 1 GHz, this is an inductance of

   $$L = 50/2\pi\,\mathrm{nH} = 7.96\mathrm{nH}.$$

   All points on circle 4 correspond to a capacitive normalized reactance of -1, which for a $50\,\Omega$ line at 1 GHz is a capacitance of

   $$C = 1/50(2\pi)\,\mathrm{nF} = 3.18\mathrm{pF}$$

   The resistance and reactance circles are orthogonal (they intersect at right angles).

5. The Smith chart is used to find impedances of loads on transmission lines from reflection coefficients, and vice versa. The scales around the periphery of the chart show distances in electrical wavelengths. This is used if you want to move down a transmission line from the generator towards the load, or up a line from the load towards a generator, shown with arrows in Fig. 2.5. The electrical distance scale covers a range from 0 to 0.5, which means that it includes the periodicity of transmission lines with half a wavelength. When you move around the chart by 2π, the impedance is the same as when you started at a point that is half a wavelength away along the transmission line. Also, if you move from a short by a quarter of a wavelength, you reach the open, and vice versa.

5. The Smith chart can also be used to directly find the SWR associated with a certain reflection coefficient by reading the amplitude of the reflection coefficient (radius of circle) off a different scale at the bottom of the chart.

Example

At the load of a terminated transmission line of characteristic impedance $Z_0 = 100\,\Omega$, the reflection coefficient is $\rho = 0.56 + j0.215$. What is the load impedance? What is the SWR?

We first need to convert the reflection coefficient to polar form: $\rho = 0.6\angle 21°$, and then plot this point on the Smith chart. You can use a compass and the scale at the bottom of the chart to set the magnitude of 0.6 with which you draw a circle centered at the origin. All points on this circle have a reflection coefficient of magnitude 0.6. Then you can specify the phase by drawing a straight line from the center of the chart to the 21° phase angle point at the outer edge of the chart. The point where the line and the circle intersect give a normalized impedance of $z = 2.6 + j1.8$, which corresponds to the actual load impedance of $Z = Z_0 z = 260 + j180\,\Omega$.

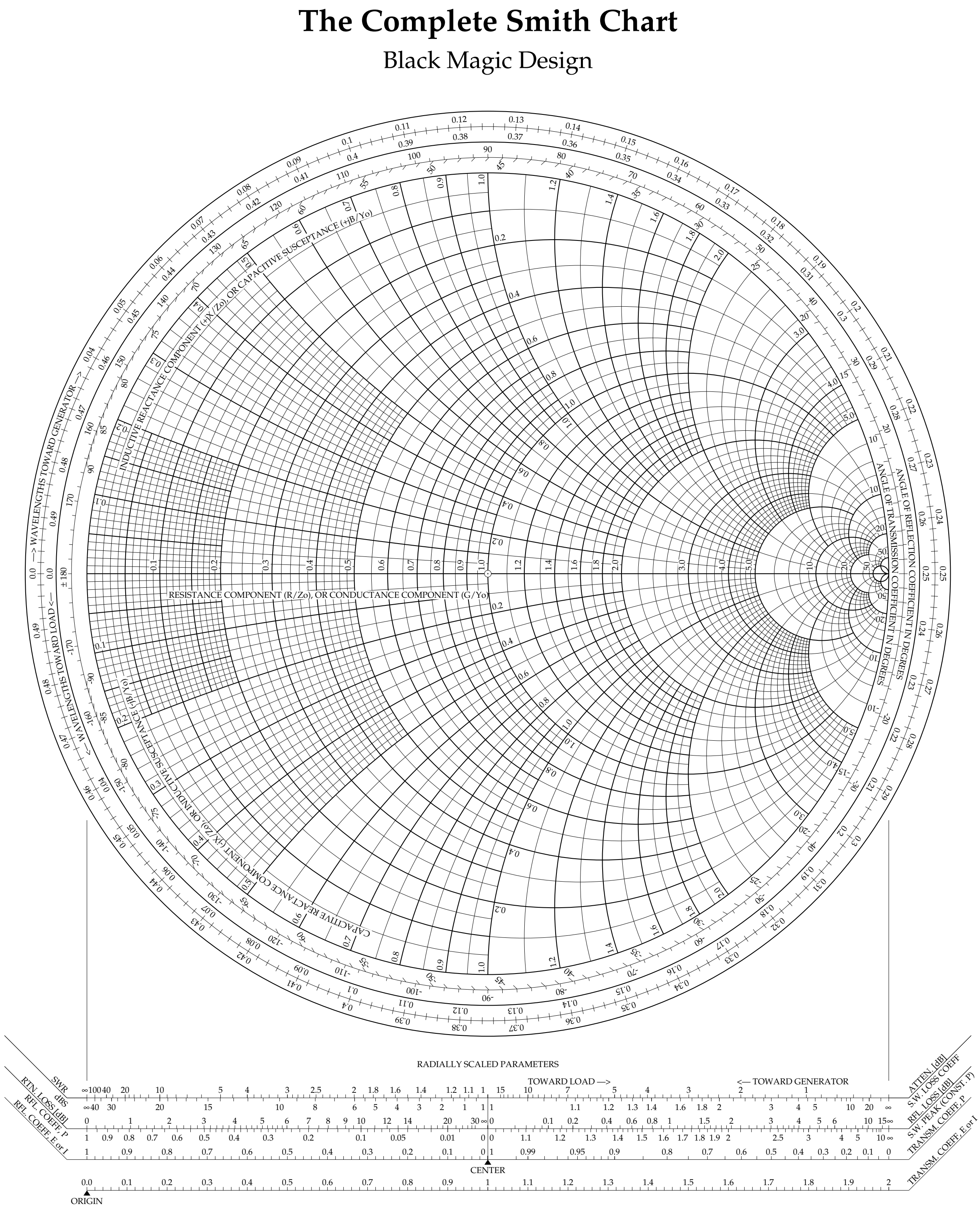


Figure 2.5: The Smith chart, with associated scales.

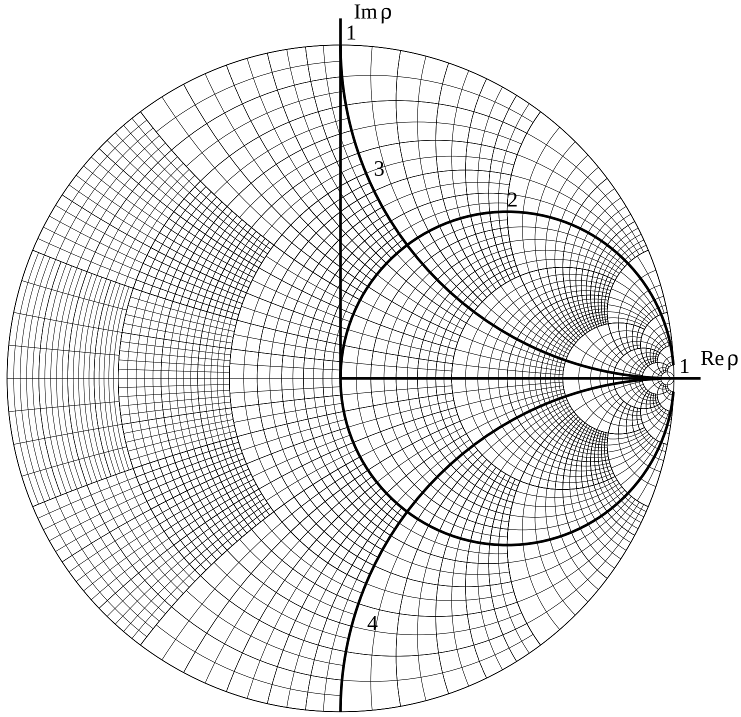


Figure 2.6: The Smith chart. Circle 2 is one of constant $r = 1$, and circles 3 and 4 correspond to lines of constant reactance $x = \pm 1$ respectively.

2.5 Admittances on the Smith Chart

So far we have looked at how the Smith chart is used with normalized impedances. In a similar way, it can be used for normalized admittances, and this comes up when you are dealing with, for example, transmission lines or other circuit elements connected in shunt. For a normalized load impedance z_L, when connected to a length $\lambda/4$ of transmission line, the input impedance is transformed to

$$z_{in} = \frac{1}{z_L}$$

and the normalized impedance is converted into a normalized admittance. Adding a quarter wave section corresponds to a 180° rotation on the Smith chart (since a full rotation corresponds to $\lambda/2$). In other words, if you flip the chart upside down, you can use it for admittances. Another way to get a normalized admittance is just to image the normalized impedance point across the center of the chart. Sometimes, an overlaid admittance and impedance chart are used, and the admittance scales are usually in a different color than those for the impedance. Which method you use is largely a matter of personal preference, so long as it is consistent and correct.

For example, let us find the input admittance to a 0.2λ long $50\,\Omega$ transmission line terminated in a load $Z_L = 100 + j50\,\Omega$. First we would locate the point $z_L = 2 + j$, Fig. 2.7, and then find y_L by 180° rotation to be $0.4 - j0.2$ (which we read off the *impedance* grid). Now we can get the input impedance by moving z_L towards the generator by 0.2λ, reading off $z_i = 0.5 - j0.49$ as shown in the figure. The admittance $y_i = 1.02 + j1.01$ can be found either by rotating z_i by 180° along the SWR circle, or by rotating y_L by 0.2λ and reading off the impedance grid.

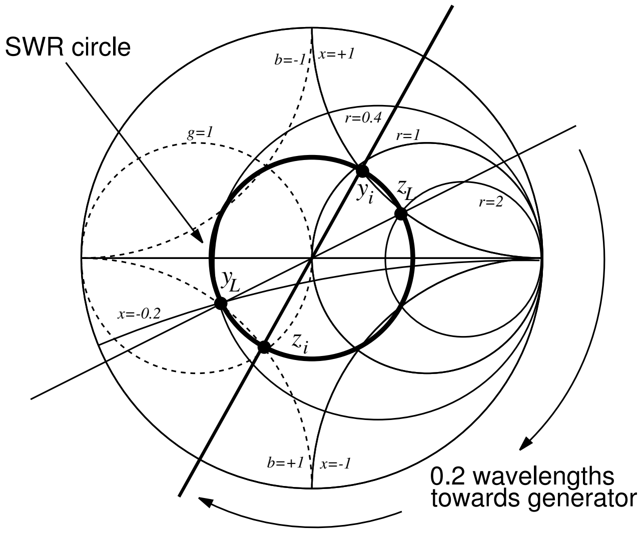


Figure 2.7: Using the Smith chart as an admittance chart. The impedance scale is shown as solid lines, and the admittance scale as dashed lines. The SWR circle is shown as a heavy line.

2.6 Impedance Matching Methods

One of the most important design tasks at microwave frequencies is impedance matching. The basic idea is illustrated in Fig. 2.8. A matching network is placed between a load and a transmission line. The matching network should be lossless so that no power is lost unnecessarily. It is usually designed so that the input impedance looking right from the plane 1-1′ is equal to Z_0. Often it is important that the matching network covers a large frequency range. If the matching network is properly designed, reflections are eliminated on the incoming transmission line to the left of the matching network, although there may be multiple reflections between the matching network and the load. The whole point of impedance matching is to maximize power delivered to the load and minimize the power loss in the feed line. It is also used for improving the signal to noise ratio of an active microwave circuit, as well as for amplitude and phase error reduction in a power distribution network (such as an antenna array feed).

Most commonly, matching is performed in one or more of the following techniques:

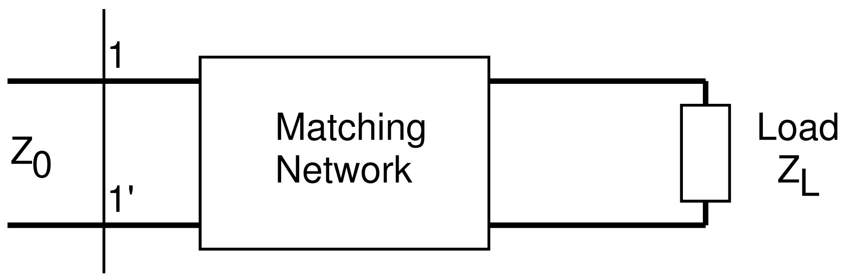


Figure 2.8: A matching network matches a load impedance to a transmission line within a certain frequency bandwidth.

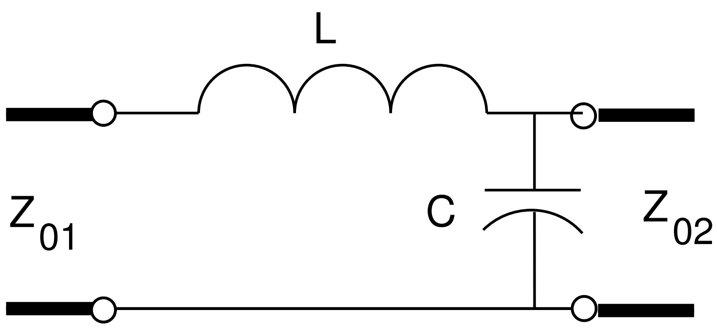


Figure 2.9: Lumped element matching example.

- matching with lumped elements
- single-stub matching
- double-stub matching
- quarter-wave section matching.

In the labs, we will look at examples of single-stub and quarter-wave section matching.

2.7 Lumped Element Matching

Matching with lumped elements is done with one or more reactive L-sections (a series and shunt reactance). A lumped element means an inductor or a capacitor, as opposed to a section of transmission line. At microwave frequencies, it is not easy to realize a pure inductance or capacitance, because what is an inductance at lower frequencies will have a significant capacitive part at higher frequencies, and vice versa. However, at lower microwave frequencies, such as those used for cellular radio, lumped element matching is common and inexpensive.

Example

Consider two lossless transmission lines with different characteristic impedances Z_{01} and Z_{02} as shown in Fig. 2.9. We wish to place a capacitor in parallel with the second line to change the real part of the overall impedance, and then put an inductor in series to cancel the imaginary part of the impedance. The result should be such that the impedance seen by the first line equals the real value Z_{01}, and therefore the reflected wave on the first line is eliminated.

It is straightforward to show that this requires:

$$Z_{01} = j\omega L + \frac{Z_{02}}{1 + j\omega C Z_{02}}$$

We multiply both sides by $1 + j\omega C Z_{02}$, and force the real and imaginary parts of the equation to hold separately. The result is:

$$\frac{L}{C} = Z_{01} Z_{02}$$

and

$$1 - \omega^2 LC = \frac{Z_{01}}{Z_{02}}$$

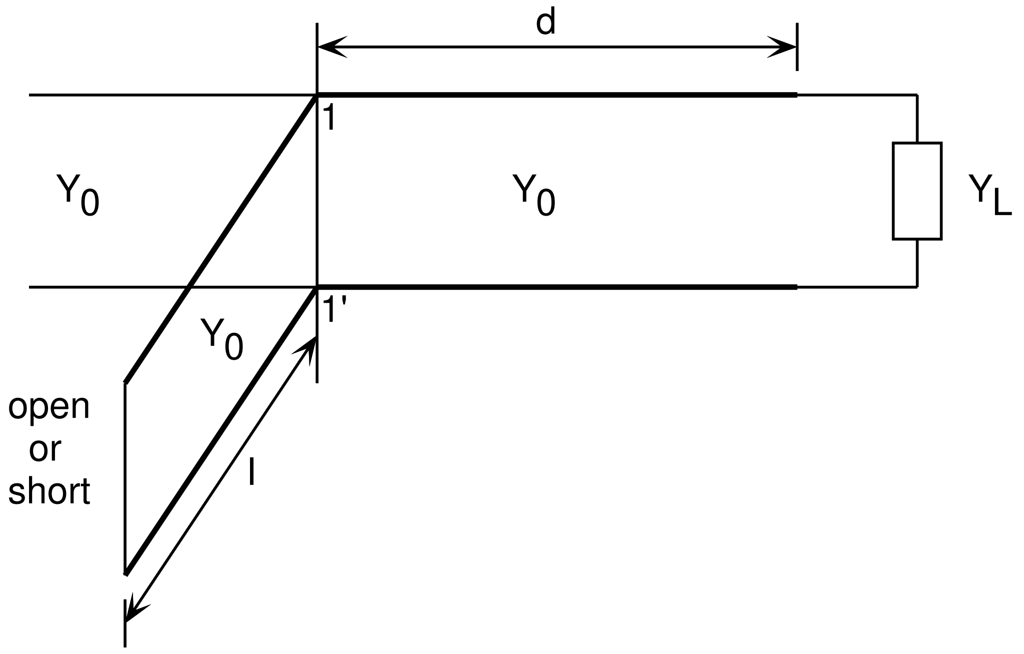


Figure 2.10: Shunt single-stub matching.

We can see from the second of these that this circuit can work only if $Z_{01} < Z_{02}$ (in the other case, we have to change the positions of the inductor and capacitor). Solving the first equation for L and substituting into the second to find C, we obtain:

$$L = \frac{\sqrt{Z_{01}(Z_{02} - Z_{01})}}{\omega}; \quad C = \frac{1}{\omega Z_{02}}\sqrt{\frac{Z_{02} - Z_{01}}{Z_{01}}}$$

Evidently, the match only works for a single frequency, if Z_{01}, Z_{02}, L and C do not change with ω. It is typical of matching networks that they only do their job at a limited set of frequencies.

2.8 Single-Stub Matching

2.8.1 Smith chart method

A shunt stub is an open or short circuited section of transmission line in shunt with the load and the line that the load is being matched to, shown in Fig. 2.10. The distance d between the stub and load needs to be determined, as well as the length l of the stub, the characteristic impedance of which is Z_0. The idea is that the distance d is selected so that the admittance $Y = 1/Z$ looking towards the load at plane $1 - 1'$ is equal to $Y = Y_0 + jB$, and then the stub admittance is chosen to be $-jB$ to tune out the reactive part of the input admittance Y, which results in a matched condition. We will solve an example of single-stub matching both analytically and on the Smith chart. Since the stub is in shunt, it is more convenient to use admittances instead of impedances.

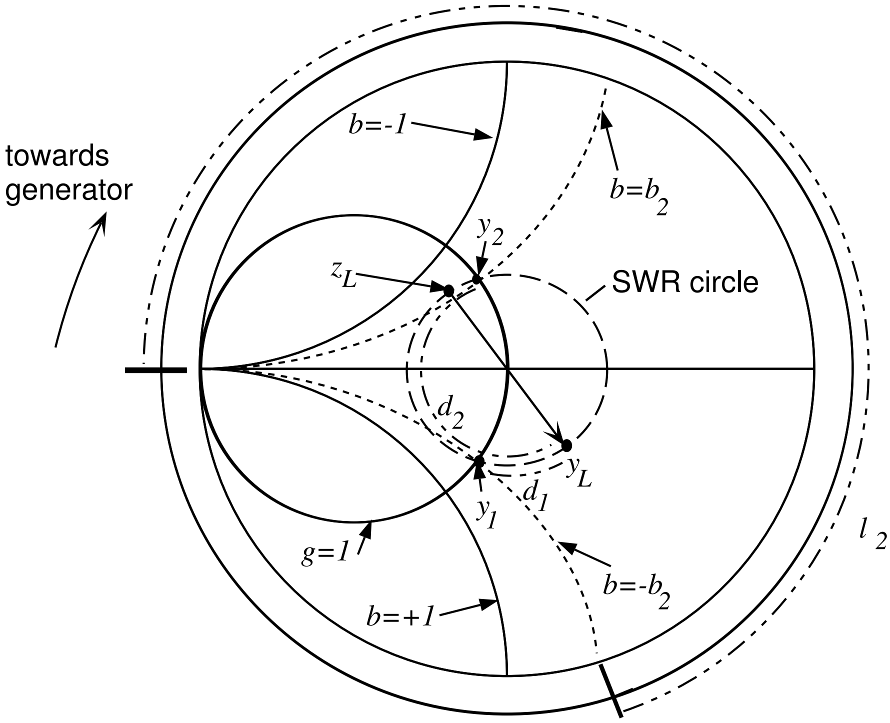


Figure 2.11: Normalized load admittance y_l.

Example

Let us match a load impedance of $Z_L = 15 + j10\,\Omega$ to a $50\,\Omega$ transmission line using a single shunt stub. We will use the admittance chart and we first plot the normalized admittance $y_L = 1/z_L$ and the corresponding SWR circle, as shown in Fig. 2.11. The two points of intersection of the SWR circle and the $g = 1$ circle give two admittance values $y_1 = 1 - j1.33$ and $y_2 = 1 + j1.33$. This means that the stub susceptance needs to be $b_1 = +1.33$ or $b_2 = -1.33$ and the stub needs to be connected at $d = d_1 = 0.325\lambda - 0.284\lambda = .041\lambda$ or $d = d_2 = (0.5 - 0.284)\lambda + 0.171\lambda = .387\lambda$ away from the load towards the generator.

The stub itself can be either open-circuited or short-circuited at its end. To find the length of an open circuited stub of susceptance b_1, we reason in the following manner: we need to end up at $y = 0$ (open circuit), so we move from $y = 0$ along the outer edge of the Smith chart ($g = 0$) to the circle corresponding to b_1 by $l_1 = 0.147\lambda$. Similarly, the second solution gives $l_2 = 0.353\lambda$.

Do these two matching networks function the same? In order to find that out, let us look at how the two matched circuits in Fig. 2.12(a) behave when the frequency changes. First we need to say at what frequency we matched the load and what the load is. Let us assume that the load is matched at 2 GHz and that it is a resistor in series with an inductor. This means that the load impedance is at 2 GHz a resistor of $R = 15\Omega$ in series with a $L = 0.796\,\text{nH}$ inductor. Now we can plot the reflection coefficient at the plane $1 - 1'$ as a function of frequency, and a sketch of this dependence is shown in Fig. 2.12(b). Solution 1 has broader bandwidth, and this makes sense since the length of the transmission lines is shorter, so there is less wavelength dependence.

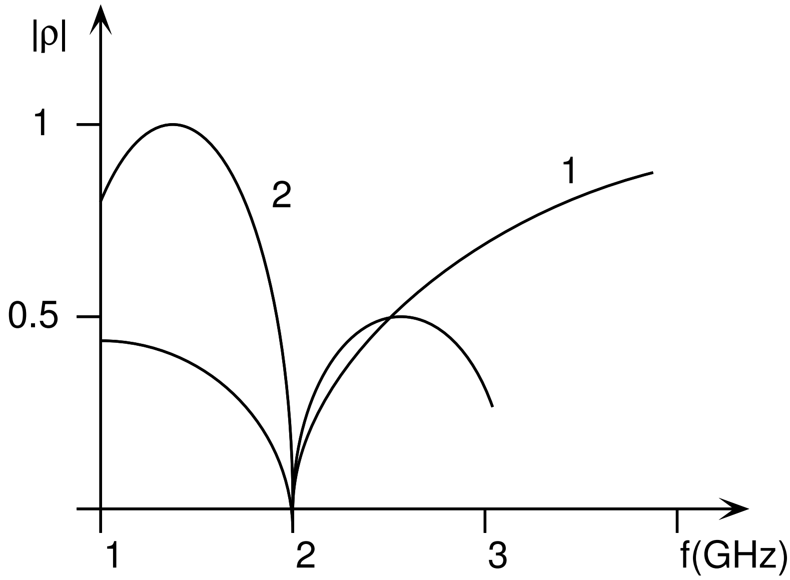


Figure 2.12: Reflection coefficient of single-stub matched circuit as a function of frequency for two different stub matching solutions.

2.8.2 Analytical method

Returning to the general stub-matching problem, the solution to this matching problem can be expressed in analytical form as follows. First, we write the load impedance as

$$Z_L = \frac{1}{Y_L} = R_L + jX_L. \tag{2.21}$$

Then the impedance Z at a distance d from the load is

$$Z = Z_0 \frac{(R_L + jX_L) + j\, Z_0 x}{Z_0 + j\,(R_L + j\,X_L)\,x}, \tag{2.22}$$

where $x = \tan \beta d$. The admittance is

$$Y = \frac{1}{Z} = G + jB \tag{2.23}$$

$$= \frac{R_L\left(1 + x^2\right)}{{R_L}^2\left(X_L + xZ_0\right)^2} + j\,\frac{R_L x - (Z_0 - X_L x)\,(X_L + Z_0 x)}{Z_0\left[{R_L}^2 + (x_L + Z_0 x)^2\right]} \tag{2.24}$$

We need $G = Y_0 = \frac{1}{Z_0}$, and from this condition we find x, and therefore d, from (2.23):

$$Z_0\,(R_L - Z_0)\,x^2 - 2X_L Z_0 x + \left(R_L Z_0 - {R_L}^2 - {X_L}^2\right) = 0 \tag{2.25}$$

or

$$x = \frac{X_L \pm \sqrt{R_L\left[(Z_0 - R_L)^2 + {X_L}^2\right]/Z_0}}{R_L - Z_0} \quad , \quad R_L \neq Z_0. \tag{2.26}$$

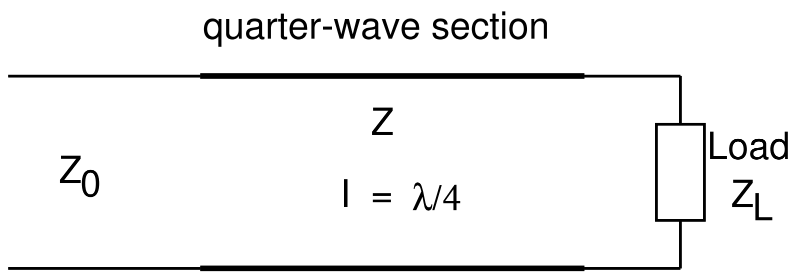


Figure 2.13: Matching with a single quarter-wave section of transmission line.

If $R_L = Z_0$, the solution degenerates to the single value $x = -\frac{X_L}{2Z_0}$. In any case, the value of d for a given solution for x is:

$$\frac{d}{\lambda} = \begin{cases} \frac{1}{2\pi} \arctan x, & x \geq 0 \\ \frac{1}{2\pi}(\pi + \arctan x), & x \leq 0 \end{cases} \tag{2.27}$$

according to whether x is positive or negative, so as to guarantee a positive value of d.

To find the length of the stub, we first plug x into (2.23) to find the susceptance $B_{\text{stub}} = -B$. For an open circuited stub,

$$\frac{\ell_{\text{open}}}{\lambda} = \frac{1}{2\pi} \arctan\left(\frac{B_S}{Y_0}\right) = \frac{-1}{2\pi} \arctan\left(\frac{B}{Y_0}\right), \tag{2.28}$$

and for a short-circuited stub

$$\frac{\ell_{\text{short}}}{\lambda} = \frac{-1}{2\pi} \arctan\left(\frac{Y_0}{B_S}\right) = \frac{1}{2\pi} \arctan\left(\frac{Y_0}{B}\right). \tag{2.29}$$

Once again, if the resulting stub length is negative, we can increase its length by $\lambda/2$ to make it positive.

2.9 Quarter-Wave Section Matching

In chapter 1 we said that quarter wave long sections of transmission line play the role of impedance transformers. In fact, they are often used for impedance matching. If a real load impedance $Z_L = R_L$ needs to be matched using a quarter-wave long section of characteristic impedance Z to a line impedance Z_0 at some frequency f, as in Fig. 2.13, the matching section will be $l = \lambda/4 = c/4f$ long, and the impedance will be equal to

$$Z = \sqrt{Z_0 Z_L}. \tag{2.30}$$

Using only a single quarter-wave section of transmission line gives a narrow bandwidth match, since it is a quarter-wavelength long only at a single frequency. This can be improved by using several cascaded sections in a matching circuit, each of which takes care of only a portion of the needed impedance change.

A drawback of this type of matching is the fact that only real impedances can be matched with sections of transmission line. A complex impedance can always be transformed into a real impedance by adding an appropriate transmission line section to it, or using a stub, but such procedures reduce the bandwidth of the match.

2.10 Measuring reflection coefficients: The Slotted Line

Time-average power is virtually the only thing about a microwave signal that can be easily measured, as you will find out in Lab 3. In this chapter, we will see how such power measurements can be used indirectly to measure voltages, impedances and reflection coefficients. One such technique uses a slotted line configuration, which enables direct sampling of the electric field amplitude (via diode detection) of a standing wave.

The slotted line is a coax or waveguide section that has a longitudinal slot into which a movable probe with a diode detector is inserted. There is a generator at one end of the line, and the unknown load terminates the line at the other end. The probe is a needle-like small post that acts as a receiving antenna and samples the electric field. (You will understand better how this works when we study antennas later.)

In slotted line measurements, we want to find the unknown load impedance (or what is equivalent, its reflection coefficient). We measure the SWR on the line and the distance from the load to the first voltage minimum l_{min}. We need to measure two quantities, since the load impedance is a complex number with both an amplitude and a phase. From the SWR, we obtain the magnitude of the reflection coefficient as

$$|\rho| = \frac{SWR - 1}{SWR + 1}. \tag{2.31}$$

We know that the voltage minimum occurs for $e^{j(\theta - 2\beta l)} = -1$; that is, when

$$\theta = \pi + 2\beta l_{min}. \tag{2.32}$$

(Note: Any multiple of $\lambda/2$ can be added to l_{min} without changing the result, since the voltage minima repeat every $\lambda/2$.) In conclusion, by measuring the SWR and l_{min}, we can find both $|\rho|$ and θ, and therefore we know

$$\rho = |\rho| e^{j\theta}. \tag{2.33}$$

From here, we can find the unknown complex load impedance to be

$$Z = Z_0 \frac{1+\rho}{1-\rho}. \tag{2.34}$$

In modern labs, network analyzers are used to measure Z's and ρ's. The slotted line is now mainly of historical interest, except at high millimeter-wave frequencies or when the expense of a network analyzer cannot be met. However, doing a few slotted line measurements in class can help you get a better feeling of some quantities than staring at the network analyzer screen.

2.10.1 Example

A $50\,\Omega$ air-filled coaxial slotted line measurement is done by first placing a short circuit at the load position. This results in a large SWR on the line with sharply defined voltage minima, as shown in Fig. 2.14(a). On some arbitrarily positioned scale along the axis of the line, the voltage minima are observed at $z = 0.2$, 2.2, and 4.2 cm. (Question: What is the frequency equal to?) This step of the procedure is actually a way of *calibrating* the measurement system and determining in what plane exactly the load is placed. The positions of the minima are called *proxy planes*.

The short is next replaced by the unknown load, the SWR is measured to be 1.5 and the voltage minima (not as sharp as with the short) are found at $z = 0.72$, 2.72, and 4.72 cm, shown in Fig. 2.14(b). From the voltage minima, knowing that they repeat every $\lambda/2$, we can find the distance $l_{min} = 4.2 -$

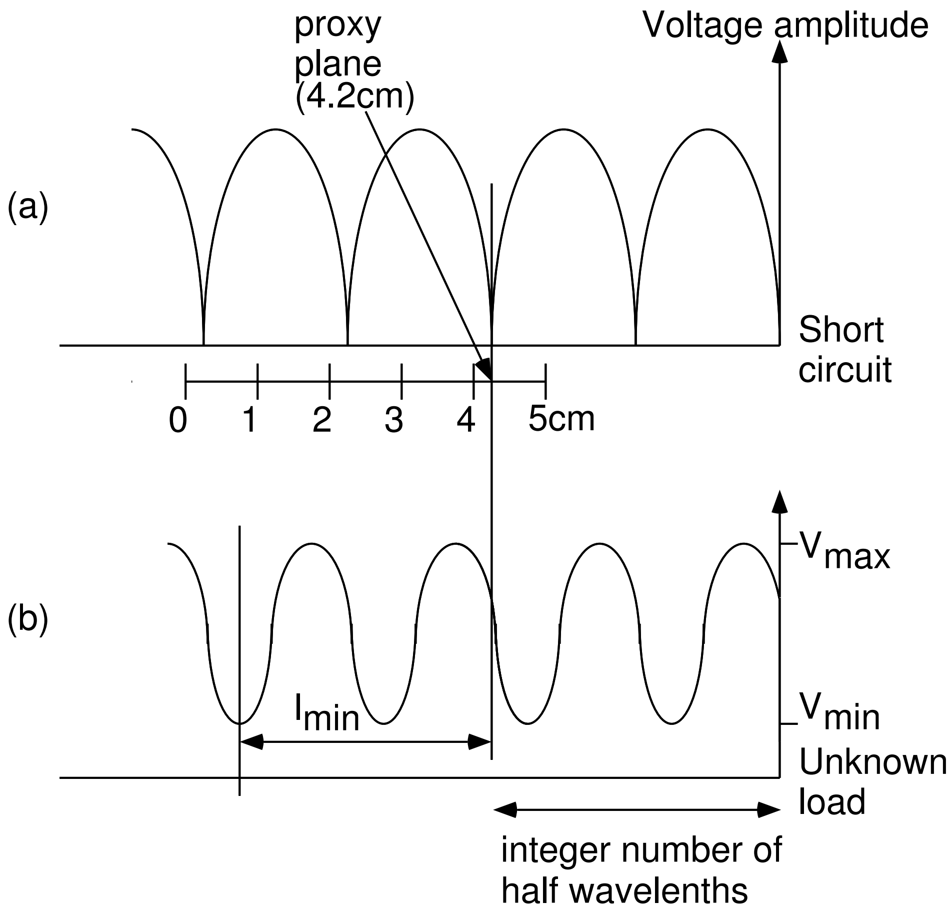


Figure 2.14: Voltage standing waves on a line terminated in a short (a) and an unknown load (b).

$2.72\,\mathrm{cm} = 1.48\,\mathrm{cm} = 0.37\lambda$. Then

$$\begin{aligned}
|\rho| &= \frac{1.5-1}{1.5+1} = 0.2 \\
\theta &= \pi + 1.48\,\mathrm{cm} \cdot 2 \cdot \frac{2\pi}{4\,\mathrm{cm}} = 7.7911\,\mathrm{rad} = 446.4^\circ = 86.4^\circ \\
\rho &= 0.2\; e^{j86.4^\circ} = 0.0126 + j0.1996 \\
Z &= 50\frac{1+\rho}{1-\rho} = 47.3 + j19.7\Omega
\end{aligned}$$

It is clear that the accuracy of this technique will be dependent on how accurately we can measure distance along the slotted line.

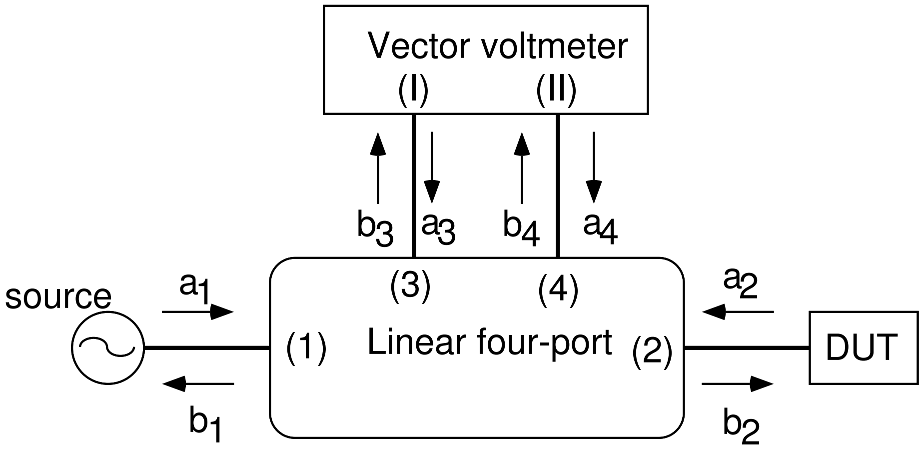


Figure 2.15: The vector voltmeter based network analyzer.

2.11 Multiport S-parameter Measurements: The Network Analyzer

The vector network analyzer is an instrument that is capable of quickly and accurately measuring S-parameters of two-port networks. It can measure impedances and reflection and transmission coefficients in the frequency domain, as well as the time domain response of microwave networks. The network analyzer is part of every modern microwave lab. Most commercial network analyzers today are made by Hewlett Packard. The most sophisticated network analyzer on the market today is the Agilent ENA and PNA network analyzer series (which have replaced the classic HP 8510C VNA). The 65-GHz version costs around $240,000, while the 110-GHz precision network analyzer costs in the neighborhood of $500,000. (2005 data). There are ways to do measurements up to 650 GHz, and the price goes up accordingly. The reason is that microwave components at higher frequencies are much harder to make with low losses, and for active devices, obtaining enough power from semiconductor devices is a problem.

The rapid variations with which the phase of a microwave signal is associated are virtually impossible for most instrumentation to follow accurately, and only power (or amplitude) is directly measurable. Phase is thus only measured indirectly: for example, a reference signal and a signal to be measured can be sent to a mixer to obtain a lower-frequency signal related to the original, and the same thing repeated with the reference signal phase shifted by 90°. The powers of the two low-frequency outputs are then enough to determine the phase of the original signal. When combined with a power measurement to determine the amplitude of the signal, we have a measurement of the complex voltage (or current). Such an instrument is called a *vector voltmeter*. Note that the vector voltmeter is really a voltage comparator for two separate voltages—the one to be measured, and a reference voltage in the present example.

A common way to perform power measurements, which is used in most commercial instruments, is to combine a linear four-port with a vector voltmeter. The vector voltmeter used in commercial network analyzers is a linear two-port device which generates a complex response that can be written in the form

$$\rho' = K_0 + K_r \frac{a_I}{a_{II}} \tag{2.35}$$

where a_I and a_{II} are the incident wave amplitudes at its two ports (I and II), and K_0 and K_r are complex constants characteristic of the vector voltmeter.

In a one-port reflection measurement, a vector voltmeter is connected to a linear four-port. The other two ports are connected to a source and the device under test (DUT), Fig. 2.15. The device described here is usually called a *reflectometer*, and the linear four-port typically consists of two directional couplers. Directional couplers are 4-ports which are not made at lower frequencies, but are often used at microwave frequencies. We will study several different implementations of directional couplers in this course.

The principle of operation of this configuration is as follows. We need to set up the equations for the wave amplitudes in the various ports, and establish the form of the relationships between them. The constants which appear in these relations will be determined by calibration, in which known loads are connected at the DUT port, and the measured values of response from the vector voltmeter are used to obtain the constants.

The source sends a wave a_1 into the four-port, which we assume to be matched to the source. Parts of this wave reach the DUT and get reflected, and parts of it get to the ports of the vector voltmeter and get reflected. The vector voltmeter samples different combinations of the waves that are incident and reflected at the DUT. Since the four-port and the vector voltmeter are linear, we can use superposition to write:

$$\begin{aligned} b_4 &= K_1 a_1 + K_2 a_2 \\ b_3 &= L_1 a_1 + L_2 a_2 \\ b_2 &= M_1 a_1 + M_2 a_2 \end{aligned} \tag{2.36}$$

where the K's, M's and L's are complex constants dependent on the s-parameters of the four-port and those of the vector voltmeter.

By definition, $a_2 = \rho b_2$, where ρ is the unknown reflection coefficient of the DUT, so we can solve for b_2 in terms of a_1 only, and b_3 and b_4 in terms of a_1 and b_2:

$$\begin{aligned} b_4 &= K_1 a_1 + K_2 \rho b_2 \\ b_3 &= L_1 a_1 + L_2 \rho b_2 \\ b_2 &= \frac{M_1}{1 - \rho M_2} a_1 \end{aligned} \tag{2.37}$$

Now from these formulas, the we can express the complex ratio b_3/b_4 in terms of the various network-analyzer-dependent constants (K_1, L_1, etc.), which left as an exercise. When you plug the result of your homework into this expression, and divide top and bottom by the coefficient of ρ in the numerator, you should get the following expression:

$$\rho' = \frac{\rho + A}{B\rho + C}. \tag{2.38}$$

where A, B and C are complex constants dependent only on the properties of the network analyzer, and not on those of the DUT. So, the final result is that the value of ρ', which the vector voltmeter gives you, is a bilinear transform of the true reflection coefficient of the device you are measuring.

The result of the measurement is independent of the level of the test signal incident at port 1, and also of the source impedance. The three unknown complex constants in (2.38) are dependent on the internal properties of the network analyzer, and can change due to component variations, changes in temperature and humidity, and so on. Thus, three known standards must be measured to determine these constants. This procedure is called *calibration* and is an essential step in the operation of the network analyzer. From the formula for ρ',

$$-A + \rho\rho' B + \rho' C = \rho. \tag{2.39}$$

This is a linear equation in the unknown constants A, B and C, so by observing three values of ρ' from three known values of ρ, we know these constants. Thus, only three calibration standards are needed for a one-port measurement in this system.

The situation is a bit more complicated when the S-parameters of an unknown two-port are to be measured. The setup for doing this is shown in Fig. 2.16. The part of the four-port connected to port 1

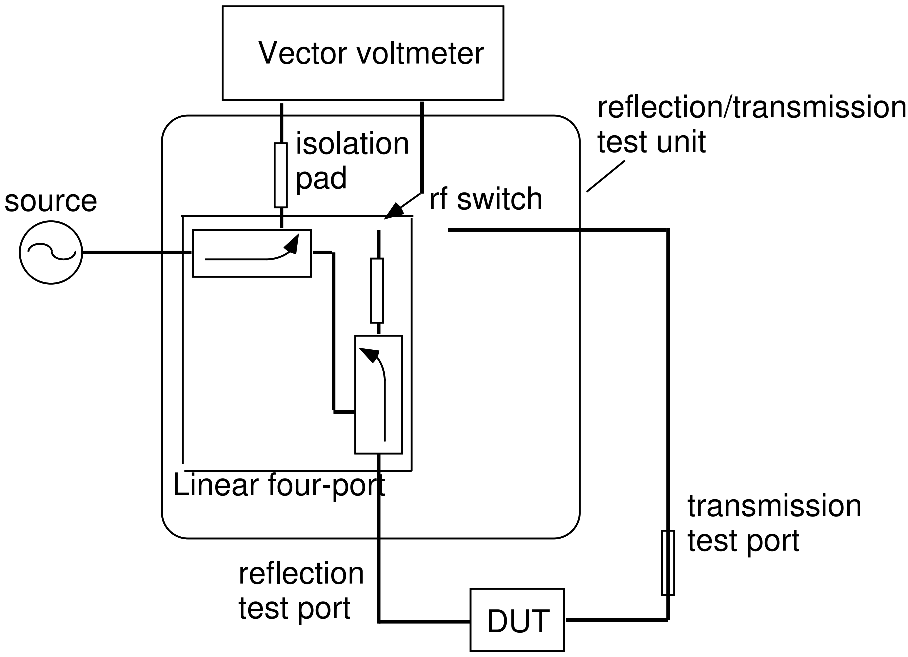


Figure 2.16: A reflection-transmission test set for full S-parameter measurements using a 4-port network analyzer.

of the DUT is the reflectometer from above (it is in the dashed box in Fig. 2.16), and another port, the "transmission-return" receives the signal from port 2 of the DUT. A coaxial switch decides which signal is to be given to the vector voltmeter for the complex ratio measurement. To get the full S-matrix, the DUT must be flipped end for end once during the measurement, and the switch switched each time to observe the reflection at both ports and the transmission in both directions through the DUT. This gives a total of four measurements. A similar derivation as in the previous case of the reflection measurement can be derived, but the expressions for the S-parameters that come out of it are very messy. It turns out that you need to measure six known standards to perform a calibration. It is important to understand that in a network analyzer, these measurements are done at many frequencies (usually swept between a lower and upper limit), so you also need to measure your standards at all these frequencies of interest.

A full S-parameter test set, such as the ones used in the lab, consists of two transmission-reflection test sets like the one discussed above, placed back to back. Another coaxial switch selects which end of the setup receives the test signal. There is no need to flip the device, and this improves measurement speed and accuracy. The network analyzer contains a computer that does all the computations necessary to determine all the S-parameters, as well as many other parameters that you might be interested in.

In summary, four-port network analyzers measure the S-parameters of two-ports by using a single vector voltmeter with an arrangement of switches and directional couplers to route the appropriate signals to the voltmeter. The S-parameters of the measured two-port can be then found from the measured data after a bit of mathematical manipulation. If you wish to measure a 3-port or 4-port, typically all but two ports at a time are terminated in matched impedances and the 2-port measurements are repeated the necessary number of times. How many times do you need to repeat a two-port measurement to characterize a 4-port network?

2.12 Practice questions

1. A measurement of a two-port gave the following S-matrix:

$$\mathbf{S} = \begin{bmatrix} 0.1\angle 0^\circ & 0.8\angle 90^\circ \\ 0.8\angle 90^\circ & 0.2\angle 0^\circ \end{bmatrix}$$

   Determine if the network is reciprocal and whether it is lossless.

2. In a common-source amplifier, define the S-parameters and relate them to quantities you have studied in circuit analysis.
3. Prove that a three port network cannot be matched, reciprocal and lossless simultaneously.
4. Show that a four port network can satisfy the matched, reciprocal and lossless conditions simultaneously.
5. What is an SWR?
6. Sketch the standing waves on a shorted and opened line. What is the SWR equal to? If the short and open are not ideal, what does the standing wave look like, and what do the SWR's become in that case?
7. A coax transmission line with a characteristic impedance of $150\,\Omega$ is $l = 2$ cm long and is terminated with a load impedance of $Z = 75 + j150\,\Omega$. The dielectric in the coax has a relative permittivity $\epsilon_r = 2.56$. Find the input impedance and SWR on the line at $f = 3$ GHz.
8. What is the Smith chart? Which quantities can you plot on it?
9. What do concentric circles centered at the middle of the chart represent?
10. What do circles of constant resistance and those of constant reactance look like?
11. Which part of the chart corresponds to real impedances, and which to imaginary ones? Which part of the chart corresponds to capacitances, and which to inductances? What if you looked at the admittance instead of the impedance chart?
12. Derive the equations for the constant resistance and reactance circles. Start by writing $\rho = u + jv$, $z = r + jx$ and $\rho = (z-1)/(z+1)$.
13. Why does a full rotation around the chart correspond to half a wavelength?
14. Pick a point on a Smith chart normalized to $50\,\Omega$. Find the value of the complex impedance associated with that point, as well as the value of resistance and inductance/capacitance at 5 GHz.
15. If a quarter-wave line is terminated in an open circuit, convince yourself that moving back towards the generator gives you a short. Do the same for a high and low impedance termination, so you can see the transformer effect of a quarter-wave line.
16. Why is impedance matching so important at high frequencies? Why did you not have to worry about it in your circuits classes?
17. When would you consider using lumped elements for impedance matching?
18. When is it convenient to use a quarter-wave transmission line section for matching?
19. Derive the expressions for the input impedance and admittance of a (a) shorted stub, and (b) open stub. Both stubs are l long and have impedances Z_0.

20. There are two solutions for single-stub matching. Which one do you chose and why? Is the answer always obvious?

16. Sketch the four-port reflectometer part of a network analyzer. Why do you need the linear 4-port?

21. Write down the calibration procedure (find the constants A, B and C) for a short-open-load calibration. How would you do a calibration if you only had one of the standards, say the short?

22. The network analyzer has 2 coaxial cables coming out of it. These are two ports. How would you measure the parameters of a three-port circuit using the network analyzer? How would you fully characterize a four-port device using a two-port network analyzer?

23. Why is the voltage at the termination Z of a transmission line with characteristic impedance Z_0 equal to

$$v = 2\frac{v_+ Z}{Z + Z_0}?$$

2.13 Homework Problems

1. Find the S-parameters of the networks (a) and (b) of Fig. 4.8. Are these networks lossless?

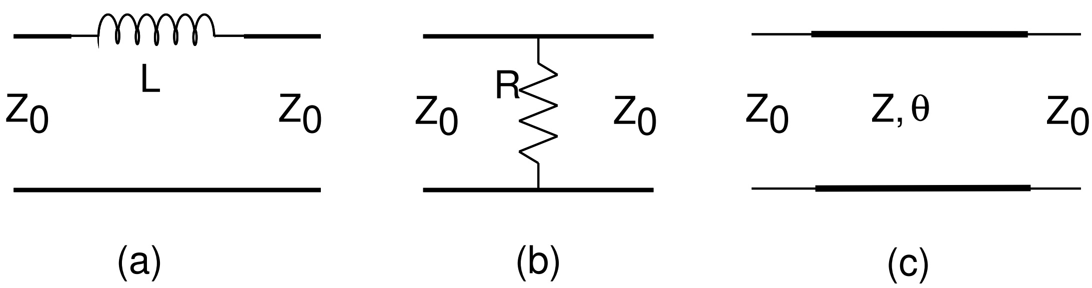


Figure 2.17: Determine the S-parameters of the two-port networks shown in (a), (b) and (c).

Reciprocal? Matched? Check for these properties using the criteria derived in the notes and your calculated S-parameters. (Do your answers make sense?)

2. Repeat problem 1 for network (c) of Fig. 4.8, which is a length l of transmission line with characteristic impedance Z connected between two lines of characteristic impedance Z_0. The "electrical length" θ of the middle section is equal to βl.

3. Repeat problem 1 for the case of the T-network shown in Fig. 2.18.

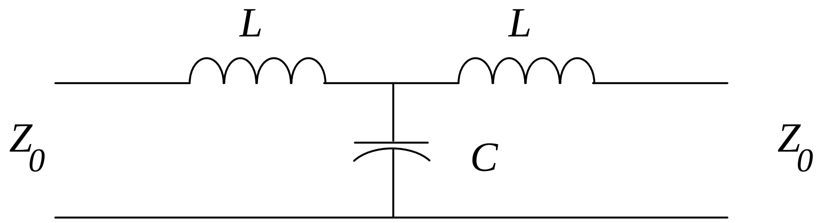


Figure 2.18: Determine the S-parameters of the two-port T-network shown.

4. Calculate and plot the magnitude and phase of S_{11} between 1 and 3 GHz for a 50 Ω line, open-circuited at the load end, that is $\theta \equiv \beta\ell = 90°$ long ($\ell = \lambda/4$) at 2 GHz.

5. Repeat problem 4 for the case of a line loaded with an impedance of $Z_L = 25$ Ω at the load end.

6. Using dependent voltage sources and passive lumped elements, model an ideal circulator (for which $S_{21} = S_{32} = S_{13} = 1$, while all other S-parameters are zero) in Designer or SPICE. Connect a source to port 1 that produces a unit incident wave, and compute the waves emanating from ports 1, 2 and 3 to demonstrate your model.

7. When we measure an S-parameter such as S_{11} in dB, should we use $10 \log |S_{11}|$ or $20 \log |S_{11}|$ to calculate it? If we measure a value of S_{11} equal to $0.5 - j0.7$, what is its value in dB? If we have an S_{12} of -15 dB, what is its value in absolute numbers (can you give a complete answer to this question or not)?

8. A low-noise amplifier has a (power) gain of 6.5 dB. An input signal of 5 mW is applied to this amplifier, while its output is fed to port 1 of an isolator whose S-parameters are: $S_{11} = 0$, $S_{21} = 0.9$, $S_{12} = 0$ and $S_{22} = 0$. Find the output power emerging from port 2 of the isolator, expressed both in dBm and in mW. Assume the input port of the amplifier is reflectionless.

9. Refer to problem 8 of Lecture/Lab1. Use 9 sections of transmission line whose electrical length is 20° each. Perform the Ansoft Designer (or SPICE) simulation of a 50 Ω transmission line terminated by a short circuit. This is to be regarded as the calibration step of a slotted line "measurement". What is the position of any proxy planes?

Now plot voltage magnitude vs. position for the case of a load impedance $Z_L = 100$ Ω terminating the line. What is the distance between the position of the voltage minimum and the proxy plane? What is the SWR for this situation? Is its value consistent with the value of Z_L you used?

10. Repeat problem 1, using a load impedance of $Z_L = 30 - j80$ Ω instead.

11. A 50 Ω air-filled coax slotted line measurement was done by first placing a short circuit at the place of the load. On an arbitrarily positioned scale along the coax, the voltage minima are observed at $z = 0.2$, 2.2, and 4.2 cm. The short is then replaced by the unknown load, the SWR is measured to be 1.5 and the voltage minima (not as sharp as with the short) are found at $z = 0.72$, 2.72, and 4.72 cm. Find the impedance of the unknown load **using a Smith chart**. Explain all your steps clearly. *Note that this same example is solved in your notes analytically. The numbers are the same so you can check your result, but you need to do this problem graphically on a Smith chart, to receive any credit.*

12. Use a shorted single stub to match a 200 Ω load to a 50 Ω transmission line. Include a Smith chart with step-by-step explanations.

13. Match a 25 Ω load to a 50 Ω transmission line at 2 GHz using (a) a single quarter-wave section, or (b) two quarter-wave sections connected in cascade. Model these circuits in Designer or SPICE, and plot the magnitude of the reflection coefficient for each from 1 GHz to 3 GHz.

14. Use an open-circuited single stub to match a $20 - j15$ Ω load to a 50 Ω transmission line at $f = 2$ GHz. There are two possible connection positions for the stub: give both solutions. After you have finished your match design, compare the two designs by modeling them in Designer or SPICE, plotting the magnitude of the reflection coefficient over the range of 1 GHz to 3 GHz. Which one has broader 10-dB bandwidth and what are the respective bandwidths? Why?

15. Use lumped elements (capacitors and/or inductors) to match an impedance of $Z_L = 100 + j75$ Ω to a 50 Ω transmission line using Designer or SPICE. Do the match at the 900-MHz cellular phone frequency. Use either a series inductor and a shunt capacitor, or a series capacitor and a shunt inductor to provide the match. Plot the magnitude of the reflection coefficient vs. frequency for 800 MHz $< f <$ 1200 MHz.

16. From the formulas (2.37) of the lecture notes, derive an expression for the ratio b_3/b_4. Then plug in expression (2.35) for ρ' and derive (2.38). What would you pick as known standards to do the calibration? For those standards, determine the expressions for the unknown constants.

Chapter 3

Microwave Power Measurement

3.1 Power Definitions

At lower frequencies, voltage and current can be measured easily, and the power is usually obtained from voltage and current. At microwave frequencies, it is hard to measure voltages and currents for several reasons: the voltages and currents change along a transmission line, and in waveguides currents and voltages do not make physical sense. On the other hand, the power flow is the same at any point in the transmission line or waveguide, and this becomes the logical quantity to measure. From an application point of view, power is critical factor. For example, for a communication system transmitter, twice the power means twice the geographic coverage (or 40% more range). Power is expensive at high frequencies, since it is difficult to make fast low-loss active devices (transistors), and tube sources are large and have limited lifetimes. It becomes therefore important to keep track of the generated power and develop means to optimize the amount that is delivered to the load (e.g. an antenna in a wireless system).

Often, microwave engineers refer to some application as a "low-power", "medium-power" or "high-power" one. Usually, what is meant is a range of power levels for each case, approximately equal to 0 – 1 mW, 1mW – 10 W and larger than 10 W, respectively. Power is most commonly expressed in dB's which are relative units. In the low-power experiments in our lab, it will be expressed relative to 1 mW: $P_{dB} = 10\log(P_{mW}/1mW)$. (Why is the logarithm (base 10) multiplied by 10 and not 20?)

Starting from fields, in terms of the electric and magnetic field, the power flow is defined by the Poynting vector:

$$P = \oint_s \left(\vec{E} \times \vec{H}\right) \cdot d\vec{s} = \oint_s \vec{S} \cdot d\vec{s}. \tag{3.1}$$

We can again notice the relationship between fields and circuits –the formula corresponds to $P = VI$.

Usually at microwave frequencies one measures and calculates *average power*. Starting from the definition of power at lower frequencies as the product of current and voltage, we notice that this product varies during an AC cycle. The average power represents the DC component of the power waveform. What is this power averaged over? The definition of power you have learned in your physics courses is rate of flow (or transfer) of energy in time, so it makes sense to average power over time. Now the question is: how long do we need to average? Usually, the averaging time is equal to many periods of the lowest frequency component of a signal we are measuring. This can be written as

$$P_{\text{av}} = \frac{1}{nT_L} \int_0^{nT_L} v(t)\ i(t)\ dt, \tag{3.2}$$

where T_L is the period of the lowest frequency component, and n is much larger than 1. In the case of a CW (continuous wave) signal, there is only one frequency, and if V and I are the RMS phasors corresponding to the voltage and current, we find that

$$P_{\text{av}} = \text{Re}\,(VI^*) \tag{3.3}$$

where * denotes the complex conjugate. In the case of an amplitude modulated signal, T_L is the period of the modulation signal, and for a pulse modulated signal, T_L is the repetition rate of the pulse.

In radar applications, pulsed modulation is used very often, and in this case it is often convenient to use the *pulse power* instead of average power. The pulse power is obtained when the energy transfer rate is averaged over the duration of the pulse, or the pulse width T_d:

$$P_{\text{pulse}} = \frac{1}{T_d}\int_0^{T_d} v(t)\ i(t)\ dt. \tag{3.4}$$

Usually, T_d is defined as the time between the one-half amplitude points. For rectangular pulses, we can measure the average power, and get the pulse power if we know the duty cycle (duty cycle = pulse width × repetition frequency):

$$P_{\text{pulse}} = \frac{P_{\text{ave}}}{\text{Duty Cycle}}. \tag{3.5}$$

How is average power measured? There are three devices for measuring average power used in modern instrumentation: *the thermistor, the thermocouple, and the diode detector*. The basic operation of all three devices is that they turn microwave power into a measurable DC or low frequency signal. Typically, there is a sensor (which contains one of the three devices from above) connected with a coaxial cable or waveguide to a power meter, which does the post-processing of the received DC signal.

3.2 The Thermistor

The thermistor falls into the category of *bolometers*. A bolometer is a device that converts RF power into heat, and this changes its resistance, so the power can be measured from the change in resistance. In the early days of microwaves, so called baretters, another type of bolometers, were used. A baretter consists of a thin metal wire (usually platinum) that heats up and its resistance changes. Baretters are usually very small, and as a consequence they are able to detect very low power levels, but they also burn out easily. A thermistor is in principle the same as a baretter, but instead of metal, a semiconductor is used for power detection. The main difference is that for a baretter, the temperature coefficient is positive, which means that the resistance grows with temperature, whereas for a thermistor it is negative.

Thermistors are usually made as a bead about 0.4 mm in diameter with 0.03 mm wire leads. The hard part in mounting it in a coax or waveguide is that the impedances have to be matched, so that the thermistor absorbs as much power as possible over the frequency range that is desired. Fig. 3.1 shows characteristic curves of a typical thermistor. You can see that the dependence of resistance versus power is very nonlinear, and this makes direct measurements hard. How are power meters then built? You will build a bulky thermistor-based power meter in the lab and calibrate the nonlinear responsivity of the thermistor.

The change in resistance of a thermistor due to the presence of RF or microwave power can be used to make a power meter for measuring that RF power level. A clever way to do this is to keep the resistance of the thermistor constant with a balanced resistive bridge circuit, as shown in Fig. 3.2. The thermistor (R_T) is DC-biased through the bridge to a known value of resistance. When there is incident RF power applied to the thermistor, its resistance tends to change, thereby unbalancing the bridge. When this happens, a voltage difference appears across the input terminals of an op-amp, whose output is used to change the DC voltage applied across the bridge in such a way as to re-balance the bridge. The new DC bias will be such that there is less DC power dissipated in the thermistor, by an amount equal to the RF power incident on it, since the balance of the bridge was maintained. This power is readily calculated from knowledge of the DC voltage across the bridge, and is translated into a power

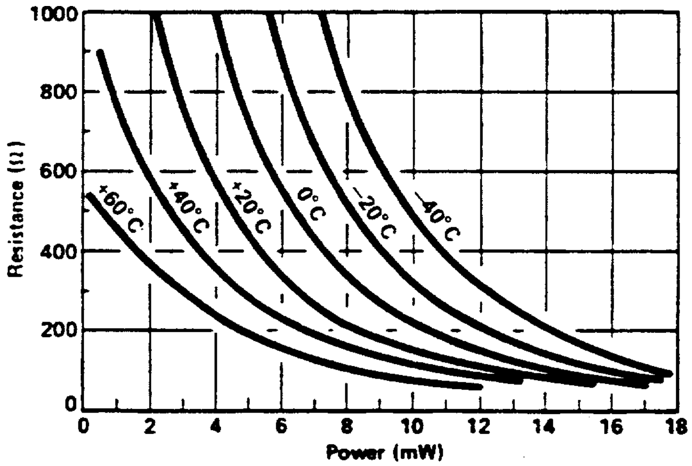


Figure 3.1: Characteristic curves of a typical thermistor bead.

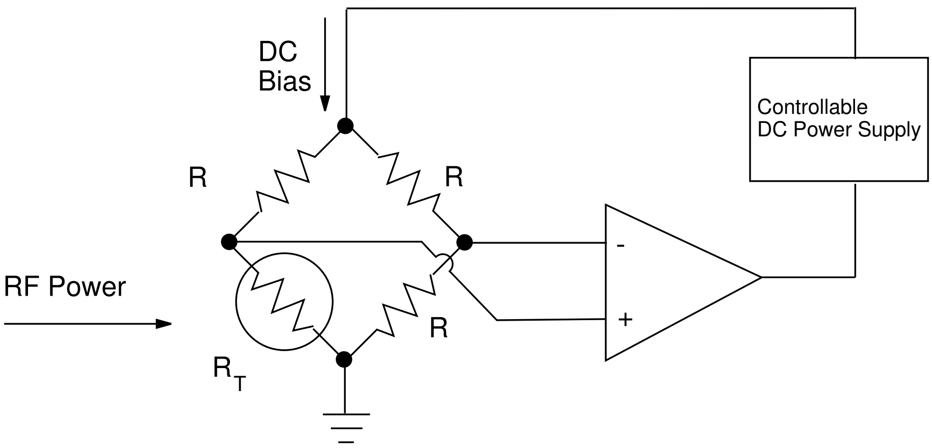


Figure 3.2: Diagram of a self-balanced Wheatstone bridge.

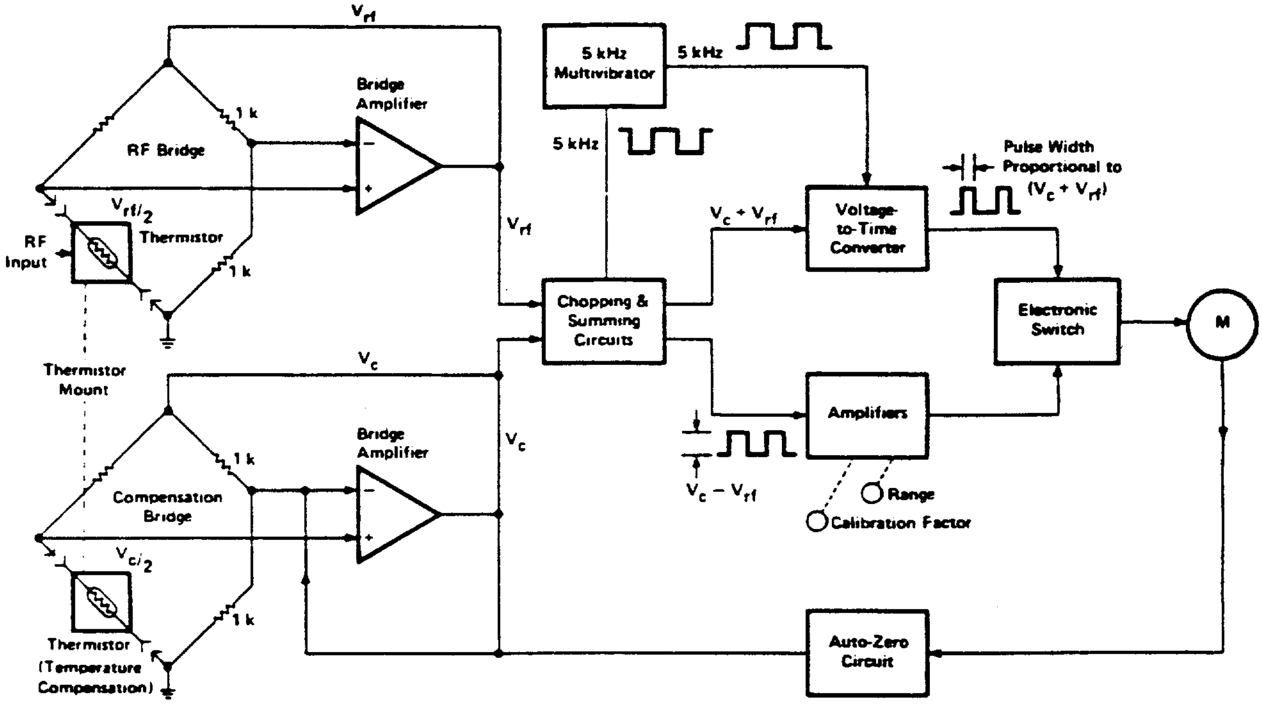


Figure 3.3: Diagram of the HP 432A power meter.

level displayed on the power meter scale. The main drawback of this scheme is that a change in ambient temperature (for example, the engineer touching the thermistor mount) would change the resistance, and make the measurement invalid. In modern thermistors this problem is solved in the mount itself, using an additional thermistor that senses the temperature of the mount and corrects the bridge reading.

A diagram of the thermistor-based HP 432A power meter is shown in Fig. 3.3. The main parts are two self-balanced bridges, a logic section, and an auto-zero circuit. The RF bridge, which contains the thermistor that detects power, is balanced by automatically varying the DC voltage V_{rf} which appears across the thermistor in this bridge. The compensating bridge, which takes care of temperature compensation, is balanced with the DC voltage V_c. If one of the bridges is unbalanced, an error voltage is applied to the top of the bridge, which causes the thermistor to change resistance in the direction required to keep the balance. The power meter is zero-set by making V_c equal to V_{rf0}, which is the value of V_{rf} in the absence of RF input power. When RF power is applied to the detecting thermistor, V_{rf} decreases, so that

$$P_{rf} = \frac{{V_{rf0}}^2}{4R} - \frac{{V_{rf}}^2}{4R} = \frac{{V_c}^2 - {V_{rf}}^2}{4R} = \frac{1}{4R}\left(V_c - V_{rf}\right)\left(V_c + V_{rf}\right). \qquad (3.6)$$

where R is the value of the fixed resistors in the bridge circuits (= 1 kΩ in Fig. 3.3). The meter logic circuit monitors the value of the right side of (3.6). Ambient temperature changes cause changes in both V_c and V_{rf}, such there is no change in P_{rf}. In the experiment, you will balance the bridge manually.

3.3 The Thermocouple

The thermocouple detector is used in the HP8481 sensor series, which you will use in the lab together with the HP437 power meters. Thermocouples generate rather low DC signals compared to thermistors. Due to progress in thin-film technology over the last few decades, they are now parts of most modern microwave power-measurement instruments. Standards are still, however, traced to thermistor power sensors, due to their stability.

A thermocouple generates a voltage due to temperature differences between its two ends. A *very* simplified physical explanation is the following: when one end of a piece of metal is heated, more electrons are made free to move. Due to diffusion, they will move away from the heated end and towards the cold end of the piece of metal, and leave extra positive charges behind. This separation of charges causes and electric field. Thermal equilibrium is reached when the Coulomb force on the charges is of equal amplitude to the force caused by diffusion. The electric field in steady-state can be integrated to find a voltage between the two ends of the piece of metal. This voltage is called the Thomson EMF. The same principle applies at a junction of two metals with different free-electron densities. Here diffusion causes the Peltier EMF. A schematic of a thermocouple, Fig. 3.4 shows that it combines the Thomson and Peltier effects. It consists of a loop of two different materials. One junction of the materials is exposed to heat, and the other is kept cold. A sensitive voltmeter is inserted into the loop. In order to have a larger value of the EMF, several thermocouples can be connected in series, or a thermopile. However, at microwave frequencies, large thermocouples also have large parasitics (inductances and capacitances), so thin film thermocouples have been developed for the microwave frequency range.

The thermocouple used in the HP8481A sensor head is shown in Fig. 3.5. It consists of a silicon chip which has one part that is n-doped and acts as one of the metals of the thermocouple. The other metal is a thin film of tantalum nitride. This thin film is actually a resistor, which can be tailored for a good impedance match of the thermocouple to the cable (50 or 75 Ω, for example). The resistor converts RF energy into heat. In the process of fabricating this thin-film sensor, a silicon-dioxide layer is used as an insulator between the silicon and the resistive film. A hole is then made in this insulator so that the resistor contacts the silicon and forms the hot junction. The cold junction is formed by the resistor and the outside edges of the silicon substrate.

As the resistor converts RF energy into heat, the center of the chip gets hotter than the outside edges (why?). Thus, there is a thermal gradient which gives rise to a thermoelectric EMF. The thermocouple chip is attached to a planar transmission line deposited on a sapphire substrate (sapphire is a good thermal conductor). The planar transmission line has a transition to a coaxial connector (outside world). The thermoelectric voltage is very low – microvolts per measured milliwatts of RF power. However, it does not change very much with ambient temperature, which is illustrated in Fig. 3.6.

Power meters which use a thermocouple, such as the HP437 which you will use in the lab, are built to detect very small voltages. The fundamental principle is the same as a lock-in amplifier: the small signal is chopped (amplitude modulated) at a low frequency in the sensor itself, (on the order of a kHz). Then the voltage is synchronously detected (i.e. demodulated) at the other end of the sensor cable (i.e. in the instrument). Because of the low DC voltages, the thermocouple does not have enough sensitivity to measure small RF power levels. For this purpose, a semiconductor diode is usually used.

3.4 The Diode Detector

A semiconductor diode is a rectifier, as you have learned in your circuits classes. At microwave frequencies, special diodes have to be used because low frequency diodes are not fast enough to detect nanosecond changes in time (corresponding to GHz frequencies). These diodes are called Schottky diodes, and they are basically a metal contact on a piece of gallium arsenide (GaAs). They can measure powers as low as -70 dBm (100 pW) up to 18 GHz. Essentially the same circuitry can be used to make a power meter with a Schottky diode as is used with a thermocouple.

A diode is a nonlinear resistor for which the current-voltage dependence obeys approximately the following equation, shown graphically in Fig. 3.7:

$$I(V) = I_S \left(e^{\alpha V} - 1 \right) . \tag{3.7}$$

where I_S is the so called saturation current, and is small (between 1 μA and 10^{-15} A), and $\alpha = n/(25\,\text{mV})$

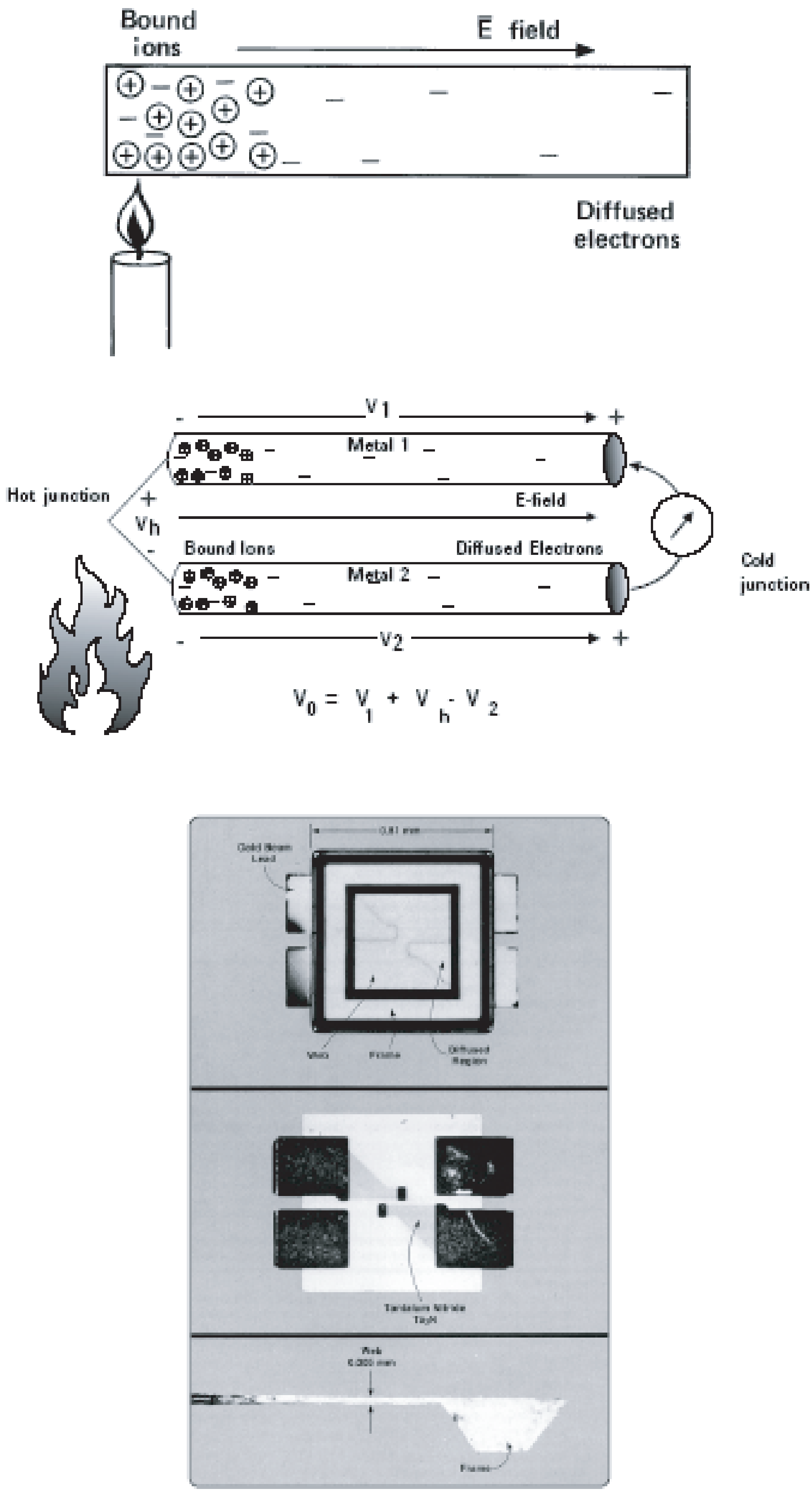


Figure 3.4: Simplified explanation of thermocouple operation. The total measured EMF is a resultant of several thermoelectric voltages generated along the circuit. The bottom figure shows a photograph of a thermocouple silicon chip.

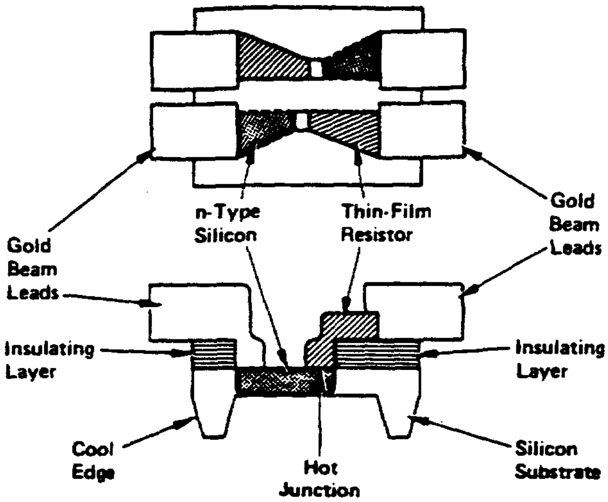


Figure 3.5: Structure of the thin-film thermocouple used in the HP8481A sensor.

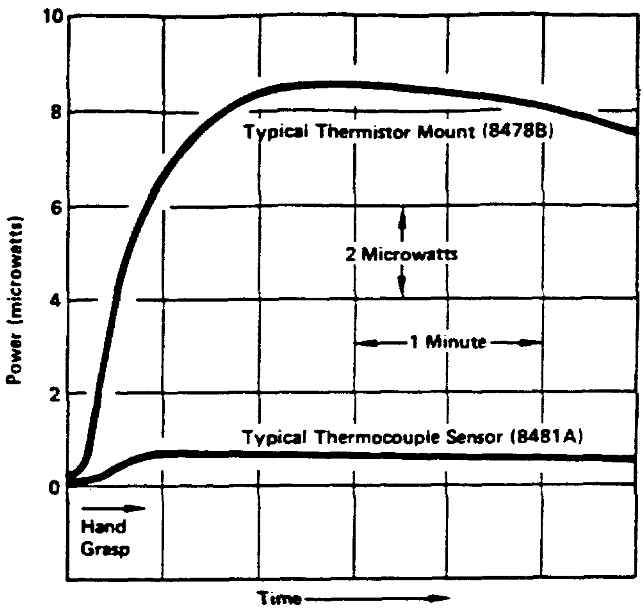


Figure 3.6: Difference in behavior of thermistor and thermocouple when grasped by the hand. (From HP Application Note 64-1.)

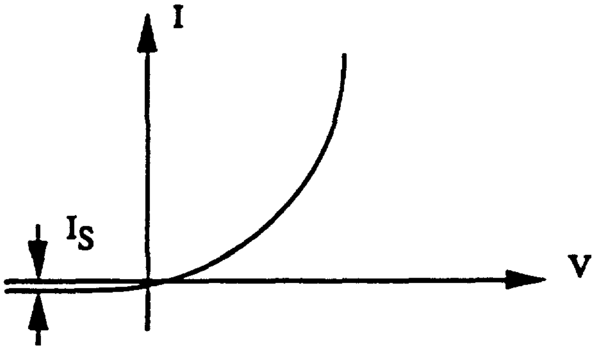


Figure 3.7: Current-voltage characteristic of a diode.

at room temperature. n is called the ideality factor, and it depends on the diode structure. For a typical Schottky diode, $n = 1.2$, and for point-contact silicon diodes $n = 2$.

When the diode is biased at a DC voltage V_0, the total voltage across the diode terminals is

$$V = V_0 + v, \tag{3.8}$$

where v is the AC voltage that we are trying to detect with this diode. The previous equation can be expanded in a Taylor series about V_0 assuming v is small compared to V_0, and the first and second derivatives evaluated:

$$\begin{aligned} I(V) &= I_0 + v\left.\frac{dI}{dV}\right|_{V_0} + \frac{1}{2}v^2\left.\frac{d^2I}{dV^2}\right|_{V_0} + \dots \\ \left.\frac{dI}{dV}\right|_{V_0} &= \alpha I_S e^{\alpha V_0} = \alpha(I_0 + I_S) = G_d = \frac{1}{R_j}, \\ \left.\frac{d^2I}{dV^2}\right|_{V_0} &= \alpha^2(I_S + I_0) = \alpha G_d = G_d{}'. \end{aligned} \tag{3.9}$$

Here $I_0 = I(V_0)$ is the DC bias current. R_j is called the junction resistance of the diode, and $G_d = 1/R_j$ is the dynamic conductance. Now the current can be rewritten as

$$I(V) = I_0 + i = I_0 + vG_d + \frac{v^2}{2}G_d{}' + \dots \tag{3.10}$$

This approximation for the current is called the *small signal* or *quadratic* approximation, and is adequate for many purposes.

In this case, the current through the diode contains a term proportional to the square of the AC voltage, or, equivalently, to the AC power. This allows the diode to work as a power detector, as we will now detail. We assume the RF voltage is

$$v = \sqrt{2}V_{RF}\cos\omega t \tag{3.11}$$

where V_{RF} is the RMS value of the RF voltage. Since, by a trig identity,

$$v^2 = 2V_{RF}^2\cos^2\omega t = V_{RF}^2\left(1 + \cos 2\omega t\right) \tag{3.12}$$

we can gather the terms from (3.10) which are constant in time:

$$I_{DC} = I_0 + \frac{1}{2}G'_d V_{RF}^2 \tag{3.13}$$

All other terms from (3.10) vary rapidly (at least as fast as the RF voltage itself) and have a time-average value of zero. They are thus not detected by ordinary low-frequency measurement devices such as voltmeters, so that only (3.13) will register on such an instrument. In particular, if the applied DC bias voltage is zero, then $I_0 = 0$, and thus I_{DC} is found to be proportional to the RF power:

$$I_{DC} = \frac{1}{2} G'_d Z_0 P_{RF} \tag{3.14}$$

where Z_0 is the characteristic impedance level of the RF circuit, and P_{RF} is the time-average RF power.

The square-law approximation is valid when v is a small signal, and that means that the input power on the diode is not too large. For large input power levels, the diode gets saturated, and behaves like a linear device. The different regimes of the diode detector with respect to the input power level are shown in Fig. 3.8.

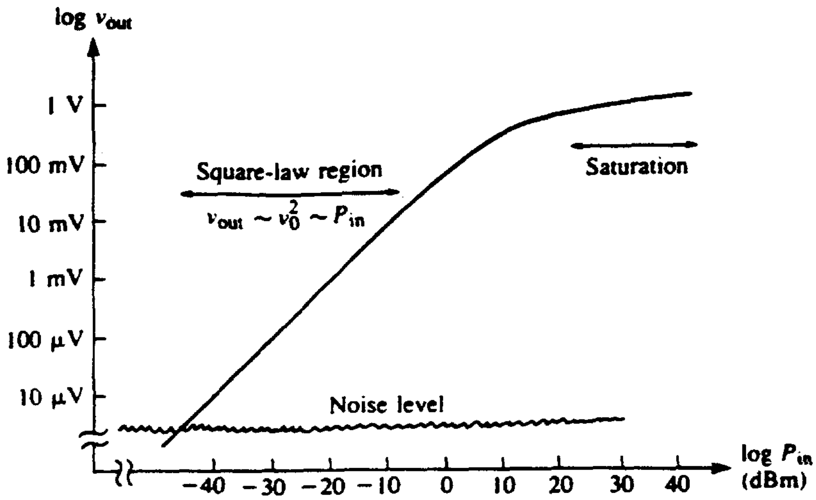


Figure 3.8: Diode detector regimes of operation.

A typical diode equivalent circuit valid at microwave frequencies is shown in Fig. 3.9. L_p and C_p are

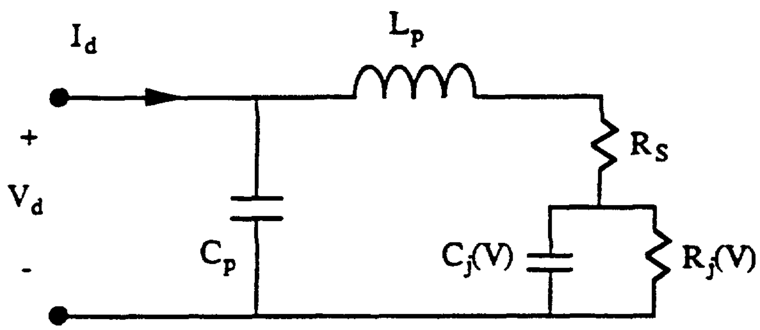


Figure 3.9: Equivalent AC circuit for a diode.

due to the diode package—L_p is a series inductance due to the leads, and C_p is a shunt capacitance of the contacts. C_j is typically a fraction of a pF. The impedance of the diode is given from the I-V curve as $R_j = 1/(\alpha I_s)$ (if $V_0 = 0$), and is usually a few kΩ, for leakage currents I_s on the order of 10 μA. Since this impedance has to be matched to the 50-Ohm impedance of the coax, the sensitivity of the diode as a power measuring device is reduced for larger impedances. The temperature performance is improved, however, for larger values of R_j, so there is a compromise involved in choosing the diode resistance for good power sensors.

3.5 Practice questions

1. You are given the electric and magnetic field vectors of a wave. How do you find the wave power flowing through a given surface? What is the equivalent for a TEM mode on a transmission line?

2. What kind of power levels, in mW, does the range between -110 dBm and +100 dBm correspond to?

3. How do bolometers measure power?

4. How does a bolometer bridge work? Why is a bridge used?

5. What is the difference between a bolometer power measurement and a diode power measurement?

6. Can the diode power measurement give you information about the phase of the wave?

7. When is the square-law approximation for the diode I-V curve valid?

8. Sketch a Schottky diode equivalent circuit and explain what the elements are. Which elements, based on physical reasoning, depend on the DC bias point?

9. In which situations would you use a thermistor, thermocouple or diode for microwave power measurements?

10. What are the advantages and disadvantages of a thermistor over a thermocouple sensor?

11. For a diode power measurement, would it be preferable to use a DC bias voltage which was positive, negative, or zero? Why?

12. How do the resistance vs. power curves (e. g., Fig. 3.1) of a thermistor affect the accuracy of a balanced bridge RF or microwave power measurement?

3.6 Homework Problems

1. Express low (¡1 mW), medium (1 mW and high power levels in dBm and dBW (referenced to 1 W).

2. In the lab, you will characterize a thermistor power detector using a bridge circuit as shown in Fig. 3.10. Assume that the resistors R and R_S are known. For a measured DC voltage V when the current through the ammeter is zero, find the expression for the DC power into the thermistor R_T.

3. The diagram for the HP8481A dual-element thermocouple sensor is shown in Fig. 3.11. The detected RF power is incident from the left, through a coupling capacitor C_c. One of the thermocouples is connected directly to RF ground, and the other is connected to RF ground through a bypass capacitor C_b. Assume each thermocouple is made to have a resistance of 100Ω. What is the input impedance seen by the RF signal, and what is the input impedance looking into the DC output side of this circuit? Is the input RF coax well matched? Do you need a choke (inductor) in the bias to prevent the RF signal from flowing into the DC out circuit? Is this circuit broadband at microwave frequencies?

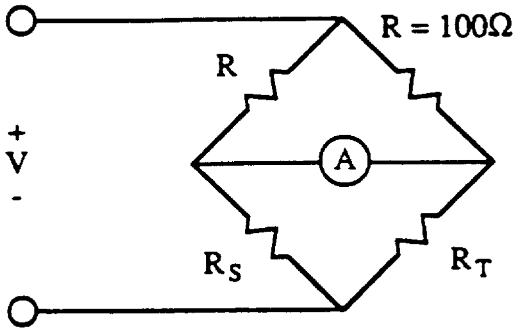


Figure 3.10: Bridge circuit for characterizing the thermistor R_T

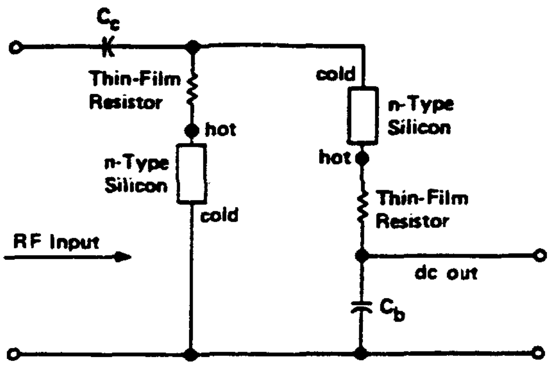


Figure 3.11: Diagram of the HP8481A dual-thermocouple coaxial mount.

4. The equivalent circuit parameters of a packaged diode, Fig. 3.9, are

$$\begin{aligned} C_p &= 0.1\text{pF} \\ L_p &= 2\text{nH} \\ C_j &= 0.15\text{pF} \\ R_S &= 10\Omega \\ I_S &= 0.1\mu\text{A} \end{aligned}$$

Use Designer or SPICE to calculate and plot the RF impedance of the diode from 4 to 14 GHz, for bias currents of both $I_0 = 0$ and $I_0 = 60\mu$A. Assume $\alpha = 1/(25\text{mV})$ and ignore the dependence of C_j on V (note that the voltage and current appearing in (3.7)-(3.10) are those across and into R_j in Figure 3.9). Is the diode well matched to $50\,\Omega$?

5. Assuming the diode equivalent circuit of problem 3, suppose a voltage wave $a_1 = 0.001$ W$^{1/2}$ is incident on the diode from the left, from a transmission line with $Z_0 = 50\ \Omega$. To what incident power in dBm does this correspond? If $f = 9$ GHz, what is the DC voltage that will be observed across the diode terminals for each of the bias conditions of problem 3? To determine this, first use SPICE to analyze the small-signal RF behavior of the circuit to find the RF voltage v. Then find the DC part of the diode current, and finally the DC part of the diode voltage V_d using the DC properties of the diode equivalent circuit.

Chapter 4

Microwave Multiport Networks

We have mentioned multiport networks in Lecture 2 when defining generalized reflection and transmission coefficients – the S- parameters. Recall that the scattering (S) matrix of an N-port device has the following basic properties:

- it has N^2 elements, so in order to determine a network completely, one needs to measure or calculate $2N^2$ numbers (a real and imaginary, or amplitude and phase, for each s-parameter in the matrix);
- it has zeros on the diagonal if the network is matched at all ports;
- it is unitary if the network is lossless, i.e. $\mathbf{SS}^* = \mathbf{I}$;
- it is symmetrical if the network is reciprocal (and therefore also linear).

The commonly used microwave multiports are briefly described in this chapter, showing their microstrip and/or coaxial and waveguide versions. They are categorized as passive, linear, non-reciprocal and active, each with more complicated properties. Strictly speaking, non-linear networks cannot be described by s-parameters, which assume linearity. However, there are cases when a network can be assumed to be linear under certain conditions, and in that case the scattering matrix can be defined and measured. For example, a transistor amplifier is nonlinear, but around a DC operating point with small-signal ac amplitudes, it can be described as linear.

4.1 Two-Port Networks

Most common two-port networks are (1) matching networks, (2) tuners, (3) attenuators, (4) filters, (5) isolators, (6) phase shifters, (7) limiters, and (8) common-terminal amplifiers (to be discussed in a later lab).

(1) *Matching networks* – you have already designed some stub matching networks, and we have discussed lumped element matching and quarter-wave matching networks. A two-port matching network is generally linear and passive, can be lossless or lossy, is not matched (why?) and is usually narrowband. Sometimes, if the load varies over time, it is desirable to have a variable matching network, in which case tuners are used.

(2) *Tuners* – these two-port networks are usually implemented in coaxial form, such as the ones you have used in Lab 2 for stub matching. A common measurement technique that uses tuners is a so called *load-pull* system. This system can present a wide range of impedances to a two-port device at both ports. It is desirable for a tuner to cover the entire passive Smith chart, so all

possible impedances. A practical difficulty in producing high-quality tuners is the fact that any realistic component will have loss, which will limit the impedance range. Namely, the very edge of the Smith chart is purely reactive and any loss does not allow a tuner to reach those impedances. Air coaxial tuners have the lowest loss and are the most common. They typically consist of two movable parts with a different characteristic impedance as shown in Fig. 4.1, and are referred to as "slug tuners". Focus Microwaves, Inc. and Maury Microwave produce tuners that have VSWRs of up to 40:1 (see, e.g. http://www.focus-microwaves.com – look under manual and automatic tuners).

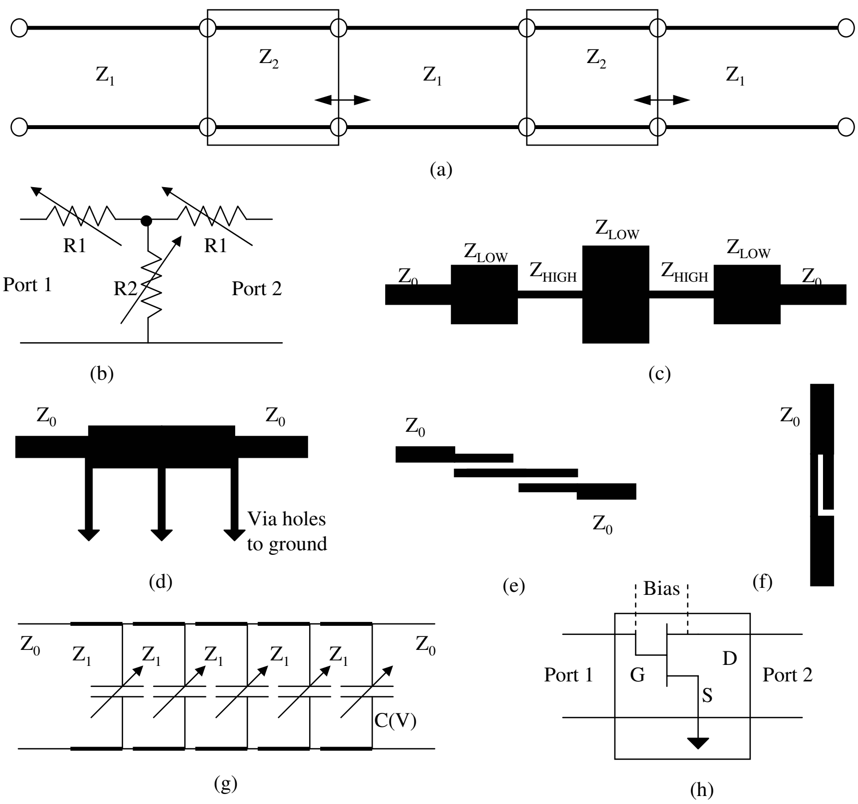


Figure 4.1: (a) Electrical model of a slug tuner. (b) Variable attenuator. (c) Low-pass filter implemented in microstrip, (d) Microstrip high-pass filter, (e) Microstrip coupled-line bandpass filter, (f) Microstrip band-stop filter, (g) Block diagram of a loaded-line phase shifter with varactor diodes, and (h) Two-port diagram of a common-source MESFET amplifier.

(3) *Attenuators* – are matched two-ports which absorb some power and are therefore lossy. They are reciprocal. Attenuators can be made to be variable, as shown in a lumped-element version, Fig. 4.1(b). They can also be implemented in waveguide using power transfer into non-propagating

modes, as you will see in the lab. The loss is defined as the ratio of the output to input power, where the input power is the incident power minus the reflected power. This allows simple power calculations for cascaded two-port networks, and loss is really like (1/gain), or in dBs, negative gain. Defined in dBs, this is referred to as *insertion loss*:

$$IL = 10 \log \frac{P_{out}}{P_{in}} = 10 \log \frac{|s_2 1|^2}{1 - |s_{11}|^2}.$$

(4) *Filters* – are two-port networks that have a designed frequency dependence. They are matched and are usually lossless and reciprocal. Low-pass filters, which at low frequencies are a ladder network of series inductors and parallel capacitors, can be designed with high-impedance lines (which behave like inductors) and low-impedance lines (which behave like capacitors). An example of such a microstrip filter layout is shown in Fig. 4.1(c). A high-pass microstrip filter is shown in Fig. 4.1(d). What is the low-frequency circuit equivalent of this filter? The main difference between the low and high frequency implementations is in their overall frequency response: in transmission-line filters, as frequency increases, the properties will repeat. E.g. when the line lengths become half-wavelength, the filter is electrically not present. Bandpass filters can be made using coupled-line sections, which effectively behave as resonant circuits. A stop-band microstrip filter, shown in Fig. 4.1(f) effectively has a parallel resonant circuit in series with the 50-ohm line. Can you identify the capacitance and inductance of the resonant circuit?

(5) *Isolators* – are non-reciprocal lossy components which transmit incident power in one direction with low loss, while the reflected power is absorbed (high loss). This is accomplished using nonlinear materials, usually ferrites. Typically, a waveguide is loaded with a ferrite so that the incident wave mode has a different field profile than the reflected wave mode, and as a result, the two modes are attenuated differently. Isolators are commonly used to protect active components, such as transmitters, from large reflections which could destroy the component. An ideal isolator has the following scattering matrix:

$$\mathbf{S} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}. \tag{4.1}$$

(6) *Phase shifters* – are two-ports that are ideally matched, reciprocal and lossless, with the following scattering matrix:

$$\mathbf{S} = \begin{bmatrix} 0 & e^{j\theta} \\ e^{j\theta} & 0 \end{bmatrix}. \tag{4.2}$$

In transmission, the phase shifter adds a phase θ at the frequency of interest. This type of device is used for steering antenna beams, for example. A simple implementation of a variable phase shifter is shown in Fig. 4.1(g), where a transmission line is loaded with variable capacitance diodes. These diodes are referred to as varactors, and their capacitance is a function of voltage, so it is a nonlinear capacitance. However, for small ac signals, the capacitance behaves as linear for a fixed bias voltage. The variable capacitance is in shunt in the line and adds to the capacitance per unit length. Thus, the phase constant becomes:

$$\beta = \omega \sqrt{L'(C' + C(V))} = \beta(V)$$

which means that it varies as a function of bias voltage V. This in turn means that the relative phase between the output and input voltage waves varies with dc bias voltage and can be controlled. This type of phase shifter is called a loaded-line shifter. Also common are switched-line digital phase shifters, in which single-pole double-throw switches are used to switch different lengths of line, and therefore different phase-shifts, into the transmission path.

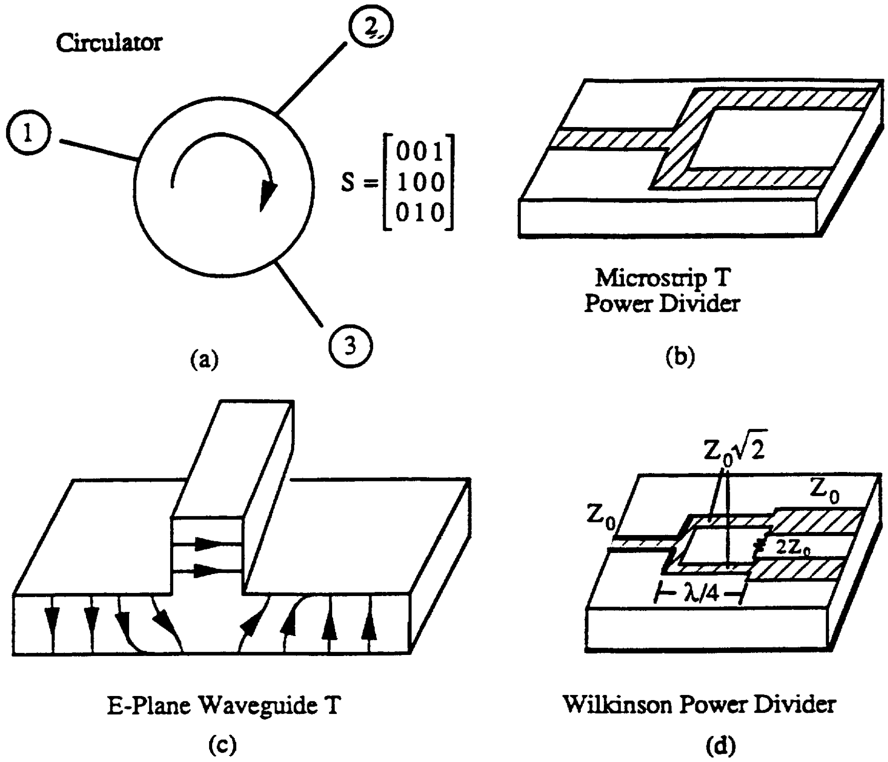


Figure 4.2: The symbol and scattering matrix for a circulator (a), and examples of lossless (b), (c) and resistive (d), power dividers.

(7) *Limiters* – are two-port networks that limit the power at the output port. They are matched, nonlinear and lossy devices. Usually, they are made with a *pin* diode, which is a high-resistance diode, in anti-parallel with a Schottky diode (which helps bias the *pin* diode). When the voltage at the input becomes too large, the resistance of the *pin* diode does not allow the voltage to track it at the output. Limiters are, like isolators, used to protect other components from high voltage stress.

(8) *Two-port amplifiers* – when one of the terminals of a transistor is grounded, a two-port amplifier results, as shown in Fig. 4.1(h) on the example of a common-source MESFET amplifier. In reality, however, this is a multi-port device: the amplification is at the expense of dc power and in reality some ac power leaks into the gate and drain bias ports. In addition, it is difficult to make an ideal rf short, so the source is never really grounded. We will investigate this more in a later lecture and lab.

4.2 Three-Port Networks

The conclusion from Lecture 2 (review of S- parameters) was that a lossless, reciprocal and matched three-port is not possible.

If one of the three conditions is dropped, a feasible device can be made. For example, if we assume the device is not reciprocal, but is matched and lossless, we can get the following scattering matrix:

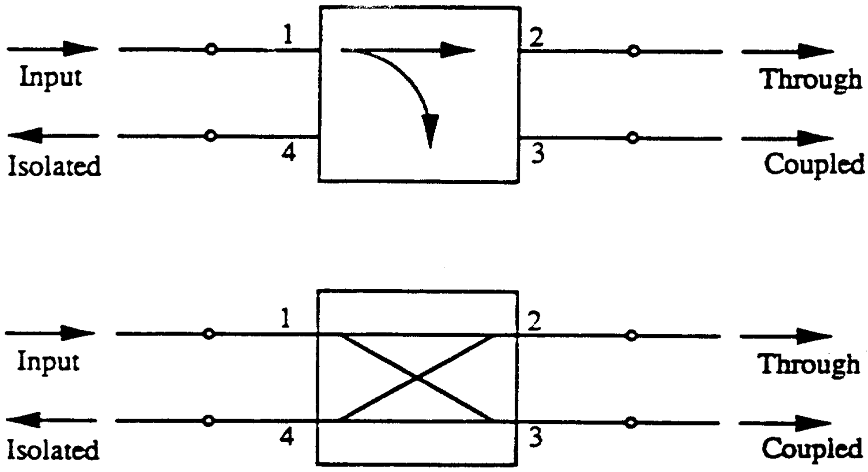


Figure 4.3: Symbols for directional couplers.

$$\mathbf{S} = \begin{bmatrix} 0 & S_{12} & S_{13} \\ S_{21} & 0 & S_{23} \\ S_{31} & S_{32} & 0 \end{bmatrix}.$$

One device that has such scattering parameters is called a *circulator*, and its symbol is shown in Fig. 4.2(a).

It has the property that power coming into port 1 will go out of port 2, and none will go out of port 3, power going into port 2 will only go out port 3, and power going into port 3 will only go out port 1. It is now obvious where the name comes from. This device is widely used in microwave engineering, and is physically realized by using ferrite materials, which give it a preferred direction by producing a static magnetic field in only one direction.

If a three port is reciprocal and lossless, but not necessarily matched at all ports, we get a power divider, such as the ones shown in Fig. 4.2(b),(c). Finally, if we relax the lossless condition, we get a resistive power divider, an example of which is shown in Fig. 4.2(d). Such resistive power dividers can be made such that $S_{23} = S_{32} = 0$ so that the two output ports are isolated. They can also be designed to have more than 2 outputs. All of these components are often used in microwave circuits.

Other three-port networks are nonlinear active circuits, such as amplifiers, mixers, oscillators, and they can often not be represented with scattering parameters.

It should be noted that sometimes a port of a 4-port network is terminated in a matched load and is then sealed and not visible. Thus, a 4-port device is effectively turned into a matched, lossless and reciprocal three-port.

4.3 Four Port Networks—The Directional Coupler

Four-port microwave devices are used very often, especially ones called *directional couplers*, and in waveguide systems the magic T (which is another special kind of *hybrid network*) is also used quite often. The different symbols used for directional couplers are shown in Fig. 4.3.

These devices have the property that a signal coming into the input port is distributed between two of the other three ports, while the fourth one is isolated. By writing down equations similar to ones we did for three ports, we can find out that a lossless, reciprocal and matched four-port does exist. If we assume symmetry of the junction (that is, we can interchange ports $1 \leftrightarrow 2$ and $3 \leftrightarrow 4$ or similarly switch

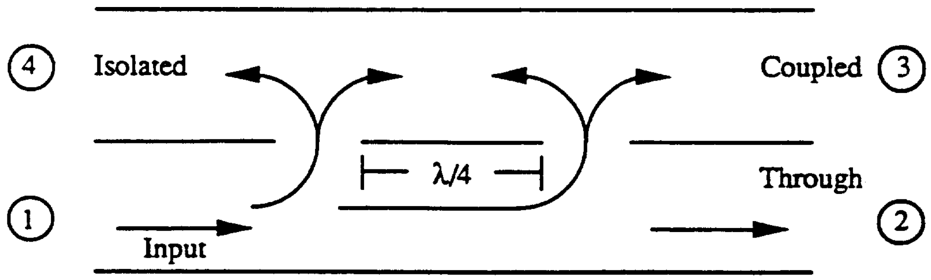


Figure 4.4: Awaveguide directional coupler with two coupling holes $\lambda/4$ apart.

the top and bottom ports without changing the network) then its s-parameters have to be in one of the two forms:

$$\mathbf{S} = \begin{bmatrix} 0 & A & jB & 0 \\ A & 0 & 0 & jB \\ jB & 0 & 0 & A \\ 0 & jB & A & 0 \end{bmatrix} \quad , \quad S = \begin{bmatrix} 0 & A & B & 0 \\ A & 0 & 0 & -B \\ B & 0 & 0 & A \\ 0 & -B & A & 0 \end{bmatrix}, \tag{4.3}$$

where A and B can be chosen as real numbers by suitable choices of reference planes.

The first type is called a *symmetrical* coupler, and the coupled port is 90° out of phase from the through port and from the input. The second type is called an *antisymmetrical coupler*, and while its through and coupled ports are either in phase or 180° out of phase with the input. However, its key property is that when ports $1 \leftrightarrow 2$ and $3 \leftrightarrow 4$ are interchanged, the coupled port response is reversed. For either type, the lossless condition tells us that A and B are not independent, and that $A^2 + B^2 = 1$. An additional design parameter is how the power splits between the two non-isolated ports. A special, and very often used case, is when it splits equally. This is called a 3 dB coupler. In this case, $A = B = 1/\sqrt{2}$, and the two matrices from above become:

$$\mathbf{S} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & j & 0 \\ 1 & 0 & 0 & j \\ j & 0 & 0 & 1 \\ 0 & j & 1 & 0 \end{bmatrix} \quad , \quad S = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{bmatrix}. \tag{4.4}$$

Three parameters are defined for a directional coupler, and this is what is usually given in spec sheets when you buy one:

$$\begin{aligned} \text{Coupling factor} &= C = 10\log\frac{P_1}{P_3} = -20\log|S_{31}|\,\text{dB} \\ \text{Directivity} &= D = 10\log\frac{P_3}{P_4} = 20\log\frac{|S_{31}|}{|S_{41}|}\,\text{dB} \\ \text{Isolation} &= I = 10\log\frac{P_1}{P_4} = -20\log|S_{41}|\,\text{dB} \end{aligned}$$

The coupling factor tells us the fraction of the input power which is coupled to the output port. The directivity and isolation tell us how well the coupler isolates forward from backward travelling waves, and is a measure of the ideality of the device. An ideal directional coupler has an infinite directivity and isolation ($s_{14} = 0$).

Let us look at a few examples of how such devices are made. If you look down a waveguide directional coupler in the lab, you will see that it consists of two waveguides with a common wall which has a slew of holes in it. Let us look qualitatively at a simple case with only two holes that are $\lambda/4$ apart (at the design frequency), Fig. 4.4. If port 1 is the input, then port 2 is the through port. A small amount of the input wave leaks out through the two coupling holes into the second waveguide, and both of these

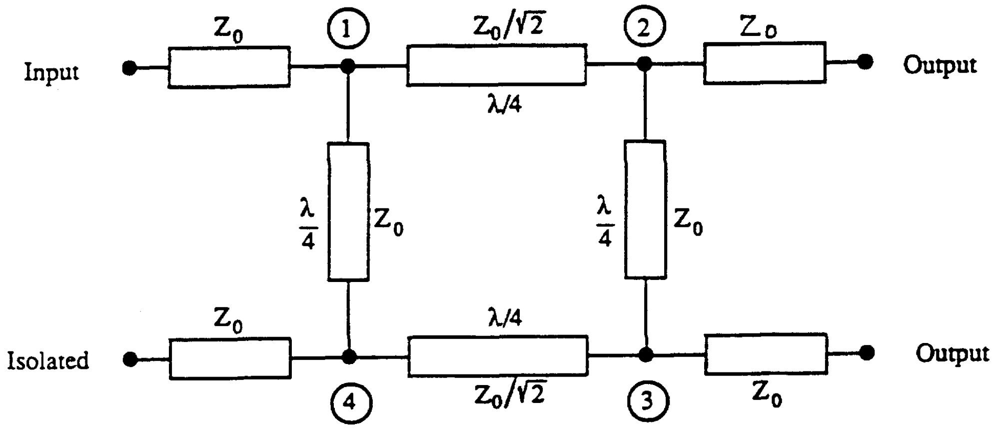


Figure 4.5: A branch line directional coupler.

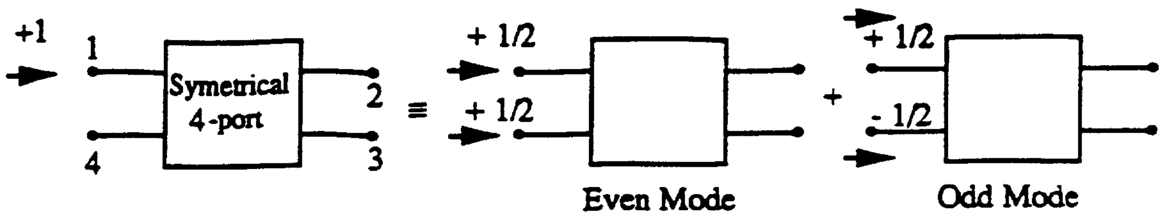


Figure 4.6: Odd and even mode decomposition for a symmetric four-port.

coupled waves propagate in both directions. The forward propagating wave coupled through the first hole adds up in phase with the forward propagating wave coupled through the second hole (since they have both travelled the same distance $\lambda/4$), so port 3 is the coupled port. The two coupled backward propagating waves, however, cancel out (they are $\lambda/2$ apart), so port 4 is the isolated port. You might guess that this type of device has a very narrow bandwidth, and the reason why the directional couplers you could see in the lab have many coupling holes is to increase the bandwidth. The design of such an optimized device is quite complicated.

4.4 Odd and Even Mode Analysis of a Branch Line Directional Coupler

A directional coupler made with sections of transmission line is shown in Fig. 4.5. This structure is highly symmetrical, and such structures are easily analyzed using odd and even modes. This is a very powerful technique, and the basic idea is that an excitation can be decomposed into a superposition of two excitations such as shown in Fig. 4.6. Then the response of the circuit to the original excitation can be found as the sum of the responses to the odd and even excitations separately (provided the circuit is linear, otherwise superposition does not hold). Let us see how this works on the example of a branch line coupler.

Fig. 4.7(a) shows a branch line coupler with normalized impedances and the transmission lines drawn as fat lines (we assume there is also a ground conductor for each of them, which is not shown). Even mode excitation, Fig. 4.7(b), gives an open circuit at the horizontal symmetry line, so the four-port network can be split into two independent two-ports. Similarly, odd mode excitation gives a short circuit along the symmetry line resulting in a different pair of two-ports. The responses of these two-ports can be

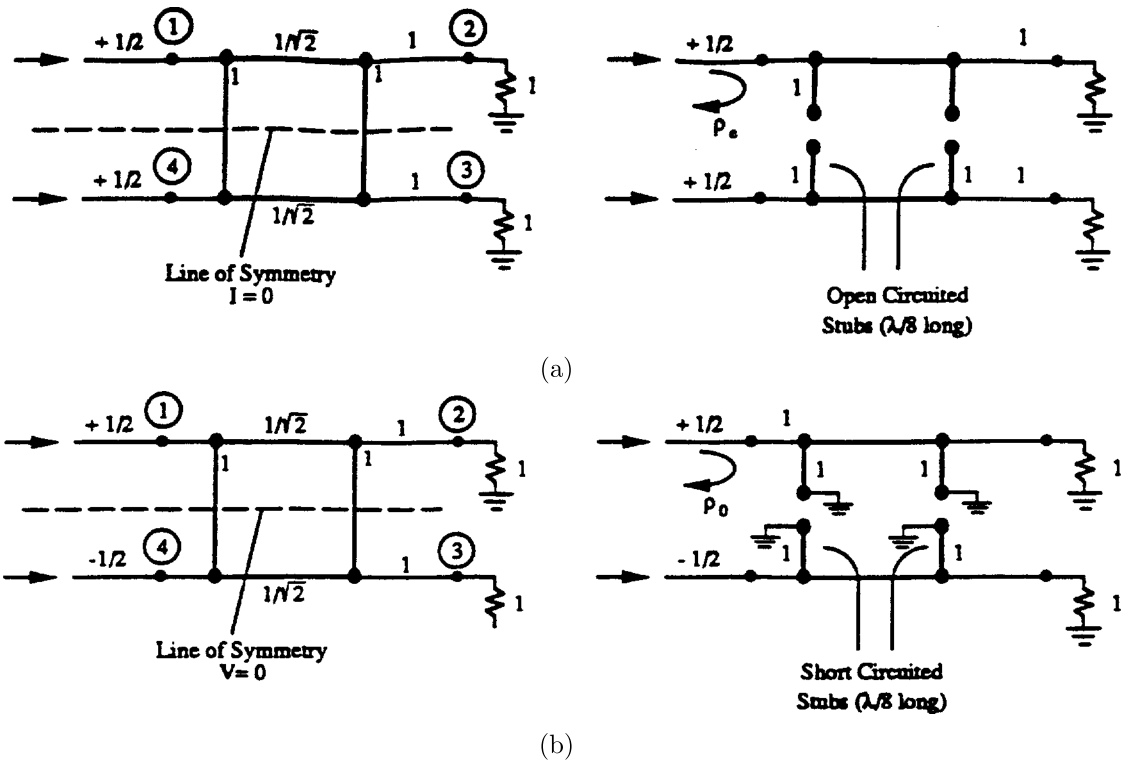


Figure 4.7: Normalized branch line coupler decomposed into even (a) and odd (b) mode excitation.

added up to get the response of the four-port coupler. If we denote the odd and even mode reflection and transmission coefficients by ρ_o, ρ_e, τ_o and τ_e, the following equations can be written for the amplitudes of the waves coming out of the four ports:

$$\begin{aligned} b_1 &= \frac{1}{2}\rho_e + \frac{1}{2}\rho_0 \\ b_2 &= \frac{1}{2}\tau_e + \frac{1}{2}\tau_0 \\ b_3 &= \frac{1}{2}\tau_e - \frac{1}{2}\tau_0 \\ b_4 &= \frac{1}{2}\rho_e - \frac{1}{2}\rho_0, \end{aligned} \tag{4.5}$$

assuming there is an incident wave $a_1 = 1$ at port 1 only ($a_2 = a_3 = a_4 = 0$), as shown in Fig. 4.7.

The even-mode and odd-mode reflection and transmission coefficients can be found to be the following:

$$\begin{array}{ll} \rho_e = 0 & \rho_0 = 0 \\ \tau_e = -\frac{1}{\sqrt{2}}(1+j) & \tau_0 = \frac{1}{\sqrt{2}}(1-j). \end{array} \tag{4.6}$$

From here, we get

$$\begin{array}{ll} b_1 = 0 & \text{(port 1 matched)} \\ b_2 = -j/\sqrt{2} & \text{(half-power, } -90^\circ \text{ out of phase from port 1)} \\ b_3 = -1/\sqrt{2} & \text{(half-power, } -180^\circ \text{out of phase from port 1)} \\ b_4 = 0 & \text{(no power to port 4).} \end{array}$$

This agrees with the first row and column of the scattering matrix for a hybrid symmetrical directional coupler.

4.5 Practice questions

1. A measurement of a two-port gave the following S-matrix:

$$\mathbf{S} = \begin{bmatrix} 0.1\angle 0^\circ & 0.8\angle 90^\circ \\ 0.8\angle 90^\circ & 0.2\angle 0^\circ \end{bmatrix}$$

   Determine if the network is reciprocal and whether it is lossless.

2. In a common-source amplifier, define the S-parameters and relate them to quantities you have studied in circuit analysis.

3. Prove that a three port network cannot be matched, reciprocal and lossless simultaneously.

4. Show that a four port network can satisfy the matched, reciprocal and lossless conditions simultaneously.

5. Show that the ideal isolator is lossy.

6 Show that the ideal phase shifter is lossless.

7. Write down the scattering matrix of an ideal lumped element low-pass filter in the (a) pass band and (b) stop band. Repeat for a transmission-line low-pass filter such as in Fig. 4.1(c).

8. Write down the scattering matrix for a 3-dB directional coupler for which the two non-isolated outputs are in phase quadrature.

9. How does odd and even mode decomposition work? What kind of a network do you have to have in order to be allowed to use odd and even mode decomposition?

10. Sketch the odd and even mode circuits of a microstrip branch line coupler.

11. Sketch the odd and even mode circuits of a Wilkinson power divider, shown in Figure 4.2(d).

12. What are the impedances of $\lambda/8$-long short and opened stubs of characteristic impedance Z_0?

13. Apply results from question #10 to calculating the odd and even mode reflection and transmission coefficients for the branch line directional coupler ($\rho_o, \rho_e, \tau_o, \tau_e$).

14. If you had a 20-dB directional coupler, how could you use it to measure a reflection coefficient?

4.6 Homework Problems

1. Find the S-parameters of the networks (a) and (b) of Fig. 4.8. Are these networks lossless? Reciprocal? Matched? Check for these properties using the criteria derived in the notes and your calculated S-parameters. (Do your answers make sense?)

2. Repeat problem 1 for network (c) of Fig. 4.8, which is a length l of transmission line with characteristic impedance Z connected between two lines of characteristic impedance Z_0. The "electrical length" θ of the middle section is equal to βl.

3. Find the S-parameters of an attenuator that has a 6-dB attenuation (in power).

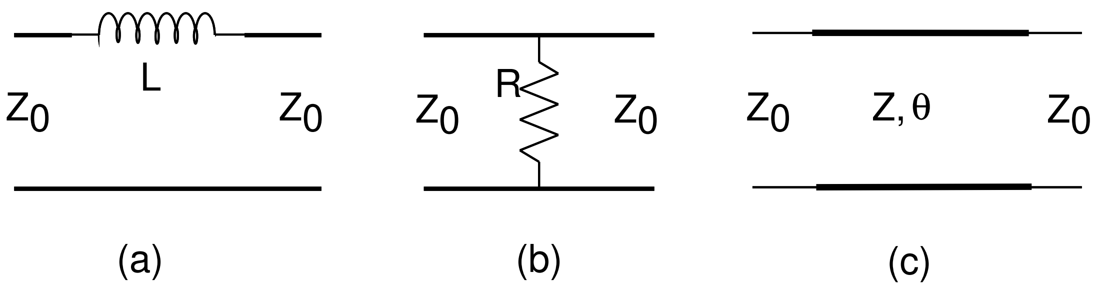


Figure 4.8: Determine the S-parameters of the two-port networks shown in (a), (b) and (c).

4. Find the s-parameters of a matched isolator that has 0.2dB insertion loss and a -25dB backward attenuation at the design frequency.

5. Using Ansoft Designer, simulate a Wilkinson power divider centered at 2 GHz. Perform a layout of the divider in microstrip, using a Duroid substrate with $\epsilon_r = 2.2$ and 0.508 mm thick. How does the layout change if you change the substrate to one that has a relative permittivity of 10.2?

6. Using dependent voltage sources and passive lumped elements, model an ideal circulator (for which $S_{21} = S_{32} = S_{13} = 1$, while all other S-parameters are zero) in Ansoft Designer or SPICE. (In SPICE, connect a source to port 1 that produces a unit incident wave, and compute the waves emanating from ports 1, 2 and 3 to demonstrate your model.)

7. Using Ansoft Designer, design a transmission-line Tee network in microstrip with no loss. Assume the network is designed to feed two antennas in phase from a common feedpoint. You want as much power to get to the antennas as possible. The antennas are designed to have a 50-ohm input impedance at and around 2 GHz. Can you use this circuit as a basis for a network that feeds 4 and 8 antenna elements, and how would you accomplish that?

8. When we measure an S-parameter such as S_{11} in dB, should we use $10 \log |S_{11}|$ or $20 \log |S_{11}|$ to calculate it? If we measure a value of S_{11} equal to $0.5 - j0.7$, what is its value in dB? If we have an S_{12} of -15 dB, what is its value in absolute numbers (can you give a complete answer to this question or not)?

9. A 50-Ω resistor is placed in series in a 50-Ω transmission line. If 1 W of power is incident from the generator:

- how much power is reflected,
- how much power is delivered to a load matched to the transmission line, and
- how much power is lost in the series resistor? What is the insertion loss of the series resistor, viewed as a two-port network?

10. A low-noise amplifier has a (power) gain of 6.5 dB. An input signal of 5 mW is applied to this amplifier, while its output is fed to port 1 of an isolator whose S-parameters are: $S_{11} = 0$, $S_{21} = 0.9$, $S_{12} = 0$ and $S_{22} = 0$. Find the output power emerging from port 2 of the isolator, expressed both in dBm and in mW. Assume the input port of the amplifier is reflectionless.

11. Use Ansoft Designer or SPICE to design a branch-line directional coupler, Fig. 4.9, at 2 GHz. Each of the four ports is assumed to be terminated in 50 Ω. From formulas (1.41)-(1.44) of the lecture

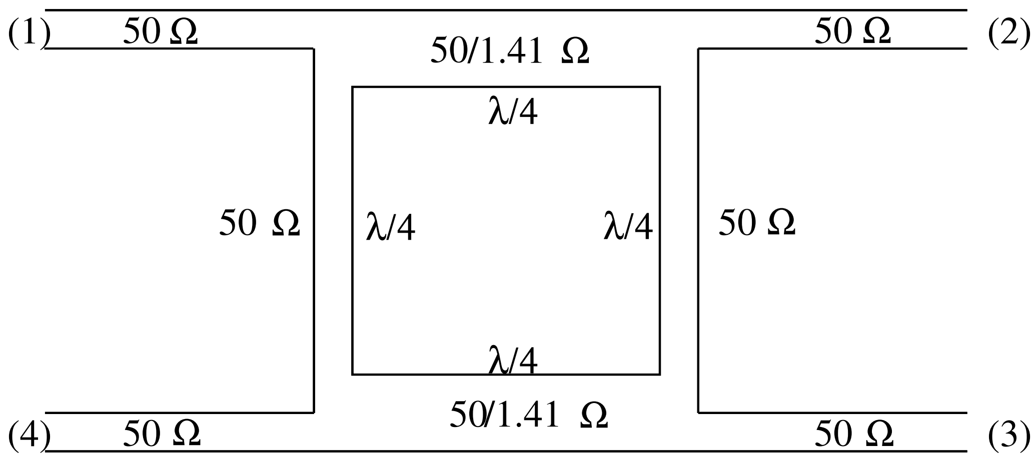


Figure 4.9: A branch-line microstrip directional coupler.

notes, find the length and width of each microstrip section in this coupler, if the substrate has permittivity $\epsilon_r = 10.2$ and thickness $h = 1.27$ mm. Print out plots of the magnitudes of the four independent S-parameters (reflection at input, transmission to the other 3 ports) between 1 and 3 GHz. Label the ports (isolated, coupled etc.). How large is the bandwidth of this circuit (make sure you state what your bandwidth definition is, since there can be more than one)?

12. Repeat Problem 11, but now make the section of line between ports 3 and 4 have a length of $3\lambda/4$, and make all the characteristic impedances in the coupler's transmission line sections equal to $Z_0\sqrt{2}$. Analyze the circuit, include your plots (same parameters as in the previous problem) and explain what kind of a circuit this is.

13. Use Ansoft Designer to design a low-pass filter using alternating low and high-impedance quarter-wave lines at 2 GHz. How does the number of lines affect the frequency response in the pass and stop bands? What happens as you increase the frequency range over which you are simulating the response? How does the ratio of the high-to-low impedance affect the response? Sketch the layout of the circuit in microstrip on a substrate with permittivity 4.5 (FR-4) which is 0.508-mm thick.

14. Use Ansoft Designer to design a microstrip high pass filter, with same substrate parameters as in the previous problem. Are there any parasitic elements you need to take into account (e.g. the inductance of a via to ground)?

Chapter 5

Time-Domain Reflectometry

5.1 Introduction

When the first transatlantic telephone cable was laid on the bottom of the ocean, engineers still did not understand losses in transmission-lines and how to make them smaller. When they received too small no signal on the other end of the Atlantic, they assumed that increasing the transmitted power would make it possible to get reception. Unfortunately, the power was too large for the cable to handle and breakdown occurred somewhere in the middle of the Atlantic. This was a financially unsuccessful project. Before the cable was laid, Oliver Heaviside, who understood transmission lines extremely well, had warned the engineers about losses, but he did not have a good reputation so nobody paid attention. Later, Mihailo Pupin, a Westinghouse engineer, and later professor at Columbia University, undertook the same task, but placed inductive coils along the cable at regular lengths to eliminate most of the loss. These are today called Pupin coils. Can you explain why adding periodically inductors along a cable reduces the losses, assuming the dominant loss is due to the imperfect conductor?

It would have been useful to have a method of determining where the first transatlantic cable broke, so that it could be possibly fixed and reused. The instrument used today to find faults in cables is called the *time domain reflectometer (TDR)*. The principle is very simple: the instrument sends a voltage impulse or step function, and measures the reflected signal. If there is a fault in the cable, it is equivalent to some change in impedance, and the voltage step wave will reflect off the discontinuity. Since both the transmitted and the reflected wave travel at the same velocity, the distance of the fault from the place where the TDR is connected can be calculated exactly. Not only can we learn where the fault is, but it can also tell us something about the nature of the fault. In the homework, you will look at what happens to a step function when it reflects off of different lumped elements in a transmission line, each representing an equivalent circuit for a specific cable fault.

5.2 Reflection From an Inductive Load

So far, we have only looked at transmission-line circuits in frequency domain and we assumed a sinusoidal waveform. Now we will look at what happens when a step function (in time) is launched down a loaded transmission line. This is done by multiplying the Laplace (or Fourier) transform of the reflection coefficient (i.e. the reflection coefficient in complex form) with the transform of a step function, and then transforming back into time domain with the inverse Laplace (or Fourier) transform. As an example, let us look at a reflection off an inductive load, shown in Fig. 5.1(a). The transmission line has a real characteristic impedance Z_0, and the incident voltage wave is $v_+(t)$. We can use the Thévenin equivalence theorem technique described to simplify the analysis. Problem 1 at the end of this chapter reviews the equivalent circuit of an open-ended transmission line. The resulting equivalent circuit is

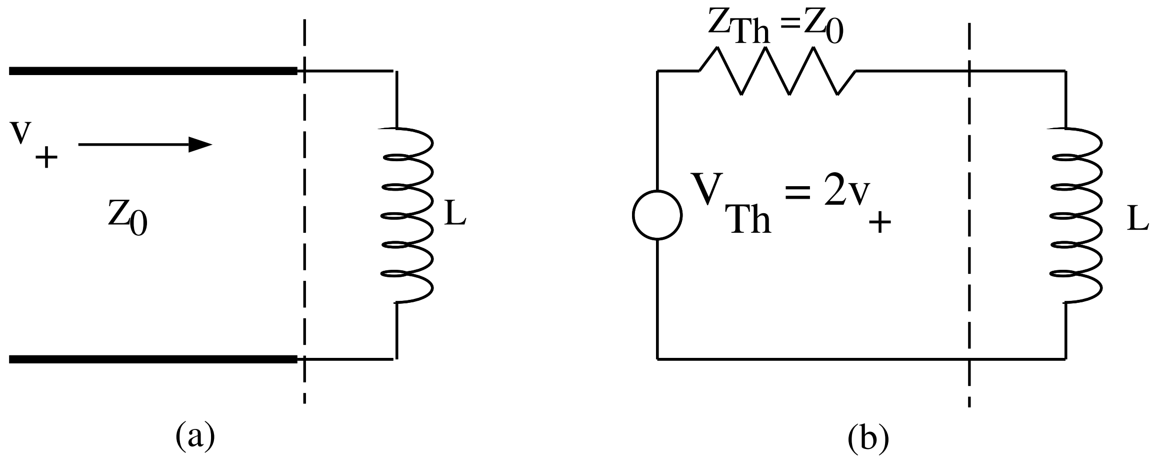


Figure 5.1: (a) Transmission line with an incident wave v_+, terminated in an inductive load. (b) Lossless transmission line is replaced by its Thévenin equivalent circuit.

shown in Fig. 5.1(b).

If we now assume that the incident voltage wave is a step function $v_+(t) = 1$, $t > 0$, the Laplace transform is

$$v_+(s) = \frac{1}{s},$$

and since the impedance function of the inductor is

$$Z = sL,$$

we find that the load voltage is equal to

$$v(s) = \frac{2L}{sL + Z_0} = \frac{2}{s + Z_0/L}. \tag{5.1}$$

We can recognize this as the Laplace transform of a decaying exponential with a time constant $t_L = L/Z_0$:

$$v(t) = 2e^{-t/t_L} \quad , \quad t > 0. \tag{5.2}$$

Since the voltage of an inductor is $v = L\mathrm{d}i/\mathrm{d}t$, we can find the current through the inductive load by integrating the voltage:

$$i(t) = \frac{1}{L}\int_0^t v dt = \frac{2}{Z_0}\,(1 - e^{-t/t_L}) \quad , \quad t > 0. \tag{5.3}$$

This is the standard build-up of current in an inductor connected in series with a resistor, which you have already seen in your circuits classes.

The reflected wave is the difference between the transmitted wave and the incident wave:

$$v_-(t) = v(t) - v_+(t) = 2e^{-t/t_L} - 1 \quad , \quad t > 0. \tag{5.4}$$

Initially, there is no current through the inductor, and the voltage is just v_+, so it looks like an open circuit and the reflection coefficient is $+1$. The current then builds up and is equal to the short circuit

Figure 5.2: The incident, reflected and total voltages for an inductive load.

current (the Norton equivalent current) and the voltage drops to zero, so the inductor appears as a short circuit. The reflected and transmitted waves are shown in Fig. 5.2.

As another example, let us look at a transmission line that is shorted at one end. If a voltage source is turned on at the other end, what will the reflected wave look like back at the source? We know that the reflection coefficient of a short circuit is -1, so the reflected wave looks as shown in Fig. 5.3(a). In TDR instruments, the reflected wave is added on to the incident step (which goes on forever in time), so in this case the display of the instrument would appear as shown in Fig. 5.3(b). The duration of the "pulse" tells us how long the line is (it corresponds to the round-trip time of the leading edge of the step).

A third, slightly more complicated example, is that of a series combination of an inductor L and a resistor R. The incident voltage is a step with unity amplitude. The voltage across the inductor is, as before,

$$v_L(t) = 2e^{-t/t_L}, \quad t > 0, \tag{5.5}$$

where the time constant is now $t_L = L/(R + Z_0)$, since the inductor sees a series connection of the characteristic impedance and the resistive load. The inductor current is

$$i_L(t) = \frac{2}{Z_0 + R}(1 - e^{-t/t_L}), \quad t > 0, \tag{5.6}$$

and the load voltage becomes

$$v(t) = v_L(t) + Ri_L(t) = 2\left[\frac{R}{R + Z_0} + \frac{Z_0}{R + Z_0}e^{-t/t_L}\right], \quad t > 0. \tag{5.7}$$

The reflected voltage wave is now

$$v_-(t) = v(t) - v_+(t) = \left[\frac{R - Z_0}{R + Z_0} + 2\frac{Z_0}{R + Z_0}e^{-t/t_L}\right], \quad t > 0. \tag{5.8}$$

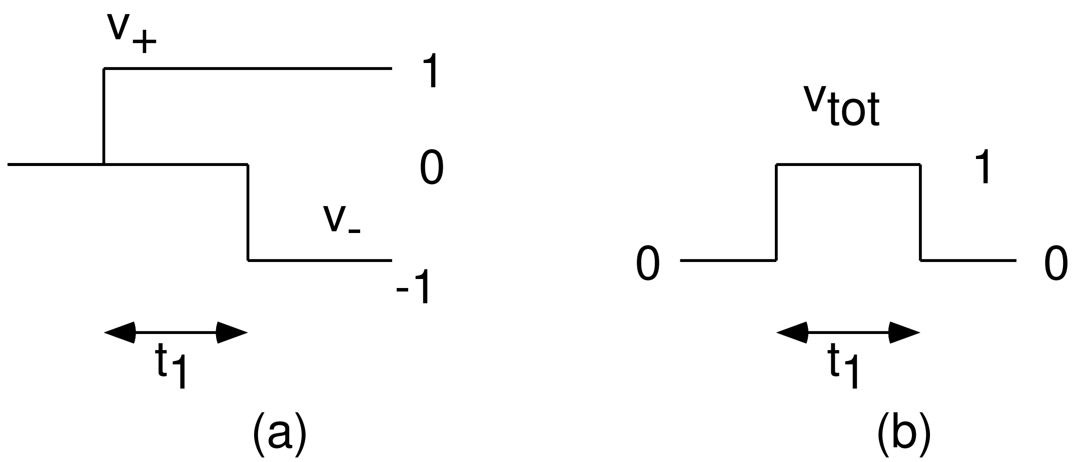


Figure 5.3: (a) Reflected voltage wave off of a short-circuited transmission line with an incident unity step function. (b) A standard TDR instrument display shows the reflected wave added on to the incident step function.

When the load consists only of a single resistor and a single inductor, or a single resistor and a single capacitor, the time response will be a single damped exponential pulse, plus a constant. In such cases, a simpler qualitative analysis can be done by just evaluating the reflected voltage at $t = 0$ (the time when the reflected wave gets back to the launching port, for example) and $t = \infty$ and assuming any transition between these two values to be exponential. In the previously analyzed case of a series RL circuit, at $t = 0$ the reflected voltage is $v_-(0) = +1$, since the inductor looks like an open circuit initially. On the other hand, as time goes by, the current through the inductor builds up and at $t = \infty$, the inductor looks like a short, so $v_-(\infty) = (R - Z_0)/(R + Z_0)$ and is determined by the resistive part of the load. The resulting plot out of a TDR (incident step plus reflected wave) is shown in Fig. 5.4.

Time domain reflectometry can also be used to determine losses in a lossy transmission line. For example, if the dominant loss in a a line is the (small) series conductor loss, the input impedance of a semi-infinite (or terminated) line can be written as:

$$Z_{in} = \sqrt{\frac{R'+j\omega L'}{j\omega C'}} = \sqrt{\frac{L'}{C'}}\sqrt{\frac{1+R'}{j\omega L'}} \approx \sqrt{\frac{L'}{C'}}\left[1+\frac{R'}{j2\omega L'}\right] = Z_0 - j\frac{R'Z_0}{2\omega}$$

This means that a line with series loss will give the TDR response of a lossless line terminated in an equivalent series capacitance, the value of which depends on frequency, series resistance per unit length and characteristic impedance of the line.

5.3 Time Constant of the Reflected Wave

The most straightforward way to measure the time constant (e.g. t_L in the inductor examples) is to measure the time t_1 which is needed to complete one half of the exponential transition from $v_-(0)$ to $v_-(\infty)$. The time for this to occur corresponds to $t_L = 0.69t_1$, where t_L is the time constant we used for

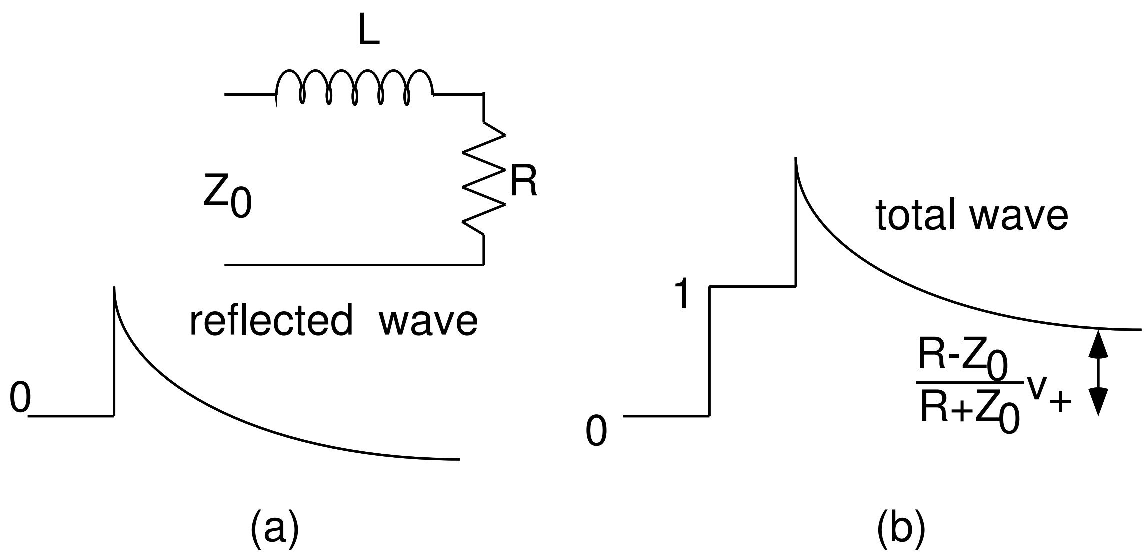


Figure 5.4: (a) Reflected voltage wave off of a series RL combination with an incident unity step function. (b) A standard TDR instrument display shows the reflected wave added on to the incident step function.

an inductive load, but the latter also holds for a capacitive load. This procedure is shown qualitatively in Fig. 5.5.

5.4 TDR Measurements in Time and Frequency Domains

A TDR instrument is really just an oscilloscope which can produce an incident wave with a very fast rise time. The fast rise time is important for a number of reasons, for example, when two discontinuities are close to each other, the instrument will not be able to separate them if the rise time is not sufficiently fast. Signal integrity as well as failure analysis often requires the ability to locate and distinguish multiple, closely spaced reflection sites. A TDR can resolve two discontinuities if they are separated by roughly half the TDR rise time. Since typical TDR oscilloscope rise time is in the 35 ps range (both step generator and oscilloscope), that limits the measurable separation between two discontinuities to approximately 17 ps. In materials with a dielectric constant near 1, this corresponds to a physical separation of about 5 mm. Typical printed circuit board material (FR4) will have a dielectric constant of approximately 4.2-4.6. The measurable separation then decreases (if the field is completely in the dielectric, such as in stripline, this becomes about 2.5 mm). Some small interconnects, such as board vias, package leads, and socket connections may be shorter than the physical distance computed above. Thus there is a need to increase the speed of the TDR step generator and the bandwidth of the oscilloscope so the combined system risetime is fast enough to resolve closely spaced reflections.

Lower quality cables and connectors can also slow down the effective system risetime and degrade resolution. It is interesting to examine the effect of the step rise time on a simple discontinuity, such as a SMA to BNC adaptor. You have already noticed that we have not used BNC connectors in the microwave frequency range, and the Agilent Application Note 1304-7 Typically, reflection performance changes with edgespeed because reflections are frequency dependent. This is easily observed with a return loss measurement on a network analyzer. When the amount of signal reflected is measured as a function of frequency, it is common to see that as frequency is increased, the magnitude of signal

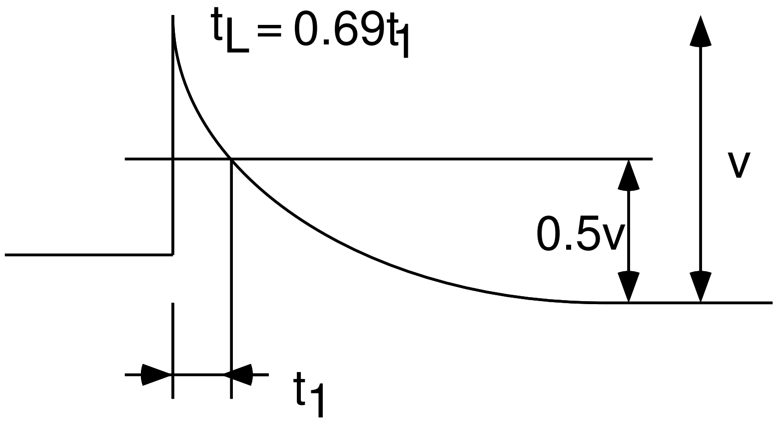


Figure 5.5: Determining the time constant of an exponential TDR response.

reflected back from a DUT increases.)

Notice in Fig. 5.6 a measurement of a 50-Ω SMA to BNC adapter, that as the risetime of the step stimulus is decreased, the nature of the reflection changes as higher data rates (frequencies) are used. At a 100 ps step speed, there is only one reflection seen corresponding to about 56 Ω. When the edge risetime is increased to 35 ps, more reflections are observed, with the dominant one corresponding to 71 Ω. At a step speed of 20 ps, the impedance discontinuity increases to over 77 Ω. In the case of the three measurements, the results obtained using a 20 ps risetime step excitation do not apply for a connector used with pulses with edges that are always slower than 100 ps in an actual application. This means that the connector might be acceptable for 100 ps edges but not for 20 ps edges. For example, systems operating at or above 10 Gb/s will involve signals with risetimes perhaps below 30 ps. Components for 40 Gb/s transmission may see edges under 10 ps. Thus a TDR with a flexible edgespeed can be useful when components used at a variety of data rates need to be analyzed.

In the lab, you will look at the reflected waveforms of a pulse off of different lumped elements using the time-domain option of the Network Analyzer. This means that the measurements are performed in frequency domain, filtered and then converted into time domain using an inverse Fourier transform. Why would one wish to use a network analyzer instead of a time domain reflectometer (basically and oscilloscope)? Consider the time domain response of a circuit with multiple discontinuities, as shown in Fig. 5.7.

The two mismatches produce reflections that can be analyzed separately. The mismatch at the junction of the two transmission lines generates a reflected voltage wave, V_{r1}, where

$$V_{r1} = \rho_1 V_+ = \frac{Z_{01} - Z_0}{Z_{01} + Z_0} V_+.$$

Similarly, the mismatch at the load also creates a reflected voltage wave due to its reflection coefficient

$$\rho = \frac{Z_L - Z_{01}}{Z_L + Z_{01}}.$$

Two things must be considered before the apparent reflection from Z_L, as shown on the oscilloscope, is used to determine ρ. First, the voltage step incident on Z_L is $(1 + \rho_1)V_+$ (not just V_+). Second, the reflected voltage from the load is $\rho(1 + \rho_1)V_+ = VrL$, but this is not equal to V_{r2} since a re-reflection

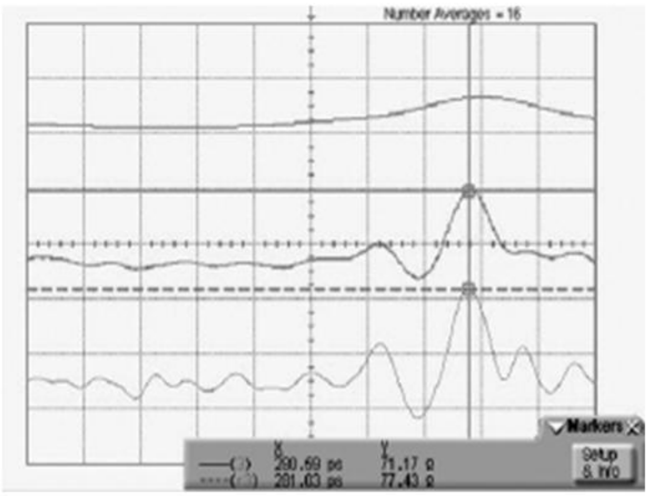


Figure 5.6: Effect of excitation rise time on TDR response of a SMA to BNC adaptor. The rise times shows are 100 ps, 35 ps and 20 ps.

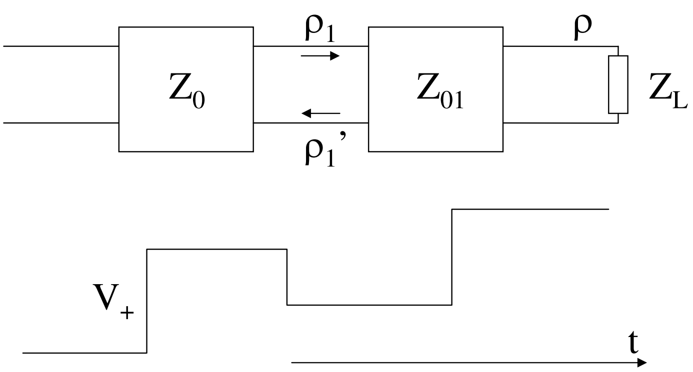


Figure 5.7: Time domain response of a circuit with multiple discontinuities, where $Z_{01} < Z_0$ and Z_{01}, Z_L.

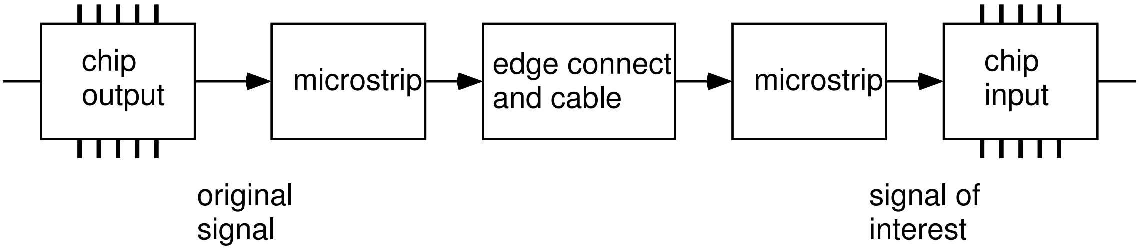


Figure 5.8: A typical digital circuit transmission path.

occurs at the mismatched junction of the two transmission lines. The wave that returns to the matched measurement instrument at the beginning of the line Z_0 is

$$V_{r2} = (1 + \rho_1')V_{rL} = (1 + \rho_1')[\rho(1 + \rho_1)V_+].$$

Since $\rho_1 = \rho_1'$, V_{r2} may be re-written as:

$$V_{r2} = [\rho(1\rho_1^2)]V_+.$$

The part of V_{rL} reflected from the junction of Z_{01} and Z_0 (i.e., $\rho_1' V_{rL}$) is again reflected off the load and travels back to beginning of the line only to be partially reflected at the junction of Z_{01} and Z_0. This continues indefinitely, but after some time the magnitude of the reflections approaches zero.

From the simple example above, it can be seen that with more discontinuities, reflections off the furthest discontinuity become smaller in magnitude because less of the initial incident wave amplitude makes it to the end of the line. Therefore, it is important to eliminate the effects of any unnecessary discontinuities, such as connectors. This is done very well with a network analyzer, which performs the measurements in the frequency domain. Thanks to calibration, the dynamic range of the measurement can be increased at the price of a more complicated measurement procedure.

5.5 TDR Considerations for Digital Circuits

Digital circuits often have printed microstrip transmission lines connecting pins of two chips, possibly through some extra interconnects, as described in Fig. 5.8. The designer wants to make sure that a 1 is indeed a 1 when it reaches the second chip, and the same for the 0 level. As edge speeds increase, depending on the logic family, transmission-line effects like overshoot, undershoot, ringing, reflections and crosstalk, can all become critical to maintaining noise margins. For example, in TTL, the values for a 1 are between 2.7 and 2 V, for a 0 they are between 0.5 and 0.8 V, and the noise margin is 0.7 V (with a 10–90% rise time of 4–10 ns). In very fast GaAs digital circuits (with a rise time of 0.2–0.4 ns), the values for a 1 are between −0.2 and −0.9 V, for a 0 they are between −1.6 and −1.9 V, and the noise margin is 0.7 V. From these numbers you can see that digital circuit margins are quite forgivable. For example, in the latter case, a 0.7-V undershoot on a 1.7-V signal is a 45% undershoot, which is considerable.

However, it is easy to have a 45% variation in a signal if there is a discontinuity (impedance mismatch) in a transmission line trace on a PC board. A special type of mismatch is coupling between adjacent traces, which are in effect parasitic inductances or capacitances. This is illustrated in Fig. 5.9 with two lines. If line 1 is excited by a step function, some of the voltage will get coupled to the next closest line even if the line is open-circuited (why?). The coupled signal will appear at both ends of the line, and this is called near and far-end crosstalk. For example, the cross talk could be as high as 25% for two parallel traces, while resistors and trace bends can cause up to 15% and 5% reflections, respectively.

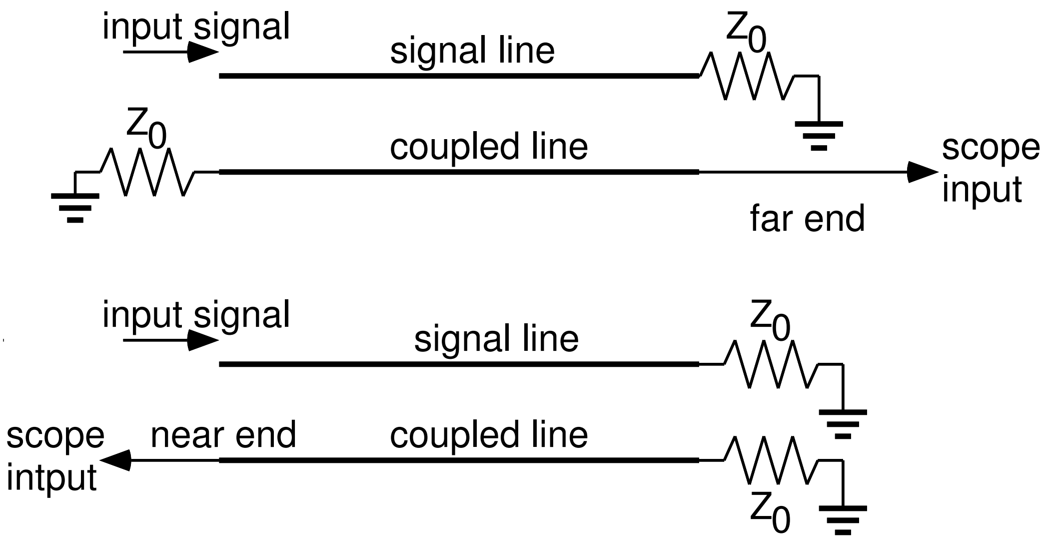


Figure 5.9: Near and far-end cross talk measurements in the case of two adjacent PC board traces.

All of these could be easily measured with TDR during the design of the digital backplane (PC-board containing all the traces for chip interconnects).

With the increase in speed of digital system design into the gigahertz region, frequency dependent effects become a bigger challenge. Yesterdays interconnects could be easily characterized by measuring the self-impedance and propagation delay of the single-ended transmission line. This was true for printed circuit board stripline, microstrip, backplanes, cables and connectors.

However, high-speed serial data formats in todays digital standards demand differential circuit topology, which also means that new measurement techniques need to be developed. For example, several new implementations of PCI Express and Infiniband reach data rates into the 4 Gb/sec range. New standards, such as XAUI, OC-192, 10-G Ethernet, and OC-768 aim even higher - up to and past 40 Gb/sec. This upward trend creates signal integrity challenges for physical layer device designers.

In order to limit unwanted effects of radiative coupling (electromagnetic interference - EMI), differential circuits are common practice, Fig. 5.10(a). Ideal differential linear passive interconnects respond to and/or generate only differential signals (two signals of equal amplitude and opposite sign). Any radiated external signal incident upon this ideal differential transmission line is considered a common signal and is rejected by the device. The so called common mode rejection ratio (CMRR) is a measure of how well a differential device rejects the unwanted common mode. The radiated common signals are usually generated from adjacent RF circuitry or from the harmonics of digital clocks. Properly designed differential devices can also reject noise on the electrical ground, since the noise appears common to both input terminals. Non-ideal differential transmission lines, however, do not exhibit these benefits. A differential transmission line with even a small amount of asymmetry will produce a common signal that propagates through the device. This asymmetry can be caused by any physical feature that is on one line of the differential pair and not the other line, including solder pads, bends, ground plane imperfections, etc. This mode conversion is a source of EMI radiation. Most new product development must include EMI testing near the end of the design cycle. There is usually very little insight as to what physical characteristic is causing the EMI problem. Mode conversion analysis provides the designer with that insight so that EMI problems can be resolved earlier in the design stage.

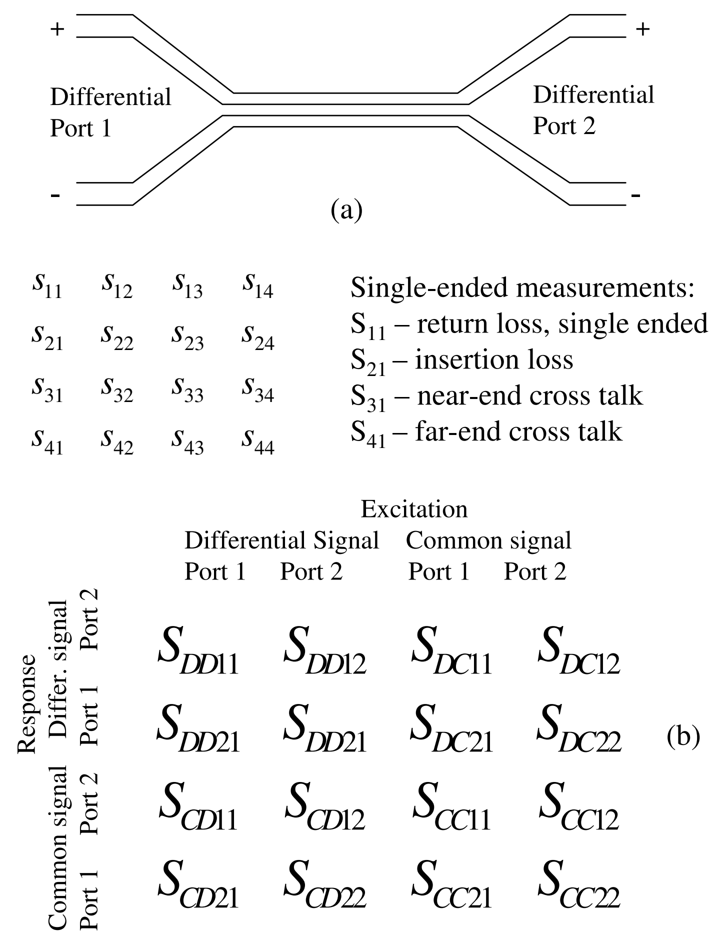


Figure 5.10: (a) Differential passive interconnect circuit (implemented in microstrip). (b) Single-ended and differential-common mode s-parameters.

To understand how odd and even (common) mode analysis applies to digital interconnects in time domain, consider a pc board layout as shown in Fig. 5.10(a). The two lines of this basic differential circuit initially have a single ended impedance of 50 Ω. The lines are physically separate with minimal coupling. The two lines then come together and the two trace widths are reduced (which would cause an increased single ended impedance). The lines are then increased in width and are spread apart. If each trace is tested individually (driven single ended), the TDR results are seen as a 50-Ω line, then a 70-Ω section, and a 50-*omega* section before being terminated in a 50 Ω load. The result is the same for each trace. In the differential measurement, the TDR system combines the results from both ports when stimulated by both steps. Thus the signal on each trace will be a combination of signal from both step generators. The result is that the differential impedance is close to 100 Ω, which was the intent of the transmission line design. The odd-mode impedance (one trace of the transmission line grounded when driven differentially) is close to 50 Ω. The result is identical to what would have been achieved with simultaneous stimulation of the four-port device, the only exceptions being if there were any asymmetry in the two pulses (in both timing or pulse shape), the imbalance would be transferred to the TDR result as measurement error.

Important insights into component behavior can be achieved through frequency domain analysis in addition to characterization in the time domain. For example, a common measurement is to determine the amount of signal that is reflected back from a component over a specific range of frequencies, perhaps from the kilohertz range through the gigahertz range. The frequency response results often yield important insight into component response, e.g., resonances are easily detected in the frequency domain. For the case of the two-port component, we are concerned with transmission and reflection at each port, a total of four complex S parameters. A differential component with just a positive and negative input and output adds two ports and four more S parameters. However, differential channels can couple to their complementary channels, doubling the eight S parameters to 16. Note that the excitation and response for these measurements are still effectively single-ended. That is, only one port is excited and one measured to construct each of the S parameters. In the following example for a differential circuit, one differential port pair is noted by ports 1 and 3, the other differential port pair by ports 2 and 4. The 16 possible measurement configurations and some physical interpretations are shown in Fig. 5.10(b).

A differential circuit can be driven in either differential or common mode, and the response measured in a differential or common mode. Thus the full S-parameter set for a two-port differential component, including single-ended, differential, common, and mixed mode configurations has 32 unique S parameters. It is important to interpret what the various differential and common mode measurement configurations provide. The differential S-parameter notation is slightly different than the single ended notation. It still follows an S “out-in” format. However, port 1 includes both the positive and negative traces of the differential input, as so does port 2.

Thus s_{DD11} indicates the reflected differential signal when stimulated differentially. Similarly, s_{DD21} indicates the differential output (at differential port 2) when a differential signal is input to differential port 1. Thus there are four basic quadrants to the 16 element differential S-parameter matrix, as displayed in Fig. 5.10. The upper left quadrant is a measurement of differential transmission and reflection for a device with two differential ports (typically differential in and differential out) when stimulated with differential signals. Similarly, the lower right quadrant gives common transmission and reflection performance when the two port device is stimulated with common mode signals.

The mixed mode parameters (combinations of differential and common mode stimulus or response) provide important information about how conversion from one mode to the other may occur which in turn provides insight into how components and channels may radiate or be susceptible to radiated signals. For example, the lower left quadrant indicates how differential input signals are converted to common mode signals. s_{CD21} would be a measure of how a differential input to port 1 is observed as a common mode signal at port 2. Common mode signals are more likely to cause radiated emissions than a differential signal, hence the SCD quadrant is useful in solving such problems. The upper right quadrant (SDC) indicates how common signals are converted to differential signals. Differential systems are intended to reduce susceptibility to spurious signals by rejecting anything that is common to both legs of the differential system. But if spurious common mode signals are converted to differential signals, they no longer are rejected. Hence the SDC quadrant measurements are useful in solving problems of susceptibility to spurious signals. For example, s_{DC21} indicates how a signal that is common mode at port one is converted to a differential signal and observed at port 2.

5.6 Practice questions

1. Why could we not use simple transmission-line analysis when calculating the step response of the inductor in Figure 5.1?
2. If you had a break in the dielectric of a cable causing a large shunt conductance, what do you expect to see reflected if you excite the cable with a short pulse (practical delta function)?
3. If you had a break in the outer conductor of a cable, causing a large series resistance, what do you expect to see reflected if you excite the cable with a short pulse (imperfect delta function)?
4. What do the reflected waves off a series inductor and shunt capacitor in the middle of a transmission line look like for a short pulse excitation, assuming that $\omega L \gg Z_0$ and $\omega C \gg Y_0$?
5. Derive the expression $t_L = 0.69 t_1$ discussed in section 5.3. This expression shows a practical way to measure the time constant of the reflected wave for the case of complex loads.
6. A PC board trace in a digital circuit is excited by a voltage $v(t)$. Derive an equation for the coupled (crosstalk) signal on an adjacent line $v_c(t)$ assuming the adjacent line is connected to a load at one end and a scope (infinite impedance) at the other end, so that no current flows through it. (***Hint:*** the coupling is capacitive and you can draw a simple capacitance between the two traces and do circuit theory.)
7. Explain how odd and even mode analysis from Lecture 4 is related to differential and common mode analysis in this lecture.

5.7 Homework Problems

1. Determine the equivalent Thevenin circuit of a matched generator feeding an open-ended transmission line, at the terminals defined by the open-circuit termination.

2. Sketch the waveforms corresponding to the reflected voltage, as well as the sum of the incident and reflected voltage waves at the source end, assuming the incident wave is a unit step function, for the cases shown in Fig. 5.11.

[Hint: Follow the examples of the short circuit termination and the inductive load that were done in the notes: (a) and (e) are free!]

3. A transmission line with characteristic impedance Z_0 is connected to a section of line with a different characteristic impedance Z, which is then terminated in a load Z_L, Fig. 5.12. Find the expression for the reflected voltage at the reference plane, which is at the junction between the Z_0 line and the Z line, considering only one round-trip bounce of the incident wave $v_+(t)$ on the Z line. If $v_+(t)$ is a unit step function, and $Z_0 = 50\,\Omega$, $Z = 75\,\Omega$ and $Z_L = 100\,\Omega$, sketch the reflected waveform.

4. A transmission line with characteristic impedance Z_0 is terminated in a load consisting of a resistor R in series with a capacitor C. Sketch the reflected voltage, as well as the sum of the reflected and incident voltage, for a unit step incident voltage wave. Explain your result.

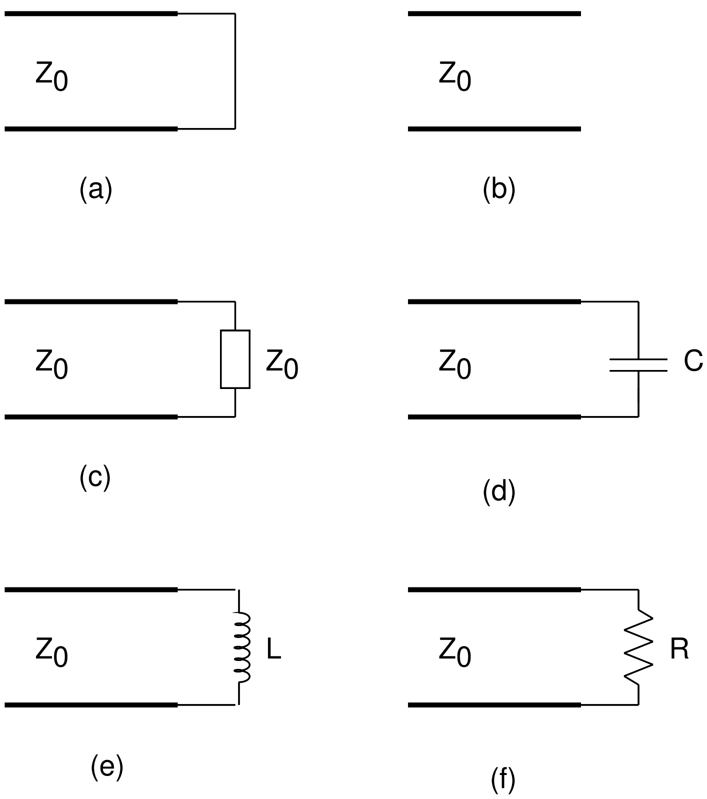


Figure 5.11: Find the waveforms of a TDR display for a short circuited line (a), an open circuit (b), a matched load (c), a capacitive load (d), an inductive load (e), and a resistive load (f).

5. Repeat problem 4, for the case of a load consisting of an inductance L in parallel with a resistor R.

6. Repeat problem 4 for the case of a load which is a parallel combination of a resistor R and a capacitor C.

7. Repeat problem 4 for the case of a load consisting of an inductance L in series with a resistor R.

8. Repeat problem 4, for the case when the load is a series combination of an inductance L with a capacitance C.

9. Using Ansoft Designer, simulate the time domain responses of circuits from problems 4-8.

10. Using Ansoft Designer, simulate in the time domain a 50-ohm source feeding a matched line of some electrical length L with a 25-ohm load at the end. Next place a section of line of different impedance between the line and the load. How does the impedance and length of the line affect the precision of the reflection measurement off the load?

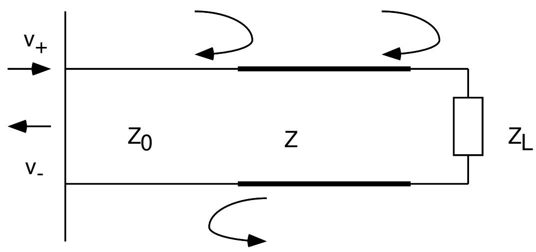


Figure 5.12: A cascade of transmission lines terminated in a load.

Chapter 6

Nonlinear Microwave Circuits

6.1 Microwave Sources

Sources are a necessary part of every microwave system. In transmitters, they supply the RF power, and in receivers, they are the so-called local oscillator (LO). The word source refers to an oscillator: DC power is supplied to it, and it produces AC power. For high power transmitters, the source usually generates very high power pulses, and where moderate or low powers are needed, the source operates in CW (continuous wave) mode, which means that it generates a sinusoidal wave. There are two types of microwave sources: *solid-state* sources, like the Gunn and IMPATT diode and various types of transistors (MESFETs, BJTs, HBTs, HEMTs, etc.) and *tubes*, like the klystron, magnetron and TWT (travelling wave tube). Tubes can give high powers and they are low noise, but they are usually expensive, need large DC power supplies (several hundreds to several thousands of volts), they generally are not very efficient (klystrons are about 10% efficient), and they have a limited lifetime (typically a few thousand hours) due to the fact that the electrodes wear out. Solid-state devices, on the other hand, are cheap, compact, need small power supplies (on the order of ten volts), are very reliable, but are usually noisy and can produce at best medium power levels. There has been a huge effort to replace tubes with solid-state sources wherever possible, but tubes are still used for very high power applications, as well as for the high millimeter-wave frequencies, where solid-state sources barely exist and can give only μW power levels.

Microwave Tubes

Historically, tubes were the first microwave sources, and they were developed during the second world war in Great Britain, and then also in the MIT radiation lab in the United States. The first tubes were the *magnetron* and the *klystron* developed for radar transmitters. Both of these tubes are still in use today: the magnetron is used in microwave ovens, and the klystron in high-power transmitters (up to 10 kW) for high orbit satellite communications (geosynchronous) and at higher millimeter-wave frequencies (all the way up to 600 GHz).

The *klystron* can be used as either an amplifier or an oscillator. Usually the so called reflex klystron is used as an oscillator, Fig. 6.1. An electron gun, consisting of an electron emitting cathode (usually made out of tungsten) and an anode at potential $+V$ with respect to the cathode, produce an electron beam with a velocity $v = \sqrt{2qV/m}$ and concentration of about 10^{12} to 10^{15}. A group of metal grids is situated in a resonator, through which the power is coupled out by means of a coax with a loop, through a waveguide, or through a semitransparent cavity wall.

The role of the grids is to perform the so called velocity modulation and bunching of the electron beam. Let us assume that the klystron is already oscillating. An RF voltage $v_g = V_g \cos\omega t$ exists between the grids, which makes the first one positive (or negative), and the second one negative (or positive) with

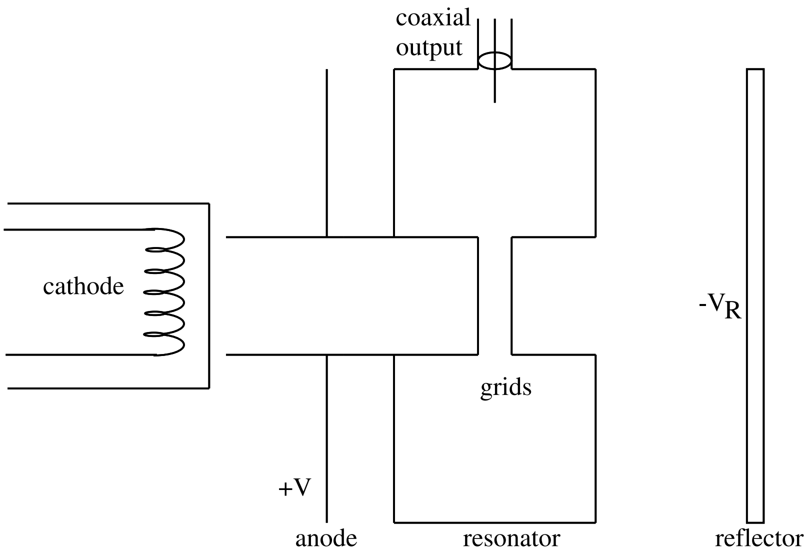


Figure 6.1: The reflex klystron.

respect to V at any given moment. What happens to the electrons that are flying between these grids? Some of them will pass through the region between the grids when $V_g \cos \omega t > 0$ and will therefore be accelerated, some will pass when $V_g \cos \omega t < 0$ and will be decelerated, and some will be unaffected, since they will pass when $V_g \cos \omega t = 0$. This is called velocity modulation. Now we have faster electrons catching up with slower electrons, and this results in so called bunching. Now the electron beam has a nonuniform density, which means that there is an AC current that induces an AC electromagnetic field. The electrons in bunches now go towards the reflector, which is at a negative potential $-V_R$ and this decelerates the electron bunches, and turns them back towards the grids. If the distances and voltages are properly designed, they will get to the space between the grids in synchronism with the RF voltage between the grids, so that positive feedback is ensured. How does this now give us microwaves?

The resonator is tuned to a microwave frequency (its dimension is a whole number of half-wavelengths). There is an electromagnetic field induced by the electron beam AC current in this cavity, but in order to get some net power in this field, somehow some energy has to be taken from the electron beam. This cannot happen if the beam is accelerated and then decelerated and so on, so that the average energy is constant. The answer is in Lenz's law: the AC current induces an EMF opposing the density variation in the electron beam. This means that, as the beam density increases, the induced EMF will decelerate the electrons, and when the beam density decreases, it will accelerate them. The electrons will be more decelerated when the beam density is greater, and less accelerated when the beam density is smaller. The end result is that many electrons are being decelerated, and few accelerated. This, in turn, means that the electron beam is giving some energy up to the electromagnetic field in the resonator.

The klystron can be modulated by changing the voltage V_R, and its frequency can be tuned by tuning the resonator. It can be operated in pulsed mode by pulsing the voltage V. It is used in ground and airborne radar, microwave relay links, and, because of its robustness and ability to operate in severe environmental conditions, in guided missiles.

The *magnetron* falls in the category of so called crossed-field tubes. It is a vacuum tube with a cylindrical cathode surrounded by a coaxial anode at a potential V, Fig. 6.2. This produces a radial electric field E between the electrodes which tends to move the electrons on straight radial paths. There is also a DC magnetic field H parallel to the tube axis, and this tends to send electrons back to the cathode

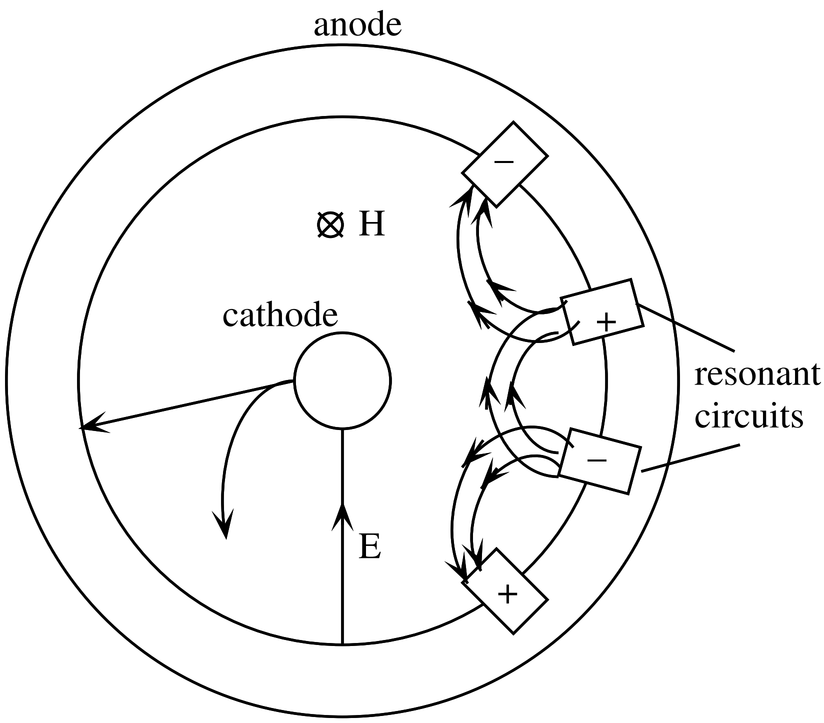


Figure 6.2: A magnetron – the center electrode is the cathode surrounded by a coaxial anode formed of an even number of resonant circuits.

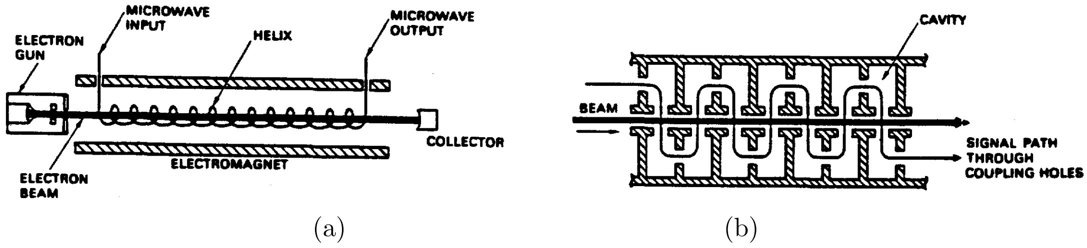


Figure 6.3: The TWT with a helix slow wave structure (a) and periodically loaded waveguide slow-wave structure (b).

in circular paths. The anode is formed of an even number of resonant circuits at a mean potential V. A high frequency potential difference is established between the cavities, so that the polarities alternate between them, as shown in Fig. 6.2. A half period later in time, the polarities will be reversed from those shown in the figure, and it will be as though the field rotated with an angular velocity equal to $2\omega/n$ where n is the number of cavities. Now the electrons rotating at an angular velocity greater than that of the field will give up some of their energy to the field. For these electrons, the electric force is greater than the magnetic force, and they will get nearer the anode. The electrons that are rotating slower than the field will be accelerated and take energy from the field. The magnetic force for these electrons becomes greater than the electric force, so they will get nearer the cathode. The first group of electrons interacts in a region of stronger fields and over a longer time than the second group, so in the overall energy balance, the electrons give energy to the field.

The anode is relatively small, and because of the high currents it gets very hot. It is almost always designed for pulsed operation, so that the anode has some time between pulses to cool down. The order of magnitude of power levels is 10 MW at 1 GHz, 1 MW at 10 GHz and 0.1 MW at 30 GHz. They are very efficient, up to even 80%. Magnetrons were the main source used for high power transmitters, until multi-cavity klystrons were made in the 60's. However, since they are cheap, robust, reliable and easy to replace, they became the standard source for microwave ovens (several kW at 2.45 GHz) and medical equipment. New types of magnetrons are also used in radar systems.

The *TWT* (Traveling Wave Tube) is a tube amplifier that falls in the category of *slow-wave devices*. As shown in Fig. 6.3(a), it consists of an electron gun, which produces the high-energy electron beam; the helix which guides the signal that is to be amplified; the collector that absorbs the unspent energy of the electrons which are returned to the gun by a DC power supply; and an electromagnet that keeps the beam from spreading. The microwave input signal is introduced at one end of the helix, and since it needs to pass a greater distance spiraling down the helix, the speed of the signal will be smaller than the speed of light. This means that the microwave signal is slowed down compared to the speed of the electrons in the beam. (This is why the helix is called a slow-wave structure.) Now the signal bunches up the electrons in the beam around its nulls, and the traveling electrons in turn induce an AC electromagnetic field. This field travels a bit faster than the signal, so the electrons give up some of their energy to the signal and amplify it. When the signal reaches the end of the helix, it ends up in the waveguide end of the tube. The remaining kinetic energy of the beam is spent as heat when the electrons hit the collector, and TWT's must be well cooled. As the power increases, the helix becomes less suitable, and often a different slow wave structure in the form of a periodically corrugated waveguide is used, Fig. 6.3(b).

The TWT is mostly used as a high-power amplifier. Maximum output powers from TWT's are, in CW mode 100 kW at 2 GHz and 1 kW at 100 GHz, and in pulsed mode, 10 MW at 2 GHz and 10 kW at 100 GHz. TWT's have efficiencies of 20 to 50%. They have a broader bandwidth than the other tubes, so they are used in terrestrial and space communication systems, radar and military jamming. Traveling wave oscillators are the so called carcinotron and BWO (Backward Wave Oscillator). They have been able to achieve operation at very high frequencies in the submillimeter-wave and far infrared range, up to 900 GHz.

Solid-State Sources

A solid-state device is a device made of a semiconductor material, such as silicon, gallium arsenide, indium phosphide, silicon carbide, etc. At lower frequencies, such devices are made out of silicon. At microwave frequencies, however, the electrons and holes in silicon are too slow to follow such fast changes in the electric field, so another material, GaAs is used instead. GaAs is much harder and more expensive to process than silicon, and that is one of the reasons why microwave solid-state devices are more expensive than the equivalent ones at lower frequencies.

Solid-state sources at microwave frequencies can be based on either two or three-terminal devices. Two-terminal devices are diodes: Gunn and IMPATT diodes are most common, but others such as tunnel diodes are occasionally encountered. Three-terminal microwave devices are all various kinds of transistors. Microwave integrated circuits are made with GaAs devices, and they are much smaller in scope than low frequency IC's. An important difference is that at high frequencies the inductive reactance of wires becomes much more important and there is more coupling between metal lines, so transmission line effects have to be accounted for. Circuits are typically several wavelengths large as a result, and thus so that microwave IC's are much larger than comparable low frequency ones. Also, since GaAs is hard to process, the yields across a wafer are small, and it becomes even more expensive to make a larger circuit with many devices in it.

A *Gunn diode* is a piece of GaAs with two contacts on it. When a voltage is applied to the two terminals, there is an electric field established across the piece of GaAs (just like in a resistor). GaAs has the property that the electrons have different velocities depending on how large the imposed electric field is. As shown in Fig. 6.4(a), as the electric field is increased, the velocity increases up to a certain point, and then the electrons slow down with a further field increase. When a bunch of electrons slow down in a piece of GaAs under DC bias, these electrons make a "traffic jam", and more electrons pile up to form a charge layer, Fig. 6.4(b). This layer now produces an electric field that decreases the original field to the left of them, and increases the field to the right, so that the charge bunch gets pushed towards the electrode on the right, forming a current pulse. When this pulse gets to the electrode, the electric field goes back to the original value, another traffic jams happens, and another pulse forms, and so oscillations form. The frequency will depend on the length of the piece of GaAs between electrodes, as well as the concentration of electrons. At 10 GHz, a typical Gunn diode will have an active region 10 μm long and a doping level of $10^{16}\,\mathrm{m}^{-3}$. They are usually biased at around 8 V and 100 mA, and have low efficiencies, only a few percent. This means that the diode dissipates a lot of heat and needs to be packaged for good heat-sinking. A MaCom Gunn diode package and basic specifications are shown in Fig. 6.5. Sometimes the chip diode is packaged on a small diamond heat sink mounted on the metal. The largest achieved power levels from commercial Gunn diodes are a few kW with less than 10% efficiency at 10 GHz in pulsed mode, and 1 W with less than 5% efficiency in CW mode. Indium phosphate Gunn diodes have achieved about 0.1 W CW at 100 GHz with about 3% efficiency.

IMPATT diodes (IMPact ionization Avalanche Transit Time) is a *pn* diode under reverse bias. The applied reverse bias voltage is large enough, 70-100 V, so that an avalanche process takes place. This means that some of the electrons get violently accelerated and as they hit the atoms, they form electron-hole pairs that are now new carriers, which in turn get accelerated and an avalanche process happens. The fact that there is a saturation velocity of the charges in every semiconductor accounts for oscillations in an IMPATT diode. The mechanism of oscillation of an IMPATT diode does not depend on the carrier mobility, so silicon IMPATT's are made as well. They dissipate a lot of heat, are usually used in pulsed mode, have higher efficiencies than Gunn diodes (about 20%), and are very noisy because of the avalanche mechanism. Commercial IMPATT's are packaged in the same way as Gunn diodes, and the maximum available powers are 50 W pulsed at 10 GHz and 0.2 W CW at 100 GHz.

The three-terminal solid state devices are different types of transistors. They are used as amplifiers, as well as oscillators, depending on what kind of a circuit they are placed in. In a transistor oscillator, the

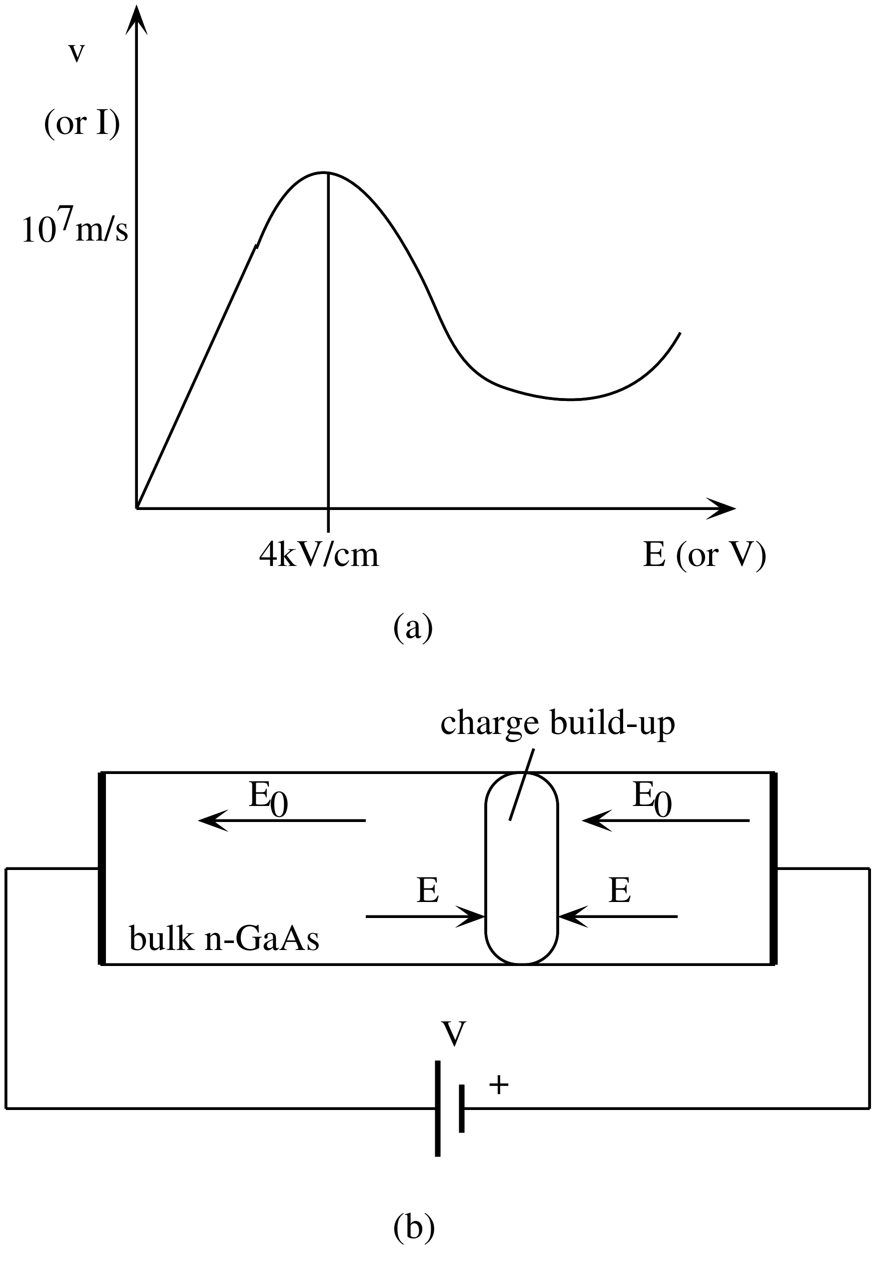
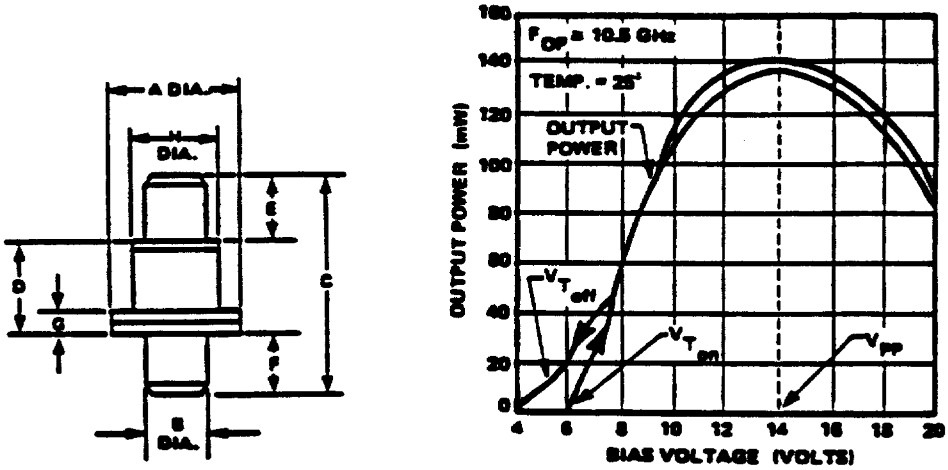


Figure 6.4: Electron velocity versus electric field magnitude for GaAs (a) and oscillation buildup in a Gunn diode (b).



(a) (b)

Figure 6.5: A commercial MaCom Gunn diode package (a) and specifications (b).

circuit provides positive feedback, so that the device is unstable. The most commonly used transistor at microwave frequencies is the GaAs MESFET (MEtal Semiconductor Field Effect Transistor), which looks like a MOSFET, but has no oxide, so that the gate contact is a Schottky diode. This makes the device very fast, and MESFET's were made to have gain up to 100 GHz. Other types of transistors used at microwave and millimeter-wave frequencies are HBT's (Heterojunction Bipolar Transistor) and HEMT's (High Electron Mobility Transistors). The fabrication of these transistors requires special technology in growing heterostructure semiconductors, and often very sophisticated photolithography. Transistors are efficient oscillators (up to 40% efficiency), but give relatively low power (a single transistor in an oscillator circuit cannot give more than a few tenths of a Watt). They have low noise, and they offer control through the third terminal, which can be used for tuning, modulation or injection-locking.

6.2 Oscillators

In the lab you will look at a Gunn-diode version of a two-port negative-conductance oscillator as shown in Fig. 6.6. Let us examine the operation of this circuit.

Quasiharmonic Description of Negative-Conductance Diodes

As we have already learned in chapter 4, a diode is a nonlinear device whose current-voltage (I-V) characteristic can be written in the form of a Taylor series:

$$I(V) = I_0 + i = I(V_0) + v_d G_d + \frac{v_d^2}{2} G_d' + \frac{v_d^3}{6} G_d'' + \dots \tag{6.1}$$

where I is the total diode current, $V = V_0 + v_d$ is the total diode voltage, $I(V_0)$ is the DC bias current, V_0 is the DC bias voltage and i is the current due to the RF signal voltage v_d across the diode. We should keep in mind that this is only a partial view of a diode's behavior, because there is a device capacitance that becomes important as frequency increases. However, the conductance behavior described by (6.1) is the basis for oscillator and amplifier operation, and we will focus our attention on it.

If the RF voltage contained only a single frequency component:

$$v_d \simeq \sqrt{2} V_1 \cos \omega t \tag{6.2}$$

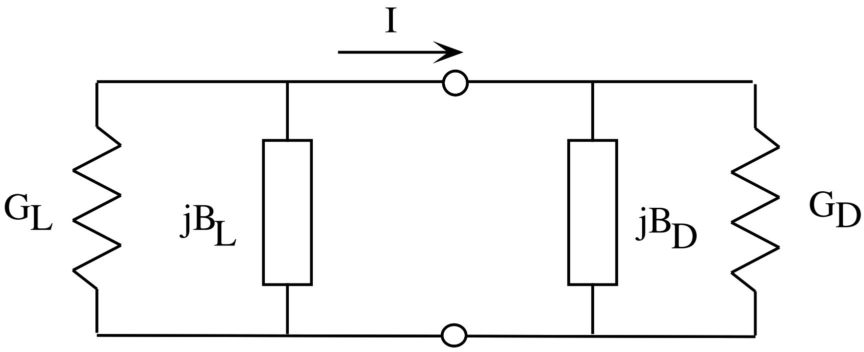


Figure 6.6: A one-port RF oscillator circuit.

where V_1 is nonnegative and real, the nonlinear nature of (6.1) would create in the current i not only a current at this *fundamental* frequency ω, but also at the second harmonic 2ω, the third harmonic 3ω and so on. Indeed, because of the trigonometric identities

$$\cos^2 \omega t = \frac{1}{2}(1 + \cos 2\omega t), \qquad \cos^3 \omega t = \frac{1}{4}(3\cos \omega t + \cos 3\omega t) \tag{6.3}$$

etc., we can write (6.1) as

$$\begin{aligned} I(V) &= \left[I(V_0) + \frac{1}{2} G_d' V_1^2 + \dots \right] \\ &\quad + \sqrt{2} \left[G_d V_1 + \frac{1}{4} G_d'' V_1^3 + \dots \right] \cos \omega t \\ &\quad + \sqrt{2} \left[\frac{1}{2\sqrt{2}} G_d' V_1^2 + \dots \right] \cos 2\omega t \\ &\quad + \sqrt{2} \left[\frac{1}{12} G_d'' V_1^3 + \dots \right] \cos 3\omega t + \dots \qquad (6.4) \\ &= I_0 + \sqrt{2} I_1 \cos \omega t + \sqrt{2} I_2 \cos 2\omega t + \sqrt{2} I_3 \cos 3\omega t + \dots \end{aligned}$$

A linear external circuit connected to such a device will of course produce voltages at these harmonic frequencies in addition to the fundamental, but it is often possible to use the so-called *quasi-harmonic* approximation, in which for purposes of solving the circuit only the fundamental frequency is taken into account. Better approximation to the circuit behavior can be obtained using the *harmonic balance* method, in which a number of higher harmonic frequencies are also taken into account.

We thus take only the fundamental components in (6.4), which results in the diode being characterized by a conductance G_D which is dependent on the amplitude V_1 of the fundamental RF voltage applied to it:

$$I_1 = G_D(V_1) V_1 \tag{6.5}$$

where

$$G_D(V_1) = G_d + \frac{1}{4} G_d'' V_1^2 + \dots \tag{6.6}$$

It is assumed that $|I_2|, |I_3|, ... \ll |I_1|$. The diode is thus represented as a nonlinear one-port device, whose input admittance is a function of the RF voltage amplitude: $Y_D = G_D + jB_D \simeq G_D(V_1)$. This admittance is usually a function of bias, frequency and temperature in addition to V_1.

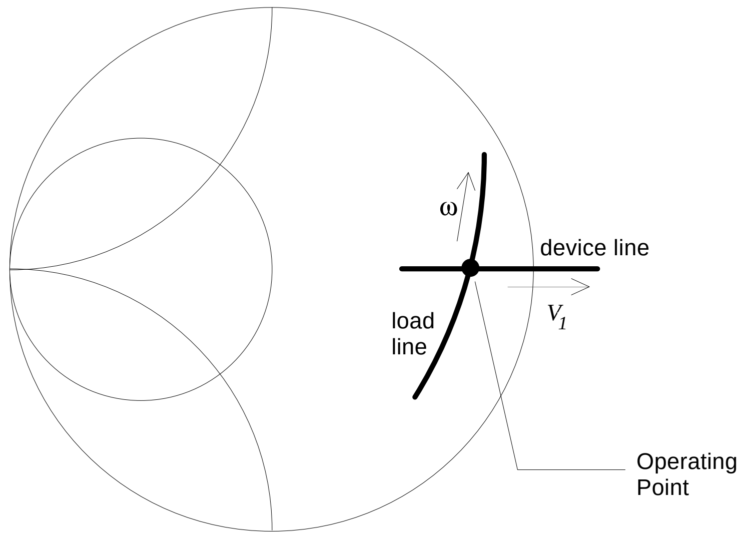


Figure 6.7: Admittance Smith chart showing intersection of device line and load line for oscillation condition.

Oscillator Resonance Condition

The diode in Fig. 6.6 is connected to a passive load admittance $Y_L(\omega) = G_L(\omega) + jB_L(\omega)$. Kirchhoff's voltage and current laws give us

$$[Y_D(V_1) + Y_L(\omega)]V_1 = 0. \tag{6.7}$$

or, if the RF voltage V_1 is not zero: $Y_D(V_1) + Y_L(\omega) = 0$. For a given RF voltage amplitude V_1, the solutions ω_r to this equation represent the complex natural (angular) frequencies of the circuit. When V_1 is such that one of these natural frequencies is real, the circuit oscillates. The two real equations that determine this condition are:

$$G_D(V_1) + G_L(\omega) = 0 \qquad \text{and} \qquad B_L(\omega) = 0. \tag{6.8}$$

Since we have neglected the susceptance B_D of the diode, the second of eqns. (6.8) will determine the frequency of oscillation. Once this has been done, the first equation is then solved to find the amplitude V_1 of the RF voltage produced. Since the load is passive, $G_L > 0$, and we must have $G_D = -G_L < 0$. The positive conductance G_L dissipates energy, while the negative conductance G_D means that energy is produced. That is why you will often read in the literature about negative-conductance (or negative-resistance) devices, meaning sources.

The solution of (6.8) can be represented graphically on the Smith chart. First, we plot a line corresponding to the variation of Y_L with ω. This is known as the *load line*. On the same plot, we superimpose a line corresponding to the variation of the *negative* of the device admittance $-Y_D$ (in our case, $-G_D$) with the RF voltage amplitude V_1. This is known as the *device line*. The intersection of the device line with the load line is the solution of (6.8) (see Fig. 6.7). This graphical approach can also be used when the susceptance B_D of the device needs to be taken into account.

The way in which oscillations build up in the circuit is of some interest. For typical negative-conductance diodes, the zero-amplitude conductance $G_D(0) = G_d$ is negative, while the coefficient G_d'' of the *I*-*V* curve is positive. This results in a device conductance vs. V_1 dependence as shown in Fig. 6.8. If there is initially no signal in the circuit, but a small amount of noise (corresponding to a small value of V_1) is present, then G_D is more negative than it will be at the oscillation point, $G_D + G_L$ is negative, and the natural frequency of the network will have a negative imaginary part (or in Laplace transform terminology, is in the right half-plane). This means that the circuit response is a growing exponential,

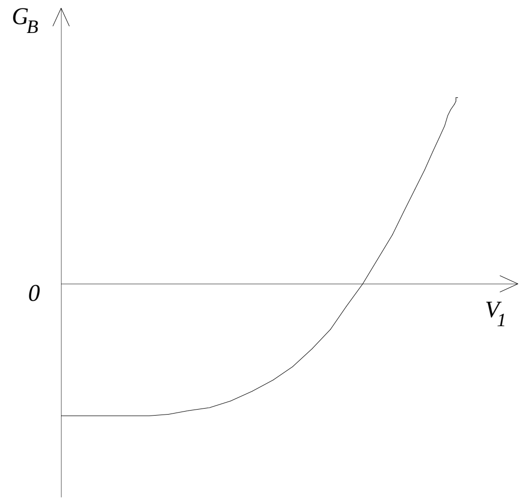


Figure 6.8: Typical G_D vs. V_1 curve for a negative-conductance diode.

causing the amplitude V_1 to increase. But this has the effect of causing G_D to become less negative, and along with it the imaginary part of the natural frequency. Finally, when the amplitude V_1 reaches its value for free oscillations, the imaginary part of the natural frequency goes to zero, and the oscillation is stabilized. In the same way, we can see that if any small deviation from the steady-state oscillation temporarily occurs, the value of V_1 will tend to return to its steady-state value again. The oscillation is stable in our example.

Injection Locking

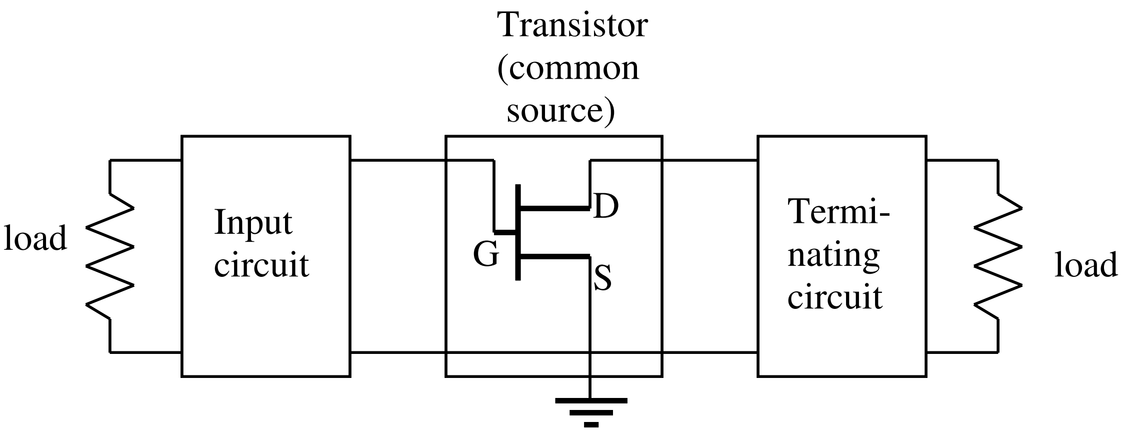
An oscillator can be coaxed into oscillating at a slightly different frequency than the free-running oscillation frequency determined by the load admittance as in (6.8). This is done by injecting a small RF signal at the desired frequency (sufficiently close to the free-running frequency) into the oscillator circuit. If the Kirchhoff circuit laws can be met at the injection frequency, the oscillation will occur at this frequency, and the oscillator is said to "lock" at the new frequency. The larger the injected voltage, the farther away from the free-running oscillation frequency we can make the circuit oscillate.

Two-port Oscillators

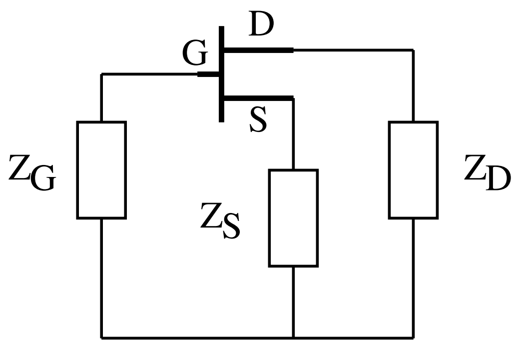
Let us next look briefly at transistor oscillators. A transistor has three terminals, therefore it is hard to visualize a transistor oscillator as a one-port oscillator. Transistors are usually thought of as two-port devices, one terminal usually being grounded (common-source, common-drain or common-gate configurations). If none of the terminals is grounded, we can think of it as a three-port device. Oscillator circuits for both cases are shown in Fig. 6.9.

6.3 Resonators

We saw that the right admittance needs to be connected to a negative conductance device in order for it to oscillate. Oscillators are often named by the type of circuit connected to them. Since this circuit usually determines the frequency of oscillation to some extent, and it acts as a tank of energy, it is called a resonator. The most common resonators are lumped elements, sections of transmission line, waveguide cavities, dielectric resonators, YIG resonators and varactor diode resonators. The last two types give



(a)



(b)

Figure 6.9: A two-port transistor oscillator RF circuit (a), and a three-port transistor circuit (b), where each of the leads are connected to transmission lines with a common ground lead.

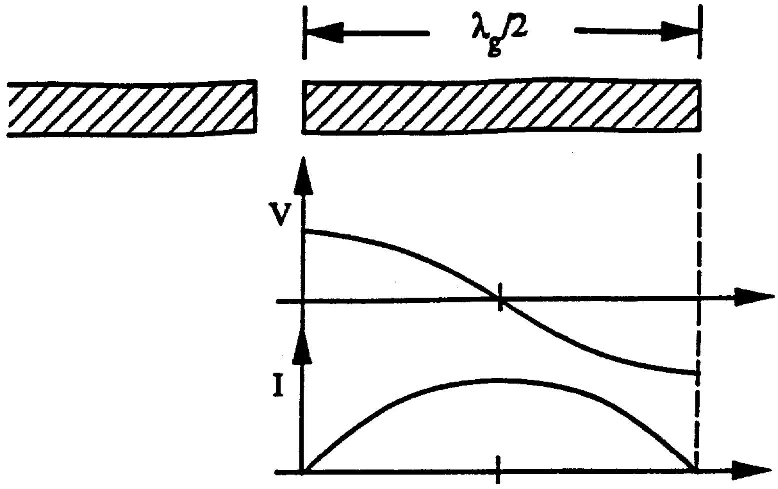


Figure 6.10: Microstrip resonator with associated voltage and current waveforms. The energy is usually coupled in through a gap (capacitive coupling).

electrically tunable oscillators, so called VCO's. All of the structures can be made to have low loss and high quality (Q) factors. For a high Q resonator ($Q > 50$), the reflection coefficient is on the boundary of the outer Smith chart, and the phase is determined by the transmission line connecting it to the device.

Lumped element resonators are high Q capacitors and inductors with associated parasitics. For example, a 1 mm chip capacitor has typically about 0.5 nH parasitic inductance.

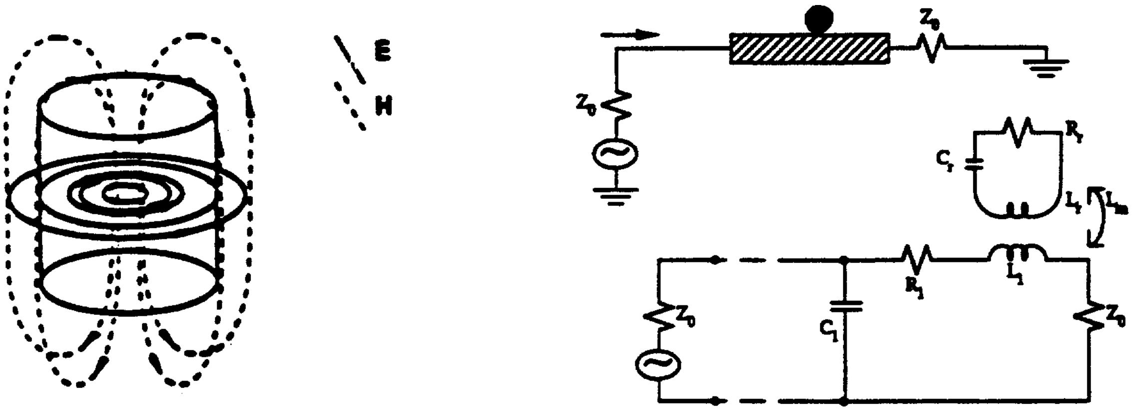
Microstrip resonators can be open or shorted sections of line that provide the right impedance for instability, rectangular $\lambda/2$ resonating line sections, circular disks, circular rings, triangular microstrip etc. We will test a microstrip resonator in the lab. The resonator looks like the one in Fig. 6.10. A Gunn diode is placed in the resonator so that the frequency of oscillation is determined approximately by the resonator half-wave resonance. (It is approximate because the Gunn diode is not infinitely small and its loading changes the resonator properties.)

Waveguide cavity resonators are usually $\lambda/4$ shorted stubs. The output can be coupled out either with a short loop (magnetic coupling) or a short monopole (electric coupling). Often there is a mechanical tuning screw near the open circuit end of the cavity. The lowest order rectangular cavity resonator is a TE_{101} mode, where the width and the length of the cavity are $\lambda_g/2$ at the resonant frequency.

A dielectric resonator looks like an aspirin pill (often called a dielectric puck). These are low cost resonators used externally in microstrip oscillators. They are made out of barium titanate compounds, with relative dielectric constants between 30 and 90. The increased dielectric constant gives high energy concentration, but also higher losses. The mode used in commercially available resonators is the so called $TE_{01\delta}$ mode, shown in Fig. 6.11(a).

The (l, m, n) subscripts usually used in waveguide resonators are modified to (l, m, δ), where the δ indicates that the rod is a bit larger than half of a period of the field variation. This occurs because in a dielectric rod, the modes are so called quasi TE and TM. When a dielectric puck is placed next to a microstrip line, the fields that "spill" out of the line are captured and stored in the resonator. The equivalent circuit for a puck placed close to a microstrip line is shown in Fig. 6.11(b). The YIG is a high Q ferrite sphere made of yttrium iron garnet, $Y_2Fe_2(FeO_4)_3$. It can be tuned over a wide range by varying a DC magnetic field, and making use of a magnetic resonance which ranges between 500 MHz and 50 GHz depending on the material and field used. YIG resonators have typically unloaded Q factors of 1000 or greater.

The varactor diode resonator can be thought of as the dual of a current-tuned YIG resonator. The varactor diode is just a Schottky or pn diode that has a capacitance that varies nonlinearly with bias



(a) (b)

Figure 6.11: The $TE_{01\delta}$ mode in a dielectric resonator, and the equivalent circuit for a dielectric resonator coupled to a microstrip line (b).

in the reverse bias mode. Silicon varactors have faster settling times, and GaAs varactors have larger Q values. The cutoff frequency for a varactor is defined for $Q = 1$. For a simple series RC equivalent circuit we have

$$Q = \frac{1}{\omega RC} \qquad \text{and} \qquad f = \frac{1}{2\pi RC}. \tag{6.9}$$

The frequency tuning range of the varactor is determined by the capacitance ratio C_{max}/C_{min} which can be as high as 12 for hyper-abrupt junction diodes. Since R is a function of bias, the maximum cutoff frequency occurs at a bias near breakdown, where both R and C have minimum values.

6.4 The Spectrum Analyzer

A spectrum analyzer is an instrument that shows the amplitude (power or voltage) of an input signal as a function of frequency. In that respect, its function is complementary to the oscilloscope; the latter has a horizontal axis calibrated in terms of time, whereas the former shows a display of spectral amplitude versus frequency. For example, if you hooked up a radio antenna to the input, you would see lines between 88 to 108 MHz, a spectral line for each station. The lines will be larger if the radio station is closer and the signal stronger. It will be narrow if the signal is clear and clean, and wide and jagged if it is noisy. A spectrum analyzer is often used to examine electronic equipment for electromagnetic emissions, or at microwave frequencies to characterize oscillators, modulation, harmonic distortion and interference effects.

A spectrum analyzer is a very sensitive receiver. Fig. 6.12 shows a simplified block diagram of a spectrum analyzer. Microwave spectrum analyzers cover frequencies from several hundred MHz to several tens of GHz. The frequency resolution is determined by the bandwidth of the IF, and is typically 100 Hz to 1 MHz. The sweeper has a variable LO that repetitively scans the receiver over the desired frequency band (horizontal axis). The input bandpass filter is tuned together with the local oscillator and acts as a preselector to reduce spurious intermodulation products. The IF amplifier provides wide dynamic range (this means that the instrument can measure both very small and very large signals). An excellent Agilent application note on Spectrum Analyzer basics is App. Note 150.

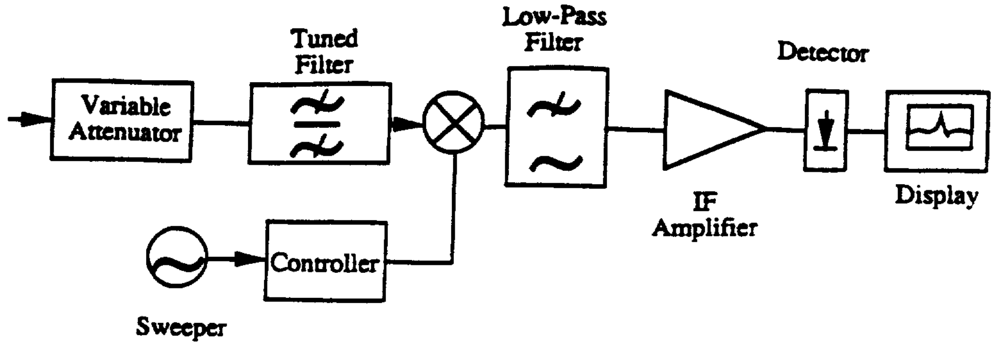


Figure 6.12: Block diagram of a spectrum analyzer.

6.5 Nonlinear Behavior of Microwave Circuits in the Time Domain

Electrical signals that are periodic in time can be represented as a sum of sinusoidal signals of different frequencies with appropriate amplitudes and phases. Let $f(t)$ be the periodic function with period T, such that

$$f(t) = f(t + nT), \quad n \text{ is an integer.} \tag{6.10}$$

If we look at the set of functions $\phi_k = e^{jk2\pi t/T} = e^{jk\omega t}$, where k is an integer, we notice that these functions are also periodic in T. By superposition of the ϕ_k's, we can represent the function f as:

$$f(t) = \sum_{-\infty}^{+\infty} a_k e^{jk\omega t}, \tag{6.11}$$

where the coefficients a_k need to be determined. This representation of a periodic function $f(t)$ is called a *Fourier series*. If you multiply both sides of the previous equation by $e^{-jk\omega t}$ and integrate over a period, you will find that the Fourier coefficients are given by

$$a_k = \frac{1}{T}\int_T f(t)e^{-jk\omega t}\mathrm{d}t. \tag{6.12}$$

An instrument called the Transition Analyzer can be used to look at the frequency representation of time periodic signals. It is tricky to make an oscilloscope at 20 GHz. You might wonder why this is so, when the spectrum analyzer gives us all the harmonics of a waveform, so we could plug them into a Fourier series to find the time waveform. But, there is a problem: the coefficients a_k are, in general, complex numbers; i. e., they have an amplitude and a phase. When you use the spectrum analyzer, you can get the amplitude of the different harmonics, but you have no idea what the phase is, and as a consequence, you cannot define the Fourier series. As an example, Fig. 6.13 shows two time waveforms that would have the same spectra shown on a spectrum analyzer. The Transition Analyzer is best described as a combination of a very fast sampling scope and a network analyzer. It can display and perform measurements in both time and frequency domain. The tricky part is how to measure the microwave phases accurately, and the instrument actually triggers of an IF signal and keeps track of its phase.

So far we have looked at both linear and nonlinear circuits in frequency domain. A linear circuit is a circuit where the output has the same frequency components as the input signal. For example, passive linear circuits are attenuators, directional couplers, loads, cables, connectors etc. Some active circuits are also linear (or approximately so), for example low-power amplifiers. We have also looked at nonlinear circuits, such as mixers and oscillators. In mixers, the frequency changes to a higher or

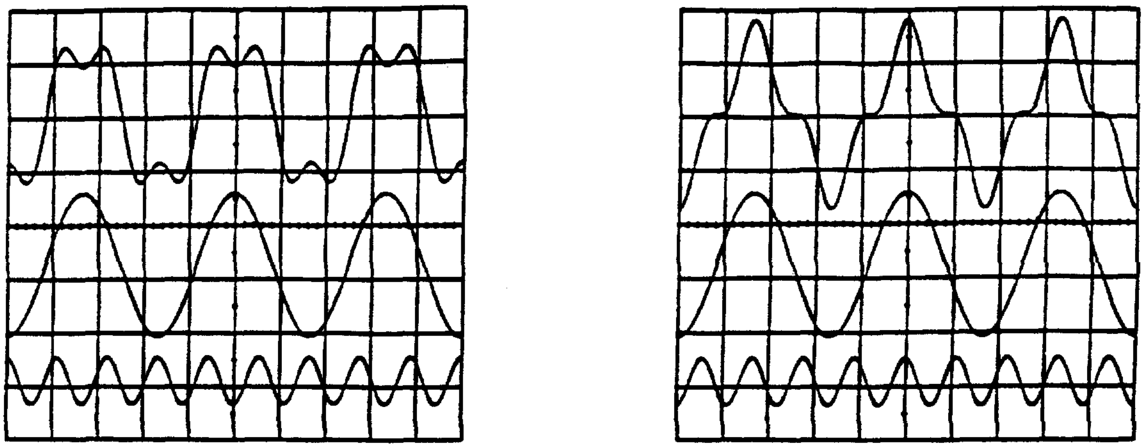


Figure 6.13: Two waveforms that give the same output on a spectrum analyzer.

lower frequency, and in oscillators, you have noticed harmonics of the fundamental oscillation frequency. Looking at linear circuits in time and frequency domain is not that interesting: you just see a sinusoidal time waveform and the associated delta function spectrum.

6.6 Practice questions

1. What are the different types of microwave sources and in what respects are they different?
2. Explain why an oscillator has a negative resistance for the example of a one-port oscillator.
3. What determines the amplitude of the RF voltage of a negative-conductance oscillator?
4. What would happen if there were more than one intersection of the device line and the load line of a negative-conductance oscillator?
5. If a negative conductance device had an RF conductance $G_D(V_1)$ such that the derivative $G'_D(V_1) < 0$ over a certain range of V_1 values, and the load was such that the oscillation point was in this range, would the oscillation be stable? Explain.
6. Sketch a microstrip resonator with the associated current and voltage waveforms. What is the impedance equal to at different points along the resonator?
7. How is the quality factor of a resonator defined?
8. How does a microwave spectrum analyzer work?
9. Can you reconstruct a time domain waveform from the output of a spectrum analyzer using Fourier analysis? Why?
10. How would you recognize an active from a passive circuit on a Smith chart plot that covers a certain frequency range?
11. How is an oscillator Smith chart plot different from an amplifier Smith chart plot?
12. What is injection-locking?

6.7 Homework Problems

1. Find the Fourier series for a train of pulses of unit amplitude and 50% duty cycle at 1 GHz, Fig. 6.14. Which are the three most important frequencies that describe the square wave?

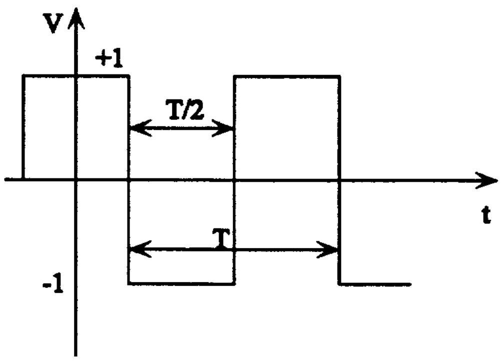


Figure 6.14: A square wave with unit amplitude. Find the Fourier series.

2. In this problem we will look at how an oscillatory circuit appears on a Smith chart. Let us look at the case of a series resonant circuit (R, L and C) fed from a generator (V and R_0) as shown in Fig. 6.15. You will examine the circuit behavior dependence on the value of the resistor R (it can

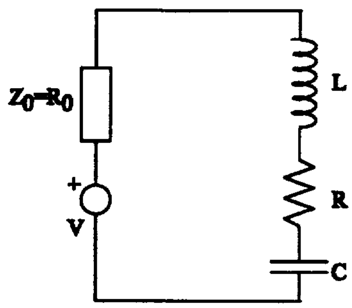


Figure 6.15: A series resonant circuit fed from a one-port generator of characteristic impedance R_0.

be positive or negative).

(a) What is the input impedance of the resonant circuit equal to? What is the resonant frequency ω_0 equal to?

(b) Derive the expression for the reflection coefficient ρ of the resonant circuit (the transmission line connecting the circuit to the generator V has a characteristic impedance R_0).

(c) Determine the value of the amplitude of ρ (whether it is larger or smaller than 1) and the phase of ρ at three frequencies around resonance, $\omega_1 = \omega_0 - \delta$, ω_0 and $\omega_2 = \omega_0 + \delta$, for the following three cases:

   (i) $R \geq 0$

   (ii) $-R_0 \leq R < 0$, and

   (iii) $-\infty < R < -R_0$.

(d) Sketch the reflection coefficient for all three cases on a Smith chart (the radius needs to be larger than 1). What is the direction of the Smith chart plot as the frequency increases in the three cases?

(e) Use Ansoft Designer (or Spice) to verify your analysis. Use $R_0 = 50\ \Omega$, a 1 nH inductor and a 1 pF capacitor, and perform the analysis for f between 4 and 6 GHz. Change the value of R as in your analysis in part (c) above.

3. An important practical problem that arises in transmission line circuits (such as microstrip) is how to supply the bias voltage to the negative resistance device. We need to provide a clear DC path for the bias voltage and current to the diode, but prevent significant RF signal from appearing along the bias path.

(a) A Gunn diode connected to a microstrip resonator is shown in Fig. 6.16. Use a substrate

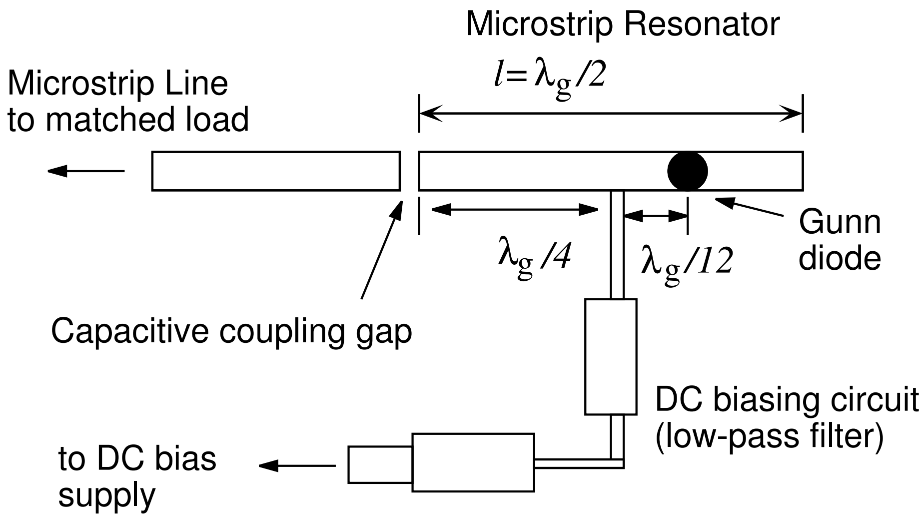


Figure 6.16: A microstrip Gunn diode oscillator.

permittivity of $\epsilon_r = 2.2$, a substrate thickness of $h = 1.27$ mm. What strip width is necessary for the microstrip line to have a characteristic impedance of $50\,\Omega$? What length l of such a microstrip will be half a guide wavelength at a fundamental oscillation frequency of $f = 5$ GHz?

(b) Model the capacitive coupling gap to a $50\,\Omega$ microstrip output line as a 0.1 pF capacitor as shown in Fig. 6.16, and assume the input microstrip output line is terminated in a matched load. Assume the Gunn diode will be connected $\lambda_g/12$ (again, at $f = 5$ GHz) from the center of the microstrip resonator as shown. Use SPICE to determine the Smith chart load line seen at the Gunn diode connection point, neglecting the effect of the bias network. Display the results for $f = 0$ to 10 GHz on a Smith chart plot.

(c) Alternating 80Ω and 25Ω quarter-wave (at the fundamental oscillation frequency) sections of line are used to connect the bias voltage to the RF circuit as shown. If two pairs (4 sections in all) of alternating line sections are used, and the DC bias supply is modeled at RF as a 0.1 nH inductor in series with a 0.5Ω resistor (connected from the end of the bias line to ground), use SPICE to determine the microwave impedance of this DC biasing circuit seen at the connection point to the microstrip resonator. Use the same frequency range as in part (b).

(d) Use Designer or SPICE to determine the Smith chart load line presented to the Gunn diode by the entire passive part of the circuit (including the biasing network). Use the same frequency range as in part (b). Compare this to the load line without the biasing network. How would you expect the biasing network to affect the behavior of the oscillator?

4. In this problem, you will use SPICE to model the buildup of oscillations in a Gunn diode oscillator. Use the circuit shown in Fig. 6.17. The Gunn diode is modeled by the SPICE element G1, a

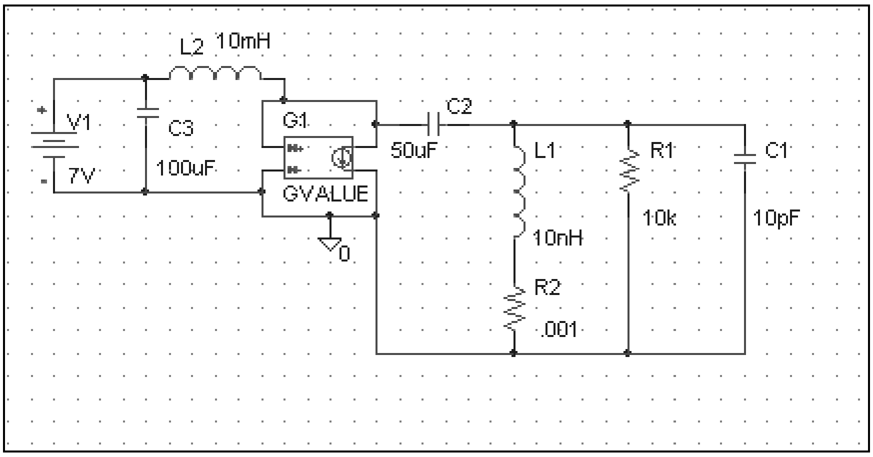


Figure 6.17: Circuit for modelling Gunn diode oscillator in SPICE.

voltage-controlled current source, in which the current source has the functional form:

$$I(V) = I_{thr}\left[\frac{V}{V_{thr}}e^{1-V/V_{thr}} + K\left(1 - e^{-V/V_{thr}}\right)\left(1 - \frac{V}{V_{thr}}e^{1-V/V_{thr}}\right)\right]$$

where $K = 0.8$, $V_{\mathrm{thr}} = 4\,\mathrm{V}$, and $I_{\mathrm{thr}} = 0.1\,\mathrm{A}$. This dependence is realized in PSPICE by using the VALUE= form of the G-element, or in the Schematics editor by using the part type GVALUE. A DC bias voltage of 7 V is provided through an RF filter as shown, and the Gunn diode is connected to a parallel RLC network through the large capacitor C2 (inserted to remove the DC voltage) to achieve oscillation.

(a) Do a *transient* analysis of this circuit covering the time interval $0 < t < 2\ \mu$s. Plot the current through the inductor L1 over this time interval. Explain what is happening.

(b) In PROBE, perform a Fourier transform of your data to display the frequency content of the waveform you obtained in part (a). At what frequency is the spectrum largest? What is the half-power bandwidth of this spectrum (i. e., find the frequencies at which the current is $1/\sqrt{2}$ of its maximum value, and take the difference)? Expand the scale of the plot to cover the frequency range from 400 to 600 MHz, and attach a copy of this plot with your solution.

(c) Change the bias voltage to 5 V. What is the effect of this on your answers to parts (a) and (b)?

Chapter 7

Antennas

7.1 Antenna Characteristics

Antennas are a part of microwave engineering where circuit theory and electromagnetic field theory meet. A transmission line circuit feeds the antenna, which radiates the power from the feed into a propagating electromagnetic wave. Most textbooks on antennas treat the electromagnetic field aspect in detail, from the radiation point of view. In this class it is more appropriate to treat antennas as elements of microwave circuits.

The vocabulary used in antenna engineering is the following:

Directivity (D) – This parameter tells us how much power the antenna radiates in a given direction, compared to the total radiated power. A high directivity would be roughly 30 dB and higher. The directivity is a function of angle, usually expressed in the spherical coordinate system. However, when it is quoted a single number without a specified angle, it refers to maximum directivity, usually at broadside (zero angle).

Efficiency (η) – An antenna is made usually of metals and dielectrics that have some loss. The efficiency is defined as the total power radiated by the antenna divided by the input power to the antenna.

Gain (G) – The gain is the product of the directivity and the antenna efficiency. This is the number you measure in the lab, and then you can back out the directivity if you know the efficiency (losses and mismatch). Again, if quoted in specification sheets as a single number, the gain refers to the maximal gain over all spatial angles.

Impedance (Z) – An antenna is fed from some circuit and represents a load to that circuit. It is important to match the impedance of the antenna to the feed so that most of the generator power will be radiated. The impedance of an antenna is usually a complex number and varies with frequency.

Polarization – The polarization of an antenna is defined with respect to the polarization of the electric field vector that is radiated. We use a unit vector $\vec{a}_p$ to indicate this direction. Usually antennas are linearly (vertical or horizontal) or circularly (left or right) polarized. For linear polarization, $\vec{a}_p$ can be taken to be a real unit vector. For circular polarization, the polarization vector is essentially complex, e. g., $\vec{a}_p = (\vec{a}_x + j\vec{a}_y)/\sqrt{2}$, where $\vec{a}_x$ and $\vec{a}_y$ are cartesian unit vectors.

For example, AM broadcasting antennas have almost linear vertical polarization, while FM antennas have linear horizontal polarization, and this determines how the receiving antennas are oriented. Some antennas are designed to radiate and receive two polarizations (used in atmospheric sensing). Circularly polarized antennas are used in satellite communication, since waves passing through the atmosphere undergo Faraday rotation due to the presence of the Earth's DC magnetic field.

Bandwidth – The bandwidth of an antenna tells us in what frequency range the antenna has its specified characteristics. Usually the impedance of the antenna changes with frequency, so the match to the circuit will change, resulting in degraded performance. Bandwidth can be quoted in many ways:

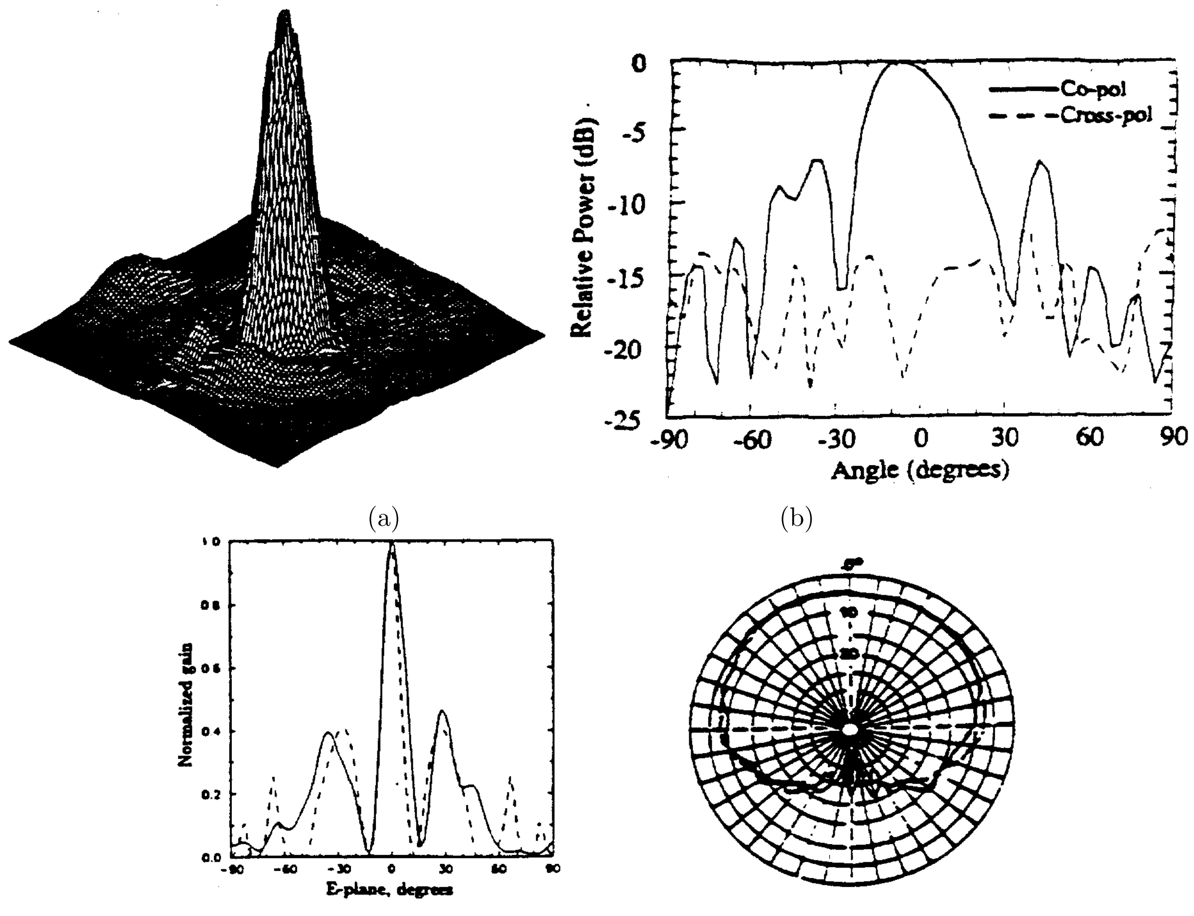


Figure 7.1: Plotting radiation patterns: a two-dimensional pattern (a), a normalized power pattern plotted in dB (b) or on a linear scale (c), and a polar pattern plot (d).

VSWR (referring to the input match), gain, beamwidth (referring to the frequency range over which the antenna main beam is preserved to a given degree), polarization (referring to the radiated wave polarization remaining as specified within the frequency range), sidelobe level, efficiency, etc.

Far-field – The electromagnetic fields of the antenna very close to it are referred to as the near-field, and they are usually very hard to know exactly. However, the most important region is far away from the antenna, and this is where the wave radiated by the antenna can be approximated by a plane wave. This region is called the far field. A commonly used rule of thumb for where the far field begins is a distance $2D^2/\lambda$, where D is the largest linear dimension of the antenna, and λ is the operating wavelength.

Radiation pattern – An antenna will radiate better in some directions, and worse in others. A graphical description of the radiation is the radiation pattern, which is usually the plot of radiated power (on a linear or dB scale) versus spatial angle in the far-field. Different ways of plotting a radiation pattern are shown in Fig. 7.1. We can plot the two-dimensional pattern, Fig. 7.1(a), or certain cuts of it. The most commonly used cuts for linearly polarized antennas are the *E-plane* and *H-plane* patterns. The E-plane is the plane determined by the radiated electric field vector and the line from the antenna to the observation point; the H-plane is perpendicular to it and determined by the radiated magnetic field vector. Very often the radiation pattern is plotted in a polar plot, like the one shown in Fig. 7.1(d). The radiation pattern is usually plotted in terms of the power contained in one polarization of the electric

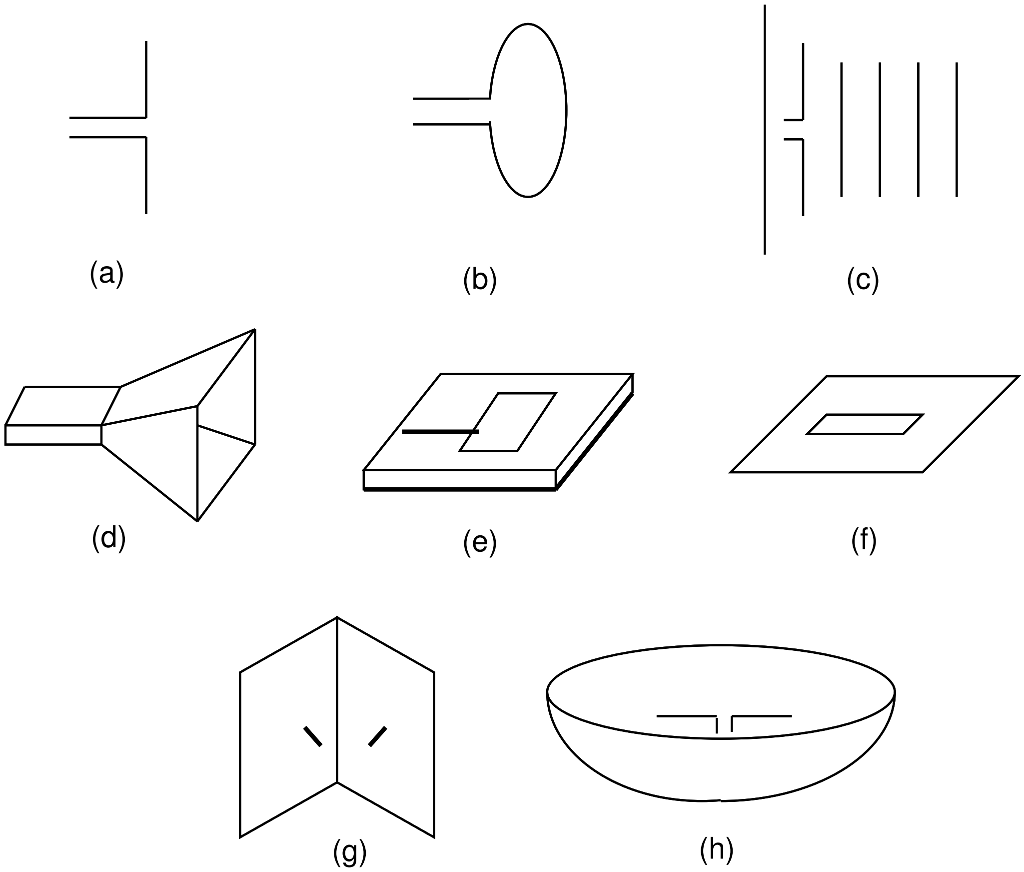


Figure 7.2: Most commonly used antenna types. A wire dipole antenna (a), a wire loop antenna (b), a Yagi-Uda array (c), a horn antenna (d), a microstrip patch antenna (e), a printed slot (f), a corner reflector (g) and a parabolic reflector antenna (h).

field. In actual antennas, the polarization is never completely pure in a single direction. The dominant (or desired) polarization is then referred to as *co-polarized*, while the polarization perpendicular to that is referred to as *cross-polarized*. Sometimes both of these patterns are plotted.

(Half-Power or 3 dB) Beamwidth – If we consider a plane cut of the radiation pattern which contains the direction of maximum radiation, the half-power beamwidth is the angle between the two directions in which the radiation power density has fallen to one-half of its maximum value. Sometimes other beamwidths are defined, for example, the 10-dB beamwidth which is the angle between directions at which the power level has fallen to 10 dB below the maximum value. For antennas which are not symmetric, beamwidths in different planes (e. g., the E-plane and the H-plane) may be different.

Feed point – An antenna is generally connected to a waveguide or transmission line in order to carry power to or from the antenna. The port at which this connection is made is called the feed point of the antenna. Here, the antenna presents the look of a circuit to the feed line, while to space it presents the look of a launcher or interceptor of wave fields.

There are many different kinds of antennas, a few most commonly used are shown in Fig. 7.2. The antennas shown in Fig. 7.2(a), (b) and (c) are so called wire antennas. They are used mostly at lower frequencies (up to UHF) and they have relatively low gains, but are cheap, lightweight and easy to make.

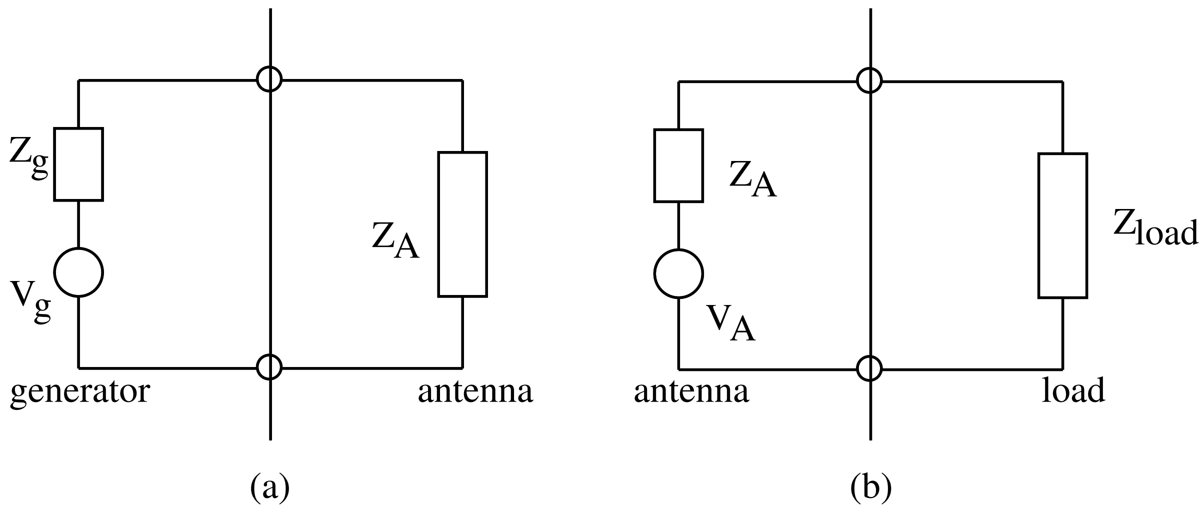


Figure 7.3: The equivalent circuit for (a) a transmitting and (b) a receiving antenna.

The antenna shown in Fig. 7.2(d) falls into the category of aperture antennas; it is very commonly used at microwave and millimeter-wave frequencies, and has a moderate gain. Fig. 7.2(e) and (f) show printed planar antennas, used at microwave and millimeter-wave frequencies. They are compatible with microwave printed circuit fabrication. They have low gain, but are cheap and lightweight and conformal to any surface. The antennas in Fig. 7.2(g) and (h) are of the reflector type, and are designed for very high gain, and are usually very large (measured in wavelengths).

7.2 Transmitting and Receiving Antennas

Antennas can be used either to send a wave (these are called *transmitting antennas*) or to "capture" one (these are referred to as *receiving antennas*). Fundamentally, if we know the properties of a certain antenna as a transmitter, we know its properties as a receiver as well. This follows directly from reciprocity, as we will see.

Transmitting Antennas

If we first look at a transmitting antenna as a circuit, the input is the power from a generator, and the output is the power radiated by the antenna, Fig. 7.3(a). From the generator's point of view, the antenna is a load with an impedance $Z_A = R_A + jX_A$. The antenna receives a power P_{tot} from the generator:

$$P_{tot} = R_A |I_A|^2. \tag{7.1}$$

where I_A is the current at the feed point of the antenna. In a real antenna, not all of this power is radiated due to losses to heat. For this reason, the resistance of the antenna can be split into two parts: the radiation resistance R_r that corresponds to the power "lost" to radiation, and R_{loss} that accounts for ohmic losses, so now we have:

$$R_A = R_r + R_{\text{loss}}, \qquad P_{tr} = R_r |I|^2, \qquad \text{and} \qquad P_{\text{loss}} = R_{\text{loss}} |I|^2. \tag{7.2}$$

The efficiency η of a transmitting antenna is defined as

$$\eta = \frac{P_{tr}}{P_{tot}} = \frac{R_r}{R_A} = \frac{R_r}{R_r + R_{\text{loss}}}. \tag{7.3}$$

The field radiated by a transmitting antenna in free space in the far-field region has the form of a *spherical wave*, described in spherical coordinates as:

$$\vec{E} \simeq \vec{\mathcal{E}}(\theta, \phi)\frac{e^{-jk_0 r}}{r} \qquad , \qquad \vec{H} \simeq \vec{\mathcal{H}}(\theta, \phi)\frac{e^{-jkr}}{r} \tag{7.4}$$

where $k_0 = \omega\sqrt{\mu_0\epsilon_0}$ is the propagation constant of a plane wave in free space, $\vec{\mathcal{E}}(\theta, \phi)$ is the directional pattern of the electric field in the far-field region, and $\vec{\mathcal{H}}(\theta, \phi)$ is that of the magnetic field. The wave propagates outward in the radial (r) direction, analogously to the propagation of a plane wave, but there is an additional amplitude decay proportional to $1/r$. The amplitude and polarization of the field may vary with the angular direction (θ, ϕ), but we always have:

$$\hat{r} \times \vec{\mathcal{E}}(\theta, \phi) = \zeta_0 \vec{\mathcal{H}}(\theta, \phi) \tag{7.5}$$

where $\zeta_0 = \sqrt{\mu_0/\epsilon_0} \simeq 377\,\Omega$ is the wave impedance of free space, and $\hat{r}$ is the unit vector pointing in the outward radial direction. In other words, the electric and magnetic field vectors are always perpendicular to each other and to the direction of propagation. In this sense, the spherical wave behaves locally like a plane wave, so that everything we know about plane waves can be taken to be approximately true of the far field of a transmitting antenna.

The power radiated by the antenna is thus distributed more in some directions and less in others. The time-average power density radiated by an antenna at a distance r is found from the radial component of the Poynting vector:

$$S(r, \theta, \phi) = \text{Re}\left(\vec{E} \times \vec{H}^*\right) \cdot \hat{r} = \frac{1}{r^2}\text{Re}\left(\vec{\mathcal{E}} \times \vec{\mathcal{H}}^*\right) \cdot \hat{r} \quad \text{W/m}^2 \tag{7.6}$$

remembering that the fields are taken to be RMS quantities. The total radiated power from a transmitting antenna can be expressed as:

$$P_{tr} = \int_0^{2\pi} d\phi \int_0^{\pi} r^2 \sin\theta S(\theta, \phi)\, d\theta \tag{7.7}$$

(although, when other definitions of the θ coordinate are used, the limits of integration over that variable may be different, as will the factor $\sin\theta$ in the integrand).

Defining the *radiation intensity* $U(\theta, \phi)$ as

$$U(\theta, \phi) \equiv \text{Re}\left(\vec{\mathcal{E}} \times \vec{\mathcal{H}}^*\right) \cdot \hat{r} = \frac{1}{\zeta_0}\vec{\mathcal{E}} \cdot \vec{\mathcal{E}}^* \tag{7.8}$$

we express (7.6) as

$$S(r, \theta, \phi) = \frac{U(\theta, \phi)}{r^2} \tag{7.9}$$

The radiation intensity is the power radiated per *solid angle*, where we define the solid angle Ω through its differential element

$$\text{d}\Omega = \frac{\text{d}S}{r^2} = \sin\theta\, d\theta d\phi \tag{7.10}$$

with respect to a reference point (the origin $r = 0$), so that from (7.7) and (7.9) we have

$$U(\theta, \phi) = \frac{dP_{tr}}{d\Omega} \tag{7.11}$$

An antenna that radiates power equally in all directions is called an *isotropic antenna*. Such an antenna is a useful ideal which we use as a comparison for other antennas. If a total power P_{tr} were radiated by an isotropic antenna, the power density would be independent of θ and ϕ:

$$S_i = \frac{P_{tr}}{4\pi r^2}, \tag{7.12}$$

where r is the distance of an observation point from the feed point at the antenna center. The radiation intensity $U(\theta, \phi) = U_i$ of an isotropic antenna is likewise a constant:

$$U_i = \frac{P_{tr}}{4\pi} \tag{7.13}$$

An antenna that is not isotropic will radiate a different power density in different directions, and of course any real antenna will lose some power P_{loss} to heat. If its actual radiated power density is $S(r, \theta, \phi)$, we define the *directivity*, $D(\theta, \phi)$, in spherical coordinates as

$$D(\theta, \phi) = \frac{S(r, \theta, \phi)}{S_i} = \frac{U(\theta, \phi)}{U_i}, \tag{7.14}$$

The *gain* $G(\theta, \phi)$ of an antenna is defined similarly, but using instead the isotropic radiation intensity that would result if the *total* power input to the antenna were all radiated:

$$S_{i,tot} = \frac{P_{tot}}{4\pi r^2}, \tag{7.15}$$

The gain is thus the directivity of the antenna multiplied (reduced) by its efficiency:

$$G(\theta, \phi) = \frac{S(r, \theta, \phi)}{S_{i,tot}} = \frac{S(r, \theta, \phi)}{S_i} \frac{P_{tr}}{P_{tot}} = \eta D(\theta, \phi) \tag{7.16}$$

so that dielectric and conductor losses are taken into account in the gain, but not in the directivity.

An approximate formula for obtaining the directivity of narrow-beam antennas (such as horns) from the maximum directivity is:

$$D_{\max} \simeq \frac{32,000}{\theta_E \theta_H}, \tag{7.17}$$

where θ_E and θ_H are the 3-dB beamwidths in the E and H-planes in degrees, respectively.

Example—The electric current element

Besides the isotropic radiator, there is another idealized antenna which has conceptual importance: the electric current element. This antenna is a very short wire of length $\Delta l \ll \lambda$ carrying a uniform current I_A, for which we define a vector called the (RMS) dipole moment $\vec{p}$ whose length is $p = I_A \Delta l$ and whose direction is that of the current in the wire. We can use superposition to study the fields of more complicated antennas by adding the fields of collections of electric current elements. If the dipole moment is oriented along the z-axis, then the (RMS) far field ($r \gg \lambda$) in free space of the electric current element is given in spherical coordinates by

$$E_\theta = \frac{j\zeta_0}{2\lambda} p \sin\theta \frac{e^{-jk_0 r}}{r}, \tag{7.18}$$

$$H_\phi = \frac{E_\theta}{\zeta_0} \tag{7.19}$$

As in usual practice, we have defined the angle θ so that it ranges from 0 to π, and the equatorial plane is at $\theta = \pi/2$. In antenna engineering, it is often standard practice to define θ so that the maximum of an antenna's pattern is located at $\theta = 0$. In that case, sin would be replaced by cos in (7.18).

For the electric current element, the radiated power density is

$$S(\theta,\phi) = E_\theta H_\phi^* = \frac{\zeta_0}{4}\frac{|p|^2}{\lambda^2}\frac{\sin^2\theta}{r^2} \tag{7.20}$$

By (7.7), the total radiated power from this element is

$$P_{tr} = \frac{2\pi\zeta_0}{3}\frac{|p|^2}{\lambda^2}$$

and thus from (8.2), the radiation resistance R_r (which is the same as R_A for this lossless antenna) of the element is

$$R_r = \frac{2\pi\zeta_0}{3}\frac{(\Delta l)^2}{\lambda^2} \tag{7.21}$$

and so finally from (8.3) and (8.4) the gain of the electric current element is

$$G(\theta,\phi) = \frac{3}{2}\sin^2\theta \tag{7.22}$$

and the gain in the direction of maximum radiation is 1.5, or 1.76 dB.

Effective length

Other kinds of antennas with more complicated current distributions will have more complicated field patterns and gain functions due to interference between the fields from current elements on different parts of the antenna. For a particular direction of radiation, we often use the concept of an *effective length* to express the transmitting behavior of an arbitrary antenna, although the concept is most practical for thin wire antennas. Suppose the electric field (in the far field) of a transmitting antenna is $\vec{E}(\theta_{tr},\phi_{tr})$ in the direction (θ_{tr},ϕ_{tr}). Now replace this antenna with an electric current element whose current is the same as that at the feed point of the original antenna, and whose dipole moment is parallel to the electric field of the original antenna (and thus is perpendicular to the line between the feed point and the far-field observation point $(r,\theta_{tr},\phi_{tr})$). The effective length of the original antenna is a vector $\vec{l}_{\text{eff}}$ in the direction of this dipole moment whose length $l_{\text{eff}} = \Delta l$ is what the electric current element would need to have to produce the same (far-field) electric field in the direction (θ_{tr},ϕ_{tr}) as the original antenna. That is,

$$\vec{l}_{\text{eff}}(\theta_{tr},\phi_{tr}) = \frac{2\lambda r\vec{E}(\theta_{tr},\phi_{tr})e^{jk_0 r}}{jI_A\zeta_0} \tag{7.23}$$

Note that $\vec{l}_{\text{eff}}$ may be a complex number in the general case. It can be easily related to the radiated power density by using (7.20):

$$S(\theta_{tr},\phi_{tr}) = \frac{\zeta_0|I_A|^2|\vec{l}_{\text{eff}}|^2}{4\lambda^2 r^2} \tag{7.24}$$

Example – An Electrically Short Dipole

An electrically short dipole l long is shown in Fig. 7.4. By electrically short, we mean that $l \ll \lambda$. The current $I(z)$ on such an antenna varies from a maximum of I_A at the feed point, linearly to zero at the ends. The far-field of this antenna is obtained by the superposition (integration) of the far-fields of the incremental electric current elements $I(z)\,dz$ distributed along the dipole. Because the dipole is electrically short, the distances r and angles (θ_{tr},ϕ_{tr}) of all these incremental current elements to the observation point are essentially the same, so from the definition of effective length above, we find easily that

$$l_{\text{eff}} = \frac{\sin\theta_{tr}}{I_A}\int_{-l/2}^{l/2} I(z)\,dz = \frac{l\sin\theta_{tr}}{2} \tag{7.25}$$

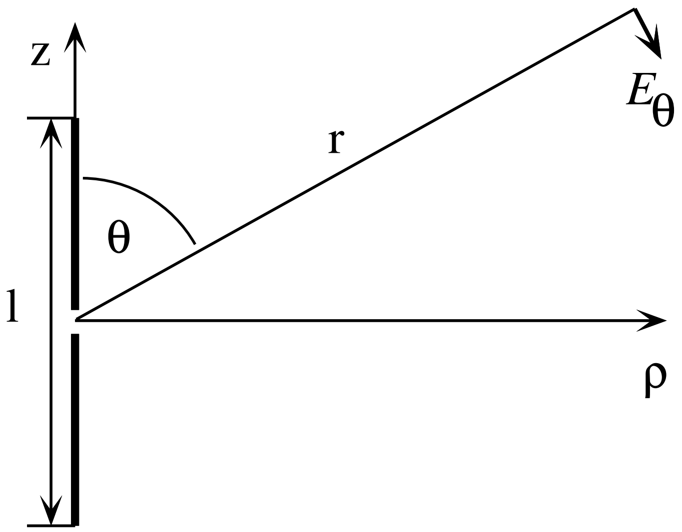


Figure 7.4: An electrically short dipole.

and that the direction of $\vec{l}_{\text{eff}}$ is parallel to the unit vector $\hat{\theta}_{tr}$, calculated at the angle θ_{tr}. We may thus write

$$\vec{l}_{\text{eff}} = \frac{l \sin\theta_{tr}}{2} \hat{\theta}_{tr} \tag{7.26}$$

Receiving Antennas

If we look at a receiving antenna as a circuit, the input is the incident power on the antenna, and the output is the power delivered by the antenna to the load. From the point of view of the load, the antenna is some generator. We can find the equivalent voltage source and impedance of this generator from Thévenin's theorem. The voltage source is just the open circuit voltage, and the impedance is found by turning off the voltage source and measuring the impedance from the terminals. This is exactly how we found the impedance of the transmitting antenna, which tells us that the impedance of the antenna is the same whether it is receiving or transmitting. The equivalent circuit for a receiving antenna is shown in Fig. 7.3(b).

The response of a receiving antenna to incident radiation is measured by the same effective length defined for the transmitting antenna (used primarily in the case of "thin" antennas) or by the *effective area* in more general situations. We can show (see section 7.4) that the Thévenin or open-circuit voltage at the feed point of the receiving antenna is

$$V_A = \vec{l}_{\text{eff}} \cdot \vec{E}_{inc}. \tag{7.27}$$

where $\vec{E}_{inc}$ is the incident electric field at the feed point of the antenna (where V is measured).

The effective area $A(\theta, \phi)$ is defined in terms of the power P_r available from the receiving antenna terminals and the power density incident on the antenna in the "correct" polarization from the direction (θ, ϕ):

$$A(\theta, \phi) = \frac{P_r}{S(\theta, \phi)}. \tag{7.28}$$

The effective length and effective area can be related in the following way. If the antenna is impedance-matched, and the polarization of the incident electric field is aligned with the direction of $\vec{l}_{\text{eff}}$, then the

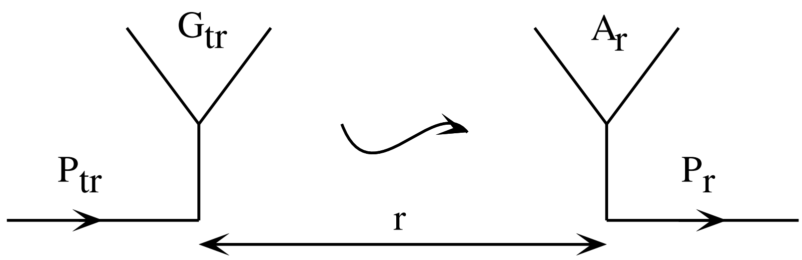


Figure 7.5: A communication system.

received power P_r is maximized, and its value is called the available power. We can write this power in terms of the effective length and incident electric field magnitudes:

$$P_r = \frac{V^2}{4R_A} = \frac{(l_{eff}E)^2}{4R_A}. \tag{7.29}$$

From the definition of the effective area, we now find

$$\frac{A(\theta,\phi)}{\zeta_0} = \frac{|l_{eff}(\theta,\phi)|^2}{4R_A}, \tag{7.30}$$

since the power density magnitude is just $S_{inc} = E_{inc}H_{inc}$ (from Poynting's theorem) and $E_{inc}/H_{inc} = \zeta_0$.

7.3 The Friis Transmission Formula

A communication system, Fig. 7.5, consists of two antennas: a transmitter and receiver that are at a distance r away (the two antennas often change roles in such a link). The transmitting antenna has a gain G, and the receiving antenna has an effective area A. The power density S incident at the receiving antenna is given by

$$S(\theta_{tr},\phi_{tr}) = \frac{ERP}{4\pi r^2}, \tag{7.31}$$

where $ERP = P_{tr}G_{tr}$ is the so called *effective radiated power* defined as the product of the gain and the transmitted power, and G_{tr} is implicitly a function of the direction (θ_{tr},ϕ_{tr}) along which the wave travels from the transmitting antenna to reach the receiving antenna. Physically, (7.31) is the power of an isotropic source that would give the same power density at the receiving antenna. From (7.28), the received power is then given by the *Friis transmission formula*:

$$P_r = \frac{P_{tr}G_{tr}A_r}{4\pi r^2}. \tag{7.32}$$

where A_r is implicitly a function of the direction (θ_r,ϕ_r) from which the wave incident from the transmitting antenna arrives at the receiving antenna.

The Friis formula (7.32) is valid if the polarizations of the transmitting and receiving antennas are the same. If they are not, we must additionally multiply (7.32) by the *polarization loss factor* PLF, which for linearly polarized antennas is equal to $\cos^2\theta_p$, where θ_p is the angle between the polarization directions of the transmitting and receiving antennas.

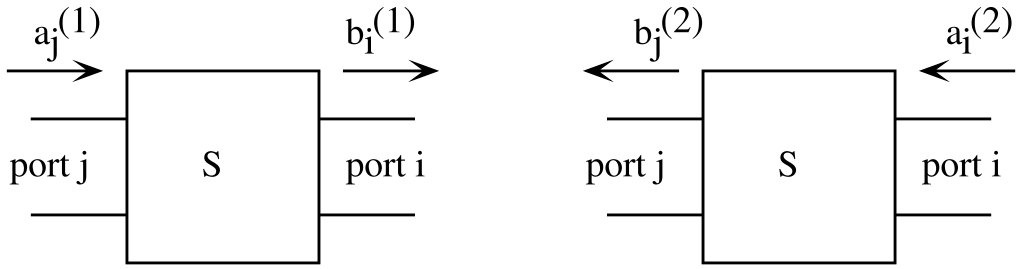


Figure 7.6: Reciprocity applied to a network.

7.4 Reciprocity and Antennas

Consider a network terminated in matched loads and a single incident wave $a_j^{(1)}$ at port j, Fig. 7.6. The output at port i is a scattered wave $b_i^{(1)}$ given by

$$b_i^{(1)} = s_{ij} a_j^{(1)}. \tag{7.33}$$

(we use lower-case letters to denote the scattering parameters in this chapter in order to distinguish them from the power density $S(\theta, \phi)$). Now consider a second situation, with a network identical to the first one, but now with an incident wave $a_i^{(2)}$ at port i, so that the scattered wave at port j is

$$b_j^{(2)} = s_{ji} a_i^{(2)}. \tag{7.34}$$

If the system is *reciprocal*, the scattering matrix is symmetric, that is $s_{ij} = s_{ji}$, so that

$$a_j^{(1)} b_j^{(2)} = a_i^{(2)} b_i^{(1)}. \tag{7.35}$$

Physically, this means that there is a simple way to predict what happens in a network if we interchange the input and output. This is a very powerful idea, and looks simple, but it implies many things about the network: that it is made of linear materials that have symmetric conductivity, permeability, and permittivity tensors.

It is easy to show from the relationship between the wave amplitudes and the port voltages and currents that reciprocity also implies the formula

$$V_j^{(1)} I_j^{(2)} = V_i^{(2)} I_i^{(1)} \tag{7.36}$$

under the condition that all ports except i (for excitation 2) or j (for excitation 1) are short-circuited (that is, their port voltages are zero). This formula can be used to deduce the equality of the effective length for a transmitting antenna to the effective length for the same antenna used in the receiving mode. Although the proof we give here holds only for electrically small antennas, it can be generalized to hold for arbitrary antennas, provided the phase shift changes along the antenna are accounted for (due to variations in the distance to the observation point for the transmitting case, and variations in the incident field phase in the receiving case).

In the transmitting case, a voltage V applied to the feed point of the antenna (which we will designate as port 1), results in a current $I(z)$ at a point z on the antenna (which we will designate as port 2, if the antenna is broken at that point to form two terminals a distance dz apart; in the present case this port is short-circuited). In the receiving case, let the feed point be short-circuited, but the small interval dz

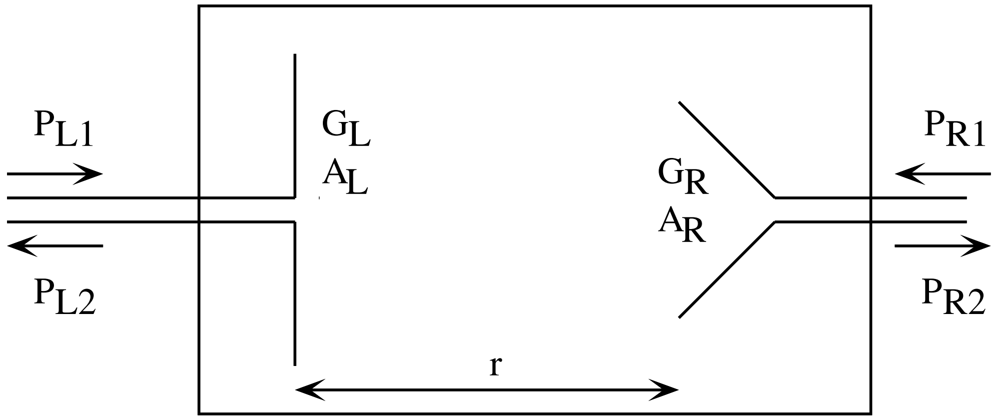


Figure 7.7: Reciprocity applied to a system consisting of two antennas.

at port 2 is excited by an incremental voltage source $E_{z,inc}(z)dz$ from the incident wave, and produces the short-circuit current $dI_{sc}(z)$ at port 1. By (7.36),

$$V dI_{sc}(z) = E_{z,inc}(z) dz I(z) \tag{7.37}$$

so that the *total* short-circuit current at the feed-point due to the incident wave is

$$I_{sc} = \int dI_{sc}(z) = \frac{1}{V} \int E_{z,inc}(z) I(z)\, dz \tag{7.38}$$

The short-circuit current must be the Thévenin equivalent voltage V_A for the receiving antenna divided by the antenna impedance. Moreover, if the antenna is electrically small, there is almost no variation of the incident field along the antenna. Thus,

$$V_A = \frac{Z_A}{V} E_{z,inc} \int I(z)\, dz = E_{z,inc} l_{\text{eff}} \tag{7.39}$$

since $V/Z_A = I_A$ is the feed-point current in the transmitting case, and $E_{z,inc}$ is the component of the incident electric field parallel to the vector effective length. This proves (7.27).

Next let us apply reciprocity to a system of two antennas, as shown in Fig. 7.7. Applying reciprocity, or interchanging the input and output, in this case implies using the receiving antenna as the transmitting antenna. Since our transmitting antenna was characterized by gain, and the receiving one by effective area, we expect to get a relationship between these two quantities when we use reciprocity. Assume that the two antennas in Figure 7.7 are different and matched to the lines feeding them, so that there is no mismatch loss (this is not a necessary assumption, because the losses can be taken into account separately anyway, since they do not affect the gain). First let us use the antenna on the left (subscript L) as the transmitter, and the right one as the receiver, and then the other way round. The reciprocity condition (7.35) in terms of transmitted and receiving powers now becomes

$$P_{L1} P_{L2} = P_{R1} P_{R2}. \tag{7.40}$$

The Friis transmission formula (7.32) allows us to rewrite the received powers P_{R1} and P_{L2} in terms of the gains and effective areas:

$$P_{L1} \left(P_{R2} \frac{G_R A_L}{4\pi r^2} \right) = P_{L1} \left(\frac{G_L A_R}{4\pi r^2} P_{R2} \right), \tag{7.41}$$

which after eliminating the common terms gives us the following:

$$\frac{G_L}{A_L} = \frac{G_R}{A_R}. \tag{7.42}$$

This means that the gain and effective area of an antenna are related by a universal constant (units are m^{-2}) which does not depend on the type of antenna or direction. Its value can be determined by computing it for any specific antenna. For the electric current element (which has no heat losses and thus has $R_A = R_r$), for example, we have from (7.22) that $G = 3\cos^2\theta/2$, while from (7.21) and (7.30) we have $A = 3\lambda^2\cos^2\theta/8\pi$. Thus we have for this (or any other) antenna that

$$\frac{G}{A} = \frac{4\pi}{\lambda^2} \tag{7.43}$$

Eqn. (7.43) allows us to rewrite the Friis formula (7.32) in terms of antenna gains only:

$$P_r = P_{tr}G_{tr}G_r\left(\frac{\lambda}{4\pi r}\right)^2 \tag{7.44}$$

or in terms of effective areas only:

$$P_r = P_{tr}\frac{A_{tr}A_r}{\lambda^2 r^2} \tag{7.45}$$

7.5 The Gain and Effective Area Integrals

In the case of a lossless transmitting antenna, the radiated power and the power received from the generator are the same, so we can write the gain as

$$G(\theta,\phi) = \frac{4\pi r^2 S(r,\theta,\phi)}{P_{tr}}. \tag{7.46}$$

where $S(\theta,\phi)$ is the power density in $\mathrm{W/m^2}$ transmitted in the direction (θ,ϕ).

When (7.46) is combined with (7.11) and integrated over all angles, the resulting integral looks like

$$\oint G\mathrm{d}\Omega = \frac{4\pi}{P_{tr}}\oint\frac{\mathrm{d}P_{tr}}{\mathrm{d}\Omega}\mathrm{d}\Omega, \tag{7.47}$$

which in turn gives us the value of the gain integral:

$$\oint G\mathrm{d}\Omega = 4\pi. \tag{7.48}$$

The *antenna theorem* says that the effective area of an antenna at a certain frequency is related to the square of the wavelength at that frequency as

$$\oint A\mathrm{d}\Omega = \lambda^2. \tag{7.49}$$

One proof of (7.49) depends only on properties which follow from Maxwell's field equations and is quite simple, making use of (7.43) and (7.48). This theorem can also be derived in a very elegant and simple way from principles of thermodynamics, but this is beyond the scope of this course.

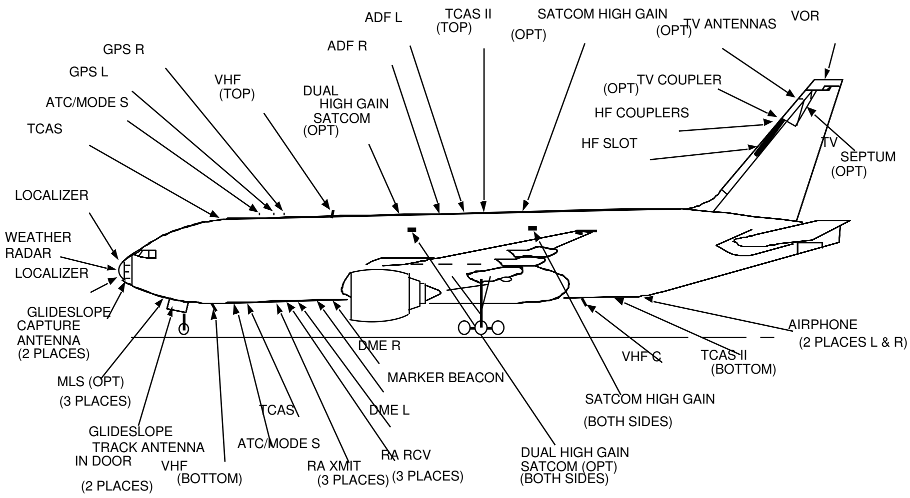


Figure 7.8: The 54 different antennas on a jet airliner used for navigation, communication and weather prediction.

7.6 Some Application Examples

If you start paying attention, antennas are all around us. Air traffic would be impossible without radio-wave communication, navigation and weather prediction. It is interesting to look at the number of different antennas used on an aircraft. As an example, a typical jet airliner has around 50-60 antennas, as shown in Fig. 7.8. The antennas cover a broad frequency range, from HF to high-microwave frequency radar and satellite communication antennas.

A smaller single-engine aircraft has also a surprising number of antennas:

- - Two ADF antennas for automatic direction finding, where one of the antennas is a directional long wire from the tail to the cockpit, and the other is an omnidirectional small antenna. This system operates at frequencies that coincide with AM radio.

- - The transponder antenna serves for altitude determination. Air traffic control allocates a channel to each aircraft, they send a pulse and the transponder retransmits the pulse, so that traffic control can determine the altitude.

- - Two UHF antennas in the several 100-MHz range serve for both voice radio and navigation. The latter is referred to as VOR (Very high frequency Omni-directional Range determination). This radio serves to provide radial information by measuring the phase of signals sent from several ground transmitters. Radial information is referenced to one of the transmitters and is plotted

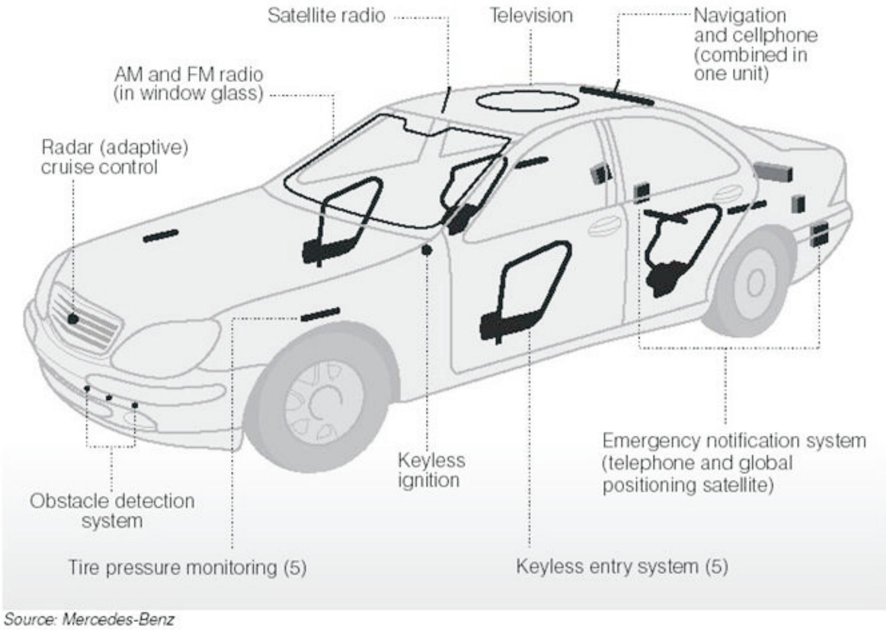


Figure 7.9: The 14 different antennas on a Mercedes Class S automobile (2005).

similarly to a compass. These UHF antennas are relatively short wires positioned at 45 degrees w.r.t. the vertical.

- - The DME stands for Distance Measuring Equipment and determines range. It is at UHF and is keyed to the navigational frequency.
- - Several GPS antennas in the commercial band (around 1 GHz). These are replacing the older low-frequency LORAN positioning system, which transmits carefully timed signals. Many aircraft have both positioning systems.
- - Satellite XM radio antennas at 2.4 GHz.
- - The antenna for the emergency locating transmitter (ELT) which is turned on automatically when the aircraft crashes. Currently this system operates at 121.5 MHz, and there is a new frequency at 400 MHz (which requires a new antenna).
- - Weather radar antennas for precipitation sensing operate in X band and often use horn antennas.
- - A separate air traffic antenna is used to receive transponder signals from other airplanes, and tells the pilot - what other airplane traffic is around.
- - A very broadband lightening detector antenna measures ratios of power levels at different frequencies to detect nearby lightening.

Many cars are equipped with a variety of antennas for communications, localization, emergencies, and even radar for collision avoidance (76/77 GHz in Europe and the US and 60 GHz in Japan) and parking, the latter usually used by trucks at X-band. Fig. 7.9 shows a sketch of approximate location and shape of the 14 antennas on a Mercedes Class S car.

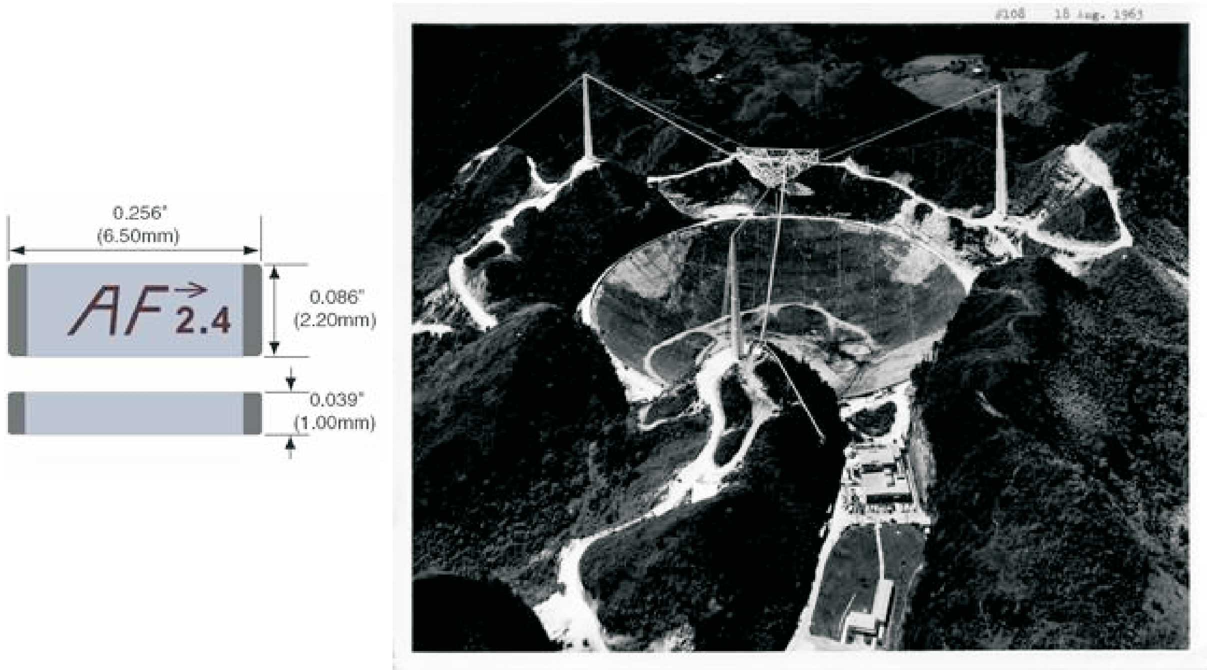


Figure 7.10: Currently marketed (2005) surface-mount “chip” antennas with dimensions indicated (left) and the largest single antenna reflector in the world during its construction in 1963 in Arecibo, Puerto Rico.

Antennas vary drastically in size, both physical and electrical. For example, a number of manufacturers currently offer “chip” antennas, Fig. 7.10. On the other hand, the largest single antenna is 305 m in diameter and fills an entire valley on the island of Puerto Rico. This antenna is used most of the time as a receiver for radioastronomy, although it can also be used as a planetary radar. Arrays of antennas that behave like one antenna can cover many square kilometers in area.

7.7 Practice questions

1. What are the efficiency, gain, directivity, polarization and radiation pattern of an antenna?
2. Where is the far field for a satellite TV dish antenna 1.5 m in radius at 4 GHz (C band)?
3. Sketch the equivalent circuit for a transmitting and receiving antenna. How is the antenna impedance defined in each case?
4. What is the definition and physical meaning of the effective length and effective area for an antenna?
5. What do the gain theorem and antenna theorem for antennas say?
6. Write down the Friis transmission formula. Can it be written down in terms of effective areas only, and also in terms of gains only (the two gains and effective areas are those of the receive and transmit antennas)?

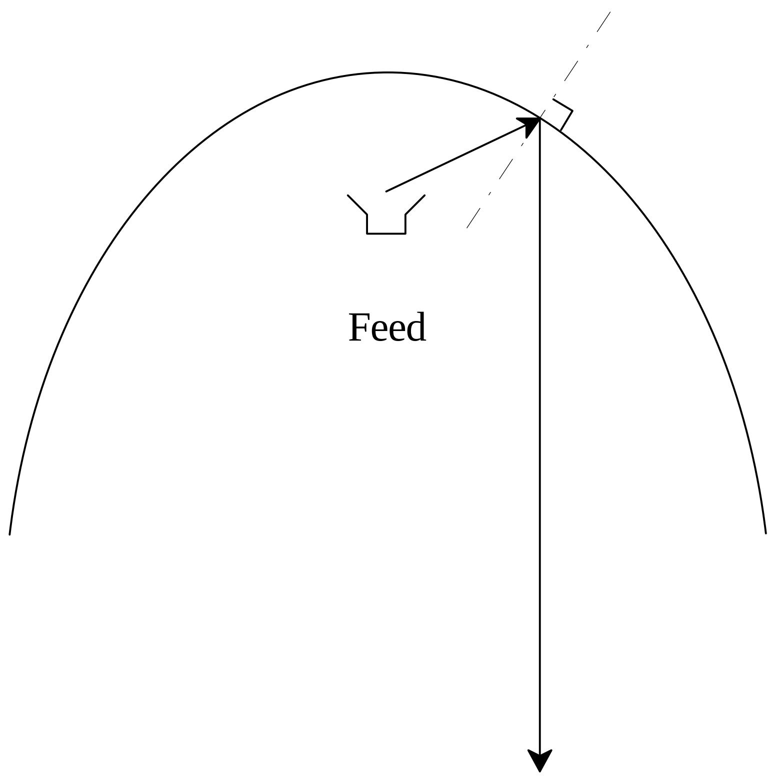


Figure 7.11: Incident and reflected rays at parabolic cylinder reflector.

7.8 Homework Problems

1. Find formulas for the radiation resistance and effective area of the short dipole shown in Fig. 7.4.

2. In a microwave relay system for TV each antenna is a reflector with an effective area of $1\,\mathrm{m}^2$, independent of frequency. The antennas are placed on identical towers 100 km apart. If the required received power is $P_r = 1$ nW, what is the minimum transmitted power P_t required for transmission at 1 GHz, 3 GHz and 10 GHz?

3. A UHF radio system for communication between airplanes uses quarter-wave monopole antennas that have a gain of 2. If the required received power is $P_r = 1\,\mathrm{pW}$, what is the minimum transmitted power P_t required for successful transmission at 100 MHz, 300 MHz, and 1 GHz?

4. A feed antenna is placed at the focus of a parabolic cylinder reflector whose equation is $y = x^2/4p$, where p is the distance from the vertex (back) of the parabola to the focus as shown in Fig. 7.11. Show that *any* ray emanating from the feed antenna will be reflected into the same direction (parallel to the axis of the parabolic cylinder) if we assume that rays are reflected according to the geometrical optics law of reflection. This means that the angle made between the incident ray and the normal to the reflector surface is the same as the angle between the reflected ray and the reflector surface.

5. An antenna is placed at some distance in front of a flat conducting surface. Draw lines emanating from the antenna in several directions towards the surface which represent radiating waves coming out of the antenna at various angles. If each of these lines (which are termed *rays*) hits the surface and produces a reflected ray according to the same law of reflection as a plane wave, draw the directions of the reflected rays that result. Under what conditions might this flat reflector focus more radiation into a narrower beamwidth, and therefore increase the gain by comparison to the antenna with no reflector?

6. Now place the antenna inside a conducting surface bent at an angle of 90° as shown in Fig. 7.12 (a so-called *corner reflector*). Repeat the exercise of Problem 5. Which reflector do you think will

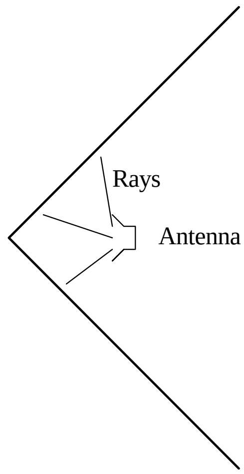


Figure 7.12: Antenna radiating towards a 90° corner reflector.

provide a larger increase (or at least a smaller decrease) in the gain of the original antenna?

7. The "chip" antenna in Fig. 7.10 is marketed to operate at (copied from manufacturer's specification sheets in 2005): (1) a center frequency of 2.45 GHz with a bandwidth of around 180 MHz, (2) an omnidirectional pattern, (3) linear polarization, (4) a VSWR less than 2, (5) a 50-ohm input impedance, (6) a gain greater than 0.5 dBi (gain with respect to an isotropic radiator), and (7) a maximum power handling capability of 3 W. What is the largest gain of this antenna according to the specifications, assuming the feed will not radiate and the antenna is 100% efficient? What would you do to make an antenna if someone handed you this chip? For more information you can look up the following web page www.antennafactor.com.

8. Calculate the directivity of the Arecibo dish assuming the dish is perfectly illuminated by the feed antenna. The feeds cover frequencies from around 300 MHz to 10 GHz (it actually goes down to 47 MHz, but rarely used that low). Arecibo is used mostly for receiving very low noise powers from distant objects, but also sometimes as a planetary radar. At 1 GHz, how much power would get to Jupiter if 1 GW was transmitted from Arecibo when the beam points exactly in the Jupiter direction?

9. Calculate the minimal far-field distance of a (1) 3-meter diameter satellite TV dish at 12 GHz and (2) a 25-meter diameter dish at a frequency of 44 GHz (military satellite frequency).

Chapter 8

Microwave Transistor Amplifiers

8.1 Microwave Three-Terminal Devices – The MESFET

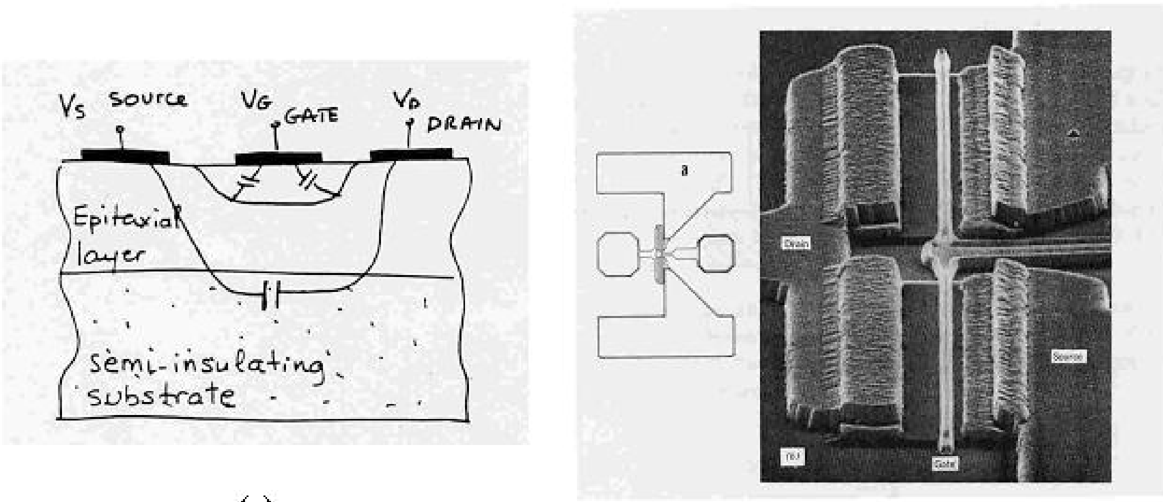
The most commonly used active device at microwave frequencies is the Metal Semiconductor Field Effect Transistor (MESFET). The MESFET is a GaAs device with a physical cross section shown in Fig. 8.1a, and a typical electrode layout is shown in Fig. 8.1b. The MESFET is a unipolar device, which means that there is only one type of carrier. The MESFET has three terminals: the source, gate and drain. The gate length is usually a fraction of a micrometer, and the width can be hundreds of micrometers. The three electrodes are deposited on an n-type GaAs epitaxial layer which is grown on a semi-insulating substrate. The epitaxial layer is on the order of $0.1\,\mu$m thick and the doping is $10^{16} - 10^{17}\,\mathrm{cm}^{-3}$. The source and drain are ohmic contacts (low resistance, usually made of a gold-germanium alloy), and the gate is a Schottky contact. Associated with the Schottky barrier is a depletion region which affects the thickness of the conducting channel. The gate is biased negatively with respect to the source, and the drain positively. When the voltage is changed on the gate, the thickness of the channel changes, and this controls the current.

When you purchase a MESFET, it can come in a package or in chip form. You will be given measured s-parameters at a few different bias points for a certain frequency range. These s-parameters are measured usually with the source terminal grounded and the drain and gate looking into $50\,\Omega$, so they are two-port parameters. The s_{21} parameter corresponds to the gain of the device in common-source configuration. The amplitude and the phase of all four parameters are given at many discrete frequencies. Another way to represent the transistor is with an equivalent circuit, like you have probably done in your circuits classes. The idea behind equivalent circuits is to model the device over a range of frequencies with invariant parameters. Let us begin with the simplest linear (small-signal) equivalent circuit, for which it is simpler to use admittance parameters, given by $\mathbf{I} = \mathbf{YV}$.

The y-parameters may be converted to s-parameters using the formulas given, e.g. in Microwave Engineering, David Pozar, page 187. At very low frequencies, say below 1 GHz, the capacitances and inductances associated with the MESFET are quite small, and we can assume they are negligible. The same is true for the resistive losses. The result is a simple low-frequency model shown in Fig. 8.2a. This model has an infinite input impedance and cannot be matched at the input. What is the order of magnitude of the elements of this circuit? Let us choose an Agilent MESFET in chip form similar to the package you will use in the lab, and at 500 MHz with $V_{DS} = 3\,\mathrm{V}$ and $I_{DS} = 20\,\mathrm{mA}$, the s-parameters are as follows:

$$s_{11} = 0.97\angle - 20^\circ \approx 1,$$

i.e., the input appears as approximately an open circuit with some capacitance (negative phase). Similarly,



(a) (b)

Figure 8.1: A cross section of a MESFET (a) and photograph and electrode layout (b) (courtesy Tom Midford, Hughes)

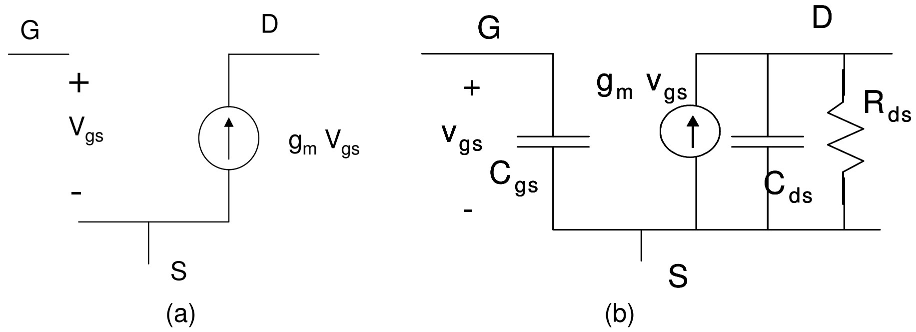


Figure 8.2: (a) Low and (b) high-frequency equivalent circuit models of a microwave MESFET. No package parasitics are included, so this model is referred to as the *intrinsic* device model.

$$s_{21} = 5.0\angle 166^\circ \approx -5,$$

which represents the gain of a common-source (inverting) amplifier. In the reverse direction, ideally a transistor does not transmit a signal, so

$$s_{12} = 0.029\angle 77^\circ \approx 0.$$

Finally, the output reflection coefficient of the transistor reduces to

$$s_{22} = 0.52\angle - 11^\circ \approx -0.5.$$

If we wish to find the values of the elements in the equivalent circuit, we would first solve for the s-parameters of the equivalent circuit in terms of the unknown elements, and then set the expressions equal to the known s-parameters, thus getting a system of equations. Finding the expressions for the s-parameters of the equivalent circuit can be quite complicated, and usually the y-parameters are found and subsequently transformed to s-parameters. For the low-frequency model from Fig. 8.2a, the y-parameters are

$$\mathbf{Y} = \begin{bmatrix} 0 & 0 \\ g_m & g_{ds} \end{bmatrix}, \tag{8.1}$$

and the s-parameters are obtained by using the conversion formulas:

$$\mathbf{S} = \begin{bmatrix} 0 & 0 \\ \frac{-2g_m}{1+g_{ds}} & \frac{1-g_{ds}}{1+g_{ds}} \end{bmatrix}, \tag{8.2}$$

From the measured scattering parameters, the voltage gain is found to be about $A_V = g_m/g_{ds} = 1.2$ and $g_m = 2.92$, $gds = 0.25$. These conductance values are normalized to $1/50 = 0.02\,\mathrm{S}$.

At higher frequencies, capacitances need to be included in the model. The circuit in Fig. 8.2b is called the "intrinsic" equivalent circuit because additional parasitics from the package are not included. The depletion capacitance of the Schottky barrier gate is represented by C_{gs} and C_{gd}, with C_{gs} usually being larger. The reason is that the positive voltage on the drain causes the depletion region on the drain side to be wider than on the source side. Also, the separation between the drain and gate contacts is usually about $1\,\mu$m larger than that between the source and gate. The capacitance between the source and drain is primarily through the substrate, and is not negligible because of the high dielectric constant of GaAs of 13. The resistance of the gate is significant because the gate contact is long and thin, and a typical value is $10 - 15\,\Omega$. The usual figure of merit for the transistor is the voltage gain A_V. Since both conductances are proportional to the gate width, the voltage gain does not depend on the width. This is important in Microwave Monolithic Integrated Circuits (MMICs), where there is complete control over gate widths, but gate lengths are fixed by the fabrication process. The gate length determines the maximum operating frequency of the device (directly, the RC time constant). An experimentally obtained formula is

$$f_{max} = \frac{33 \times 10^3}{L}\,\mathrm{Hz},$$

where L is the gate length in meters. Several cutoff frequencies are commonly used. f_T is the cutoff frequency when the short-circuited current gain of the device drops to unity. This parameter is often used, but never measured, since a microwave transistor tends to oscillate with a short-circuit load. There is a very simple rule: if you wish to make a device with a high cutoff frequency, you need to increase the saturation velocity and decrease the gate length. The saturation velocity in bulk GaAs is limited, and to overcome that the semiconductor material under the gate must be modified. This is done in High Electron Mobility Transistors (HEMTs). Bipolar transistors at lower microwave frequencies are made in silicon and silicon-germanium, and they can also be made to operate into the millimeter-wave range using GaAlAs and InP, referred to as Heterojunction Bipolar Transistors, or HBTs. The name comes from a complicated layered semiconductor structure. The details of operation of the latter two devices is beyond the scope of this course.

8.2 Transistors as Bilateral Multiports

Transistor s-parameters are usually given as common-source two-port parameters. Since there is feedback between the output and input port of a realistic transistor, the two-port is considered to be bilateral. In that case, the input scattering parameter is affected by the load impedance through this feedback (i.e. it is different from s_{11} of the device), and the output scattering parameter is affected by the generator impedance (i.e. it is different from s_{22} of the device). Referring to Fig. 8.3, the input coefficient of the two port, with given s-parameters, terminated in an arbitrary load is found from the reflection coefficient definition to be (see homework problem):

$$s_{in} = \frac{b_1}{a_1} = s_{11} + \frac{s_{12}s_{21}s_L}{1 - s_{22}s_L}. \tag{8.3}$$

A similar expression can be derived for the output reflection coefficient s_{out} looking into port 2, by replacing s_L with s_g and the appropriate s-parameters of the two port:

$$s_{out} = \frac{b_2}{a_2} = s_{22} + \frac{s_{12}s_{21}s_g}{1 - s_{11}s_g}. \tag{8.4}$$

When can we assume that a transistor is unilateral? The answer is: when s_{12} is small enough. What does that mean? We will find out when we learn about transistor amplifier gain.

8.3 Stability of Microwave Transistor Circuits

If the magnitudes of s_{in} or s_{out} from Fig. 8.3 are greater than unity at some frequency, the circuit is unstable at that frequency and potentially an oscillator. When designing an amplifier, it is important to make sure that this does not happen, by properly designing the matching networks. An amplifier is *unconditionally stable* when $|s_{in}| < 1$ and $|s_{out}| < 1$ for all passive source and load impedances, and is *conditionally stable* if this is not true in some range of generator and load impedances. Usually the following necessary and sufficient conditions are used for unconditional stability:

$$K = \frac{1 - |s_{11}|^2 - |s_{22}|^2 + |\Delta|^2}{2|s_{12}s_{21}|} > 1 \qquad \text{and} \qquad \Delta = s_{11}s_{22} - s_{12}s_{21} < 1.$$

Manufacturers will also give stability circles on a Smith chart for a given device. These circles bound regions of impedances presented to the device that would cause instabilities. These areas of the Smith chart should be avoided when designing an amplifier, and selected when designing an oscillator.

When a transistor is mounted into a microstrip hybrid circuit (such as what you will do in the lab), RF connections to ground are made using metallized via-holes. In a transistor amplifier circuit, where the source is connected to ground, the via hole is not a perfect short circuit, but rather presents a inductance between the source terminal of the device and ground. This will result in different s-parameters of the device when it is considered to be a two-port circuit. These new parameters might not satisfy the stability criteria. You should start worrying if your or parameters exceed 0 dB at some frequency when you analyze a circuit. This does not have to mean oscillation, but it is a good indication of a possible one.

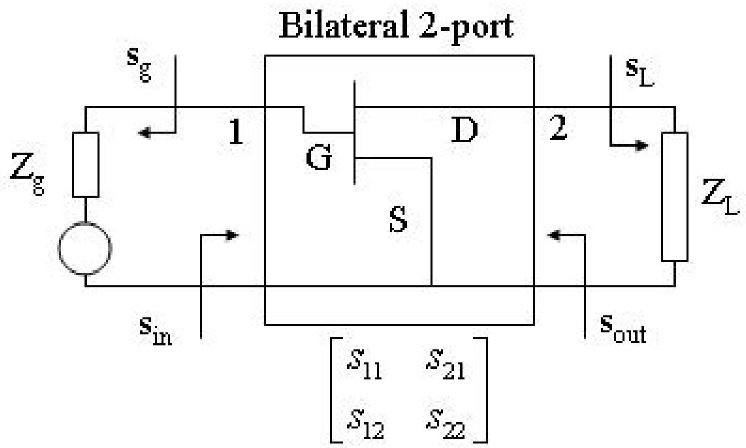


Figure 8.3: Input and output scattering parameters of a bilateral two-port transistor network.

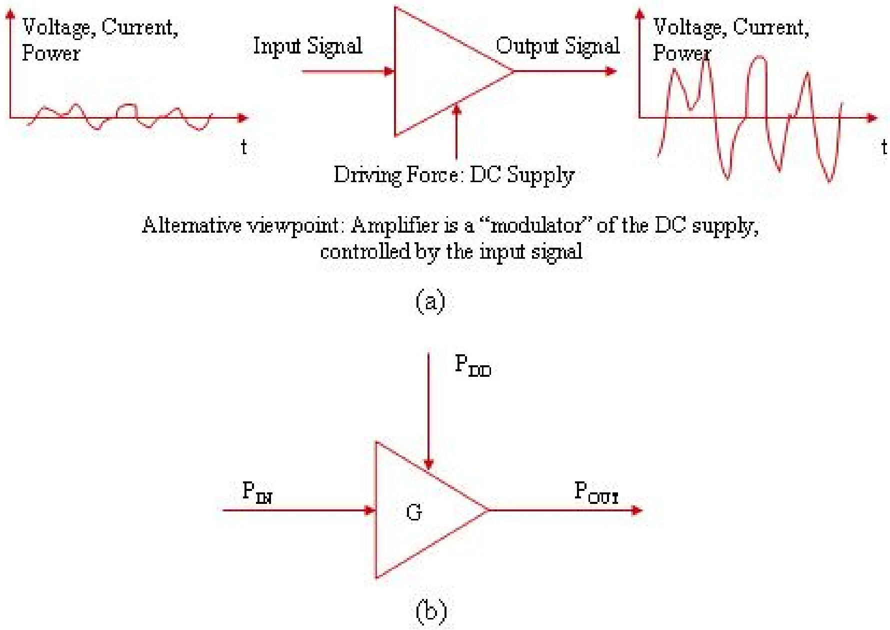


Figure 8.4: (a) General block diagram of an amplifier. (b) Power budget for defining efficiency.

8.4 Microwave Transistor Amplifiers

A simple graphical explanation of an amplifier is shown in Fig. 8.4a. An amplifier is frequently represented by a triangle which includes the transistor, input and output matching sections, and an applied DC bias.

Whether one is interested in buying or designing an amplifier, the following questions need to be answered:

1) What is the required output power?

2) What is the maximum and minimum required gain?

3) What is the operating frequency and bandwidth?

4) Is the amplifier matched and stable?

5) What sort of heat-sinking is required?

To answer these questions we need to define a set of parameters that describes the amplifier. These parameters will be defined for linear amplifiers with time-harmonic input signals:

Gain — The power gain, G, can be given as the ratio of output to input power (the gain is almost always specified in terms of dB). Voltage and current gains can similarly be defined for the ratios of the voltage or current magnitudes.

There are several definitions of the power gain in terms of the amplifier circuit. One definition useful in amplifier design is the *transducer gain*, G_T:

$$G_T = \frac{P_{load}}{P_{source}}.$$

This definition of gain is a ratio of the power delivered to a matched load divided by the power that the source would deliver to a matched load in the absence of the amplifier. If the amplifier is not matched to the source, then reflected power is lost and a second definition, the *power gain*, G, can be defined as:

$$G = \frac{P_{load}}{P_{source} - P_{refl}} = \frac{P_{load}}{P_{in}}.$$

Input and Output Match — A typical transistor with an applied DC-bias has s_{21} with magnitude greater than 0 dB, and with s_{11}, s_{22} magnitudes near 0 dB (poorly matched to 50 Ω), and a small value of s_{12}. The point of doing impedance matching is to ensure the stability and robustness of the amplifier, achieve a low VSWR (to protect the rest of the circuit and conserve input power), achieve maximum power gain, and/or to design for a certain frequency response (such as flat gain or input match over a certain frequency band).

Power Consumption — If one is buying an amplifier, the required DC voltage and current are usually given. If designing and fabricating an amplifier, does the design call for a single or dual voltage supply? Is the bias circuitry external or does it need to be integrated into the amplifier circuit? There is a direct relationship between the amount of DC power consumption and RF output power. Therefore, we can decide how much DC voltage and current to use based on the amount of output power we are interested in.

Heat Issues — Remember that an amplifier dissipates energy as it performs amplification. Therefore, we must ask how much heat is generated and see how this compares to environmental regulations and to the maximum allowed temperature of the device itself. If necessary, we may need to use some sort of temperature regulation external to the amplifier. A heat sink is often used, but its size (determined by the amount of heat it needs to dissipate) can become quite large compared to the amplifier circuit itself. Forced cooling can be used in the form of a fan (or fans), or even liquid cooling.

Amplifier Efficiency — Parameters describing the amplifier efficiency quantify the power budget of the amplifier system. The efficiency definitions below refer to Fig. 8.4b.

1. *Overall (Total) Efficiency:* from a conservation of energy standpoint, the overall efficiency makes the most sense because it is the ratio of the total output power divided by the total input power (RF and DC). It is given by
$$\eta_{total} = \frac{P_{out}}{P_{DD} + P_{in}}.$$

2. *Drain (or collector) Efficiency:* if we are only interested in the amount of output RF power compared to the amount of input DC power, then we can use the drain (or collector, in the case of a BJT or HBT) efficiency given as

$$\eta_{D,C} = \frac{P_{out}}{P_{DD}}.$$

3. *Power Added Efficiency (PAE):* the power added efficiency is very useful in that it relates how well the power from the DC supply was converted to output RF power assuming that the input power is lost. Note that if the gain is very large, then the PAE converges to the value of the drain efficiency. PAE is given by

$$PAE = \frac{P_{out} - P_{in}}{P_{DD}}.$$

4. *Power dissipated:* finally, the total power dissipated is the power that is not accounted for by the input RF, output RF, or input DC power, and can be found from

$$P_{DISS} = P_{DD} + P_{in} - P_{out}.$$

Loss Mechanisms:— The dissipated power is usually mostly power lost as heat, but several other mechanisms of power loss are possible. The following are the five main types of amplifier power loss:

1) DC power converted to heat (i.e. I^2R losses in the resistive elements of the transistor).
2) The input RF used to control the device.
3) Radiation from amplifier components (transmission lines or lumped elements) - especially if the amplifier is mismatched.
4) Conversion of the output power to harmonics of the fundamental frequency.
5) Loss in the DC and/or control circuitry.

Amplifier Stability and Main Causes of Instabilities — If you followed the standard amplifier design "recipe" as outlined in the next section, the amplifier might still oscillate. Some commonly encountered reasons for oscillations are discussed briefly bellow.

- *Source-lead inductance:* in a microstrip circuit, the ground is at the bottom of the substrate, so the source (emitter) lead needs to be connected to it with a vertical interconnect, for standard device packages. In monolithic microwave integrated circuits (MMICs), the same issue arises when microstrip is used. The source leads are grounded in both cases with metalized via holes. Even though the substrates are often quite thin (0.75 mm at most, or in the case of MMICs, 100 μm at most), the via inductance is not negligible. If it is not taken into account in the device s-parameters, the grounding inductance can cause instabilities. It provides a feedback voltage between the gate and the drain which can cause the input and output reflection coefficients to become larger than unity. This inductance is between 0.5 and 1 nH for standard substrate thicknesses, and the reason most transistor packages have 2 source leads is to half the total grounding inductance by connecting two such inductors in parallel.

  For the general purpose MESFET we have been using as an example, the stability factors for the given 2-port s-parameters are as follows: $K > 1, \Delta < 1$ in the range between 4 and 12 GHz. When a 1-nH inductance is connected to the source, the stability factors become approximately: $K < 1, \Delta < 1$ in the range between 4 and 8 GHz and $K < 1, \Delta > 1$ from 8 GHz to about 12 GHz.

- *Poor RF ground:* This effect is harder to model. However, you should be aware of the fact that a good continuous RF ground is not easy to achieve. For example, the substrate could bend and not be connected at all points to the package ground. Another common problem is the connector ground not being well connected to the package ground. Keep this in mind if a circuit that you simulated as stable ends up oscillating when you build it.

- *Bias-line oscillations:* When non-ideal bias lines are part of the amplifier circuit, since the transistor has gain at frequencies other than the operating frequency, the bias lines can present impedances that drive the amplifier into instability at some frequency other than the design frequency. These can be far away from the design frequency and in that case easily taken care of by adding capacitors that short the bias-line oscillations.

This concludes the short discussion of basic amplifier definitions and parameters. In the next part of the lecture we discuss some basic amplifier design procedures.

8.5 Amplifier Design Procedures

While concerns about output power, bandwidth, efficiency, and heat sinking were brought up in the previous section, the basic linear amplifier design discussed next is concerned only with the problem of matching the input and output of the transistor.

Amplifier design (whether linear or nonlinear design) almost always begins with transistor analysis (IV curves or equivalent circuit transistor parameters), then with Smith chart analysis, ending with computer aided design. For this discussion on linear amplifier design, we start with the transistor to remind ourselves what configuration it is to be used in. Next we use the Smith chart to design the input and output matching sections.

Linear (small signal) Unilateral Design Example

The word linear in this case means that the output power is a linear function of the input power (the gain at a given frequency is constant and does not change with output power). Linearity is a consequence of operating the device in small signal mode. The graphical meaning of small signal can be seen below as related to the IV curves of the device, Fig. 8.5. The term unilateral means that power flows in only one direction through the device, i.e. $s_{12} = 0$.

The design begins by taking note of some basic properties of the transistor used as the active device. The IV curves for a BJT (HBT) are shown below (IV curves for a FET are similar and the discussion here applies). The quiescent point, Q, indicates the chosen base and collector (gate and drain) DC bias points. The RF input of the common emitter (common source) design creates a current swing on the base, which is translated into a voltage swing on the drain through the amplification process.

Given that the bias and input drive level have been chosen for us, our design task is to find input and output matching sections for the transistor. The assumption is that the manufacturers of the transistor were kind enough to give us the s-parameters of the device for the chosen bias point and input drive. For example let us assume that we are given the following s-parameters at 5 GHz:

$$f = 5\,\text{GHz} : s_{11} = 0.7\angle - 100^\circ, s_{12} = 0, s_{21} = 2.5\angle 60^\circ, s_{22} = 0.7\angle - 30^\circ.$$

If we now assume 50-Ω source and load impedances, the task becomes to transform both s_{11} and s_{22} into 50 Ω. To perform a conjugate match, we simply want design matching sections such that

$$s_g = s_{11}^* \qquad \text{and} \qquad s_L = s_{22}^*$$

This means that for the source we must see, instead of 50 Ω, the complex conjugate of s_{11}, so $s_g = 0.7\angle 100^\circ$. Similarly, $s_L = 0.7\angle 30^\circ$. By conjugate matching the input and output of the device to

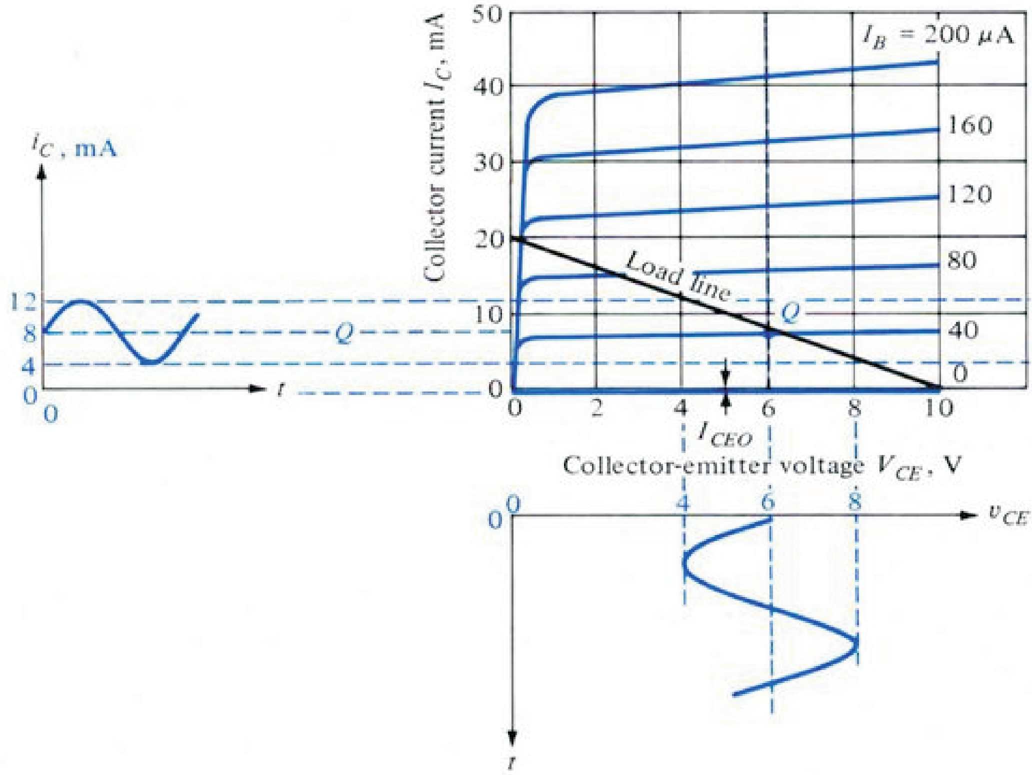


Figure 8.5: IV curves of a transistor showing a load-line, operating point Q and input current and output voltage waveforms.

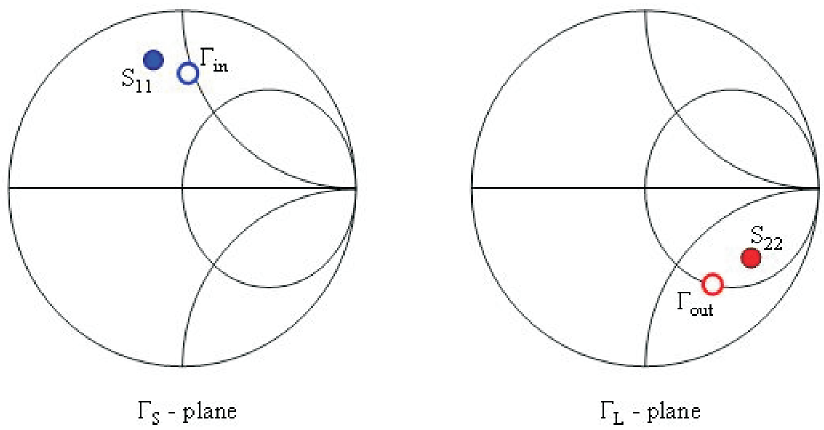


Figure 8.6: Smith chart comparison of the difference in the transistor reflection parameters and the circuit reflections given by s_{in} and s_{out} when $|s_{12}| > 0$.

the source and load, we have guaranteed the maximum transducer gain possible for this amplifier. This matched transducer gain is a function of s_{21} and the matching sections and is given by

$$G_{Tmax} = \frac{1}{1-|s_g|^2}|s_{21}|^2\frac{1-|s_L|^2}{|1-s_{22}s_L|^2}.$$

We may stop here with the design of the unilateral case. We have achieved the best possible match and gain of the device at 5 GHz. But we have neglected many things in this simple example. So now we ask, what if s_{12} is not zero or negligibly? Later we will ask in addition the following questions: what happens at other frequencies, what happens when we exceed the small signal approximation, and what do we do if concerned with noise or efficiency?

Linear (small signal) Bilateral Design Example

If the s_{12} parameter is non-zero, the analysis becomes more complex, but we can still follow the strategy in the previous example. If $s_{12} > 0$, it is mostly due to the capacitance between the gate and drain discussed earlier in the equivalent circuit model of the device. The important result of this input-to-output coupling is that we can no longer say that $s_{in} = s_{11}$ as in the unilateral case. Similarly, s_{out} is no longer simply s_{22}, but now depends on also on s_{21}, s_{12} and s_g.

Therefore, we now have to perform a simultaneous match of the input and output (in other words, what we do at the input affects the output match, and vice versa). Note that the denominators in the expressions for s_{in} and s_{out} and are less than unity. This indicates that for some values of the transistor s-parameters, s_g and s_L, it may be possible for s_{in} and s_{out} to be greater than 1, which indicates that the transistor amplifier circuit is giving power back to the source (i.e. it is an oscillator).

A simultaneous match can be found analytically by the simultaneous equations, given in most microwave texts. An excellent reference for amplifier design is the book "Microwave Amplifiers" by Gonzales. Fig. 8.6 shows how s_{in} and s_{out} differ from s_{11} and s_{22}. In this particular example it can be seen how the matching can differ when the feedback between drain and gate s_{12} is taken into account. In practice the difference in gain can be 2 dB or more.

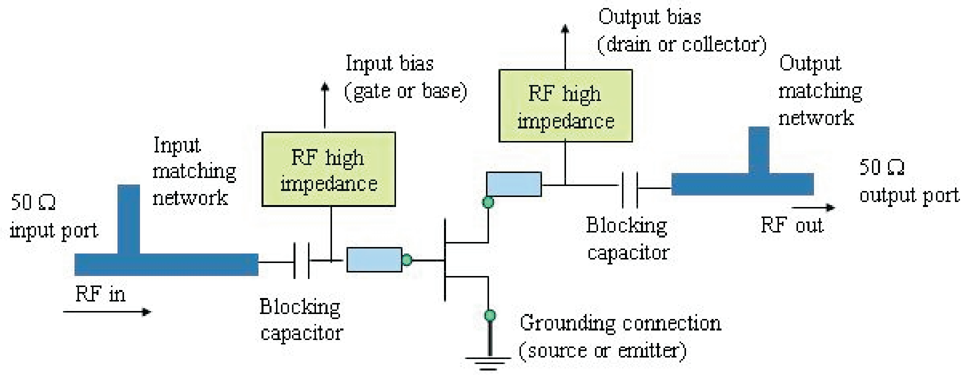


Figure 8.7: Complete circuit schematic of a microwave amplifier including the bias circuits in the gate and drain.

8.6 Biasing the Transistor in an Amplifier

In the block diagrams in Fig. 8.4, there is no obvious place where power to the device is added. An amplifier has AC gain at the expense of input DC power. In the beginning of this lecture, we mentioned that a MESFET needs a positive drain-to-source DC voltage and a negative gate-to-source voltage. This means that in general the input and output of the amplifier need to be connected to a bias supply, which should not change the input and output reflection coefficients. A full amplifier circuit is therefore shown in Fig. 8.7.

Designing good biasing circuits is half of amplifier design, and the following are critical design parameters:

1) the bias network needs to be "invisible" to the RF waves, i.e. as close to an open circuit as possible. The reason is that we cannot afford any of the RF power to be lost in the biasing circuit and power supply.

2) the DC bias needs to be isolated from the RF circuit, i.e. we do not want the DC voltage to be present at the RF input (e.g. we might be dealing with a 2-stage amplifier, and the previous stage requires a different voltage).

3) the DC bias circuit should behave properly over the entire frequency range where the device has gain, so as not to cause instabilities.

In order to satisfy the first criterion, the DC bias lines need to be inductors. It is difficult to make an inductor at microwave frequencies due to parasitics. Another option is that the bias lines have the characteristics of a low-pass filter, as shown in Fig. 8.8. In order to satisfy the second requirement, a DC blocking capacitor needs to be added to the circuit, and needs to be taken into account in the design. Capacitors are not ideal shorts at microwave frequencies (they have parasitics), and they need to be taken into account at microwave frequencies.

The bias is supplied through a biasing circuit that needs to present a high impedance to all present RF signals so as not to present an additional (usually not well characterized) load. This is relatively straightforward in a narrowband amplifier design, but becomes a challenge for the broadband case. RF capacitors are needed to block the DC signal to the RF input and output. Usually the source (or emitter) terminal of the active device is connected to RF (and often DC) ground. In the case of microstrip, one or more metalized via holes are used for grounding, and they present an equivalent inductance between the

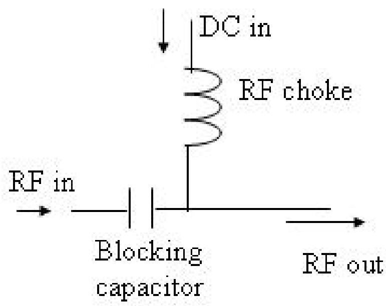


Figure 8.8: A bias Tee equivalent circuit.

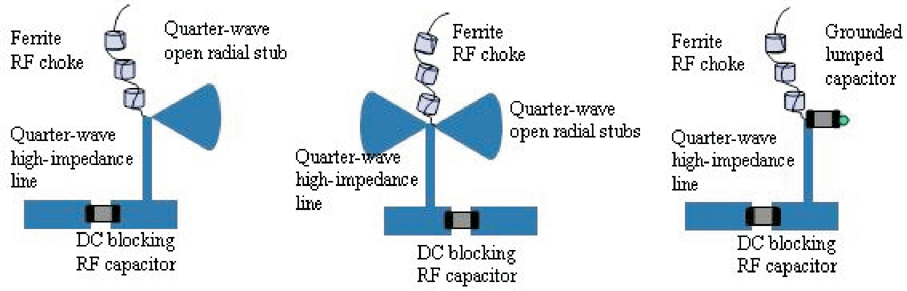


Figure 8.9: Examples of microstrip biasing circuits.

terminal and ground. You will examine the effect of this inductance in your homework and project. In the case of coplanar waveguide (CPW) circuits, the connection is more straightforward and the parasitic reactance can be minimal.

Commercial bias Tees are fairly large and have typically SMA connectors at the two RF ports. Often, biasing circuits are part of amplifier design, and some examples are shown in Fig. 8.9.

The DC biasing circuit should be taken into account when analyzing stability, i.e. it is part of the input and output network. Even though it is designed to present a high impedance to the RF signal at the design frequency (convince yourself why this is so), it is not a real open circuit. For example, at a frequency other than the design frequency, the quarter-wave shorted line is not a quarter-wavelength long, and therefore is not an open circuit to the RF signal. There is also some loss in the blocking capacitor - the DC blocking capacitor has lead inductance and some resistance, and it will not be perfectly matched to the input RF 50-ohm line. In the grounded capacitor implementation (righthand side of Fig. 8.9, the capacitor and via hole have inductance that is usually not well characterized, and this also determines the quality of the open circuit presented to the RF signal.

The ferrite choke is an inductor at lower frequencies and is effective at choking frequencies up to a few hundred MHz. This is important, since the bias lines can be good antennas for broadcast signals. At microwave frequencies, however, the ferrite is just a large resistor (the material is very lossy), so the RF currents will be very attenuated and will not reflect back into the circuit. However, the power is lost and any power flow into the ferrite lines should be minimized. A difficult problem is a broadband bias line. In principle, a good inductor with several hundred nH inductance would solve the problem, but microwave inductors typically do not work above a few GHz due to parasitic capacitance. If loss is not an issue, however, the Q factor of the inductor can be reduced by adding resistors or ferrites and very broadband bias networks can be made. One company that makes tiny cone-shaped inductors with ferrite loading is called Piconics (see their web page www.piconics.com for more on these inductors). The idea is that the Q is greatly reduced at high frequencies, so the resonance due to the parasitic capacitance is not relevant.

8.7 Practice questions

1. How is a MESFET different from a MOSFET?

2. Why is GaAs used for transistors at microwave frequencies?

3. How large (order of magnitude) is the gate length of a microwave FET, and why does it need to have that size?

4. What are Y parameters of a network and why are they useful in determining an equivalent circuit from measured parameters?

5. How is a MESFET biased?

6. What are the input and output impedances of a MESFET at low frequencies equal to? What changes at high frequencies?

7. What are the voltage gain, transition frequency and maximum frequency of a transistor?

8. Why is a circuit unstable if $|s_{ii}$ at any port i is larger than unity?

9. Prove that if in a two-port network $|s_{11}| > 1$, then $|s_{22}|$ has to be greater than unity at the same frequency. Use the bilateral two-port network equations.

8.8 Homework Problems

1. Derive expressions (8.3) and (8.4) starting from Fig. 8.3.

2. Derive the Y-parameters of the high-frequency MESFET equivalent circuit shown in Fig. 8.2b. (You might get an over-determined system of equations, in which case you need to make a well-informed choice about the correct approximate expression.)

3. For the device with specifications as given in the file MESFET-parameters.pdf on the class web page, compare the s_{11} and s_{22} with the s_{in} and s_{out}, respectively, at 1 GHz and 4 GHz for three different load and generator impedances.

4. For the device with specifications as given in the file FET-parameters.pdf on the class web page, determine the stability at 1 GHz and 5 GHz.

5. An amplifier has a small signal gain of 12 dB. The amplifier is biased at 5 V and 40 mA. When the input power reaches 12 dBm, the amplifier saturates and the gain of the amplifier compresses by 3 dB (this is very high compression). The output power at that compression is 21 dBm. Find the drain (collector) efficiency of this PA, the power-added efficiency and the total efficiency. How much of the DC power is converted to heat?

6. An amplifier has a power-added efficiency $PAE_1 = 30\%$. Compare the battery power usage and heat dissipation for an amplifier with the same bias point, same output power, same gain, same input match, but $PAE_2 = 70\%$.

7. Choose a transistor available in Ansoft Designer. Plot the s-parameters of the transistor in 2-port common source configuration and answer these questions: (1) what is the f_T of the device? (2) What is the largest value of s_{21}? (3) Does the transistor have any potential instabilities - how can you tell? (4) How well is the device matched to 50 Ω at input and output? (5) What would happen to the gain if you matched the input? (6) How do the s-parameters change if the bias point is decreased and/or increased ($V_{DS}I_{DS}$ larger)?

8. For the transistor chosen in problem #7 design a unilateral match for input and output at some frequency between 1 and 8 GHz based on the s_{11} and s_{22} parameters you read off the plot. Connect your matching circuits to the device and simulate the response over a range of frequencies. Discuss and attach plots.

9. For the problem above, perform a bilateral match to improve gain, and input and output match while maintaining stability over the range where the transistor has gain.

10. Design a bias Tee network for 2 GHz on FR4 substrate ($\epsilon_r = 4.5$, 0.508-mm thick) using Ansoft Designer. If needed look up available values for capacitors and inductors (a good company for capacitors is ATC www.atceramics.com, and for inductors CoilCraft www.coilcraft.com). Plot the match at RF ports, and coupling between RF and DC ports. What is the bandwidth of your bias Tee?

Chapter 9

Microwave Communication Links

9.1 Transmitters and Receivers

At microwave frequencies, two types of communication systems are used: guided-wave systems, where the signal is transmitted over low-loss cable or waveguide; and radio links, where the signal propagates through space. It turns out that guided-wave systems are much lossier over long distances than free-space propagation, and this is illustrated in Fig. 9.1. You can see that even standard fiber-optical cable introduces much more attenuation than does free-space propagation. The reason for this is that in a guided wave system, the attenuation of power with distance approximately follows an exponential function $e^{-2\alpha z}$, and the power radiated from an antenna falls off as $1/z^2$, as we saw in the previous lecture. (Recently, propagation in fibers has been shown to have extremely low loss when irbium-doped amplifiers are parts of the optical link. This is not shown in Fig. 9.1.)

In principle, microwave links could be made directly at a microwave frequency and detected at that frequency. The problem in this case is that it would be hard to simultaneously transmit many channels, which is needed for every communication system. The reason is that it is difficult and expensive to make amplifiers and filters for each channel. The solution is a so called *superheterodyne* system, in which the microwave signal is the carrier that travels through free space, and it is modulated with a signal of much lower frequency. At the receiving end, the microwave frequency is converted to a lower frequency, at which it is easy and cheap to make amplifiers and filters. For example, an analog TV channel is about 6 MHz wide, and if the microwave carrier is 4 GHz, 66 TV channels can be transmitted in a microwave link that has only a 10% bandwidth, and the same link could transmit as many as 100,000 voice channels.

A typical heterodyne transmitter and receiver are shown in Fig. 9.2(a) and (b). The input signal (voice, video or data) is at some frequency f_m, called the *baseband frequency*. A microwave signal, called the *local oscillator*, or LO, is modulated with the baseband signal in the mixer (a nonlinear device: see the next section). The result is a double-sideband signal, which means that a signal of frequency $f_{LO} - f_m$ and one at $f_{LO} + f_m$ are transmitted. This process is called *up-conversion*. The power amplifier is designed to amplify at the microwave frequency f_{LO}, and this power is then radiated by the transmitting antenna. At the receiver end, the signal received by the receiving antenna is first amplified by a low-noise amplifier at the microwave frequency. Then it is demodulated by a mixer whose frequency is offset from the transmitter LO exactly by the value of f_{IF}, which is called the *intermediate frequency*. The IF signal is then filtered, amplified with a high gain amplifier, and then the original baseband signal at f_m is restored after detection. For satellite communication, typically the microwave carrier is at 4 GHz (or 12 GHz) and the IF frequency is between 10 and 100 MHz, where low-frequency circuits can be used. For radar applications at millimeter-wave frequencies, say at 40 GHz, the IF is a few gigahertz, and then microwave low-noise circuits are designed for the IF signal, which is then usually further down-converted.

There are many variations to this scheme. Many different modulation techniques are used: single

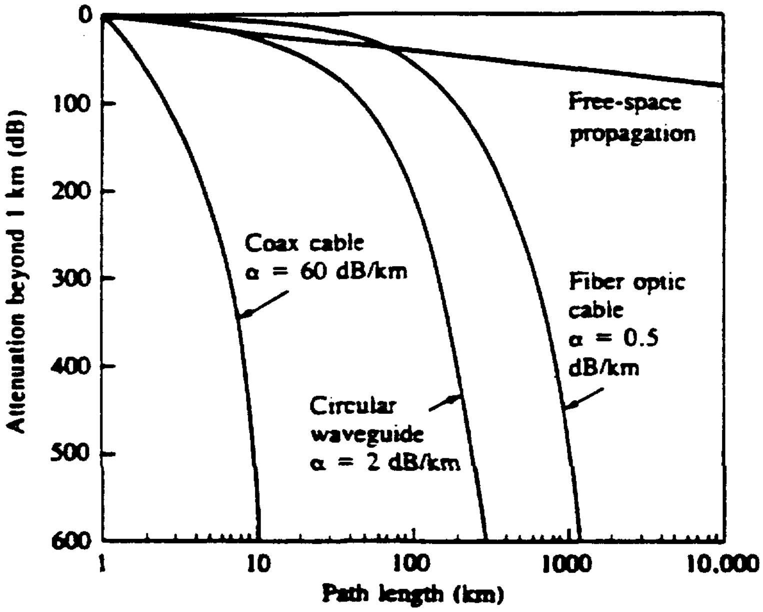


Figure 9.1: Attenuation for different transmission systems.

sideband, which allows more channels, since each one occupies less bandwidth; frequency modulation (FM) which has a higher signal-to-noise ratio, as well as digital modulations of different types. In the lab, you will look at both amplitude modulation (AM) and frequency modulation (FM). The first one means that the amplitude of the microwave carrier is changed to carry information, for example it can be turned on and off, which would represent a 1 or a 0. Frequency modulation means that the frequency of the microwave carrier is changed to carry information. Examples of the time waveforms and spectra of AM and FM signals are shown in Fig. 9.3.

9.2 Mixers

Mixers shift the frequency of a signal, either to a higher or lower frequency. This means that a mixer has to use a nonlinear device, as linear devices never give frequency components other than those that are input to them. The mixer you will use in the lab is just the waveguide-mounted detector diode operated in a different way. As opposed to power detectors, in a mixer detector the phase and amplitude of the microwave signal waveform are retained. We already mentioned that when using a mixer, a microwave link can have many channels, and this is why it is so widely used in telecommunications. Some disadvantages of a mixer are that it adds noise to the signal, and that it produces power at many frequencies, since it is a nonlinear device. We will look at the operation of the simplest type of mixer, that contains a single nonlinear device, usually at microwave frequencies a Schottky diode, Fig. 9.4.

As we have learned in chapters 4 and 10, the diode current is a nonlinear function of voltage. To analyze the operation of a mixer, we use the small-signal (quadratic) approximation:

$$I(V) = I_0 + i = I_0 + v_d G_d + \frac{v_d^2}{2} G_d' + \dots \tag{9.1}$$

The current in this approximation has a DC component (I_0), a component proportional to the incident AC voltage, and one proportional to the square of the incident voltage. In a Schottky-diode mixer used in a receiver, the incident low-level (attenuated after long-range propagation) high-frequency carrier (RF signal) carries information which need to be received. How is this done? The low-frequency modulated

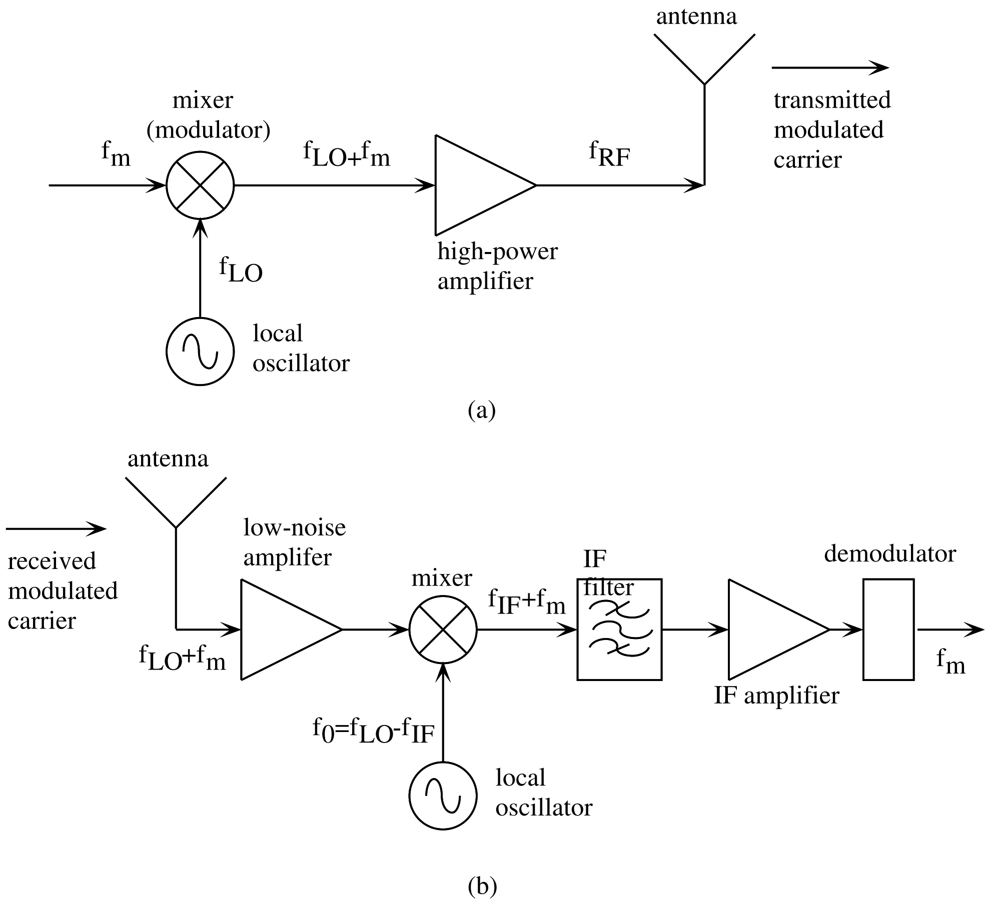


Figure 9.2: Block diagram of an AM microwave transmitter (a) and receiver (b).

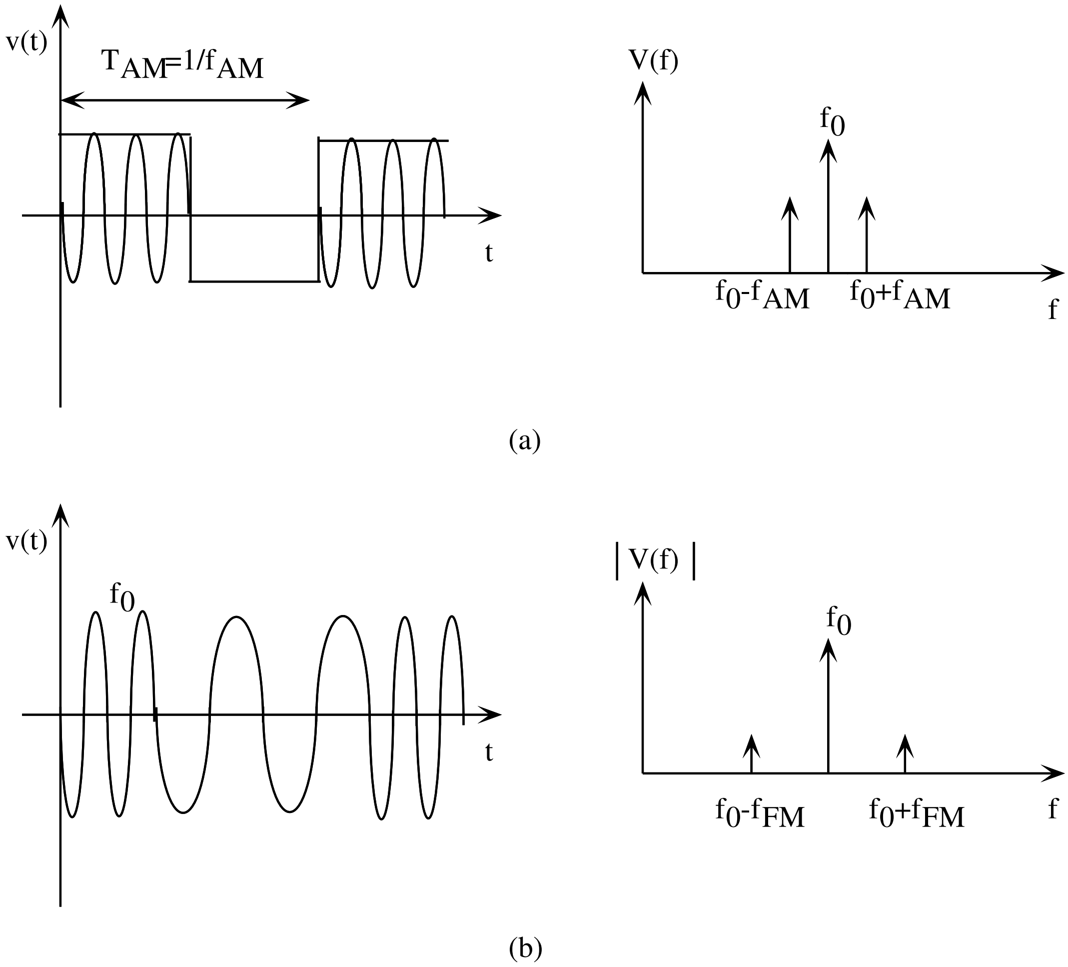


Figure 9.3: Time domain waveforms and spectra of an (a) AM and (b) FM modulated sinusoidal signal of frequency f_0.

signal is retrieved by adding the radio frequency carrier (RF signal) and a local oscillator (LO) signal close in frequency to the RF signal, and applying their sum to the diode:

$$v_d = v_{LO} + v_{RF} \tag{9.2}$$

The LO and RF voltages can be written as:

$$v_{LO} = \sqrt{2}V_0 \cos\omega_0 t \quad \text{and} \quad v_{RF} = \sqrt{2}V \cos\omega_{RF} t. \tag{9.3}$$

where usually the received signal voltage is much smaller than the LO voltage: $|V| \ll |V_0|$.

The diode current expression (9.1) will now have a constant DC bias term, linear terms at both frequencies ω_{RF} and ω_0, and the v_d^2 term of the current which gives the following frequency components:

$$\begin{aligned} i &= G_d' \left(V \cos\omega_{RF} t + V_0 \cos\omega_0 t\right)^2 \qquad (9.4) \\ &= \frac{G_d'}{2} \left\{V^2 + V_0^2 + V^2 \cos 2\omega_{RF} t + V_0^2 \cos 2\omega_0 t + 2VV_0 \cos\left[(\omega_{RF} - \omega_0)t\right] + 2VV_0 \cos\left[(\omega_{RF} + \omega_0)t\right]\right\} \end{aligned}$$

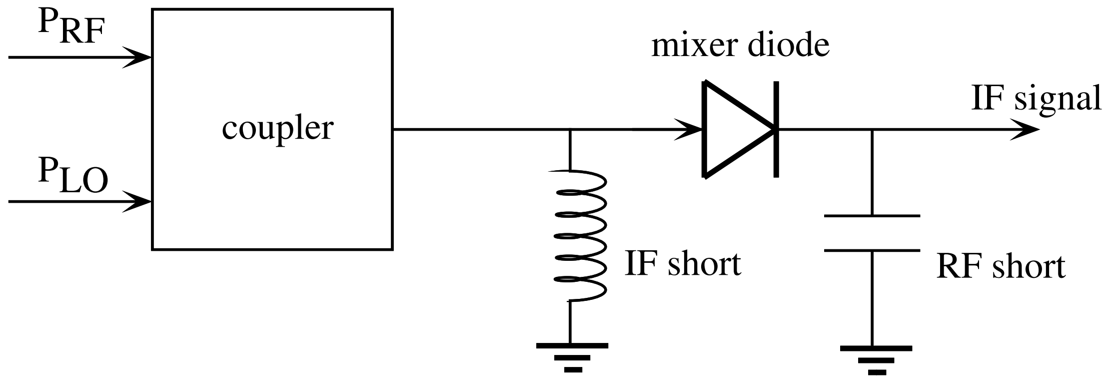


Figure 9.4: A single-ended mixer circuit.

In this expression, the DC terms are not important, and the 2ω and $2\omega_0$ terms can be filtered out. The most important terms are the ones at $\omega_{RF} + \omega_0$ and $\omega_{RF} - \omega_0$. In case of a receiver, the frequency $\omega_{RF} - \omega_0$ is the intermediate (IF) frequency.

At the transmitter end, the mixer performs the up-conversion. In this case, the IF signal at ω_i mixes with the LO signal at the diode. The resulting frequency components at the mixer output are at $\omega_0 \pm \omega_i$, where $\omega_0 + \omega_i$ is called the upper sideband, and $\omega_0 - \omega_i$ is the lower sideband. Both sidebands are close in frequency to the LO frequency, and are the transmitted RF signal.

A disadvantage of this single-ended circuit is that the diode treats the RF and the LO the same way. This means that if the LO is noisy, and its noise extends to the RF frequency, the noise will be converted down to the IF frequency and will make the output noisy. To see this, replace V_0 by $V_0 + V_n(t)$, where V_n is of similar magnitude to the received voltage V (and hence small compared to V_0). Then, $V_0^2 \to V_0^2 + 2V_0V_n(t)$, and a new term appears in (9.4) that is the same order of magnitude as the IF term already present. Since the typical RF signal has a very low level when it gets to the mixer, any additional noise makes reception much harder.

There are several ways to deal with this problem. For example, the IF frequency can be increased, because the noise is typically smaller further away from the center frequency of the noisy signal. Another way is to use a very clean LO. A third way is to make a different mixer circuit, called a *balanced mixer* shown in Fig. 9.5. Again let us assume that the LO voltage has an AM noise component $V_n(t)$, so that

$$v_{LO} = \sqrt{2}\,[V_0 + V_n(t)]\cos\omega_0 t \tag{9.5}$$

Once again, let

$$v_{RF} = V\cos\omega_{RF} t \tag{9.6}$$

and assume that $V \ll V_0$, and $V_n(t) \ll V_0$. Now we can write the voltages across the two diodes as

$$\begin{aligned} v_1(t) &= \sqrt{2}\,[V\cos(\omega_{RF}t - \pi/2) + (V_0 + V_n)\cos(\omega_0 t - \pi)] \\ v_2(t) &= \sqrt{2}\,[V\cos(\omega_{RF}t - \pi) + (V_0 + V_n)\cos(\omega_0 t - \pi/2)] \end{aligned} \tag{9.7}$$

We will now assume that the diodes are identical, and we will look only at the quadratic (nonlinear) terms in the diode currents, $i_1 = Kv_1^2$ and $i_2 = -Kv_2^2$, where K is some constant. After plugging in the voltage expressions, and after low-pass filtering, we are left with DC terms, noise and IF ($\omega_i = \omega_{RF} - \omega_0$)

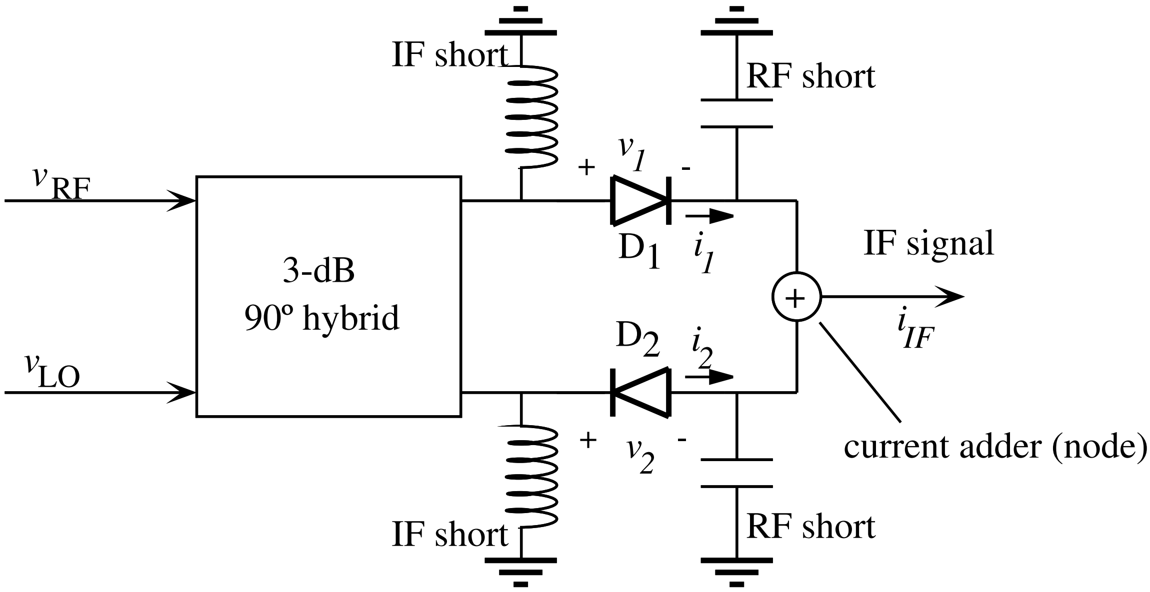


Figure 9.5: A balanced mixer circuit, using a 90° hybrid.

terms:

$$\begin{aligned} i_1(t) &= K\Big[V^2 + (V_0+V_n)^2 - 2V(V_0+V_n)\sin\omega_i t\Big], \\ i_2(t) &= -K\Big[V^2 + (V_0+V_n)^2 + 2V(V_0+V_n)\sin\omega_i t\Big]. \end{aligned} \tag{9.8}$$

The output current is now

$$i_{IF}(t) = i_1 + i_2 = -4KV(V_0+V_n)\sin\omega_i t \simeq -4KVV_0\sin\omega_i t, \tag{9.9}$$

since we said that $v_n \ll v_0$. This means that the noise terms are canceled to first order, but that the desired IF terms add up in phase.

Since the mixer is usually one of the first stages of a receiver, it is very important that it has low noise. Also, it is important that it is not lossy. Two quantities by which a mixer is usually described are its *noise figure* (typically 5 to 10 dB) and its *conversion loss*, or *conversion gain*. The conversion loss is defined as

$$L = \frac{P_{IF}}{P_{RF}}, \tag{9.10}$$

and is usually quoted in dB. For a typical commercial Schottky diode mixer, the conversion loss is 5 dB. A diode mixer has conversion loss, since it cannot produce power. Mixers can also be made with transistors (transistors are also nonlinear), in which case we talk about conversion gain. At microwave frequencies, the commonly used transistor is the MESFET (MEtal Semiconductor Field Effect Transistor), and the ones used for mixers are made with two gates, so that the RF signal comes in through one gate terminal, and the LO through the other. For very low-noise applications, where the RF signal is very small, for example in radio astronomy, people use superconducting diodes at liquid helium temperatures, which reduces the noise.

Two RF frequencies $\omega = \omega_0 + \omega_i$ and $\omega = \omega_0 - \omega_i$ will give the same IF frequency ω_i when mixed with an LO at ω_0. Usually, only one of these RF frequencies is wanted, and the other one is called the

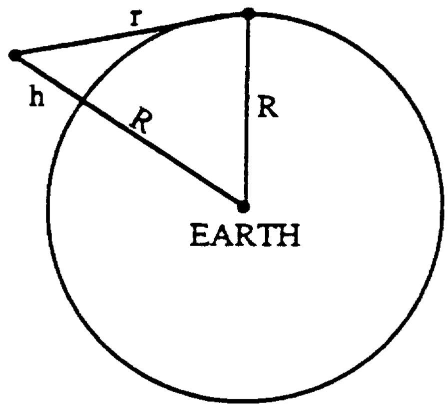


Figure 9.6: Line-of-sight path limit on curved Earth.

image and is undesired. In this case an image-rejection mixer circuit is used. Another commonly used mixer circuit is the double-balanced mixer, which suppresses even harmonics of the LO and the RF. This mixer looks like a bridge circuit and uses four diodes.

9.3 Propagation

9.4 Line-Of-Sight Transmission Paths

AM broadcasting systems rely on surface wave transmission between two points on the Earth's surface. Short wave radio systems use bounces off of the ionosphere. The UHF and VHF radio used by airplanes, as well as microwave radio relay links, propagate along a direct path. This is called *line-of-sight* propagation, illustrated in Figure 9.6. You can see from the figure that the range will be limited by the curvature of the Earth. That is why almost all radio relay stations are put up on high peaks, even though the weather conditions at these places often complicate the design (and very few people want to work there). From the figure, we can write:

$$(R+h)^2 = R^2 + r^2, \tag{9.11}$$

where R is the radius of the Earth, h is the height of the antenna above ground, and r is the range of the communication link. If we solve for r and expand, we can write approximately:

$$r = \sqrt{h^2 + 2Rh} \simeq \sqrt{2Rh}. \tag{9.12}$$

This formula predicts shorter ranges than the ones achievable in reality. The reason is the change in the refractive index of the atmosphere, so that the waves really follow a curved, not a straight path. This is shown in Figure 9.7. The waves bend towards the denser layers, and this gives longer ranges. This effect varies quantitatively depending on the place on the Earth and the time. A reasonable approximation is usually to use an effective radius for the Earth R_{eff}, which is 4/3 of the actual radius, and is usually taken to be 8500 km. If both the receiving and transmitting antennas are above ground, the line-of-sight approximate range formula becomes:

$$r = \sqrt{2R_{eff}h_t} + \sqrt{2R_{eff}h_r}. \tag{9.13}$$

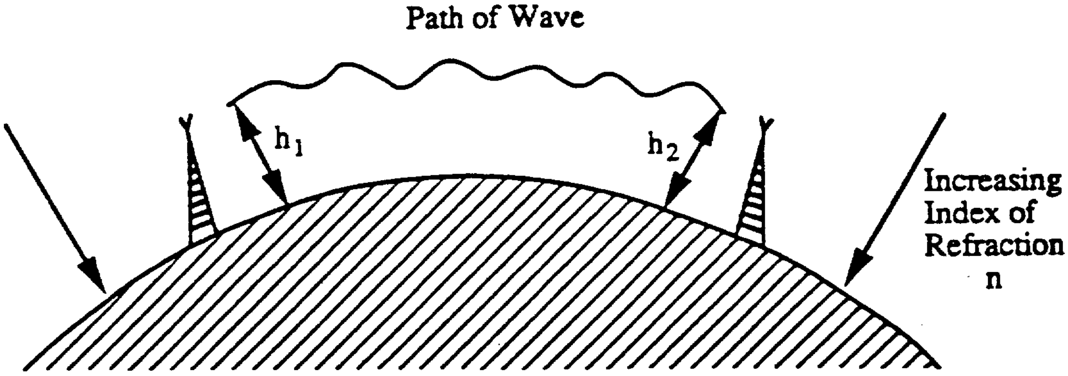


Figure 9.7: The variation of the refractive index of the atmosphere makes the paths of the waves longer.

9.5 Propagation, Reflection and Transmission in the Ionosphere

The upper layers of the Earth's atmosphere, about 50 to 500 km above the surface, are in the form of an ionized gas. This layer is called the ionosphere. Let us look at a plane wave propagating through the ionosphere. There are both positive and negative ions, as well as free electrons in the ionosphere, but the dominant effect on wave propagation comes from the electrons, since their relative charge to their mass is the largest. The equation of motion for a charged particle moving in a sinusoidally varying electric and magnetic field is:

$$m\frac{\mathrm{d}v}{\mathrm{d}t} = q\vec{E}\cos\omega t + q\vec{v}\times\mu_0\vec{H}\cos\omega t, \tag{9.14}$$

and since $E = H\sqrt{\mu_0/\epsilon_0}$, the second term on the right hand side of the previous equation is c_0/v smaller than the first term, so we will neglect it. Now we have, approximately:

$$\vec{v} = \frac{q}{\omega m}\vec{E}\sin\omega t, \tag{9.15}$$

$$\vec{J} = Nq\vec{v} = \frac{Nq^2}{\omega m}\vec{E}\sin\omega t, \tag{9.16}$$

$$\nabla\times\vec{H} = \vec{J} + \epsilon_0\frac{\partial\vec{E}}{\partial t}, \tag{9.17}$$

$$\nabla\times\vec{H} = -\omega(\epsilon_0 - \frac{Nq^2}{\omega^2 m})\vec{E}\sin\omega t. \tag{9.18}$$

In the case of pure vacuum, $N = 0$, so from the above formula we see that ions reduce the permittivity of the ionosphere, and this effect is proportional to the ratio q/m. We can define an *effective permittivity* of the ionosphere as

$$\epsilon' = \epsilon_0(1 - Nq^2/\epsilon_0\omega^2 m) = \epsilon_0(1 - \frac{f_c^2}{f^2}), \tag{9.19}$$

where f_c is called the *cutoff frequency* of the ionized gas. Everything we have said so far about wave propagation holds for propagation through ionized gases, provided we use the effective permittivity ϵ'. The propagation constant is

$$\beta = \omega\sqrt{\epsilon'\mu_0} = \frac{\omega}{c_0}\sqrt{1 - f_c^2/f^2}. \tag{9.20}$$

For frequencies $f < f_c$ the wave will not propagate.

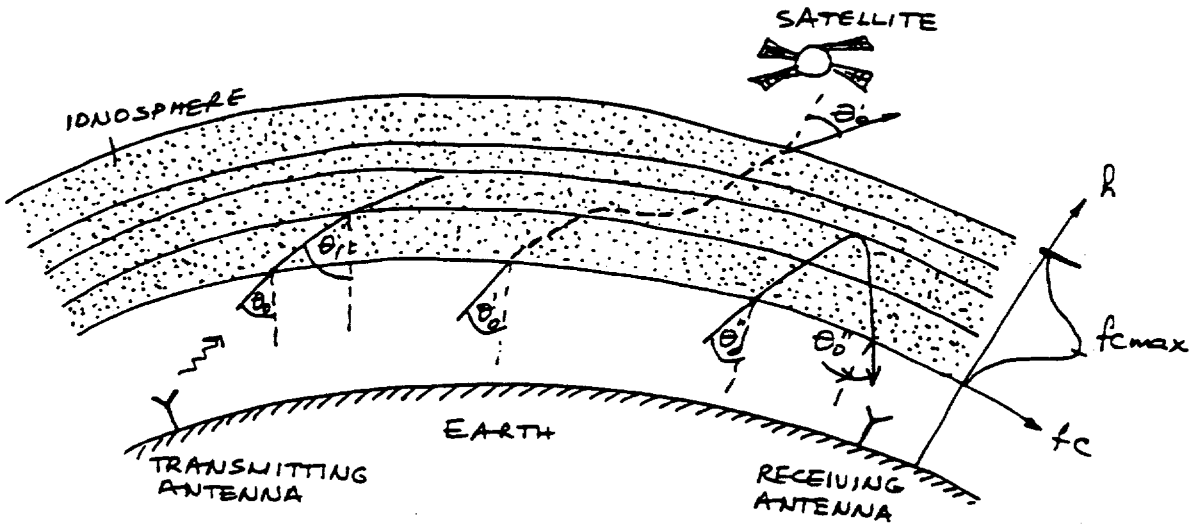


Figure 9.8: Propagation of waves through the ionosphere.

What happens to a wave transmitted from the ground when it gets to the ionosphere? The ionosphere does not have a uniform concentration of charges, and the concentration also depends on the time of the day, as well as the season, latitude, weather conditions and the intensity of the radiation of the Sun. There are more layers during the day, and some disappear during the night. Some layers of the ionosphere have a higher effective permittivity, which affects the cutoff frequency. We can plot the approximate dependence of the cutoff frequency versus the height above ground, as shown in Fig. 9.8. The frequency range is about 3 to 8 MHz.

Now we can use Snell's law to see what happens to a wave entering the ionosphere at some angle θ_0:

$$\frac{\sin\theta_0}{\sin\theta_1} = \sqrt{\frac{\epsilon_1'}{\epsilon_0}} \qquad \frac{\sin\theta_1}{\sin\theta_2} = \sqrt{\frac{\epsilon_2'}{\epsilon_1'}} \tag{9.21}$$

So, at any point in the ionosphere:

$$\frac{\sin\theta_0}{\sin\theta'} = \sqrt{\frac{\epsilon'}{\epsilon_0}}, \tag{9.22}$$

$$\sin\theta'\sqrt{1 - f_c^2/f^2} = \sin\theta_0. \tag{9.23}$$

If f and the angle θ_0 are such that $\theta' = \pi/2$ at a height where $f_c = f_{cmax}$, the wave will bend, but it will come out of the ionosphere at the same angle θ_0 at which it entered. If f and θ_0 are such that θ' reaches $\pi/2$ before f_c reaches f_{cmax}, the wave reflects off of the ionosphere and again comes back at θ_0. In this case it reaches a height where:

$$\sin\theta_0 = \sqrt{1 - f_c^2/f^2} \quad \text{and} \quad f_c = f\cos\theta_0. \tag{9.24}$$

So, a wave of 3 MHz cannot pass through the ionosphere, and one of about 8 MHz can pass if it is at almost normal incidence. If we wish to communicate with a satellite that is above the ionosphere, and want to be able to send a wave at smaller incidence angles, we have to use much higher frequencies. On the other hand, we can use the ionosphere and surface of the Earth as a large waveguide for lower frequencies, and this is used in radio communication.

9.6 Practice questions

1. Why are microwave frequencies used for communications?

2. Derive the mixing products (all the frequency components) on the receiver end of a superheterodyne link. Assume a quadratic diode I-V curve approximation. Which terms are the ones you would keep and how do you get rid of the rest?

3. The same diode can be used to detect power of an unmodulated signal, as a mixer, but also to detect power of an AM modulated signal. Show this by finding the frequency components of a Schottky diode current, if the AM signal can be written as

$$v(t) = \sqrt{2}V_0(1 + m\ \cos\omega_m t)\cos\omega t,$$

where ω_m is the low modulation frequency, ω is the high microwave frequency (usually called the carrier frequency), and m is the modulation index (it describes how strongly the carrier is modulated).

4. Draw diagrams of a single-ended and balanced mixer and explain the advantage of one over the other.

5. Which quantities are important for describing a mixer?

6. Why is it important for a mixer to have low added noise?

7. What is a superheterodyne receiver?

8. What different modulation methods do you know? Where in the superheterodyne receiver block diagram does the modulation method become important for the physical implementation of the components?

9. You have a 1-Watt transmitter. At the receiver, your noise is at -40 dBm. How big is your range (distance between transmitter and receiver) if you are allowing a signal-to-noise ratio of 2:1? What does the range depend on?

10. Why are microwave relay link towers placed on top of hills?

11. Derive the range for a line-of-sight link between two towers of heights h_1 and h_2.

12. Why are microwave frequencies used for satellite communications?

13. Why does the Earth's atmosphere form a waveguide with the Earth's surface?

9.7 Homework Problems

1. Consider a balanced mixer with a 90° hybrid, like the one in Fig. 9.5. We will look at the reflected voltage waves from the diodes. Assume the sources of P_{RF} and P_{LO} are matched. If V is the RF phasor input voltage, the reflected voltages from the two diodes can be written as

$$V_{R1} = \rho V_1 = \rho V/\sqrt{2} \qquad \text{and} \qquad V_{R2} = \rho V_2 = -j\rho V/\sqrt{2}$$

where ρ is the reflection coefficient of the diodes. Write down the expressions for the two reflected waves, V_R^{LO} and V_R^{RF}, when they combine at the RF and LO input ports of the mixer. Is the input RF port well matched? Does the reflected RF signal affect the LO signal?

2. Consider the diode equivalent circuit from Fig. 3.9 and use the element values given in Problem 4 of Chapter 3. Input the element values for the biased diode (based on $I_0 = 60\,\mu$A) into SPICE. Make a plot of the response from 1 to 12 GHz. At what frequency is the diode inherently matched to $50\,\Omega$? Match the diode impedance at 5 GHz to a 50-Ohm coax RF feed using an appropriate impedance matching circuit.

3. Two antennas are placed on identical towers 100 km apart. What is the minimum height required for these towers in order to achieve line-of-sight communication?

4. Two airplanes are flying at an altitude of 10 km, with antennas mounted on each one. What is the maximum range (horizontal distance) between the airplanes for which line-of-sight communication can occur?

Chapter 10

Radar Fundamentals

Radar can be viewed as a type of communication link, where the transmitting and receiving antennas are located at the same coordinate in the case of a monostatic radar, or at coordinates close to each other in the case of a bi-static radar, as shown in Fig.radar1. The transmitter sends a wave, which eventually reflects off some object, called a "target" or "scatterer". The reflection is partly in the direction from which the wave came, and this portion is received by the transmitter. By analyzing this received signal, some conclusions can be made about the target.

Radar was invented for military purposes by the British in the Second World War and contributed greatly to the Allied forces victory. The word radar is an acronym for RAdio Detection And Ranging. Today, there are a number of military and commercial applications, some ground-based, and most air or space born. The first aircraft with a fully-operational radar was a Bristol Beaufighter. It was flown by Flight Officer Ashfield and achieved its first radar-assisted target hit on November 7, 1940. The range of the radar was three to four miles. Modern space-born (satellite) SAR radar used for mapping have ranges of thousands of miles with resolution comparable to that of a radar very close to the earth.

10.1 The Radar Equation

The basic principle of radar is as follows. The radar transmitter sends a wave with power P_T toward a target. At the target, the power density is $S(\theta,\phi) = P_T D(\theta,\phi)/(4\pi r^2)$, where r is the distance to the target, and $D(\theta,\phi)$ is the radar antenna directivity in the direction (θ,ϕ). The radar can either use the same antenna for transmit and receive (more common), or different antennas (sometimes for very high power radar). The target scatters the wave proportionally to a quantity called the radar scattering cross section, usually denoted by $\sigma(\theta,\phi)$, which is essentially the effective area of the target acting as a receiving antenna. When it reflects the wave, the target acts as a transmitting antenna with a directivity of $4\pi\sigma/\lambda^2$. Now the Friis formula can be applied one more time to obtain the power received by the radar receiver:

$$P_R = P_T \frac{D^2(\theta,\phi)\sigma^2(\theta,\phi)}{16\pi^2 r^4}.$$

Note that the power received at the radar is proportional to the inverse of the distance to the fourth power. This means that the signals that the radar receives are highly attenuated by the time they return to the radar receive antenna, therefore limiting the useful operating range. The range at which the signal is strong enough to be detected depends on several factors, most importantly:

- the transmitted power
- fraction of time during which the power is transmitted

- size of antenna - reflecting characteristics of the target(s)
- length of time the target is in the antenna beam during each scan
- number of search scans in which the target appears
- wavelength of the radio waves
- strength of background noise of clutter
- carrier frequency and atmospheric properties at that frequency.

In most cases it is not enough just to detect that a target is present, but also to find out where the target is and, if it is moving, how fast it is. The position of the target is determined by measuring the range (distance) and direction (angle).

The range can be determined by sending pulses with a certain repetition rate and measuring how long it takes for them to get back. This type of operation is referred to as pulsed radar. Since electromagnetic waves travel at the speed of light, the time difference gives directly twice the range. However, if there are more than one target, the pulses can overlap and become impossible to differentiate. In that case, a non-pulsed waveform is used, referred to as CW-FM ranging. To find the direction of a target, the radar antenna beam is continuously scanned over a range of angles (referred to as azimuth and elevation). In tracking radar, once a target is found, it can be tracked while the radar is scanning for other targets.

Common military radar applications include: missile guidance, strategic and blind bombing, surveillance, early warning, target identification etc. Common commercial radar applications are: law-enforcement, meteorology (wind and rain profiling), mapping, vehicular (anti-collision and parking), aircraft guidance and traffic control, storm avoidance, wind-shear warning, environmental monitoring, meteor detection, etc.

10.2 Doppler Radar

When the target is moving, the frequency of the scattered wave is shifted according to the Doppler effect. By measuring the Doppler shift, the speed of the target can be detected with a radar, Fig. 10.1b. A wave of frequency f is transmitted, and the wave reflected off a target moving at velocity v will have a frequency

$$f \pm f_D \qquad \text{where} \qquad f_D = \frac{2v}{\lambda} = \frac{2vf}{c}.$$

This frequency shift is positive or negative depending on the direction of motion of the target with respect to the radar: if the target is moving towards the radar, the frequency is shifted up, and if it is moving away, it is shifted down.

10.3 CW-FM Ranging Radar

In one type of ranging radar, the frequency of the transmitter is linearly frequency modulated (FM), as shown in Fig. 10.2. If f is transmitted, by the time this wave returns to the radar, the radar frequency is f', and the difference between f and f' is an indicator of the time it took the wave to travel to the target and back. Therefore, this difference frequency Δf, often referred to as the "beat frequency", is directly related to the distance from the target. Since the signal is transmitted continuously, this type of ranging is called CW-FM, which stands for Continuous Wave Frequency Modulated.

It is simple to see how this works for a stationary target. However, most targets move and have associated Doppler shifts. In order to compensate for Doppler shift, the frequency modulation can be changed to a triangle instead of a ramp, Fig. 10.2a, enabling measurement of the Doppler frequency

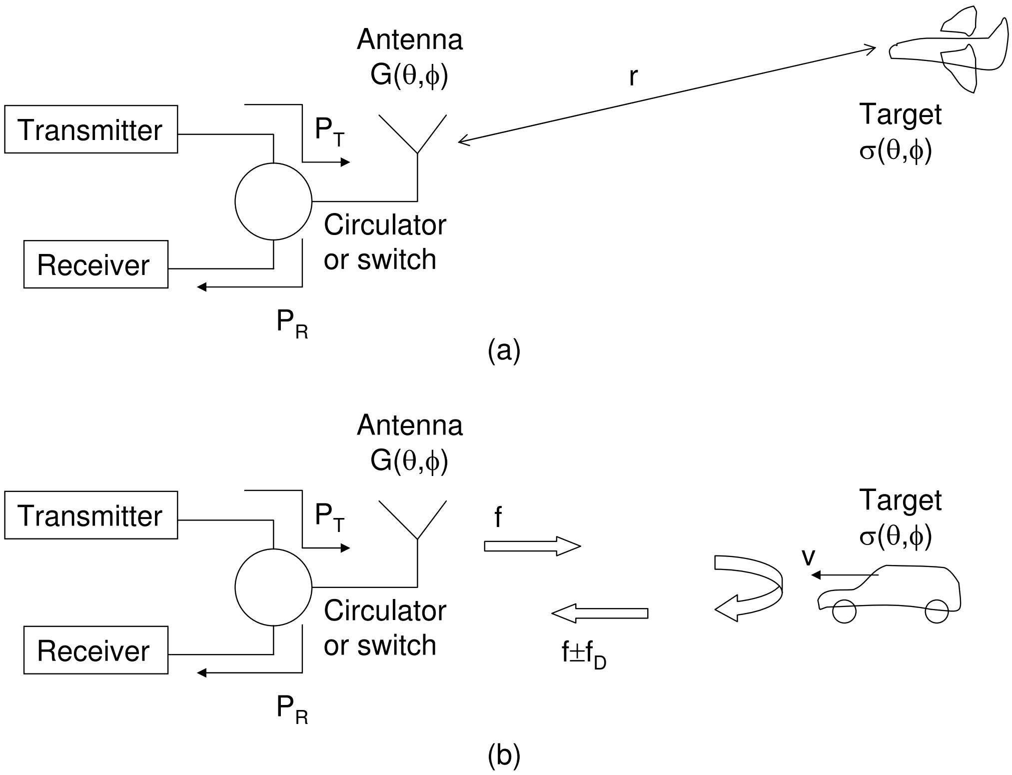


Figure 10.1: (a) Basic radar operation.(b) Basic Doppler radar schematic.

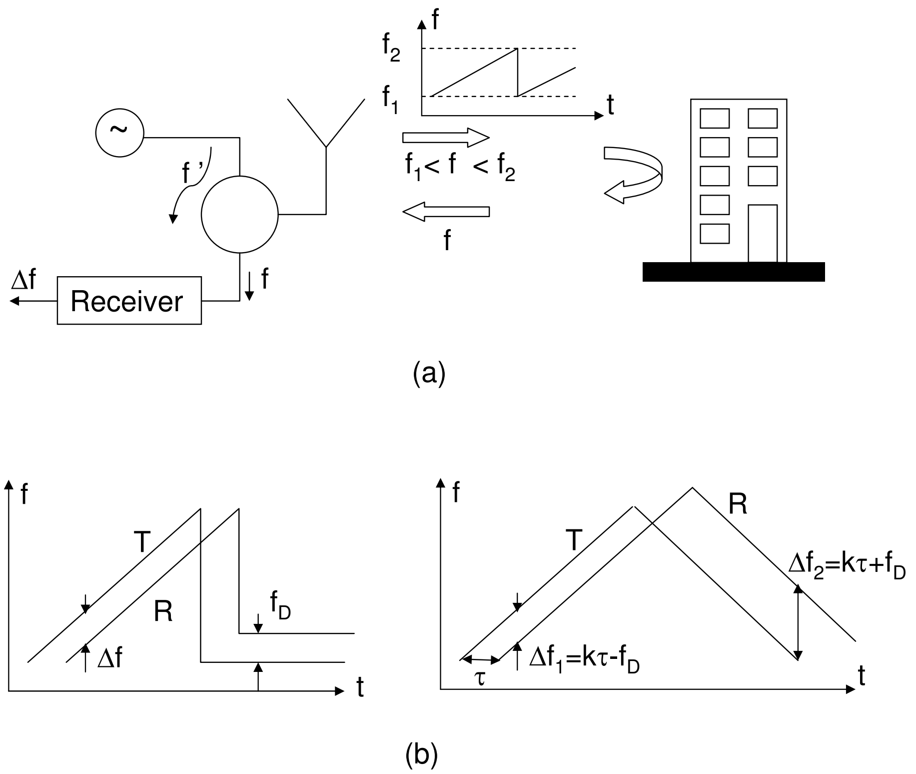


Figure 10.2: (a) Operational principle of CW-FM ranging radar. (b)Correcting for Doppler shift in CWFM radar: left – transmitting a fixed frequency for some time after the FM part, and right – triangular frequency modulation.

shift that can be taken into account in the post-processing. Alternatively, the frequency modulation can be changed to a triangle instead of a ramp, as on the righthand side of Fig. 10.2b, in which case a positive Doppler frequency for the rising frequency part will be a negative one for the part of the FM curve where the frequency is falling off. Therefore, the range can be determined by averaging the two measured frequency differences. Clearly, in all cases, the precision with which the range can be measured depends on the ability to accurately measure frequency.

10.4 Practice questions

1. Why are microwave frequencies used for radar?
2. What does the word “radar” stand for?
3. Explain the operational principle of pulsed radar for ranging.

4. Explain the operational principle of continuous-wave (CW) radar for ranging.
5. Explain the principle of Doppler radar.
6. What are the different factors that can prevent ideal operation of a radar?
7. What is ground clutter?
8. Explain how a CWFM radar works.
9. Explain how a monopulse direction finder works.
10. What are the different frequencies used for radar and why?

10.5 Homework Problems

1. Derive the radar equation if the radar has two antennas: one for receiving and another one for transmitting.

2. Assuming a 10-GHz police radar uses an antenna with a directivity of 20 dB (standard horn), and your car has a scattering cross section of 100 squared wavelengths, plot the received power as a function of target distance, for a transmitted power of 1 W. If the receiver sensitivity is 10 nW, how close to the radar would you need to slow down to avoid getting a speeding ticket?

3. Calculate expected Doppler shifts at 10 GHz for (a) a person running at 15 km/h, (b) a car speeding at 120 km/h, (c) an airplane flying at 600 km/h, assuming the direction of motion is directly towards a Doppler radar antenna.

4. For a CW-FM ranging-only radar operating at 10 GHz, calculate the required linear frequency modulation (ramp) bandwidth for a ramp period of $2\,\mu$ s, if it is required that a target range be detected when the target is 150 m away from the radar. Assume that the system bandwidth is at most 10%. Note that you have to choose one of the parameters (there is more than one solution given the constraints).

5. For a 10-GHz CW-FM ranging radar with a waveform as on the righthand side in Fig. 10.2b, with the same slope as in the previous problem, determine the difference in the beat frequency that would be recorded if the target were moving towards the transmitter at 100 km/hour.

Lab L1

The Artificial Transmission Line

An artificial transmission line is a cascaded connection of lumped circuit elements (or sometimes, of lumped elements periodically connected into an actual transmission line). It can have a variety of applications, including as a filter, or an impedance transformer. Here, it will be used here to provide a model for studying the behavior of voltages and currents on an actual transmission line, using component values scaled so that:

a) measurements can be made with ordinary low frequency laboratory equipment,

b) a wavelength is manageably small at the low operating frequency,

c) the length Δz of a single section is much less than a wavelength, and

d) impedance levels are such that connections to the laboratory instruments produce negligible disturbance (loading) to the circuit under test.

The artificial line to be used in this experiment is shown in Fig. L1.1. Each section of it consists of

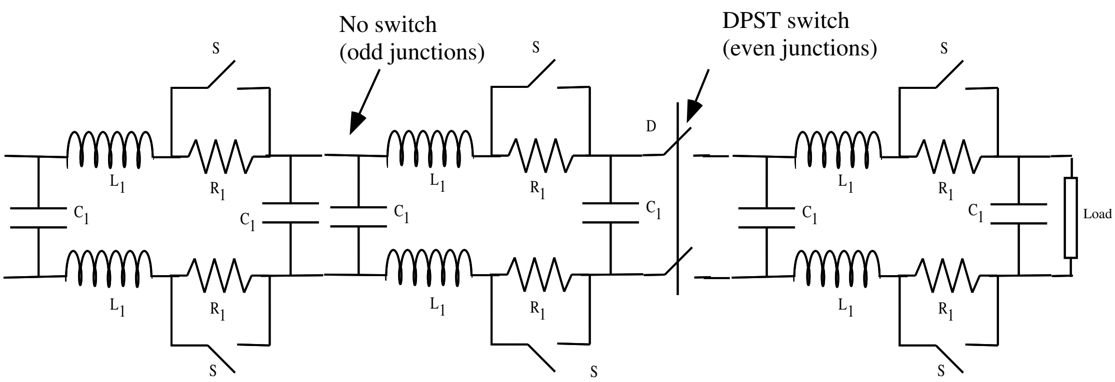


Figure L1.1: The artificial transmission line.

a rearrangement of the *L-C* section shown in Fig. 1.3 which models a short section of transmission line. The line in Fig. L1.1 is said to be *balanced*, because the series impedances and shunt admittances are evenly divided between its two conductors. That of Fig. 1.3 is said to be unbalanced. A coaxial line is an example of a continuous unbalanced transmission line, while "twin-lead" (300 Ω line used for television and FM radio antenna connections) is an example of a balanced line. Mathematically, these two types of line are identical, but in practice they behave differently. While unbalanced lines can provide better shielding of signals from outside interference, balanced lines minimize interaction with the ground and undesirable operation of antennas connected to them.

The series resistances R_1 (which represents losses in the metal of an actual line) can be switched out of the circuit (bypassed) using the single-pole single-throw (SPST) switches S. The line can be open-circuited after any section by opening one of the double-pole single-throw (DPST) switches D.

The inductance per unit length of the transmission line of which a length Δz is represented by one section of the artificial line is

$$L = \frac{L_\Delta}{\Delta z} = \frac{2L_1}{\Delta z};$$

the series resistance per unit length is

$$R = \frac{R_\Delta}{\Delta z} = \frac{2R_1}{\Delta z};$$

and the capacitance per unit length is

$$C = \frac{C_\Delta}{\Delta z} = \frac{2C_1}{\Delta z}$$

For the artificial line you will use in this experiment, $L_\Delta = 1.2\,\text{mH}$, $R_\Delta = 24\ \Omega$, and $C_\Delta = 20\,\text{nF}$.

Connect the experimental setup as shown in Fig. L1.2. The DPST switches are all set closed for this experiment.

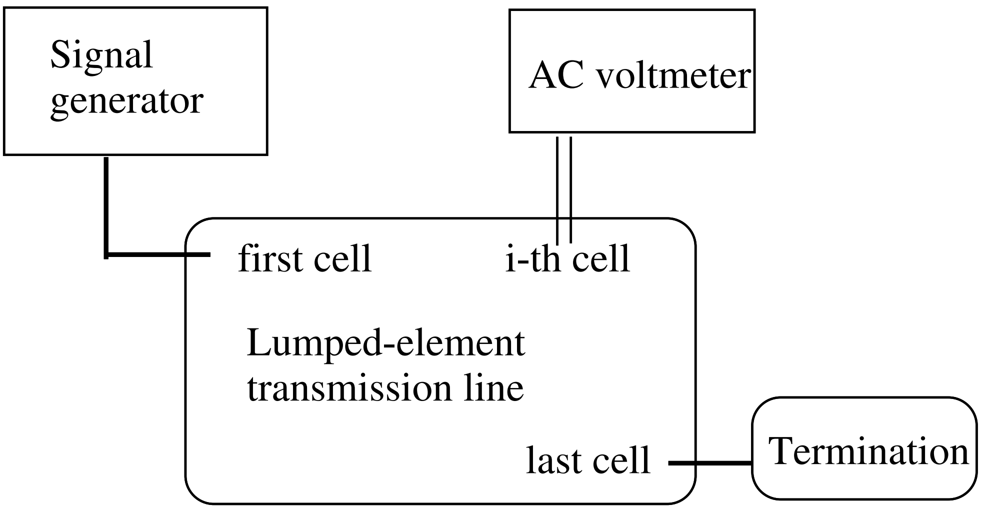


Figure L1.2: The setup for the artificial transmission line experiment.

Part I.

For this part, close all switches S (short the resistors). Set the frequency on the signal generator to 7960 Hz.

Q1: Suppose that each section of the artificial line represents a length $\Delta z = 5\,\text{cm}$ of an actual transmission line. What is the wavelength of a wave on this line? How many wavelengths long is this transmission line at $f = 7960\,\text{Hz}$? What is the value of the ratio $\Delta z/\lambda$? How do these answers change if each segment represents a length $\Delta z = 1\,\text{km}$ of actual transmission line?

Q2: What is the wavelength of the wave on an actual, air-filled transmission line at this frequency? How long would it have to be to be the same number of wavelengths long as the artificial line?

Terminate the line with a matched load for this lossless case. In order to choose the right load, you need to answer the following:

Q3: What is the characteristic impedance of this transmission line? Does it depend on the value of Δz selected for the actual transmission line it represents?

Using an AC multimeter, measure the line voltage and plot it as a function of distance along the line. The easiest way to do this is to start off by adjusting the amplitude of the signal generator to make the RMS voltage in the first section of the line equal to $1\,\text{V}$. In this way, the results you plot will be a so-called *normalized graph*, which means that the maximal value is equal to unity.

Q4: Why is the output voltage (the voltage at the load) smaller than the input voltage?

Q5: Was your line well matched? Explain.

Q6: Is the line really lossless? If you think not, where do you think the losses come from?

Part II.

Now open the SPST switches S to include the resistors in the artificial line.

Q7: What is the characteristic impedance of the line equal to now, assuming that the shunt conductance per section of the artificial line is $G_\Delta = 2G_1 = 0\,\text{S}$?

Terminate the line in the impedance you calculated above (you will now have to use both a decade resistor and a decade capacitor—should they be in series or in parallel?). Repeat the measurement of the line voltage versus distance and plot it on the graph in Part I.

Q8: Calculate the attenuation coefficient α from your measurement and from theory. How do they compare?

Q9: From your measurements, find the current along the line.

Part III.

We will now look at what happens when the line is not well matched. Set the switches for a lossless line, and terminate it with an impedance equal to $2Z_0$, where Z_0 is the characteristic impedance of the lossless line. Select the frequency of the signal generator so that the line is half a wavelength long.

Q10: What frequency is this? Measure the voltage along the line and plot it on a new graph.

Now select the frequency so that the line is a quarter of a wavelength long. Terminate the lossless line in a (1) short and (2) open circuit and plot the voltage versus position along the line.

Q11: For these two terminations, what would happen to the load voltage if the $\lambda/4$ transmission line was a perfect one?

Q12: How do the currents and voltages compare in the case of a $\lambda/2$ and $\lambda/4$ long transmission-line for an open circuit termination?

Q13: What standard circuit element does the quarter wavelength long line remind you of? Explain.

Lab L2

Basic Network Analyzer Operation

In this lab, you will learn how to use a modern network analyzer. You will measure S-parameters, reflection coefficients and impedances in the frequency domain. In another lab, you will learn how to use the time domain option of the network analyzer. There is a short user manual with the network analyzers, and the parts you will need to use are labeled. In this lab you will have to be a bit more independent, for several reasons. First, this is what you will need to be when you get a job and need to use a new instrument (no detailed instructions will ever be spelled out for you). Second, there are three different kinds of network analyzer in our lab, and although their operation is similar, there are enough differences in detail that you will need to figure them out for yourself.

At the end of the lab, using what you have learned about the network analyzer, you will examine the operation of the classic single-stub matching network.

Part I.

Before you start any measurements, you need to learn how to calibrate the network analyzer. As we saw in class, and you saw in the homework, calibration involves measuring some known standards. In the first part of the lab, you will do only one-port (reflection only) measurements, so you only need three standards. These are: a short circuit, an open circuit and a matched ($50\,\Omega$) load. In fact, the "short circuit" is really a small length of transmission line terminated in a short circuit, and the "open circuit" is really a small, known, capacitance. The precise information about these standards is stored in the network analyzer, and used to set the calibration constants in its internal calculations. We will use coaxial 3.5 mm (sometimes called SMA, informally) calibration standards. Calibration standards also come in sizes to fit other coaxial connectors (7 mm, 2.4 mm, N, GR874, etc.), and can even be made of microstrip, waveguide or any other transmission medium.

First read the following section on safety (to the instrument). This is a very expensive piece of equipment, donated by HP, which does not mean they fix it for us for free, so you need to be very careful. There are many ways to damage the instrument. Please use it according to the following rules:

- Do not bend cables (the dielectric inside is not very flexible, and the cables cost around $2,000).
- Put on the wrist strap (the input stage is very static sensitive) at all times when you touch the instrument.
- Do not touch the inside pin of the connectors. (The input impedance of microwave instruments is matched to the coax $50\,\Omega$ impedance. This is very different from the practically infinite impedance of most instruments you have used so far at lower frequencies, so it is much easier to damage the low input impedance microwave instruments. Why?)

- Handle the connectors carefully: do not force them to connect, align them properly, and use a torque wrench if necessary to tighten them. The torque wrench should not be pushed past the point at which it clicks once when tightening the connectors; most often less force is necessary (hand tight is okay).

In this part you will use only one of the ports of the network analyzer; ordinarily port 1 is used. Before you do the calibration, you have to give a few orders to the network analyzer: go to the STIMULUS block, and set the number of points to 201, the start frequency to 10 MHz and the stop frequency to 3 GHz. Use the default values for the other items.

Q1: How often in frequency does the network analyzer perform the measurement? How many times does it do the measurement at each frequency (explore the menus of the network analyzer to see if you can find the answer)?

Do a full S_{11} one-port calibration using the HP 3.5 mm calibration kit (short, open, matched load). Refer to the manual.

Q2: Where is the reference plane located after you have performed your calibration?

When you are done with the calibration, as a general rule it is useful to check how you did. You can do this by measuring your calibration standards (again, in the frequency domain).

Q3: Once you perform a calibration, what should the measurement of your standard open circuit look like? Sketch the magnitude and phase you would theoretically expect. What about the short and load? Sketch what you expect to see for them on a Smith chart.

Include plots of the magnitude and phase of S_{11}. What is its magnitude for the open circuit cal standard? What about for the short? The load? Are these plots correct? Why? Spend some time looking at the Smith chart displays of your standards to get a better feel for where things are on a Smith chart. Print out the Smith chart for the open circuit standard.

Part II.

Now measure the microstrip open circuit, and include a Smith chart plot. Use the brass 2 inch square jig. Release the thumb screws on the bottom, slide the dielectric substrate in, and then lightly tighten the thumb screws. The circuit is cut to be a bit narrower than the jig, so you need to push it as close as possible to the pin of the connector going into port 1 of the network analyzer. Do a swept frequency measurement, then tighten the thumb screws a bit more, until you get an unchanged measurement. The thumbscrews give the ground contact, and at microwave frequencies it is very important that it is a good one. Be sure not to overtighten the thumbscrews; the center pins of the SMA connectors should not be bent.

Q4: Why are the open circuit Smith chart plots from the coax open cal standard and the microstrip open different?

Part III.

Now measure the S_{11} of the capacitor and inductor using the General Radio (GR) component mount with adapters to SMA connectors. Include the plots of S_{11}. What affects the accuracy of these measurements?

The SMA to GR874 adapters fitted to the component mount have an effect on what you measure. Since calibration of the network analyzer has been done with SMA cal standards, your measurements on the component mount included the effect of these adapters. If you want to know what the capacitor and inductor look like without the adapters, you will first have to calibrate the network analyzer properly. Attach SMA-to-GR874 adapters "permanently" to the network analyzer cable, and perform a one-port calibration using the GR874 calibration standards to eliminate the effect of these adapters.

Q5: Where is the reference plane located now?

Now measure again the S_{11} of the capacitor and inductor using the General Radio (GR) component mount, subject to the new calibration. Include the plots of S_{11}.

Q6: What differences do you see between the measurements made using the different calibrations?

Part IV.

In this part of the experiment you will match a load to a transmission line using a short circuited single stub with a variable length. The other varied parameter is the distance from the stub to the load, and this is obtained with a variable length of coaxial line. You will verify the quality of the match with the network analyzer.

We will be doing matching at 500 MHz. Use a frequency range from 450 to 550 MHz and the GR874 calibration standards, and perform a one-port calibration. Note how, when the frequency range of the network analyzer is changed, calibration has to be performed all over again (why?). After you have done the calibration, look at the Smith chart display of S_{11} with the GR coaxial 100 Ω load connected at the reference plane. You can read the impedance directly from the top of the screen.

Q7: What is the value of the load impedance you read off the display? How does it compare to the DC value you measure with an ohmmeter connected to the load? What is the measured SWR of the load equal to?

Now place the connecting T and the shorted stub and adjustable line length in the system as shown in Fig. L2.1. Use the GR coaxial 100 Ω load at the right end of the adjustable length line. Since the

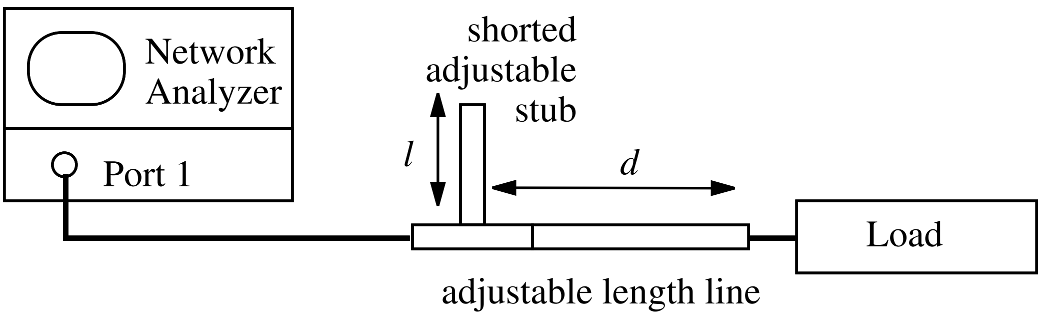


Figure L2.1: Coaxial stub matching: the quality of the match is measured with a network analyzer.

theory of a single stub shunt match is best done in terms of admittances, we will set the Smith chart display to read in terms of admittance rather than impedance (which is the default). Push the MKR key on the front panel of the network analyzer, then in turn the soft keys MARKER MODE MENU, SMITH MKR MENU, G + jB MKR, RETURN and RETURN. Now hit FORMAT → SMITH CHART to see the admittance chart display.

With the load placed at the right end of the line, observe the Smith chart display and adjust the stub and line lengths to obtain a match at 500 MHz. Proceed as follows. First, push the MKR key, then enter 500 MHz to set the marker frequency. Temporarily remove the tuning stub. Adjust the length of the adjustable line until the marker lies on the $g = 1$ (or $G = 20$ mS) circle of the admittance chart (you may need to add one of the fixed lengths of coax to achieve this). There are two places where this happens: one in the top half, and the other in the lower half of the Smith chart. Do the rest of this part of the lab for each of these cases.

Now reconnect the tuning stub. You should in theory be able to adjust the length of the tuning stub until the marker is at the center of the Smith chart, indicating a perfect match. Again, it may be necessary to insert a fixed length of coax line to get this match. As the stub length is varied, you should in principle see the marker position move along the $g = 1$ circle on the admittance chart. However, because of the nonzero lengths of line and other parasitic effects in the T-junction, this will not be exactly the case. You can compensate for this when the marker gets close to the center of the Smith chart by re-adjusting the adjustable length line until the marker is once again on the $g = 1$ circle, and then perform the final stub length adjustment.

For the lab write-up, do a theoretical stub match on a Smith chart to find the distance d from the load to the center of the T-junction, and the distance l from the center of the T-junction to the short-circuit that you need for a match. Use the measured value of load impedance found in **Q7**. Use a ruler to measure the distances you found from the experimentally achieved match. This procedure can give you a rough guess at how to measure the stub and line length with a ruler, knowing ahead of time what the lengths are.

Part V.

Next, place the $50\,\Omega$ load at the end of the line. Set the stub extension to a quarter wavelength at the operating frequency.

Q8: What kind of SWR do you expect to measure? Are you getting what you expected? If not, what could the reasons be (qualitatively)?

Use the same procedure as in the previous case to match the "matched" load. How different is the stub length you are estimating from a quarter wavelength?

Now connect one of the GR component mount jigs with a lumped element load inside. Match it by adjusting the stub and line lengths and observing the Smith chart as before (either of the two solutions is permissible). Measure the lengths as well as you can. Use a Smith chart and do a matching procedure backwards, i.e. start from the short, move a stub length l towards the generator, find the intersection of the constant reactance line with the $g = 1$ circle, move a line length d towards the load.

Q8: What is the impedance of the load? If you connect the load directly at the reference plane, do you measure approximately the same impedance?

Lab L3

Microwave Power Measurements

In this experiment, we will be doing measurements in the microwave frequency range, in X-band (8.2-12.4 GHz). You will be using the HP 8350 Sweeper and the HP 437B Power Meter with the HP 8481A Sensor. We will measure microwave power with a thermistor and a detector diode, and we will calibrate the two devices with the HP 437B Power Meter.

Part I.

First you will measure the DC properties of a thermistor mounted in a X-band waveguide (HP X487B). You will do this using a bridge, shown in Figure L3.1, as discussed in the lecture.

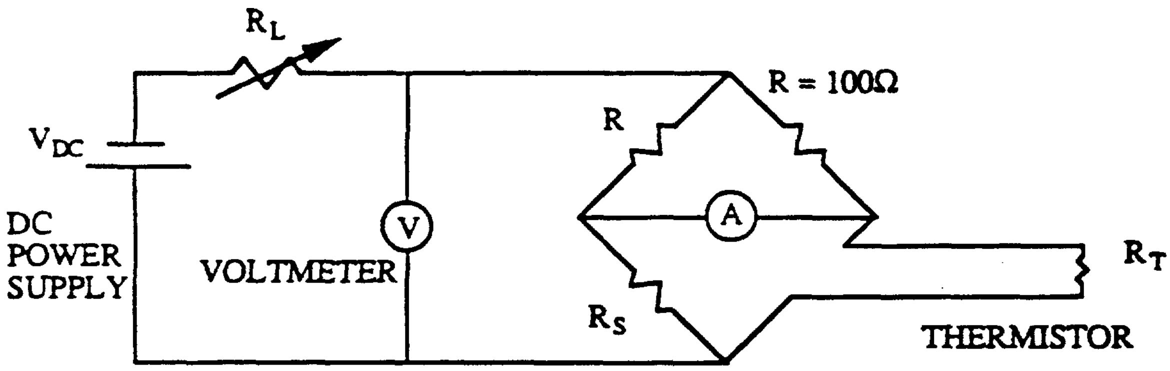


Figure L3.1: Bridge for measuring the DC characteristics of a thermistor.

Q1: If there is no current flowing through the ammeter, what does that tell you about the values of the resistors in the bridge? Explain.

Q2: For a known value of R_s in a balanced bridge, you measure a voltage V with the voltmeter. What is the power in the thermistor equal to?

By changing R_L, adjust the voltage measured by the voltmeter so that there is no current through the ammeter. Start with a value of $1000\,\Omega$ for R_S. After the current meter has settled (it takes a few seconds), record the voltage. Keep decreasing R_S by $100\,\Omega$ and repeat the measurement, always recording the resistance of the thermistor and the voltage. Make a table with columns R_T, V and P_{DC} (calculate from your answer to Q2). Plot the resistance of the thermistor (in Ω) versus power (in mW) on a linear scale. This is called the DC characteristic of the thermistor. It tells you how the resistance of the thermistor changes with different incident DC power levels.

Part II.

If we wish to use the thermistor for measuring microwave power, we have to know how the resistance changes with incident microwave power. The procedure used to determine this dependence is called *calibration.* To calibrate the thermistor, we will use the setup shown in Figure L3.2. Set the sweeper at

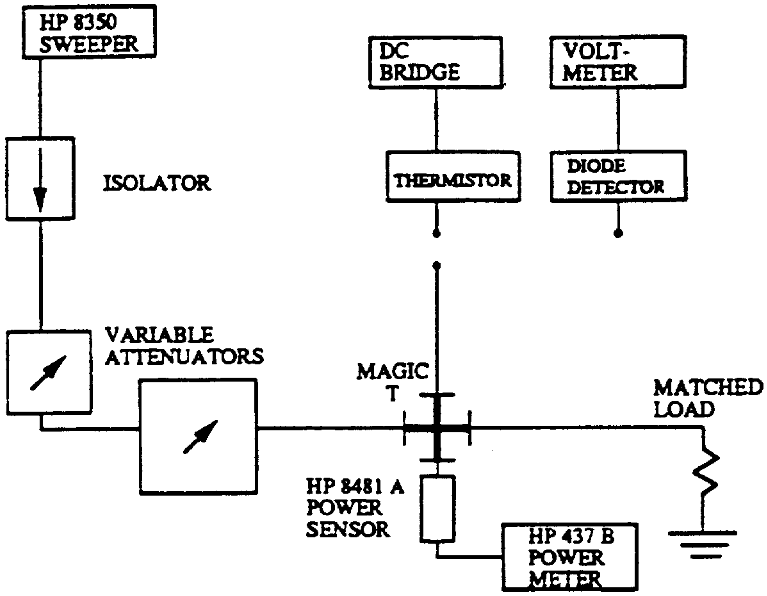


Figure L3.2: Setup for measuring microwave power with a waveguide mounted thermistor or diode detector.

9 GHz CW (continuous wave), and the power initially to 10 dBm. Pick a high power point on the DC curve of the thermistor, for example 200 Ω. The power from the sweeper goes through an isolator and two variable attenuators to a magic T.

An isolator is a two-port nonreciprocal device, which means it looks different from the two ends. It has a ferrite element inside that produces a magnetic field in one specific direction. Its purpose is ideally to prevent any power from going from the output port back through the input port, while it lets everything pass the other way.

A variable attenuator introduces different amounts of power attenuation to a wave passing through it. Both the isolator and attenuator you will use in this experiment are waveguide components, but they all have their coaxial counterparts.

The magic T is a special kind of so-called hybrid waveguide junction with four ports. If one of the ports is the input, and another one is terminated with a matched load, the power will split equally between the two remaining ports. We will use this fact to monitor the power going into the thermistor connected to one port, with the HP 437B power meter and sensor connected to the other.

A waveguide magic T is shown in Figure L3.3(a). If a TE_{10} mode is incident at port 1, there is an odd symmetry about guide 4, and the field lines of E_y are shown in Figure L3.3(b). Since the field lines of a TE_{10} mode in guide 4 would have to have even symmetry, we conclude that there can be no coupling between ports 1 and 4. On the other hand, coupling to parts 2 and 3 is in-phase and the power is split equally between them. If a TE_{10} mode is incident at port 4, port 1 will be decoupled, and the power

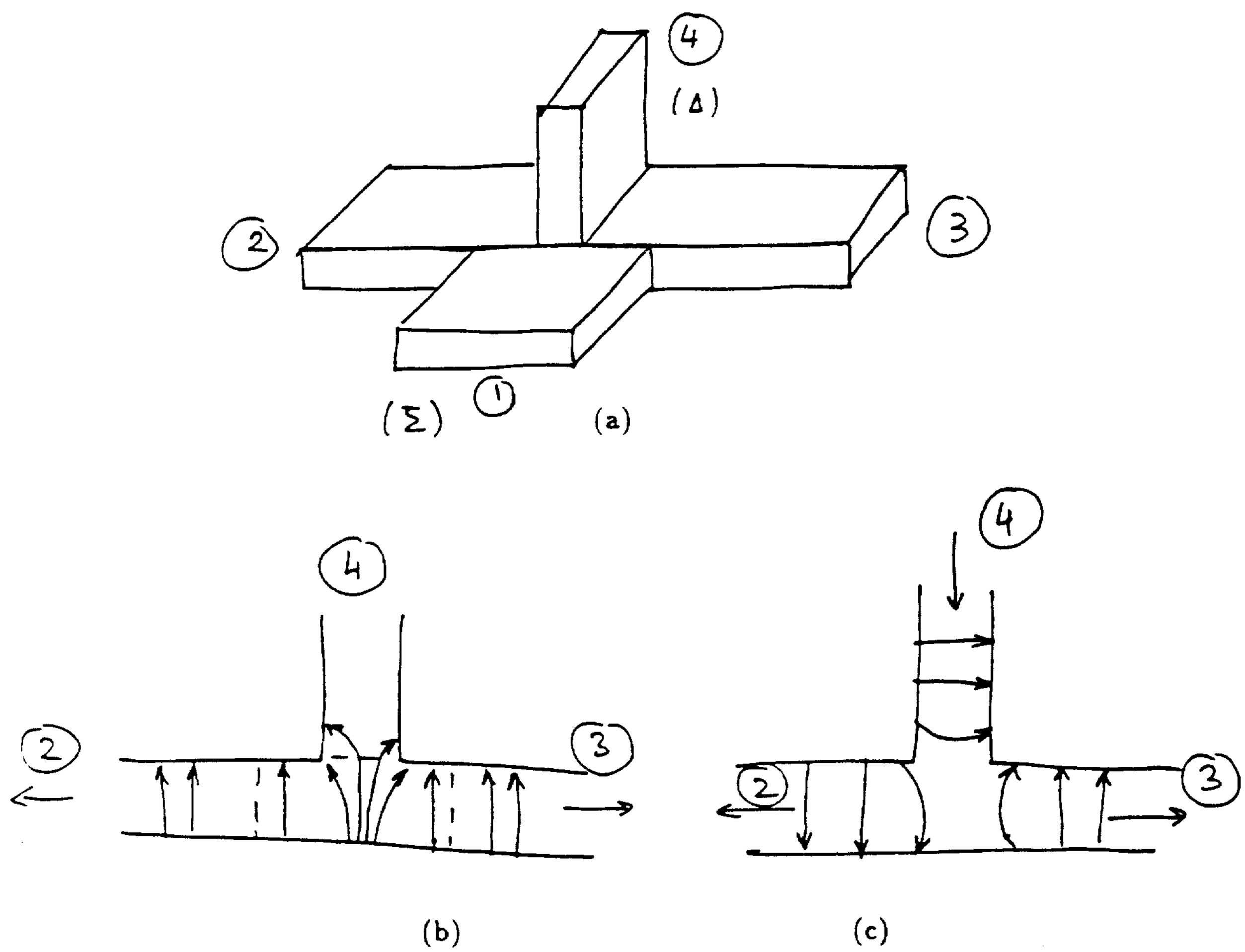


Figure L3.3: The magic T (a) and field lines for TM_{10} excitation on ports 1 (b) and 4 (c).

will split equally between 2 and 3, but with a 180° phase shift. The field lines for this case are sketched in Figure L3.3(c).

Connect the setup shown in Figure L3.2. Adjust the bridge resistors for a thermistor resistance of 200 Ω with no RF power input (RF power off on sweeper). Now turn the RF power on at about 10 dBm, and balance the bridge by changing R_L. Record the DC voltage and the power measured by the HP437B power meter.

Q3: In this measurement, you have both RF and DC power incident on the thermistor. From what you measured, how can you find the RF power alone?

Repeat this measurement with different RF power levels from the sweeper. Change the power from 1 to 10 mW with a 1 mW step. Fill in a table as shown in Table L3.1.

Plot, on a linear scale, the power measured by the power meter versus the power measured by the thermistor bridge. Include your graph.

Power from Power Meter (mW)	DC Voltage V Across Bridge (V)	Calculated RF Power (mW)

Table L3.1: Thermistor RF power measurements.

Part III.

In this part of the experiment, you will use a diode detector to measure power. You use the same setup as in the previous part, but connect the waveguide mounted diode instead of the thermistor. The diode output goes directly to a digital voltmeter. The diode is a square-law device, as we said in class, and the power detected will be proportional to the measured DC voltage. Change the RF power incident on the diode and fill in a table as the one shown in Table L3.2. Start at 9 dBm and go down in 3 dB

Power from Power Meter (mW)	DC Voltage Across Diode (V)	DC Diode Voltage in dBV

Table L3.2: Diode detector RF power measurements.

steps until you get no reading (depending on the power meter, this will be somewhere between about -35 dBm and -45 dBm). The sensor you are using with the HP 437B power meter is not very sensitive at low power levels. Probably the best way to measure the power delivered to the power meter is to first set the RF power on the sweeper for a normal sensor reading, and then decrease the power with the variable attenuator. Plot the DC voltage across the diode (in dBV) versus the incident power in dBm (this is equivalent to plotting the absolute numbers on a log-log scale). How closely does this plot follow a linear relationship? What is its slope?

Lab L4

Microwave Multiport Circuits

In this lab you will characterize several different multiport microstrip and coaxial components using a network analyzer. Some of these components do not have a counterpart at low frequencies, while others have analogues in the low-frequency region. The purpose of this lab is to learn how to perform multiport network analysis, how to diagnose a circuit functionality and how to judge the quality of your measurement.

Part I.

We begin by performing a two-port calibration of the network analyzer (refer to the analyzer instructions for details). Use a frequency range of 1 GHz to 3 GHz. This calibration consists of a one-port (or reflection) calibration for both ports, as well as a transmission calibration in which a "thru" standard is connected between ports 1 and 2 of the network analyzer. There is a third stage, called an isolation calibration, which can be omitted for purposes of this class. After you complete the calibration, check the calibration by observing all S-parameters when each of the calibration standards is connected. Save the calibration so that you can observe each device using both calibrated and uncalibrated measurements, while not having to repeat the calibration each time.

Part II.

In this part of the lab, you will measure the performance of some two-port, three-port and four-port microstrip circuits using the network analyzer. Make plots of the relevant S-parameters between 1 GHz and 3 GHz – it is up to you to choose which plots are relevant. (Including plots of all parameters in the lab can result in a reduced grade.)

Q1: Circuit #1 is a two-port circuit. At what frequency does the transmission coefficient have a dip? Explain why - how long is the open stub at that frequency? Measure the stub and, assuming the dielectric is $30\,\text{mils} = 30 \cdot 24.5\,\mu\text{m}$ thick, calculate the relative permittivity of the substrate. Use Eq.(1.42) in the notes. *Note:* This circuit was designed by a student to operate at 2 GHz, but the student did not know the permittivity of the substrate. Did he/she underestimate or overestimate the permittivity w.r.t. the real value.

Q2: What is the functionality of Circuit #2? Explain how the circuit works based on physical principles. Can you check if the circuit is lossless? Try calculating $|s_{11}|^2 + |s_{21}|^2$ – do you get close to unity? Remember that the s-parameters are defined w.r.t. voltage, i.e. the value in dB is $20 \log s$. Explain any errors in your measurement.

Q3: Circuit #3 is a low-pass filter. Draw the lumped-element low-frequency analogue of the circuit. How would you find the element values? Compare uncalibrated and calibrated measurements when answering the following questions: (1) What is the corner frequency? (2) How large is the insertion loss in the pass-band? (3) How large is the stop-band attenuation? (4) At what frequency does the stop band start?

Q4: What is the function of Circuit #4? Compare with Circuit #3 in terms of (1) – (4) from the previous question.

Q5: What is the function of Circuit #5? Draw the lumped-element low-frequency analogue of the circuit. What can you comment on the measured operating frequency versus what was probably designed? (i.e. what was probably not taken into account in the design, and why?)

Q6: What is the function of Circuit #6? How many measurements do you need to perform to characterize this circuit? How symmetric is this circuit? Is it lossless? Is it matched? Using an ohmmeter and a coaxial BNC to SMA (3.5 mm) adaptor, measure the impedances of the loads connected to ports 2 and 3 (for port numbers refer to circuit labels). What impedance would be seen at port 1 with matched loads connected to ports 2 and 3 if the matching circuit were absent? Make a plot of S_{11} from 1 GHz to 3 GHz. What frequency is the matching circuit designed for? How big is the -10-dB (2:1 VSWR) bandwidth of the matching circuit? *Note:* the -10-dB bandwidth is defined as the difference between two frequency points on a response curve that are 10 dB below the maximum response amplitude (or above the minimum, as appropriate).

Q7: Remove the matched loads from ports 2 and 3, and connect the test port cables from the network analyzer to them. Place a matched load at port 1 of the circuit. Instead of splitting the power supplied to port 1 between ports 2 and 3, while maintaining a match at port 1, the circuit is now connected as a power combiner. Look at the magnitudes of S-parameters S_{22}, S_{23} and S_{33} on the display. What should these S-parameters be if the circuit is to function well as a power combiner? Does this circuit perform this function well?

Q8: Connect ports 1 and 2 of the network analyzer to ports 2 and 3 of circuit #3, and terminate port 1 of this circuit with SMA coaxial matched loads. Measure the isolation (S_{23}) between ports 2 and 3, and the input match of each (S_{22} and S_{33}). Does the circuit work well in reverse, i. e., can it efficiently combine the power input at ports 2 and 3 for output at port 1?

Part III.

Use the network analyzer to characterize Circuit #7, the four-port microstrip hybrid directional coupler (as studied in your prelab homework). Measure the S-parameters for this circuit between 1 GHz and 3 GHz. Make use of all the symmetries you can find to reduce the number of plots. Include plots of the amplitudes in your notebook, and look at the phases on the network analyzer. Compare un-calibrated and calibrated measurements.

Q9: There are only two ports on the network analyzer. What do you do with the rest of the ports in the circuit while you are doing the measurement? Trick question: if you had no matched loads, could you still characterize the circuit, and if *yes*, what would you do?

Q10: At the operating frequency, how large is the through, coupled, reflected and isolated-port power assuming 1mW (0 dBm) incident at port 1?

Q11: The directional coupler was designed to operate at 2 GHz, but the student that designed it did not learn about different microstrip discontinuities. Explain what effect was not taken into account. How could you verify your hypothesis?

Q12: Sketch the circuit and label the ports as they are labeled on the circuit. What are the relative phases at all the ports? Does this agree with the analysis from your prelab?

Q13: How large is the 3-dB bandwidth of the coupler? What is the -10-dB bandwidth equal to?

Part IV.

In this part of the lab, you will characterize a few coaxial components. Set the network analyzer for a frequency range of 300 kHz to 3 GHz. Perform a full two-port calibration. Circuit #8 is a coaxial directional coupler. Circuits #9 and #10 are unspecified component whose functions you will determine using the network analyzer.

Q14: What range of frequencies is circuit #8 designed for? What are its coupling factor and isolation?

Q15: What does circuit #9 do, and why do you think so?

Q16: What does circuit #10 do, and why do you think so?

Lab L5

TDR Using the Network Analyzer

In this experiment, you will be using the Network Analyzer to verify the time domain waveforms from your homework. This instrument has a time domain option in which it sends a short pulse instead of a CW (continuous wave) signal to the test port. The amplitude and phase of the reflected signal off of the load are displayed. The main difference between the network analyzer and a real time domain reflectometer is that a network analyzer performs the measurement in frequency domain (looks at the amplitude and phase of the reflected signal as a function of frequency), and then uses its internal computer to perform an inverse Fourier transform to give us the time domain response, similar to what we did in class in the example of an inductive load (we went from the Laplace domain, which is frequency, to time domain).

Q1: What should the voltage and the reflection coefficient look like on the screen for a shorted network analyzer port and an open port?

To get to the time domain option of the HP 8702 Network analyzer, use the following procedure:

(0) USE THE GROUNDING STRAP!

(1) Hit the NORMAL OPERATION screen button.

(2) Choose the SYSTEM key in the INSTRUMENT STATE group.

(3) Choose TRANSFORM MENU

(4) Choose LOW PASS STEP softkey

(5) Press XFORM ON

(6) Do CAL using device type 2-PORT ELECTRICAL/ELECTRICAL

(7) In the CAL MENU, pick FULL 2-PORT

(8) Perform the (one-port) calibration

(9) Go to LOCAL MENU

(10) SYSTEM CONTROLLER, HP-DIAG ON

(11) Go to FORMAT MENU, REAL

Similar procedures apply for the other network analyzers.

Part I.

In this part of the experiment, you will measure the electrical length (time delay divided by an assumed propagation velocity) of a GR cable. What velocity does the network analyzer assume for this purpose? You can both short and open the cable at the other end. You should get the same result, so use this as a check. Sketch the time waveforms from the display. Make sure to label your graphs (units and scales).

Q2: From your sketch, what is the physical length of the cable, if it is filled with polyethylene ($\epsilon_r = 2.26$)? Could you have just measured it with a yardstick?

Part II.

In this part of the experiment, you will measure the length of an adjustable stub. This is just a piece of air coax terminated with a short that can slide along the coax to shrink it or stretch it.

Q3: What are the shortest and longest lengths of the stub? Sketch the corresponding waveforms you measured, and label the time and amplitude axes. Is the stub termination perfect?

Q4: Zoom in on the edge of the steps. Sketch what they look like. Where do you think the "dirt" comes from?

Part III.

Now you will find the reflection coefficients by measuring the response corresponding to different terminations. Sketch the screen display for each of the terminations, labeling the amplitude and time axes. Repeat this for a short, open, $100\,\Omega$, and $50\,\Omega$ loads.

Q5: Calculate the reflection coefficient from a $50\,\Omega$ transmission line terminated in a $100\,\Omega$ resistor. Does it agree with your measurement?

Part IV.

In this part you will convince yourself as to the importance of a good connection between lines. Connect a cable to the port you are using. Expand the display on the time range near the end of the cable (i. e., where the reflected wave appears). Now take the matched load and *slowly* connect it to the open end of the cable. Sketch the waveforms that you see in the process. What do you see when the connector is loose?

Part V.

Next, we will look at different lumped element terminations, like the ones in your homework. Use the provided metal jig (GR component mount) for connecting the components to the transmission line.

Q6: Calculate the reflection coefficient from a $50\,\Omega$ transmission line terminated in a $20\,\Omega$ resistor.

(A) Use three different resistors as terminations: one equal to $50\,\Omega$, the second one smaller, and the third one larger than $50\,\Omega$. Sketch the waveforms, labeling them properly. Do the measurements agree with your calculations in **Q5** and **Q6**? Do these loads behave like the GR loads, and if not, what might account for it? How closely do the measured reflection coefficients match the expected ones?

(B) Now connect the capacitor to the jig. Sketch the display and label it carefully.

Q7: From the display, and your homework results, calculate the capacitance of the capacitor. The characteristic impedance of the coax is $50\,\Omega$.

(C) Wind three inductors of your own on a piece of straw or on a pen (these may already be made for you in the lab). Wind them as a tight solenoid with different numbers of windings.

Q8: What is the inductance of each solenoid proportional to? Why? (*Hint:* Remember the flux definition of the inductance $L = \Phi/I$.)

(D) Connect your inductors one by one to the jig, and sketch the resulting displays.

Q9: Calculate the inductances for your inductors from the results of your homework and the measured waveforms. Do they agree qualitatively with your answer to the homework?

Part VI.

Finally, we will examine the transmitted wave of a two-port circuit in the time domain. Connect the GR component mount with an inductor connected inside it to the side arm of a coaxial T-connector. Connect the remaining ports of the "tee" to ports 1 and 2 of the network analyzer. Display S_{11} and S_{21} in the time domain, sketch or print it out and label it carefully.

Q10: Use the measured plot of S_{11} to compute the value of the inductor as in Part V above. What should the plot of S_{21} look like theoretically? Does your measured plot agree with the theory?

Lab L6

The Gunn-diode Waveguide-mounted Oscillator; Spectrum Analysis

In this lab you will study the use of a Gunn diode as an oscillator (to generate microwave signals) and as an amplifier. Its configuration as an oscillator is shown in Figure L6.1. The Gunn diode cathode

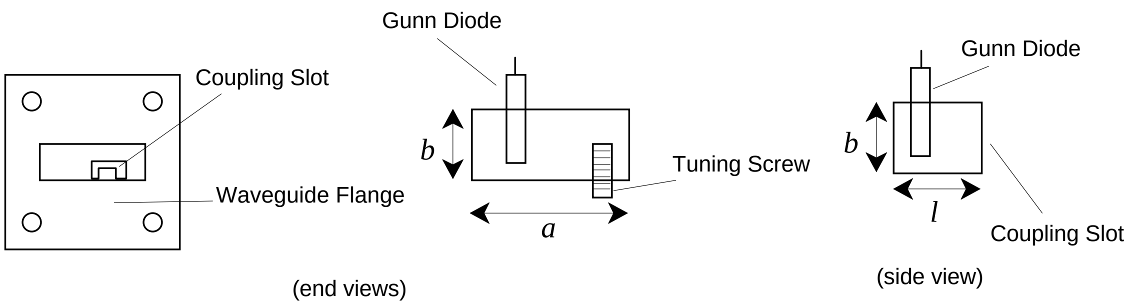


Figure L6.1: A waveguide cavity Gunn diode oscillator.

is secured by a threaded metal cap into the metal wall of a short length $l = 0.5$ in. of rectangular waveguide, whose cross-sectional dimensions are a little smaller ($a = 0.831$ in., $b = 0.266$ in.) than those of WR-90 waveguide. The DC to RF conversion efficiency of Gunn diodes is low no matter what amount of RF power is generated, so they get quite hot. Thermal conduction to the metal waveguide wall is improved with thermally conductive paste, so that the waveguide metal acts as a heat sink. The anode of the diode is connected to a probe that protrudes into a cavity formed by closing off two ends of the piece of waveguide.

There is a small U-shaped slot cut into one of these end walls in order to couple energy out of the cavity via WR-90 waveguide. A tuning screw that penetrates into the waveguide cavity will allow for adjustment of the resonant frequency of the cavity, and therefore of the frequency of the oscillator.

Inside the cylindrical housing which contains the Gunn diode, there is a coaxial choke—a section of air-filled coaxial transmission line $0.235''$ long—put in place between the diode package and the outer wall of the housing (which is at "ground" potential). Refer to Fig. L6.2.

Q1: What do you think is the purpose of the coaxial choke?

The output of the oscillator is coupled through a section of WR-90 waveguide (an isolator, to be

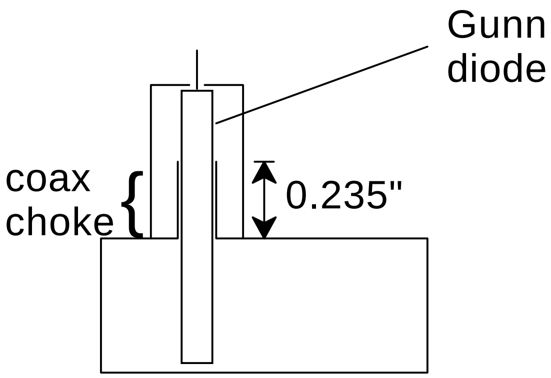


Figure L6.2: Waveguide Gunn diode housing showing coax choke.

specific) to a spectrum analyzer, on which you can observe whether, and at what frequencies, the Gunn diode is oscillating. Set the start frequency of the spectrum analyzer to 8 GHz, and the stop frequency to 22 GHz.

Part I.

Measure the DC *I*-*V* curve of the Gunn diode (CAUTION: When applying DC voltage to the Gunn diode, make sure the voltage of the DC power supply is initially turned all the way down first. When increasing the bias voltage, do so gently so as not to damage the Gunn diode with transient voltage spikes). Go up to 10 V (there is a protection circuit which includes a 15 V Zener diode placed across the bias voltage to protect the Gunn diode from accidental overvoltages). Notice that the diode gets very hot when DC bias voltage is applied. Include the *I*-*V* plot in your notebook. Mark the points on the DC curve for which you see oscillations on the spectrum analyzer.

Q2: Plot the DC power dissipated in the diode vs. the DC bias voltage.

Q3: Are you able to get good readings of the DC current through the diode at all bias voltages? What is happening on the spectrum analyzer display when unusual DC current readings are encountered?

Part II.

Set the bias voltage to one at which you have a stable, single-frequency oscillation (around 7-8 V). Note the value of the oscillation frequency. With a small screwdriver, adjust the tuning screw, and observe the effect on the resonant frequency.

Q4: What range of oscillation frequencies can be produced by adjusting the tuning screw?

Adjust the tuning screw to obtain the original oscillation frequency.

Part III.

Set the bias voltage to the lowest value for which you have a stable oscillation; you should see more than one frequency on the spectrum analyzer. This should occur at around 4 V.

Q5: What are the oscillation frequencies? Does any of these correspond to the length of the resonator? Record the relative amplitudes of all frequencies you observe. What kind of a waveform do you think the diode is generating? Is the quasi-harmonic approximation valid for these operating conditions?

Q6: Compare the power in the fundamental frequency to the DC power delivered to the diode. What is the efficiency of this device under these operating conditions?

Q7: What do you think is responsible for the presence of the other frequencies in the output of the oscillator?

Set the start and stop frequencies of the spectrum analyzer to frequencies 500 MHz below and above the oscillator frequency.

Q8: Vary the DC bias voltage from 0 to 10 V as in **Q2**, and record the value of the fundamental (lowest) oscillation frequency f_{osc} vs. V. Does f_{osc} depend strongly on the bias voltage?

Reset (increase) the DC bias to obtain a stable, single-frequency oscillation. This should occur at around 7-8 V.

Q9: Repeat **Q6** for this operating point.

Part IV.

In this part, you will modulate the oscillator signal by varying the bias voltage. Set the spectrum analyzer to display a 1 MHz span of frequencies centered around the oscillator frequency. Connect a signal generator, set to produce a 1 kHz square wave, to the terminals of the Gunn diode, using a large capacitor in series to prevent the DC bias from damaging the signal generator. Increase the amplitude of the 1 kHz signal until you observe a change in the spectrum analyzer display.

Q10: What do you observe? Narrow the frequency scale and look at the effect of bias modulation on the spectrum of the oscillator signal. Repeat this observation using a 1 kHz sine wave. What differences do you see?

Q11: Repeat **Q10** using modulation frequencies of 10 kHz and 100 kHz. Report and comment on your observations.

Part V.

In this part, you will attempt what is known as *injection locking* of the oscillator. Connect a 10 dB directional coupler between the Gunn diode resonator and the HP 8350 sweep oscillator as shown in Fig. L6.3. Set the bias on the Gunn oscillator to a low level as in Part III. Set the power level of the sweep oscillator to -20 dBm, and set it to CW operation. Manually tune the sweep oscillator frequency close to and past the frequency of the Gunn oscillator.

Q12: What do you observe? Switch the sweep oscillator signal on and off. What happens to the fundamental oscillation? What happens to the oscillations at other frequencies? Reduce the frequency span to examine the neighborhood of the fundamental oscillator frequency in detail, and observe the effect of injection locking on the purity of the signal.

You can actually "pull" the oscillator frequency by adjusting the sweeper's frequency. How much you can pull it is called the injection-locking bandwidth of the oscillator. Move the frequency of the sweep oscillator up and down as far as you can while maintaining a single oscillation line on the spectrum analyzer. The injection-locking bandwidth is the difference between the highest and lowest of these frequencies.

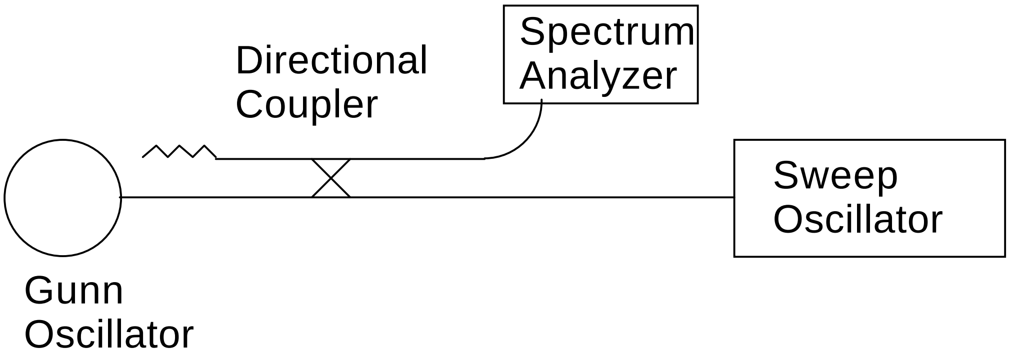


Figure L6.3: Injection locking of a waveguide Gunn diode oscillator.

Q13: What is the injection-locking bandwidth? Repeat this observation with sweeper power levels of -15, -10, -5 and 0 dBm. What is the rough functional dependence of injection-locking bandwidth vs. injected power?

Q14: Reset the bias voltage to a higher level to obtain a stable, single-frequency oscillation as in Part IV. Repeat **Q12** and **Q13** for this case. What differences, if any, do you see?

Lab L7

Antenna Measurements

In this lab you will study antennas from both the circuit and the electromagnetic fields point of view. You will measure their E and H-plane radiation patterns in the anechoic chamber, and on the bench with a parabolic reflector to enhance the gain of the antenna. Since we have only one anechoic chamber, one group will start on it and then do the reflector and network analyzer measurements in the second half of the lab; the other group will start with the bench measurements and then perform the antenna measurements in the second half of the lab.

In a horn antenna, the mouth of the horn can be viewed as a patch of a uniform plane wave which radiates. Since an electric field can be replaced by equivalent currents, you can imagine the mouth of the horn as a patch of uniform surface current. Horns are built by flaring waveguide ends, so they often have apertures of rectangular or circular shape.

Q1: How would you make a horn with a high gain?

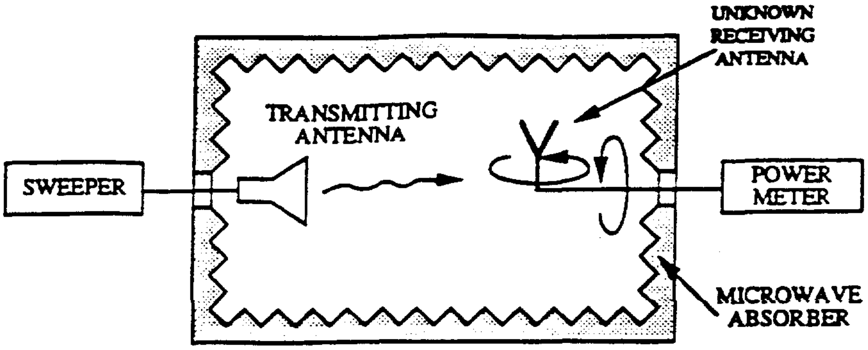
Q2: Define the E and H-planes for a rectangular horn antenna such as the one shown in Fig. 7.2(d) of the lecture notes. Sketch them in your report.

Part I.

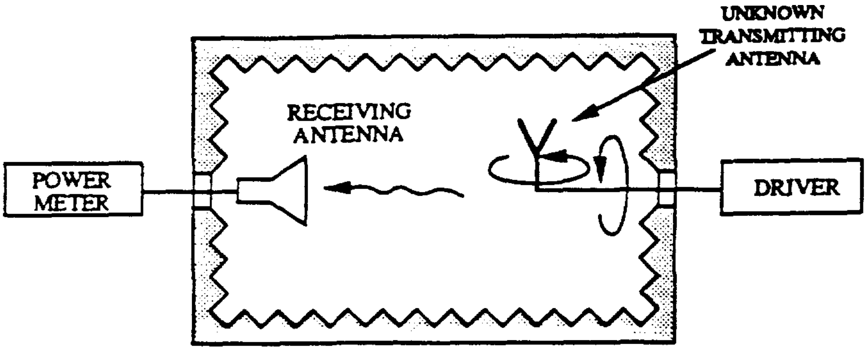
At microwave frequencies, antenna measurements are usually done in a special room called an *anechoic chamber*, Fig. L7.1(a,b). The purpose of the anechoic chamber is to absorb any reflected waves that could enter the receiving antenna and cause wrong power measurements. The walls, floor and ceiling are lined with microwave absorber, which is an array of cones made of a composite material that absorbs waves at microwave frequencies. Because of the shape of the absorber, the small portion that is reflected is scattered in all directions (*diffuse reflection*) rather than reflected all in one direction (this is called *specular reflection*). Typically, a wave reflected off of a piece of this material is 20 dB or more smaller in power than the incident wave. Sometimes the floor of the chamber is a sheet of metal to simulate the effect of the surface of the Earth. The two antennas need to be far enough away from each other to ensure far-field measurements.

Q3: The chamber in the lab is 4 m long wall to wall, and the two antennas are about 3 m apart. The horn at one end is a wideband horn used from 2 to 18 GHz. How large of a square-shaped aperture antenna could be properly measured in this chamber at 3 GHz, assuming the effective area to be the same as the geometric area?

Q4: We will use the HP 437B power meter with the HP 8141 sensor for power measurements. How much power would you need to transmit at 3 GHz in order to achieve adequate power levels at the power meter? Assume that the gain of the wideband horn is 9 dB, and that the effective area of the antenna under test is that determined in **Q3**.



(a)



(b)

Figure L7.1: Antenna pattern measurement for a receiving (a) and transmitting (b) antenna in an anechoic chamber.

In this part of the lab you will measure the E and H-plane patterns of a microwave horn antenna at 10 GHz. This is the center of the X-band. There is a rotation stage on the cart in the anechoic chamber. The motion of the rotation stage is controlled by a stepper motor driven from the PC. Antenna radiation patterns can be measured in transmitting or receiving mode; you will use the receiving mode, with an incident wave provided by a standard-gain horn antenna at the back of the chamber, fed by a sweep oscillator. The antenna under test is mounted on the rotation stage in either E-plane or H-plane orientation, and connected to the power meter.

Choose your resolution and measure the E and H-plane patterns for incident waves of both vertical and horizontal polarization (four patterns in all) between -90° and +90°. The results show up on the screen in one of three formats: linear plot, log plot and polar plot. You can toggle between them using the arrow keys. These plots are all normalized to the peak received power. The actual received powers appear on the display of the power meter: make a note of the maximum value of this power for each antenna. The data is saved in a file that you can save on diskette if you wish. You can do a print screen, or use the file with your favorite plotting program later. Include plots in your notebook.

Q5: How do you need to turn the transmitting and receiving antennas to measure the E-plane, and how to measure the H-plane? Sketch this in your notebook as clearly as you can indicating the direction of the electric field vector on both antennas.

Q6: How wide are the 3 dB beamwidths of the patterns?

Q7: Calculate approximately the directivity of the waveguide horn from its E and H-plane co-polarized patterns using (7.17).

Q8: Which of your measurements measure the cross-polarized patterns? What is the cross-polarization ratio equal to at the co-polarized pattern peak? (The cross-polarization ratio is defined as the ratio of the co-polar power to the cross-polarized power, and is usually quoted in dB.)

Repeat **Q5** to **Q8** for an "antenna" consisting of a straight length of WR-90 waveguide with no taper. How much higher or lower is the gain of this antenna than that of the waveguide horn?

Part II.

In this portion of the lab, you will look again at the reflection coefficient of the horn antenna on the HP/Agilent 8510 or Agilent 8719 network analyzer. This time, an SMA-to-waveguide transition is connected to the horn, so it may have a different reflection coefficient than the horn itself. The TA or professor will help you use the network analyzer, since the antenna is designed for frequencies above 3 GHz, and thus we cannot use the HP 8702 machines.

Q9: Plot the S_{11} coefficient for the waveguide horn from $f = 2$ GHz to 16 GHz, and include the plot in your notebook. What do you think the design frequency (frequencies) of this antenna is (are)? What is the bandwidth of the antenna? How well is it matched at the design frequency (frequencies)? In the same way, measure S_{11} for the length of WR-90 waveguide with no taper, connected via an SMA-to-waveguide transition to the network analyzer. Comment on the differences you observe.

Q10: Repeat the above for several antennas available in the lab: a linearly-polarized patch antenna, a dual-polarized patch, a patch with a parasitic stacked radiator, a Vivaldi tapered slot antenna and others.

Part III.

In this part of the experiment, we will examine the effects of changing the distance and relative polarizations of two antennas. On the bench, connect the horn antenna under test through an isolator to the output of an HP 8350B sweep oscillator set to CW operation at 10 GHz. At the other end of the bench, a distance of 1 to 2 m away, place a second (receiving) horn antenna connected to a power meter. Set the sweep oscillator to a power level of 10 dBm, and measure the received power in the second horn.

Q11: If the horns are identical, and stray reflections and power sources in the room can be neglected, what is the gain of each of the horns?

Move the two antennas as far apart as you can on the lab bench, and set the transmitter power high enough to put the received signal comfortably above the noise level. Measure the distance between the horn apertures, and record this together with the received power. Now move the antennas closer together in increments of 20 cm, and record the received power at each new separation distance until the antennas are touching.

Q12: When the antennas are touching, if the horns are well matched to the waveguides, you can assume $P_{tr} = P_r$, which serves to determine the value of P_{tr}. Now plot the received power vs. distance between the antennas. Do your results obey the Friis transmission formula? At what distance are the antennas no longer in each other's far field?

Now set the distance between the antennas at a value such that far field conditions are valid. Support the isolator of the transmitting antenna with the block of wood that has polarization angle markings on it. Record the received power. Now rotate the polarization of the transmitted signal, 10° at a time, recording the power of the received signal at each angle.

Q13: Plot the received power vs. angle of polarization rotation. Can you find a simple mathematical law that correctly fits the dependence you see?

Next, point the receive and transmit antennas directly at each other, and move the receive antenna until maximum power is measured. Now move it away from this position first laterally then vertically, keeping its orientation otherwise constant. To achieve vertical offset, it is most convenient if you rotate both the transmitting antenna and the receiving antenna by 90° and offset the receiving antenna horizontally. Note the distances from the axis at which the received power is 3 dB and 10 dB lower than its maximum value.

Q14: Calculate the 3-dB and 10-dB beamwidths in the E-plane and the H-plane from these measurements. How do your computations compare with those of **Q6** for the 3-dB beamwidths?

Part IV.

Finally, we will examine the effect of concentrating the radiation from the horn more into a single direction through the use of a parabolic reflector. Since radiated fields emerge to some extent in all directions from any antenna, including a horn, we can redirect much of that side radiation into a main direction by bouncing it off a surface such that all reflected waves propagate in the same direction.

We will now consider the effect of bouncing the radiation from the horn off a reflecting surface. A reflector in the shape of a parabolic cylinder will redirect all rays coming from an antenna in a single plane perpendicular to the cylinder axis into the same direction if the antenna is placed at the focus of the parabola, as shown in homework problem 4 of chapter 7.

Place the horn antenna under test so that it points toward the parabolic reflector, but pointed slightly upwards so that the reflected rays will not bounce back towards the horn itself (if they do, this is called feed blockage, and negates much of the improvement in gain that we hope to get from the reflector). Note that the focal distance p of this reflector is around 23 cm. Place the receiving horn antenna at least 2 m from the reflector surface, in a direction you think likely to be one of maximum radiation from the waves bounced off the reflector.

Energize the sweep oscillator, and observe the strength of the signal received in the receiving horn. Adjust the position of the transmitting horn until the received signal strength is a maximum.

Q15: Is the received signal strength larger than that measured for **Q9**?

Q16: What is the effect of adjusting the upwards angle of the feed horn (remembering to re-orient the reflector position for maximum signal strength)?

Q17: Can you think of other possible enhancements to increase the gain of the horn antenna?

Lab L8

Microstrip Microwave Transistor Amplifier Measurements

In this lab you will characterize one or two microwave transistor amplifiers using the network analyzer, as well as using a sweeper and a power meter. The steps of designing the amplifiers you will be testing were similar to what was done for the pre-lab. It is a conjugate matched MESFET and/or BJT amplifier designed for 2 GHz, with external bias networks. You will also examine the bias tee networks that were designed for the amplifier.

Part I. Network Analyzer Calibration

First, perform a two-port calibration of the network analyzer, as has been done for previous labs (refer to the analyzer instructions for details). Use a frequency range of 1 to 3 GHz. This calibration consists of a one-port (or reflection) calibration for both ports, as well as a transmission calibration in which a "thru" standard connected between ports 1 and 2 of the network analyzer. There is a third stage, called an isolation calibration, which can be omitted for purposes of this class. After you complete the calibration, check the calibration by observing all S-parameters when each of the cal standards is connected. Make sure the output power of the network analyzer is 0 dBm, or less (STIMULUS MENU, then POWER).

Part II. Bias Network Characterization

In this part of the lab, you will measure the performance of the bias tee network. The bias tee network uses a radial stub as part of the low pass filter for the DC input.

Q1: What other kinds of bias Tee networks could you design for this amplifier? Sketch the circuits you have in mind.

Q2: Connect the bias tee to the network analyzer. Leave the DC bias lines unconnected for the time being. Draw $|s_{11}|, |s_{22}|, |s_{21}|$ (all on a single dB scale) for the bias Tee.

Q3: What is the bandwidth (define) of the bias Tee? How would you make a more broadband biasing circuit at 2 GHz?

Q4: After disconnecting the bias Tee from the network analyzer, verify that the DC blocking cap in line with the RF transmission line is indeed doing its job, and explain in your report.

Q5: Connect the ground and DC input wires to a power supply and measure and record the DC voltages at each of the RF ports using a voltmeter.

Part III. Amplifier Characterization - Small-Signal

MESFET biasing — The required biasing for the MESFET is: $V_{GS} = -0.5\,\mathrm{V}$ to -1.5 V (NEGATIVE!) Do not increase V_{GS} above -0.5 V (especially, do not make it positive). $V_{DS} = 1 - -4\,\mathrm{V}$ (positive). Do not exceed 4 V!!! Connect the bias tee circuits to the amplifier circuit. The ports corresponding to the gate and drain are labelled on the amplifier board. The ports on the bias lines to be connected to the amplifier are labelled with the same letter as the port to which they connect. Before connecting the power supply to the circuit, set the power supply voltage level to -1 V for the gate bias and 0 V for the drain bias and turn it off. Connect an ammeter to monitor the drain current. You can use the current limiting feature on the power supply (if it exists) to be sure that too much current is not drawn (set this to 80 mA).

BJT biasing — Before connecting the power supply to the circuit, set the power supply voltage level to 3.0 V and turn it off (Note: Do not set the level greater than 3.5 V to avoid burning the BJT). You can use the current limiting feature on the power supply (if it exists) to be sure that too much current is not drawn (set this to 40 mA). Connect the bias tee circuits to the amplifier circuit. The ports corresponding to the base and the collector are labelled on the amplifier board. The ports on the bias lines to be connected to the amplifier are labelled with the same letter as the port to which they connect.

Connect the circuit to the network analyzer, Fig. L8.1. Connect the power to the bias tees (positive to the top wires and negative to the wires connected to the ground plane). Turn on the power and observe the $|s_{21}|$ parameter for gain.

For either amplifier, you will get 20% of the grade for this lab taken off if you burn the transistor - call the TA before turning on your amp.

Obtain and sketch the s-Parameters for the amplifier. Record values at the design frequency (2 1GHz).

Q6: Is the amplifier working and how do you know? Explain. What is the "best" operating frequency? What is the gain at that frequency? How much current is being drawn from the power supply?

Q7: What are the (1) drain (collector), (2) power-added and (3) total efficiencies of the amplifier?

For MESFET amp:— Obtain and record the S-Parameters and the current drawn by the amplifier at the design frequency as you vary the drain bias voltage between 1.0 V and 4 V while keeping the gate voltage at -1 V (change the voltage slowly and make sure you don't go over 4 V for the bias voltage!).

For BJT amp: — Obtain and record the S-Parameters and the current drawn by the amplifier at the design frequency as you vary the bias voltage between 1.0 V and 3.5 V (change the voltage slowly and make sure you don't go over 3.5 V for the bias voltage!).

Q8: What trends do you see as you change the bias voltage? Explain what you see.

Q9: What trends do you see in efficiency when you change the bias voltage?

For MESFET amp: — Fix the drain bias to 3.5V and vary the gate bias between -0.5 and -1.5V.

Q10: What trends do you see as you change the gate bias voltage? Explain what you see.

Part IV. Amplifier Characterization - Power

Connect a frequency synthesizer (sweep oscillator) to the input of the amplifier (gate side of MESFET and base side of BJT) and a power meter to the output side, Fig. L8.1. Bias the circuit with (3-4 V, -1 V) on the drain and gate for the MESFET, and 3.0 V for the BJT. Starting at -35 dBm, increase the power delivered to the amplifier until the 1-dB compression point. Record the output and input power levels for the power steps (warning, the power meter cannot take more than 20 dBm, but the compression point should be at an output power less than 15 dBm). Plot (1) Output power vs. input power and (2) the power gain versus power out, noting the 1-dB compression point.

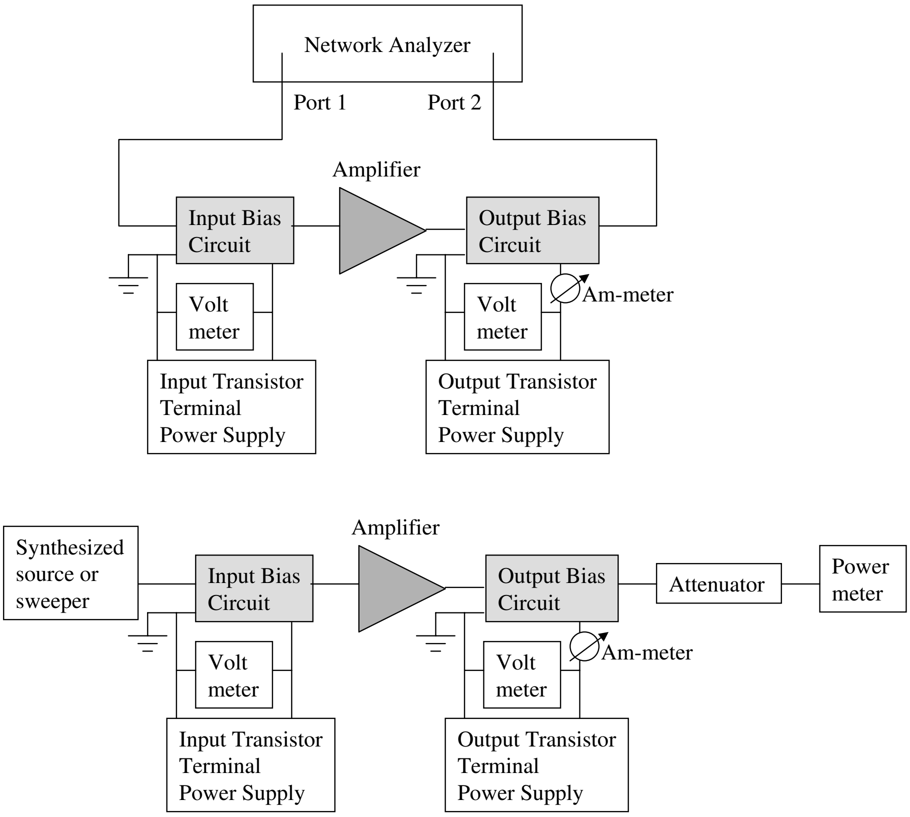


Figure L8.1: Experimental setup for small-signal amplifier characterization (top) and power characterization (bottom).

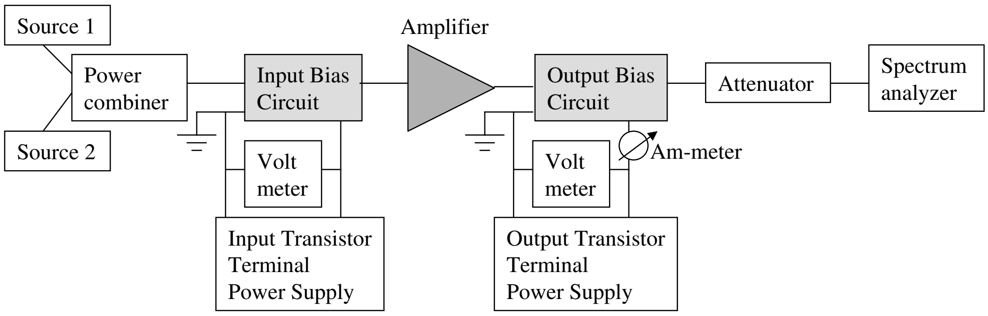


Figure L8.2: Experimental setup for two-tone nonlinearity amplifier characterization.

Part V. Two-tone Intermodulation Measurement

Connect two sweepers through a combiner circuit to the input of the amplifier, Fig. L8.2. Offset the frequencies on the two sweepers by a few MHz (1 MHz is ok). Connect the output of the amplifier to a spectrum analyzer. Bias the amplifier and set the sweepers to equal small-signal levels (say, -10 dBm). Sketch the output of the spectrum analyzer. Next increase the sweeper power levels, while keeping them the same, and observe the spectrum analyzer output. Plot the level of output power at the fundamental and the intermodulation products as a function of input power A FEW (2) dB ABOVE THE 1-dB COMPRESSION POINT. If you extrapolate the intermodulation curve, it will intersect the fundamental power curve. How many dB above the 1-dB compression point does this happen at?

Lab L9

Microwave Heterodyne Link

In this lab you will look at the processes of mixing and modulation, and at transmitters and receivers in a microwave superheterodyne link. The experimental setup is shown in Fig. L9.1. The transmitter,

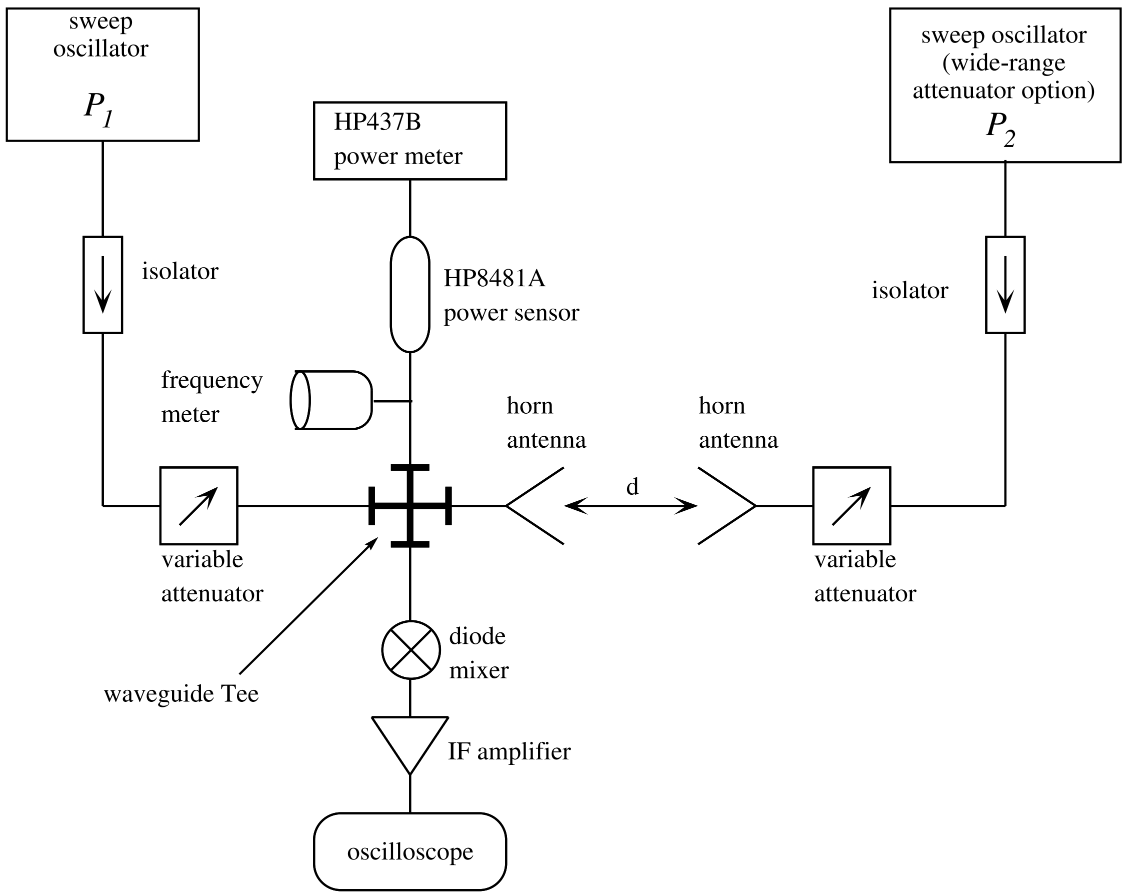


Figure L9.1: The experimental setup for the superheterodyne microwave link.

on the right hand side, is a sweep oscillator (the wide-range attenuator option means that it has a very large dynamic range, i. e., that it can generate very small as well as medium level powers). The sweeper is connected through an isolator and a variable attenuator to a standard gain horn antenna, which acts as a transmitting antenna. The receiver, on the left hand side, consists of a receiving horn, identical to the transmitting one, connected to a magic T. One port of the T is connected, through a frequency meter, to the power sensor and power meter. The other port is connected to a waveguide-mounted diode mixer which generates the IF frequency with the help of a local oscillator (the other sweep oscillator). The IF signal is then amplified and detected by the IF amplifier.

Part I.

In the first part of the lab, you will look at an amplitude modulated (AM) guided wave link, with no antennas or free-space propagation. Remove the horn antennas and connect the output of the transmitting variable attenuator directly to the magic T on the receiver side. The sweepers have the possibility of internal amplitude modulation with a 1 kHz square-wave. Set the transmitter sweeper at 9 GHz with the internal AM on. Set both attenuators at 0 dB. Adjust the power of the LO sweeper P_1 so that the power meter is reading 0 dBm (or as high as possible) for the receiver attenuator set at 0 dB. Then set the output power of the transmitting sweeper so that -20 dBm is incident from it at the magic tee.

Now you need to adjust the frequencies of the LO and RF to get exactly the right IF frequency, since without recent calibration the sweep oscillators do not display the output frequency accurately enough for the purposes of this lab. The IF amplifier is a narrow band amplifier/demodulator that will amplify only signals close to 30 MHz before detection. Set the frequency of the transmitter at 9 GHz, and the frequency of the LO on the receiver end to 9.03 GHz. You should do this by deactivating the RF from one of the sweepers, and fine tuning the other one until you observe a dip at the output of the frequency meter, which should be set to the desired frequency (9 GHz in the case of the transmitter, 9.03 GHz in the case of the LO). You could also adjust the frequency of the LO by maximizing the signal at the output of the IF amplifier, since you know it is narrow band and will respond well only at around 30 MHz. Bear in mind that the output of the IF amp is in fact the demodulated version of the amplified IF signal (the 1 kHz square wave) and not the 30 MHz IF signal itself.

Q1: What do you expect to observe on the oscilloscope (explain why)? Sketch a plot (on graph paper) of the scope screen, including all the scales.

Part II.

Now you will measure the dependence of the IF output power on the LO power. The meter on the IF amplifier measures *relative* the power level of the 30 MHz mixer product. If a coaxial low-pass filter is available, you can temporarily disconnect the diode output from the IF amplifier input, pass the mixer output through the low-pass filter to a power meter, and use this reading to calibrate the readings from the meter on the IF amplifier in terms of absolute power levels. Otherwise, you should simply regard the IF amplifier's meter readings as relative power levels.

Q2: Knowing that the mixer is a Schottky diode mixer, what do you expect the dependence of the IF power versus LO power to look like, and why?

Assuming that the input impedance of the oscilloscope is 50 Ω, adjust the LO and transmitter power levels until the voltage on the oscilloscope shows 1 V peak (in other words, 2 V peak-to-peak).

Q3: How much IF power is this? Make a note of the dB reading on the IF amplifier meter so that you can infer IF power levels from other meter readings later on.

Set the transmitter power on the sweeper to -30 dBm, and measure how the output IF power changes when you change the LO power by changing the attenuation on the variable attenuator on the receiver side. Fill in a table following the format of Table L9.1. Plot your result on a graph of IF power versus

Receiver attenuator setting	IF amplifier reading	IF power	LO power
0 dB etc.			

Table L9.1: Table for determining the IF versus LO power dependence.

LO power, on graph paper. Comment if you think something unusual is happening.

Part III.

Now you will do the same measurement, but you will keep the LO power constant and change the RF power using the attenuator on the transmitter side. Set the LO power to 0 dBm, and the RF at -20 dBm. Reduce the RF power until the signal and the noise on the scope display have the same amplitudes (this will be about -50 dBm of RF power). Fill in a table similar to Table L9.1, and sketch the scope display for the smallest value of the RF power.

Q4: Is there a difference in the behavior of the IF power in Parts II and III? Why do you think this is so?

Part IV.

Now we will look at an AM microwave link. Connect the horn antennas as in Fig. L9.1.

Q5: What is the minimum distance d between the horn apertures needed for far field operation at 9 GHz?

Q6: How far apart can the horns be if you have -30 dBm RF power radiated out of the transmitting horn, and you need at least -60 dBm of RF for a reasonable signal-to-noise (S/N) ratio at detection?

Pick a value of power output for the transmitting sweeper. Set the sweeper power to this value, and then keep increasing the attenuation of the transmitter attenuator until you observe the signal and the noise amplitudes on the scope as being equal.

Q7: Based on the values you have chosen, what is the maximum value of the distance between the horns d_{max} for S/N=1?

Q8: What would you get for d_{max} if you replaced the horns by a parabolic reflector antenna? What would you get if the RF power increased?

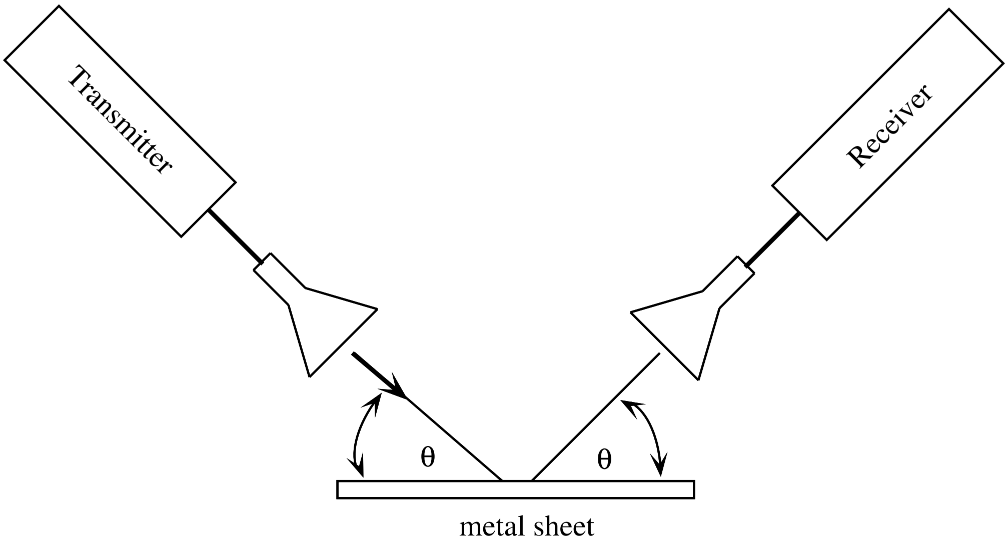


Figure L9.2: Position of horns for **Q9**.

Part V.

Now you will look at, or rather, listen to, a microwave radio link. There are a number of audio cassette tapes provided by the TA or the instructor containing fine examples of high culture. You can also bring your own, so long as they meet our strict standards of artistic merit. Ask the teaching assistant for one of them. The tapes are identified in various ways—some with numbers and some with a title, and have music recorded on them. The TA and instructor know what is recorded on the tapes we provide.

Connect the tape deck to the AM input at the back of the transmitting sweeper. Disable the internal 1 kHz modulation of the transmitting sweeper. Check that the sweeper frequencies are still set 30 MHz apart (some thermal drifting may have occurred since the beginning of the lab period). You will be using the IF amplifier as an AM radio to receive the 30 MHz mixer product. With the antennas connected, connect the (AM demodulated) output of the IF amplifier to the input of the audio amplifier. Establish a radio link.

Q9: What is the identification of your tape? What did you hear through the speakers?

Q10: Displace the two horns at an angle, as shown in Fig. L9.2 (view looking from the top). Bounce the signal from the transmitting horn to the receiving horn using a metal reflecting sheet. Can you think of a way to receive a larger signal with the horns kept in this configuration and the transmitter output power kept the same? Try it out, sketch your solution and explain your results. (You can use anything in the lab that the instructor or TA is willing to let you use.)

Lab L10

Doppler and CWFM Ranging Radar

In the lab, you will put together a Doppler radar, very similar in operation to the ones used by the police to measure vehicles speeding. We will use robust waveguide components, and measure the velocity of a moving target in the lab. You will construct a Continuous Wave Frequency Modulation radar using WR-90 waveguide components. CW-FM radar utilizes FM ranging and Doppler shifts to determine the distance to and speed of a detected object. FM ranging uses a frequency difference between transmitted and received waves to determine a round trip time, and hence distance, to a detected object.

Part I. Doppler Radar

Q1: The block diagram of the setup is shown in Fig. L10.1a. Assemble a Doppler radar, and tune the tuner for maximum signal response by moving a piece of metal with your hand in front of the antenna. How much difference in the signal amplitude can you get for different tuner positions? What is the tuner function and how does it change the level of the received signal?

Q2: Next, use a target made of a heavy cylinder (coffee can filled with rocks, e.g.) that you let roll down an inclined plane, as shown in Fig.e10-1b. Measure the Doppler frequency.

Q3: Calculate the velocity based on the Doppler measurement. Compare it to a calculation based on the inclined plane for several starting positions of the target.

Q4: Measure the tangential velocity of a "target" mounted on a bike wheel that is rotated.

Q5: Measure the velocity of either target above for velocity with a 8-GHz and 12-GHz (at edges of X-band) Doppler radar.

CW-FM Radar

Q6: Explain the functionality of all waveguide components used in the setup showed in Fig. L10.2a. What is the cutoff frequency of the WR-90 waveguide? What is the gain and far field distance of the horn antennas at f= 10 GHz?

Q7: Will the frequency range and the period of the VCO sweep waveform affect the output waveform of the diode detector for a given distance? How?

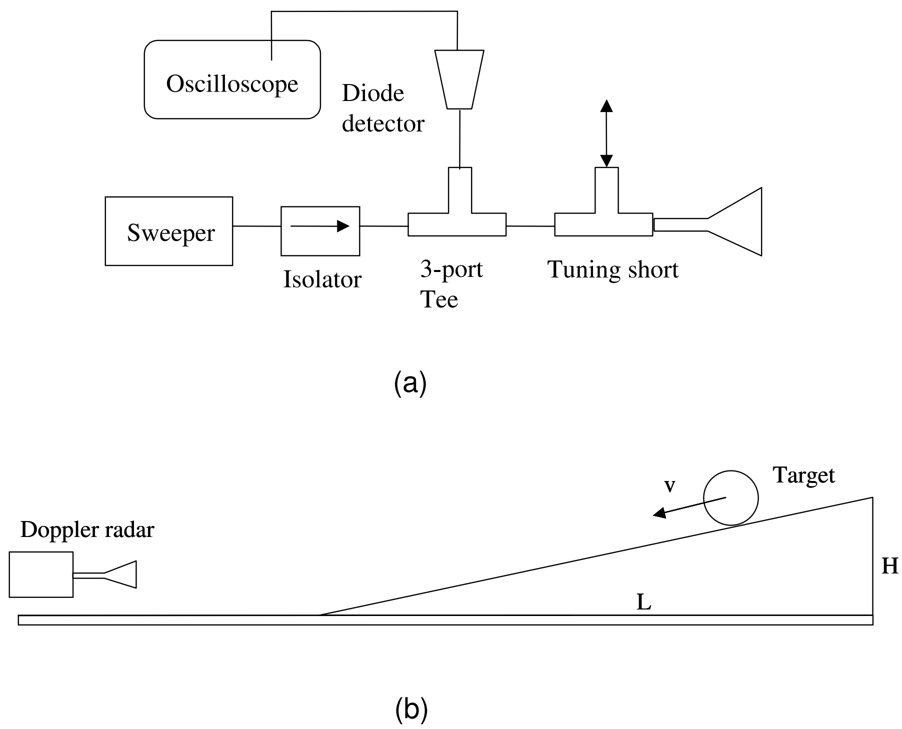
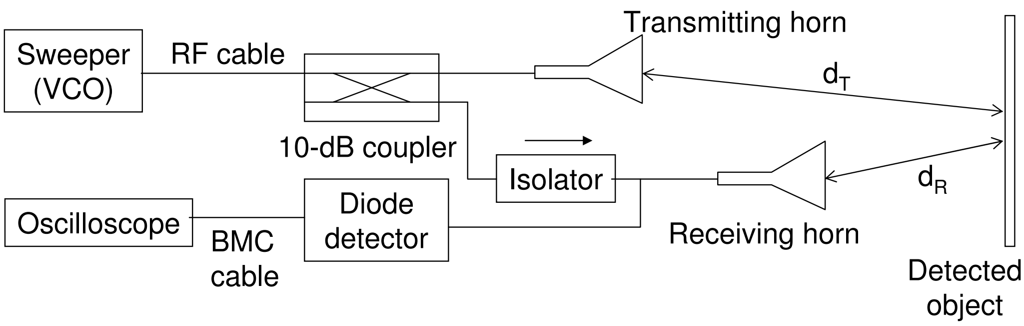
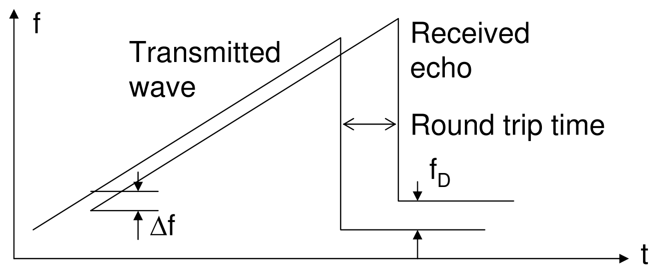


Figure L10.1: (a) Setup for Doppler radar experiment. Use an X-band sweeper, waveguide isolator, waveguide 3-port Tee, waveguide shorted stub tuner, waveguide mounted diode detector (mixer), horn antenna and oscilloscope for detecting the Doppler frequency shift. (b) Inclined plane and cylindrical target. What is the velocity of the target at the foot of the "hill"?



(a)



(b)

Figure L10.2: (a) Setup for CW-FM radar experiment. The RF signal is generated with a HP8350B (or equivalent) sweeper in internal controlled sweep mode (Sweep Trigger - INT, Sweep - INT). Use a sweep period of 0.020 seconds, a start frequency of 8 GHz and a stop frequency of 9 GHz. Waveguide components include: two horn antennas, 10 dB directional coupler, ferrite isolator, 3-port Tee, adapters and a diode detector (mixer). Use an oscilloscope for viewing the output of the mixer, trigger from the function generator. (b) CW-FM radar signals displaying simultaneous FM ranging and Doppler shift measurements?

Q8: Using a start frequency 8 GHz, a stop frequency of 9 GHz and a sweep period of 0.020 seconds on the VCO, what is the time rate of change of the transmitted frequency (df/dt)? What will the expected frequency difference between transmitted and received waves be for an object detected 50cm meters away (1 meter round trip)?

CW-FM Radar Calibration

Assemble the CW-FM radar setup shown in Fig. L10.2a. Sweep the transmitter signal frequency over the range from 8 to 9 GHz at a period of 0.020 seconds and a power of 20 dBm; monitor the diode output. It is best to store the output waveform on the oscilloscope, and then take a measurement over several periods to average the frequency.

Q9: Face the horn antennas toward each other slightly outside the far field distance. Draw a period of the diode output signal. Explain the origin of the frequency components of this signal? Which frequency component represents the distance to an object? How do you know?

Q10: What frequency difference do you expect at the diode output if the transmit and receive horn antennas are pointed toward each other and touching (i.e. $P_T = P_R$)?

The difference in frequency between the received and transmitted signals with zero distance between the two horn antennas is a calibration constant. This value corresponds to the difference between the received and transmitted signals' path lengths within the waveguide components and 3.5-mm cables.

Q11: Measure the calibration constant discussed above. What value do you get? Which signal travels further within the waveguide and by how much?

CW-FM Radar – Stationary Target Range Determination

Next, hold a flat copper plate 40 cm from the radar. Point the transmitter and receiver horn beams at the plane; making sure each is the same distance from the target. In this part of the lab, only the position (distance) of the detected object will be measured.

Q12: Measure the frequency difference between the transmitted and received signals. Calculate the distance that the value you measured for the frequency difference corresponds to? Remember to use the measured value of the calibration constant.

Q13: Repeat this measurement at 5 cm increments to a distance of 1 m; plot frequency vs. distance in your lab report. What do you expect the plot to look like? Explain any discrepancies.

CW-FM Radar – Moving Target Measurements

In the last part of the lab both the distance and speed of a detected object will be measured using the CW-FM radar setup, as illustrated in Fig. L10.2. Using analog to digital converters and digital signal processing both of these measurements can be processed simultaneously. However, using an oscilloscope to view the output visually does not allow this type of functionality due to the Doppler shift depending on the instantaneous frequency. Therefore, you will measure speed first, then distance to the moving object.

Q14: Configure the sweeper to output a constant radio frequency between 8 and 9 GHz. Measure the Doppler shift of the copper plane as you move it toward the radar. Does the calibration constant you found earlier apply to this measurement? Why or why not? How fast is the plane moving?

Q15: Return the sweeper to internal controlled sweep mode. Once again use a sweep period of 0.020 seconds, a start frequency of 8 GHz and a stop frequency of 9 GHz. Move the plane toward the radar in the same manner as before (i.e. try to move it at the same rate). Store the oscilloscope output in order to get a measurement. How will the speed of the plane need to be accounted for in the distance measurement? How far from the radar was the plane when you took the measurement?

Q16: How could the CW-FM radar setup from Fig. L10.2 be reduced in size and number of parts? What new parts would you use? Can you think of any ways to jam the radar? Explain.