# Principles of RF and Microwave Measurements

## (Lecture Notes and Experiments for ECEN 4634/5634)

by
Zoya Popović and Edward F. Kuester

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

# Contents

## Lectures and Labs


**1 Introduction and Review** **1**
1.1 What Are Microwaves? . . . 1
1.2 History . . . 3
1.3 Transmission Lines — Review . . . 6
1.3.1 Transmission Lines in the Time Domain . . . 6
1.3.2 Power Flow and Decibels . . . 8
1.3.3 Time-Harmonic Steady State . . . 10
1.4 Losses in Transmission Lines . . . 12
1.5 Loaded Transmission Lines . . . 13
1.6 Return Loss and Standing Wave Ratio . . . 15
1.6.1 Example . . . 15
1.7 Artificial Transmission Lines . . . 15
1.8 Microstrip Circuits . . . 20
1.9 Waveguides . . . 22
1.10 Practice questions . . . 24
1.11 Homework Problems . . . 25

**Lab 1: The Artificial Transmission Line** **29**

**2 Scattering Parameters, the Smith Chart and Impedance Matching** **33**
2.1 Scattering Parameters ($S$-Parameters) . . . 33
2.2 Reciprocal and Lossless Networks . . . 35
2.3 The Smith Chart . . . 38
2.3.1 Example 1 . . . 41
2.3.2 Example 2 . . . 42
2.4 Admittances on the Smith Chart . . . 42
2.5 Parasitics . . . 42
2.6 Impedance Matching Methods . . . 44
2.7 Lumped Element Matching . . . 45
2.8 Single-Stub Matching . . . 46
2.8.1 Smith chart method . . . 46
2.8.2 Analytical method . . . 47
2.9 Quarter-Wave Section Matching . . . 49
2.10 Vector Network Analyzers (VNAs) . . . 49
2.10.1 SOLT Calibration . . . 51
2.10.2 TRL Calibration . . . 52
2.11 Practice questions . . . 52
2.12 Homework Problems . . . 53

**Lab 2: Network Analyzer Operation** **57**

**3 Microwave Power Measurement** **63**
3.1 Power Definitions . . . . 63
3.2 The Thermistor . . . . 64
3.3 The Thermocouple . . . . 66
3.4 The Diode Detector . . . . 70
3.5 DC Biasing Networks . . . . 74
3.6 Measuring Reflection Coefficients: The Slotted Line . . . . 74
3.6.1 Example . . . . 75
3.7 Practice questions . . . . 76
3.8 Homework Problems . . . . 77

**Lab 3: Microwave Power Measurements** **80**

**4 Time-Domain Reflectometry** **87**
4.1 Introduction . . . . 87
4.2 Reflection from Simple Loads . . . . 87
4.2.1 Example: Reflection From an Inductive Load . . . . 87
4.2.2 Apparent Impedance . . . . 89
4.2.3 Example: Short Circuit . . . . 90
4.2.4 Example: Series *RL* Load . . . . 91
4.2.5 Example: Lossy Lines . . . . 92
4.2.6 Determination of the Time Constant of the Reflected Wave . . . . 92
4.3 TDR Measurements in Time and Frequency Domains . . . . 94
4.4 TDR Considerations for Digital Circuits . . . . 96
4.5 Practice questions . . . . 99
4.6 Homework Problems . . . . 100

**Lab 4: TDR Using the Network Analyzer** **103**

**5 Nonlinear Microwave Circuits** **107**
5.1 Microwave Sources . . . . 107
5.2 Oscillators . . . . 113
5.3 Resonators . . . . 117
5.4 The Spectrum Analyzer . . . . 119
5.5 Nonlinear Behavior of Microwave Circuits in the Time Domain . . . . 120
5.6 Practice questions . . . . 121
5.7 Homework Problems . . . . 121

**Lab 5: The Gunn-Diode Waveguide-Mounted Oscillator** **127**

**6 Microwave Multiport Networks** **131**
6.1 Two-Port Networks . . . . 131
6.2 Three-Port Networks . . . . 134
6.3 Four Port Networks—The Directional Coupler . . . . 136
6.4 Odd and Even Mode Analysis of a Branch Line Directional Coupler . . . . 138
6.5 Operating Principles of the VNA . . . . 142
6.6 Multiport Reflectometry . . . . 144
6.7 Practice questions . . . . 151
6.8 Homework Problems . . . . 152

**Lab 6-1: Microwave Multiport Circuits** **156**
**Lab 6-2: The Slotted Line** **159**
**Lab 6-3: Four-Port and Six-Port Network Analysis** **163**

**7 Antennas** **169**
7.1 Antenna Characteristics . . . 169
7.2 Transmitting and Receiving Antennas . . . 174
7.3 The Friis Transmission Formula . . . 179
7.4 Reciprocity and Antennas . . . 180
7.5 The Gain and Effective Area Integrals . . . 182
7.6 Some Application Examples . . . 182
7.7 Practice questions . . . 184
7.8 Homework Problems . . . 185

**Lab 7: Antenna Measurements** **188**

**8 Microwave Transistor Amplifiers** **193**
8.1 Microwave Three-Terminal Devices – The MESFET . . . 193
8.2 Transistors as Bilateral Two-ports . . . 196
8.3 Stability of Microwave Transistor Circuits . . . 199
8.4 Microwave Transistor Amplifiers . . . 199
8.5 Amplifier Design Procedures . . . 203
8.6 Biasing the Transistor in an Amplifier . . . 206
8.7 Practice questions . . . 207
8.8 Homework Problems . . . 208

**Lab 8: Microstrip Microwave Transistor Amplifier Measurements** **210**

**9 Microwave Communication Links: Superheterodyne Systems** **215**
9.1 Transmitters and Receivers . . . 215
9.2 Mixers . . . 217
9.3 Practice questions . . . 221
9.4 Homework Problems . . . 221

**Lab 9: Microwave Heterodyne Link** **223**

**10 Radar Fundamentals** **229**
10.1 The Radar Equation . . . 229
10.2 Doppler Radar . . . 231
10.3 FM-CW Ranging Radar . . . 232
10.4 Practice questions . . . 233
10.5 Homework Problems . . . 233

**Lab 10: Doppler and FM-CW Ranging Radar** **238**

# Appendices

**A Capturing data from the Agilent Vector Network Analyzers** **241**

**B Capturing data from the HP/Agilent Spectrum Analyzers** **243**

**C Capturing data from circuit modeling software** **245**

**D Notes on SPICE Usage** **247**
D.1 Transmission Lines . . . . . 247
D.2 Voltage-controlled sources . . . . . 248
D.3 Current-controlled sources . . . . . 250

**E Usage Notes on the DAMS Software** **251**

# Preface

In this course you will need to perform analyses of various microwave circuits using appropriate software tools. Although SPICE has the capability of modeling circuits containing transmission lines (and free versions of it are widely available), it is usually most convenient to use a program dedicated to RF and microwave applications, such as Ansoft Designer or AWR Design Environment (Microwave Office). Far simpler dedicated RF and microwave design programs like Puff or ViPEC are also available, but do not have as much power or flexibility. The freeware circuit simulator Qucs also has some capability for handling S-parameter and other microwave analyses. These programs all have differing abilities to plot results or lay out schematics, so one may be preferred over the others for reasons of convenience rather than anything else. In a few cases involving time-domain simulation or behavioral modeling for nonlinear devices, it may be necessary to use SPICE, and Linear Technology's LTspice is recommended for this purpose. Whenever a homework problem says to use "microwave design" software in the solution, you may use any of the ones mentioned here that are suitable. However, because AWR makes the use of their software available to you while you are a student here, in most cases you will probably want to use it rather than any of the others.

The reader will note that some of the practice problems and homework problems are somewhat open-ended. They are meant to stimulate creative thinking on the part of the student, since in real-world design contexts not all of the constraints are presented in an explicit or quantitatively precise way. In some cases the problems will require independent research in outside sources. In such cases, you should use *primary sources* rather than secondary ones. In other words, just because someone on a message board on the Internet says something is true, that does not constitute an adequate reference. You must use a book, a technical paper or report, a data sheet, etc., to support your solution.

A valuable supplementary reference for more details on the topics covered in these lecture notes is the book

> D. M. Pozar, *Microwave Engineering* (third edition). Hoboken, NJ: Wiley, 2005.

Much other information is covered in this text as well, so it is one of the books that all RF and microwave engineers should own. Other important sources that cover certain special topics at greater depth are:

> G. H. Bryant, *Principles of Microwave Measurements*. Stevenage, UK: Peter Peregrinus, 1993 [for general measurement techniques].
>
> J. P. Dunsmore, *Handbook of Microwave Component Measurements*. Chichester, UK: Wiley, 2012 [in-depth coverage of measurement techniques, especially with VNAs].
>
> J. D. Kraus, *Antennas* (second edition). New York: McGraw-Hill, 1988 [for antennas].
>
> G. Gonzalez, *Microwave Transistor Amplifiers*. Englewood Cliffs, NJ: Prentice-Hall, 1984 [for amplifiers].
>
> M. I. Skolnik, *Introduction to Radar Systems*. New York: McGraw-Hill, 1962 [for radar].

# 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 frequency spectrum of electromagnetic waves is depicted in Fig. 1.1. The microwave frequency region is divided into bands, as shown in Table 1.1.

Microwave networks are harder to analyze than their lower-frequency counterparts. The reason is that the size of a typical microwave circuit is comparable the wavelength, so phase variation along a portion of the 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 a full-wave electromagnetic 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 fractional 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 transmitters is avoided. Fig. 1.2 shows attenuation of a wave as it passes through the atmosphere as a function of frequency for two different altitudes.

| Band Designation | Frequency Range (Approximate) |
|---|---|
| AM broadcast band (medium wave) | 525-1705 kHz |
| Shortwave radio | 3-30 MHz |
| VHF TV (channels 2-4) | 54-72 MHz |
| VHF TV (channels 5-6) | 76-88 MHz |
| FM broadcast band | 87.8-108 MHz |
| Aircraft radio | 108-136 MHz |
| Commercial and public safety | 150-174 MHz |
| VHF TV (channels 7-13) | 174-216 MHz |
| UHF TV (channels 14-69) | 470-806 MHz |
| Wireless (shared with UHF TV) | 698-806 MHz |
| Public safety | 806-940 MHz |
| Cell phones | 824-849, 869-894, 876-960 MHz |
| L-band (IEEE) | 1-2 GHz |
| Wireless | 1.71-1.78, 1.8-1.91, 1.93-1.99 GHz |
| S-band (IEEE) | 2-4 GHz |
| Microwave ovens | 2.45 GHz |
| C-band (IEEE) | 4-8 GHz |
| X-band (IEEE) | 8-12 GHz |
| Ku-band (IEEE) | 12-18 GHz |
| K-band (IEEE) | 18-26 GHz |
| Ka-band (IEEE) | 26-40 GHz |
| V-band (IEEE) | 40-75 GHz |
| W-band (IEEE) | 75-110 GHz |
| Millimeter-wave | 110-300 GHz |

Table 1.1: Some RF and microwave frequency bands. Band designations differ according to the organization defining the standard (IEEE, NATO, EU, etc.). In addition, several frequency bands are shared by more than one category of user.

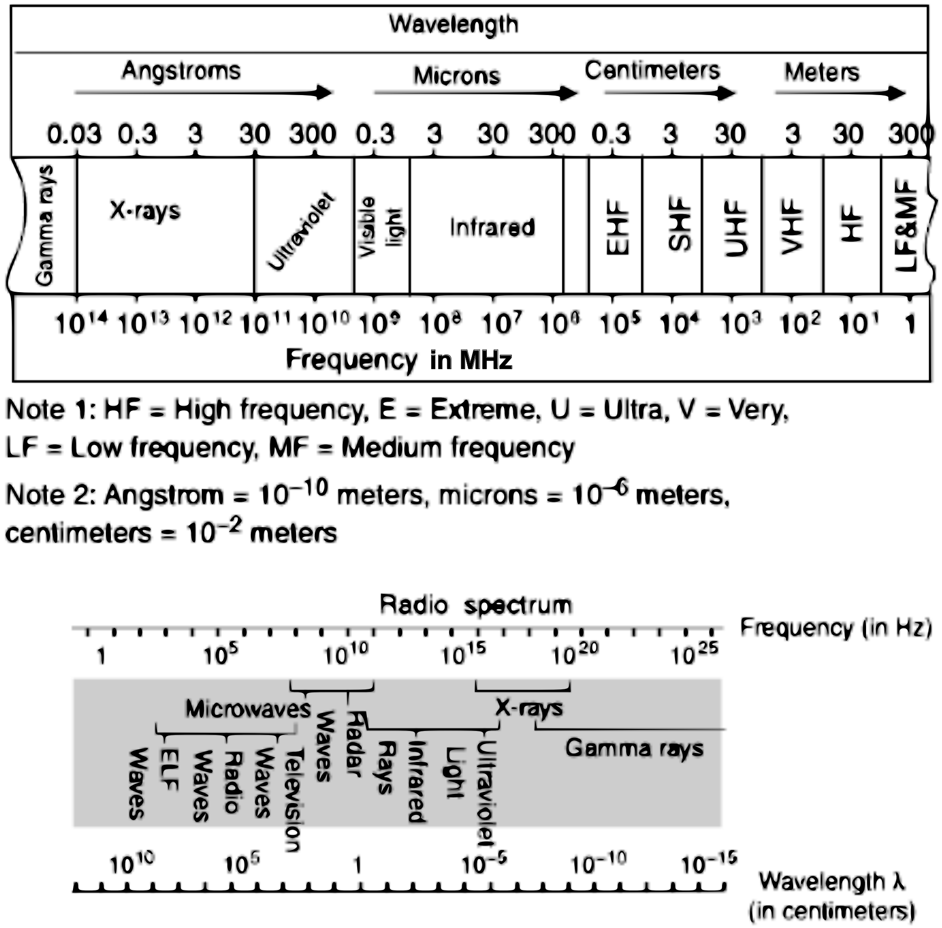



Figure 1.1: Two diagrams of the electromagnetic wave spectrum.

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

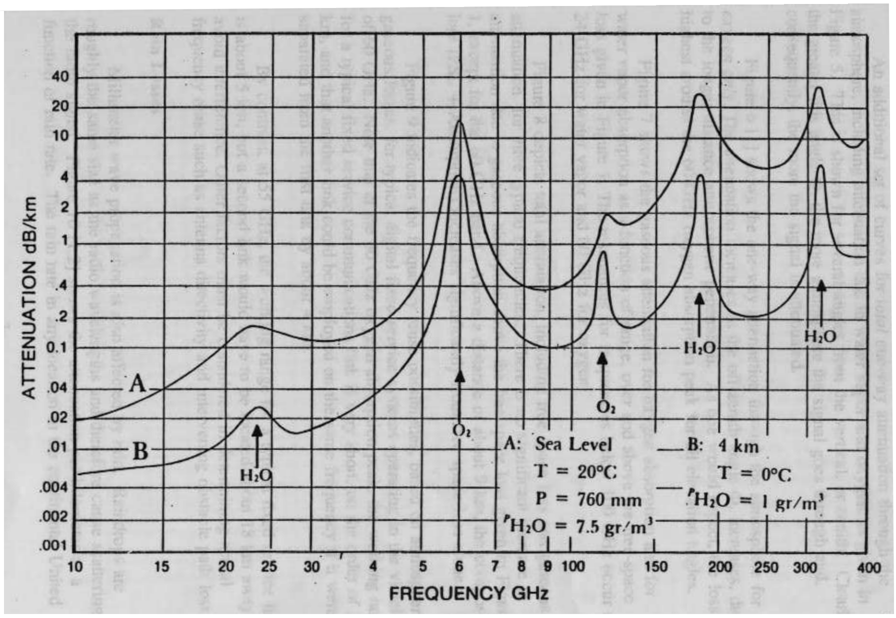



Figure 1.2: Attenuation of the atmosphere at sea level and 4 km altitude at microwave and millimeter-wave frequencies.

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 ("coax"), 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 Germany, 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 $\lambda$=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

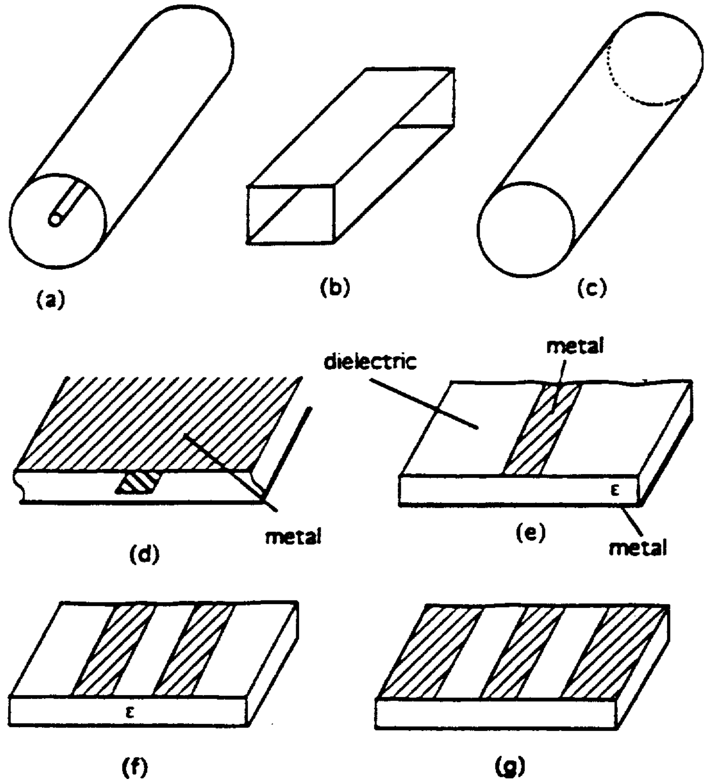



Figure 1.3: (a) coaxial line, (b) rectangular waveguide (c) cylindrical waveguide, (d) stripline (e) microstrip, (f) coplanar strips (CPS) and (g) coplanar waveguide (CPW).

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 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. 1.3 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 (Traveling 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 $\Delta z$ of cable, Fig. 1.4(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 $C_\Delta$ and a series inductor $L_\Delta$ 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. 1.4(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. 1.4(b), we implicitly assumed that all of the short sections $\Delta z$ 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:

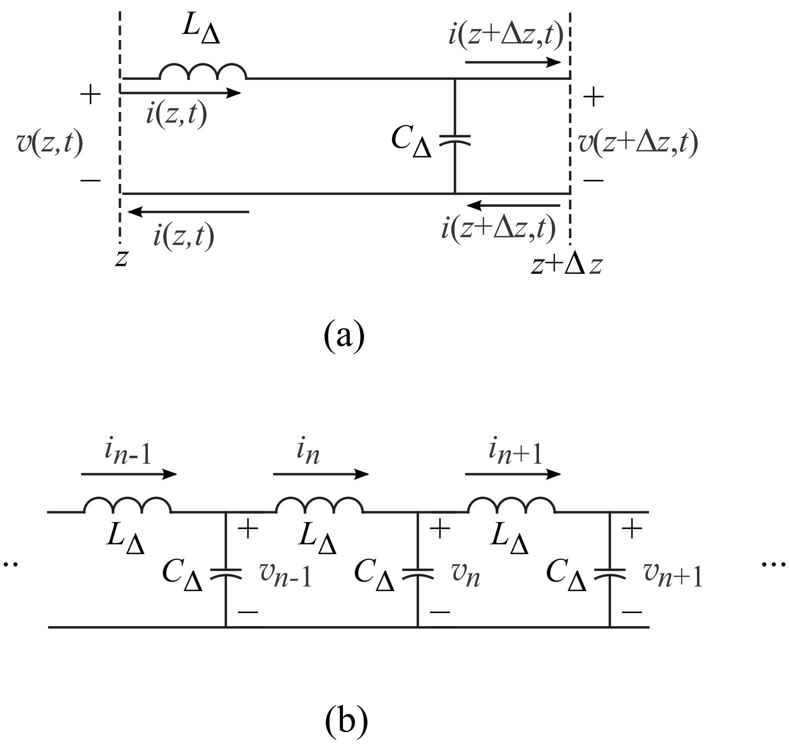



Figure 1.4: (a) A very short piece of lossless cable ($\Delta z$) can be represented as a lumped circuit consisting of a series inductor and a shunt (parallel) capacitor. (b) A longer piece of cable can be represented as many cascaded short sections, each of length $\Delta z$.

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

$L$ 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:

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

As $\Delta z$ 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 a function generator is at the beginning of the cable. We can now use (1.3) to write Kirchhoff's voltage and current laws 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 traveling through a homogeneous dielectric medium of relative permittivity $\epsilon_r$ 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* $\epsilon_e$, defined to be:

$$\epsilon_e = \left(\frac{c}{c_0}\right)^2 \tag{1.8}$$

so that $c_0$ is the same as the velocity of a plane wave in a hypothetical uniform dielectric whose relative permittivity is $\epsilon_r = \epsilon_e$.

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* $f(z + c_0 t)$.

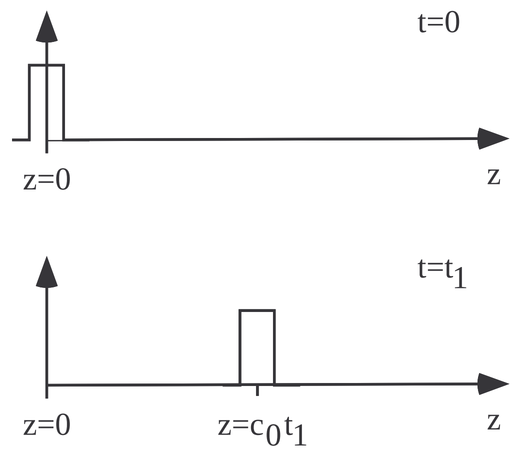



Figure 1.5: A voltage wave along a cable changes in time.

For example, let $f$ be a rectangular pulse which starts at $t = 0$, Fig. 1.5(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 $\epsilon_e$ 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

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

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

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

and $Z_0$ is called the *characteristic impedance* of the transmission line. The plus and minus signs apply to the forward and backward waves. The total voltage (or current) at any point along a linear transmission line is the sum of the incident and reflected voltages (or currents) at that point:

$$v(z,t) = v_+(z,t) + v_-(z,t) \qquad \text{and} \qquad i(z,t) = i_+(z,t) + i_-(z,t) \tag{1.12}$$

If we assume that the voltage between the lines has the same sign for each of 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. 1.6. 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

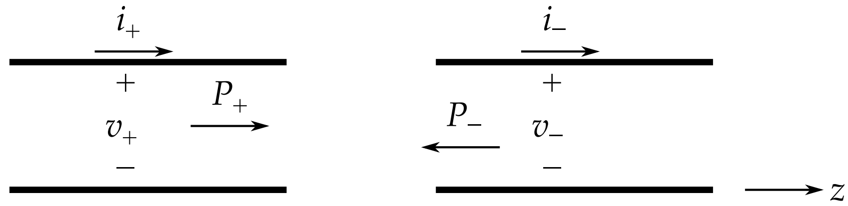



Figure 1.6: Forward and backward voltage and current waves in a transmission line.

the $+z$ direction:

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

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:

$$P_- = v_- i_- = v_-(-v_-/Z_0) = -v_-^2/Z_0 \tag{1.14}$$

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:

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

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

$$P_{\mathrm{dB}} = 10\log\left(\frac{P}{P_{\mathrm{ref}}}\right), \tag{1.15}$$

where $P$ is the power we are measuring or calculating, and $P_{\mathrm{ref}}$ 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:

$$P_{\mathrm{dBm}} = 10\log\left(\frac{P}{1\,\mathrm{mW}}\right) \tag{1.16}$$

A positive number of decibels corresponds to a ratio greater than 1 (gain), while a negative number of decibels 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 decibel quantities corresponds to multiplication of the corresponding absolute quantities, 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

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

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 dB$\mu$V: 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:

$$v(z,t) = \mathrm{Re}\left[\sqrt{2}V(z)e^{j\omega t}\right], \tag{1.17}$$

The presence of the factor $\sqrt{2}$ in (1.17) (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 (1.13) above, for example), we will have

$$\left[v^2(t)\right]_{\mathrm{av}} = |V|^2 \tag{1.18}$$

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:

$$V' = -j\omega LI, \qquad I' = -j\omega CV \tag{1.19}$$

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

$$V'' = -\omega^2 LCV \tag{1.20}$$

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 traveling in the $+z$ direction (forward) is $_+$, and for the one traveling in the $-z$ direction (backward) it is $_-$. Notice that the *forward-traveling* 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

$$\lambda_g = \frac{\lambda_0}{\sqrt{\epsilon_e}} = \frac{c}{f\sqrt{\epsilon_e}} \tag{1.21}$$

where $\epsilon_e$ is the effective permittivity defined in (1.8). Often, when there is no risk of confusion, the guided wavelength will be denoted simply by $\lambda$. Finally, the phase constant is related to the wavenumber $k_0 = \omega/c$ of a plane wave in free space by

$$\beta = k_0\sqrt{\epsilon_e} \tag{1.22}$$

In order to find the corresponding currents (forward and backward currents), we substitute the expression for the voltage back into (1.19), and obtain

$$I_+ = V_+/Z_0 \quad \text{and} \quad I_- = -V_-/Z_0, \tag{1.23}$$

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 traveling wave, but opposite signs for a backward traveling 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*:

$$\begin{aligned} V(z) &= V_+(z) + V_-(z) = V^+ e^{-j\beta z} + V^- e^{j\beta z} \\ I(z) &= I_+(z) + I_-(z) = \frac{V^+}{Z_0} e^{-j\beta z} - \frac{V^-}{Z_0} e^{j\beta z} \end{aligned} \tag{1.24}$$

where $V^+$ and $V^-$ are complex constants that specify the strength of the forward and backward traveling waves.

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. 1.7. A forward, or

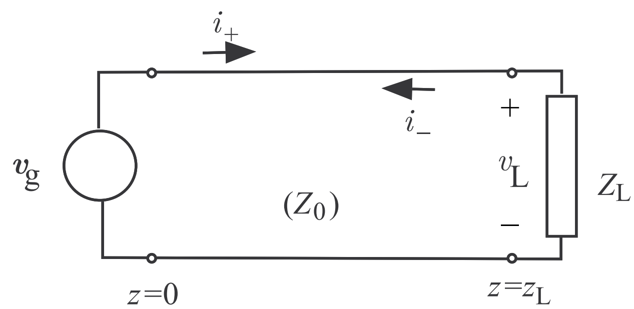



Figure 1.7: A transmission line of characteristic impedance $Z_0$ terminated in a load $Z_L$ at a distance $z_L$ from the generator.

incident, wave travels from the generator, reaches the load, and some of the power is delivered to the 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*:

$$\rho(z) = \frac{V_-(z)}{V_+(z)} \tag{1.25}$$

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

$$\tau = \frac{V_L}{V_+(z_L)} \tag{1.26}$$

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 $\tau$ the (voltage) *transmission coefficient*.

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:

$$Z(z) = \frac{V(z)}{I(z)} = \left[\frac{V_+(z)}{I_+(z)}\right] \frac{1 + V_-(z)/V_+(z)}{1 + I_-(z)/I_+(z)} = Z_0 \frac{1+\rho(z)}{1-\rho(z)} \tag{1.27}$$

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

$$\frac{Z(z)}{Z_0} = \frac{1+\rho(z)}{1-\rho(z)} \qquad \text{and} \qquad \rho(z) = \frac{Z(z)-Z_0}{Z(z)+Z_0} \tag{1.28}$$

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 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 (1.26) and (1.27), we obtain

$$\tau = 1 + \rho = \frac{2Z_L}{Z_L + Z_0} \tag{1.29}$$

Please keep in mind that (1.29) is *not* universally true, but was specifically derived for the situation shown in Fig. 1.7. Other circuits, particularly ones which contain series circuit elements in between the transmission line and the load, have a different relation between $\tau$ and $\rho$.

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

$$z = r + jx = \frac{1+\rho}{1-\rho} \qquad \text{and} \qquad \rho = \frac{z-1}{z+1} \tag{1.30}$$

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 (1.28) 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. 1.5 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. 1.8. as a distributed series resistance per unit length $R$ in $\Omega$/m, and a shunt conductance $G$ in S/m, respectively. For such a transmission line, the characteristic impedance is given by:

$$Z_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}}, \tag{1.31}$$

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

$$\gamma = \alpha + j\beta = \sqrt{(R + j\omega L)(G + j\omega C)} \tag{1.32}$$

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,

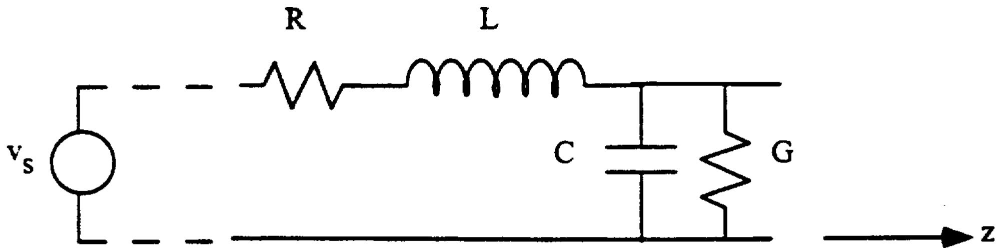



Figure 1.8: Schematic of a short section of transmission line with distributed losses included.

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 $\beta$ (phase constant), and its attenuation by $\alpha$, which is called the *attenuation constant*. Since

$$\begin{aligned} \gamma &= \sqrt{j\omega L\, j\omega C \left(1+\frac{R}{j\omega L}\right)\left(1+\frac{G}{j\omega C}\right)} \\ &= j\omega\sqrt{LC}\sqrt{1-j\left(\frac{R}{\omega L}+\frac{G}{\omega C}\right)-\frac{RG}{\omega^2 LC}} \end{aligned} \tag{1.33}$$

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

$$\gamma \approx j\omega\sqrt{LC}\sqrt{1-j\left(\frac{R}{\omega L}+\frac{G}{\omega C}\right)} \approx j\omega\sqrt{LC}\left[1-\frac{j}{2}\left(\frac{R}{\omega L}+\frac{G}{\omega C}\right)\right] \tag{1.34}$$

So, we find that

$$\alpha \approx \frac{1}{2}\left(R\sqrt{\frac{C}{L}}+G\sqrt{\frac{L}{C}}\right) \quad , \quad \beta \simeq \omega\sqrt{LC} \tag{1.35}$$

Since the power that is transmitted along the transmission line is equal to the product of the voltage and the current, it varies with distance from the generators as $P(z) = P(0)e^{-2\alpha z}$.

The unit for $\alpha$ 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:

$$\alpha_{\text{dB/m}} = 8.685\,\alpha_{\text{Nepers/m}} \tag{1.36}$$

## 1.5 Loaded Transmission Lines

Consider a transmission-line of length $l$ with a load $Z_L$, Fig. 1.9, and assume that the load has a reflection coefficient $\rho$ and is located at $z = 0$. Moving back along the line, we can write the voltage and current at some distance $z$ as

$$\begin{aligned} V(z) &= V_+(z)+V_-(z) = V_+(z)(1+\rho(z)) = V^+e^{-j\beta z}(1+\rho(0)e^{2j\beta z}) \\ I(z) &= I_+(z)+I_-(z) = I_+(z)(1-\rho(z)) = I^+e^{-j\beta z}(1-\rho(0)e^{2j\beta z}) \end{aligned} \tag{1.37}$$

where $I^+ = V^+/Z_0$. The impedance at point $z$ is

$$Z(z) = \frac{V(z)}{I(z)} = Z_0\,\frac{1+\rho(0)\,e^{j2\beta z}}{1-\rho(0)\,e^{j2\beta z}} \tag{1.38}$$

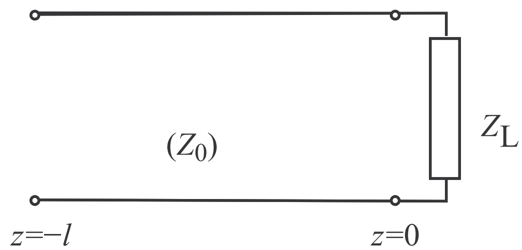



Figure 1.9: Reflection coefficient at different points along a transmission-line terminated with a load at $z = 0$.

We also know that

$$\rho(0) = \frac{Z_L - Z_0}{Z_L + Z_0}, \tag{1.39}$$

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:

$$Z(z = -l) = Z_0 \frac{Z_L + jZ_0 \tan\beta l}{Z_0 + jZ_L \tan\beta l} \tag{1.40}$$

Often, the "electrical length" $\beta l$ of a section of transmission line is identified and given the notation $\theta$. Sometimes the electrical position $\beta z$ along the line is also denoted by $\theta$. 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:

$$|\rho(-l)| = |\rho(0)| \qquad \text{and} \qquad \angle\,\rho(-l) = \angle\,\rho(0) - 2\beta l \tag{1.41}$$

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

$$\rho\left(l = \frac{\lambda}{4}\right) = -\rho(0), \tag{1.42}$$

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 (1.40) when $l = \lambda/4$ ($\beta l = \pi/2$):

$$Z\left(\frac{\lambda}{4}\right) = \frac{Z_0^2}{Z_L} \tag{1.43}$$

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-$\Omega$ load to a 50-$\Omega$ 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 (1.40) when $l = \lambda/2$ ($\beta l = \pi$). 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.

## 1.6 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

$$\mathrm{RL} = -20\log|\rho| \quad \mathrm{dB} \tag{1.44}$$

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

$$|V(z)| = \left|V^+\right| \left|1 + \rho e^{-j2\beta l}\right| = \left|V^+\right| \left|1 + |\rho| e^{j(\theta - 2\beta l)}\right|, \tag{1.45}$$

where $l = -z$ is the positive distance measured from the load at $z = 0$, and $\theta$ 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

$$V_{\max} = \left|V^+\right| \ (1 + |\rho|), \tag{1.46}$$

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

$$V_{\min} = \left|V^+\right| \ (1 - |\rho|), \tag{1.47}$$

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 $\rho$ 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

$$\mathrm{SWR} = \frac{V_{max}}{V_{min}} = \frac{1 + |\rho|}{1 - |\rho|}. \tag{1.48}$$

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

### 1.6.1 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.

From the formula for the reflection coefficient,

$$\rho = \frac{Z - Z_0}{Z + Z_0} = \frac{30 - j40}{130 - j40} = 0.297 - j0.216 = 0.368 e^{-j36^\circ}$$

The SWR is found from (1.48) to be SWR=2.163, and the return loss from (1.44) is RL = 8.692 dB.

## 1.7 Artificial Transmission Lines

A ladder network built out of repeated sections of lumped elements like those in Fig. 1.4, but without taking the limit $\Delta z \to 0$, is known as an artificial transmission line. Such lines can also be made from a combination of short sections of transmission line with lumped elements, a printed-circuit example of

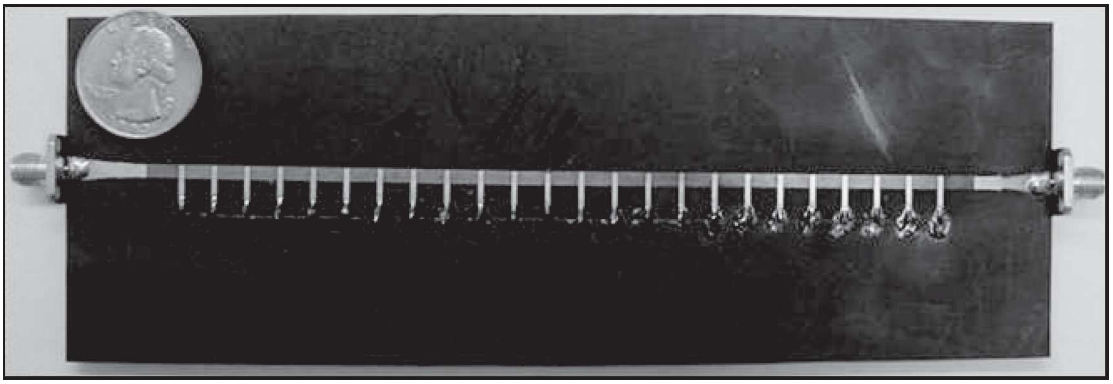


Figure 1.10: One-dimensional artificial transmission line on a substrate. A ground conductor (not shown) is on the bottom of the substrate, the stubs are shorted to ground through the substrate, and there are series capacitors connected between each pair of stubs.

which is shown in Fig. 1.10. The idea can be extended to two-dimensional ladder networks such as that shown in Fig. 1.11 One-dimensional artificial lines can function as filters, as we will see below, while two-dimensional artificial lines can perform as frequency-selective reflecting and transmitting surfaces, lenses, etc. These structures are related to artificially engineered media known as *metamaterials*.

It is most convenient to consider each segment, or unit cell, of an artificial transmission line as symmetric, as shown in Fig. 1.12; otherwise we are faced with definitions of characteristic impedances that are complex, and different for waves propagating forward than for waves traveling backward. A series impedance per unit cell $Z_\Delta$ is connected at either side to half of the total shunt admittance $Y_\Delta$ per unit cell as shown.

A wave traveling on this discrete lattice obeys

$$V_{n+1} = e^{-\Gamma} V_n; \qquad I_{n+1} = e^{-\Gamma} I_n \tag{1.49}$$

where $\Gamma = A + jB$ is the lattice propagation constant. The characteristic impedance of the artificial line is defined as

$$Z_0 = \frac{V_n}{I_n} \tag{1.50}$$

If the unit cell is intended as an approximate representation of a length $\Delta z$ of a continuous transmission line, then we put $Z_\Delta = Z\Delta z$ and $Y_\Delta = Y\Delta z$, where $Z$ and $Y$ are the series impedance per unit length and shunt admittance per unit length respectively. The lattice propagation constant is in this case related to the ordinary propagation constant of the continuous line by

$$\begin{aligned} \Gamma &= \gamma\Delta z \\ A &= \alpha\Delta z \\ B &= \beta\Delta z \end{aligned} \tag{1.51}$$

Applying Kirchhoff's voltage and current laws to Fig. 1.12 yields

$$\begin{aligned} V_n &= Z_\Delta\left(I_n - \frac{Y_\Delta}{2}V_n\right) + V_{n+1} \\ I_n &= \frac{Y_\Delta}{2}\left(V_n + V_{n+1}\right) + I_{n+1} \end{aligned} \tag{1.52}$$

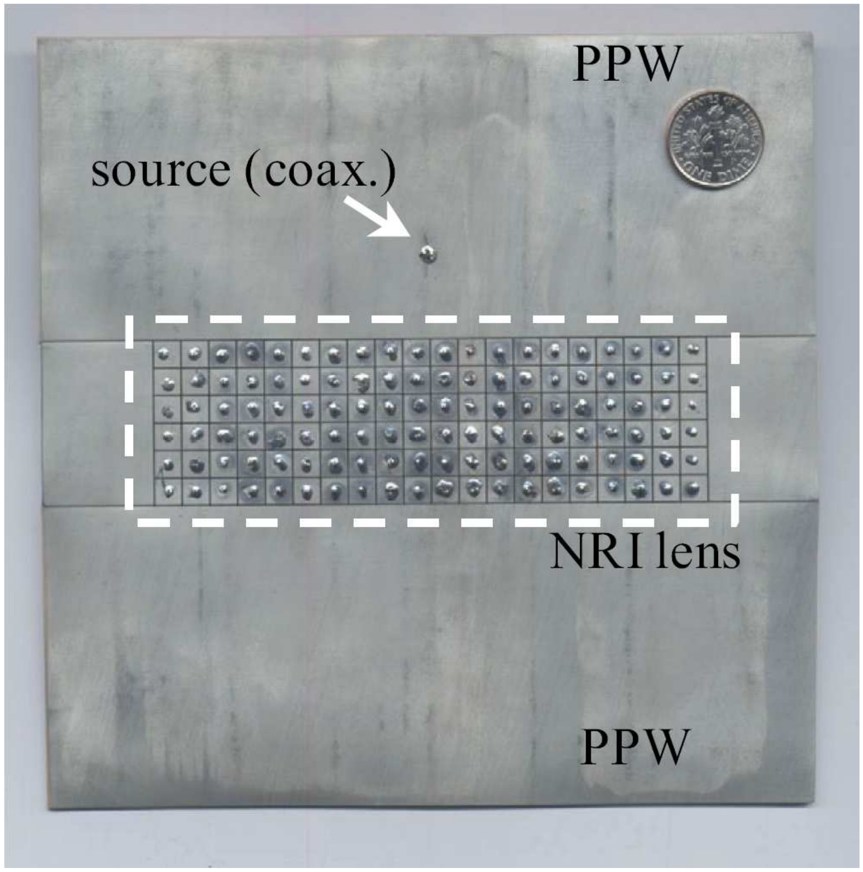



Figure 1.11: Two-dimensional artificial transmission line on a substrate.

Eliminating $V_{n+1}$ and $I_{n+1}$ from these equations using (1.49) gives

$$\begin{aligned} \left(1+\frac{Z_\Delta Y_\Delta}{2}-e^{-\Gamma}\right)V_n &= Z_\Delta I_n \\ \left(1-e^{-\Gamma}\right)I_n &= \frac{Y_\Delta}{2}\left(1+e^{-\Gamma}\right)V_n \end{aligned} \tag{1.53}$$

For a nontrivial solution ($V_n$ and $I_n$ are not zero), we find after some algebra that

$$2\sinh\frac{\Gamma}{2} = \pm\sqrt{Z_\Delta Y_\Delta} \tag{1.54}$$

and

$$Z_0 = \pm\frac{\sqrt{Z_\Delta Y_\Delta}}{Y_\Delta \cosh\frac{\Gamma}{2}} \tag{1.55}$$

If $|Z_\Delta Y_\Delta| \ll 1$, then (1.54) implies that $|\Gamma| \ll 1$, and hence the hyperbolic sine on the left side of (1.54) can be replaced by its small argument approximation $\sinh\frac{\Gamma}{2} \simeq \frac{\Gamma}{2}$ and (1.54)-(1.55) simplify to

$$\Gamma \simeq \pm\sqrt{Z_\Delta Y_\Delta} \tag{1.56}$$

and

$$Z_0 \simeq \pm\frac{\sqrt{Z_\Delta Y_\Delta}}{Y_\Delta} \tag{1.57}$$

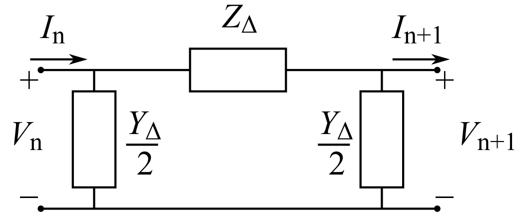



Figure 1.12: Symmetric unit cell of an artificial transmission line.

which is what would be expected for a continuous transmission line with series impedance $Z$ and shunt admittance $Y$ per unit length. It is under these conditions that an artificial line can be used as an approximation to a continuous one.

In many ways, artificial lines function analogously to continuous lines. In addition to propagation or attenuation, there can be reflection and transmission at discontinuities of impedance. Sections of artificial line can also be used as impedance-matching devices similar to the quarter-wave transformer or stub-matching circuits. Because of the flexibility in the choices of $Z_\Delta$ and $Y_\Delta$, circuit behavior not attainable with ordinary continuous transmission lines can be achieved.

As a first example, consider the special case of a conventional artificial line, for which $Z_\Delta = j\omega L_\Delta$ is a series inductance and $Y_\Delta = j\omega C_\Delta$ is a shunt capacitance as in Fig. 1.4(b). Then

$$\sinh\frac{\Gamma}{2} = \pm j\omega\frac{\sqrt{L_\Delta C_\Delta}}{2} = \pm\frac{j\omega}{\omega_c} \tag{1.58}$$

and

$$Z_0 = \pm\sqrt{\frac{L_\Delta}{C_\Delta}}\frac{1}{\sqrt{1-\frac{\omega^2}{\omega_c^2}}} \tag{1.59}$$

where

$$\omega_c = \frac{2}{\sqrt{L_\Delta C_\Delta}} \tag{1.60}$$

is called the *cutoff frequency* of this artificial line. If $\omega < \omega_c$, the lattice propagation constant is seen to be purely imaginary: $\Gamma = jB$, with

$$\sin\frac{B}{2} = \pm\frac{\omega}{\omega_c} \tag{1.61}$$

No attenuation of the voltage and current occurs from one unit cell to the next—only phase shift. Moreover, the characteristic impedance is real. Tracing the $\pm$ signs back to the original equations (1.54)-(1.55), we find that we must choose the upper (+) signs in (1.58)-(1.59) to get $Z_0 > 0$, which guarantees power flow in the $+z$-direction. On the other hand, if $\omega > \omega_c$, the real part of $\Gamma$ must be greater than zero, and attenuation occurs as each unit cell is traversed, the more so the larger $\omega$ is. From a plot of the magnitude of the transfer function $e^{-A}$ across a unit cell (Fig. 1.13), we see that this artificial line is essentially a low-pass filter. The characteristic impedance $Z_0$ is imaginary for $\omega > \omega_c$.

A second example is furnished by the so-called backward-wave artificial line shown in Fig. 1.14. In this case, we have

$$\sinh\frac{\Gamma}{2} = \pm\frac{\omega_c}{j\omega} \tag{1.62}$$

and

$$Z_0 = \pm\sqrt{\frac{L_\Delta}{C_\Delta}}\frac{1}{\sqrt{1-\frac{\omega_c^2}{\omega^2}}} \tag{1.63}$$

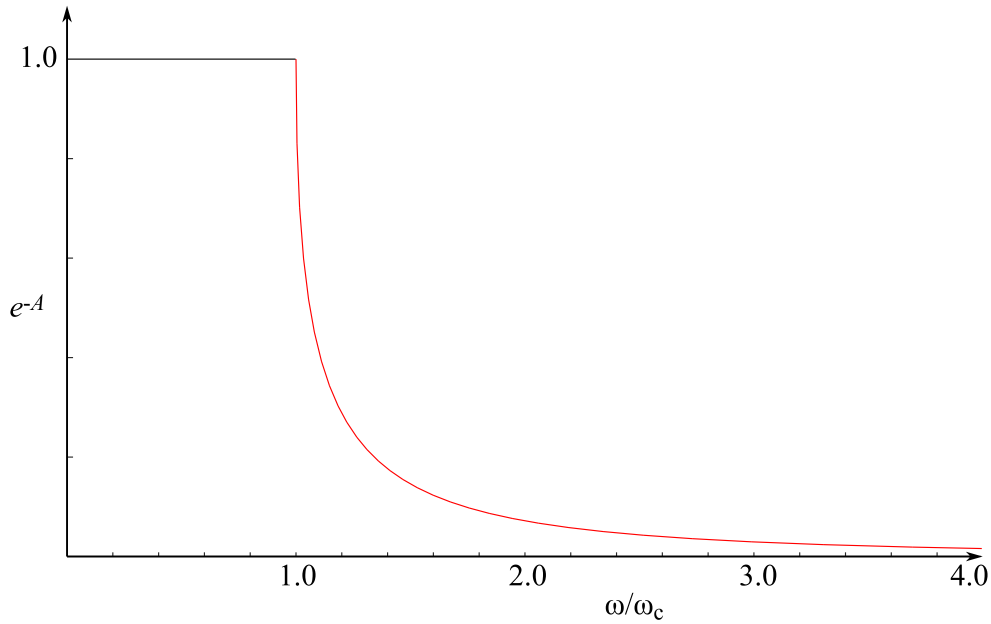



Figure 1.13: Unit cell transmission coefficient of a low-pass artificial transmission line.

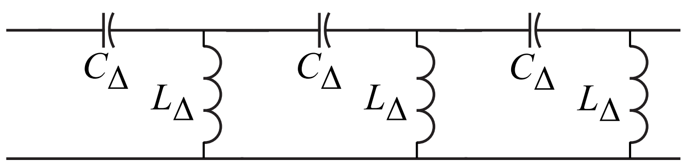



Figure 1.14: Backward-wave artificial transmission line.

where

$$\omega_c = \frac{1}{2\sqrt{L_\Delta C_\Delta}} \tag{1.64}$$

Here, if $\omega > \omega_c$, the lattice propagation constant is seen to be purely imaginary: $\Gamma = jB$, with

$$\sin\frac{B}{2} = \mp\frac{\omega_c}{\omega} \tag{1.65}$$

and it is for high frequencies that no attenuation occurs, and the characteristic impedance is real. Tracing the $\pm$ signs back to the original equations (1.54)-(1.55), we find that we must now choose the upper (+) signs in (1.58)-(1.59) to get $Z_0 > 0$, which again guarantees power flow in the $+z$-direction. Now, however, we must have $B < 0$: the phase constant is negative rather than positive. This means that phase fronts "move" in the backward $z$-direction, although power flows toward positive $z$. When $\omega < \omega_c$, the real part of $\Gamma$ must again be greater than zero, and attenuation occurs as each unit cell is traversed, the more so the smaller $\omega$ is. From a plot of the magnitude of the transfer function $e^{-A}$ across a unit cell in Fig. 1.15, we see that this artificial line is essentially a high-pass filter. The characteristic impedance $Z_0$ is imaginary for $\omega < \omega_c$.

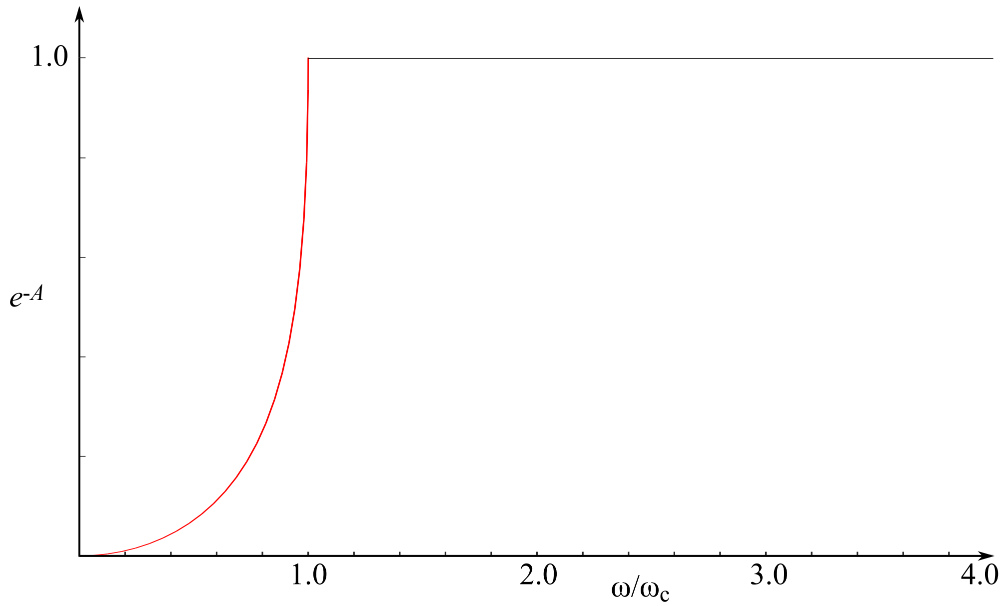



Figure 1.15: Unit cell transmission coefficient of a high-pass artificial transmission line.

## 1.8 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. 1.16. 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:

$$\epsilon_e = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \frac{1}{\sqrt{1 + 12\ h/w}} \tag{1.66}$$

There are many approximate closed-form expressions that have been obtained for microstrip line impedance and propagation constants; they give results usually within one percent of each other, and you can find them in almost any modern microwave textbook. We will just list one such set here for easy reference:

$$\begin{aligned} \beta &= k_0 \sqrt{\epsilon_e} \\ Z_0 &= \begin{cases} \dfrac{60}{\sqrt{\epsilon_e}} \ln\left(\dfrac{8h}{w} + \dfrac{w}{4h}\right), & \frac{w}{h} \leq 1 \\[2ex] \dfrac{120\pi}{\sqrt{\epsilon_e}\left[\frac{w}{h} + 1.393 + 0.667 \ln\left(\frac{w}{h} + 1.444\right)\right]}, & \frac{w}{h} > 1 \end{cases} \end{aligned} \tag{1.67}$$

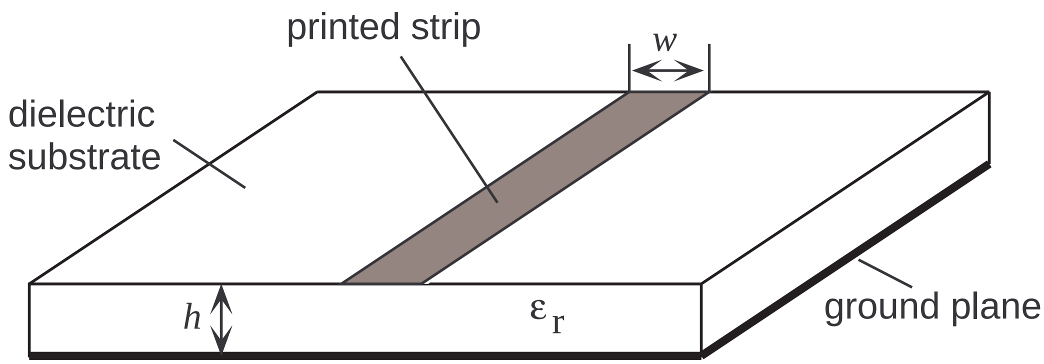



Figure 1.16: A microstrip transmission line printed on a grounded dielectric substrate.

For a given characteristic impedance $Z_0$, and permittivity $\epsilon_r$, the $w/h$ ratio can be found as

$$\frac{w}{h} = \begin{cases} \dfrac{8e^A}{e^{2A}-2} & \text{if } \dfrac{w}{h} \le 2 \text{ (or } A \ge 1.4928) \\ \dfrac{2}{\pi}\left[B - 1 - \ln(2B-1) + \dfrac{\epsilon_r - 1}{2\epsilon_r}\left\{\ln(B-1) + 0.39 - \dfrac{0.61}{\epsilon_r}\right\}\right] & \text{if } \dfrac{w}{h} > 2 \text{ (or } A \le 1.4928) \end{cases} \tag{1.68}$$

where

$$\begin{aligned} A &= \frac{Z_0}{60}\sqrt{\frac{\epsilon_r+1}{2}} + \frac{\epsilon_r - 1}{\epsilon_r + 1}\left(0.23 + \frac{0.11}{\epsilon_r}\right) \\ B &= \frac{377\pi}{2Z_0\sqrt{\epsilon_r}} \end{aligned} \tag{1.69}$$

A number of software packages include the capability of computing these quantities without the need to do so "by hand."

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. 1.17 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.

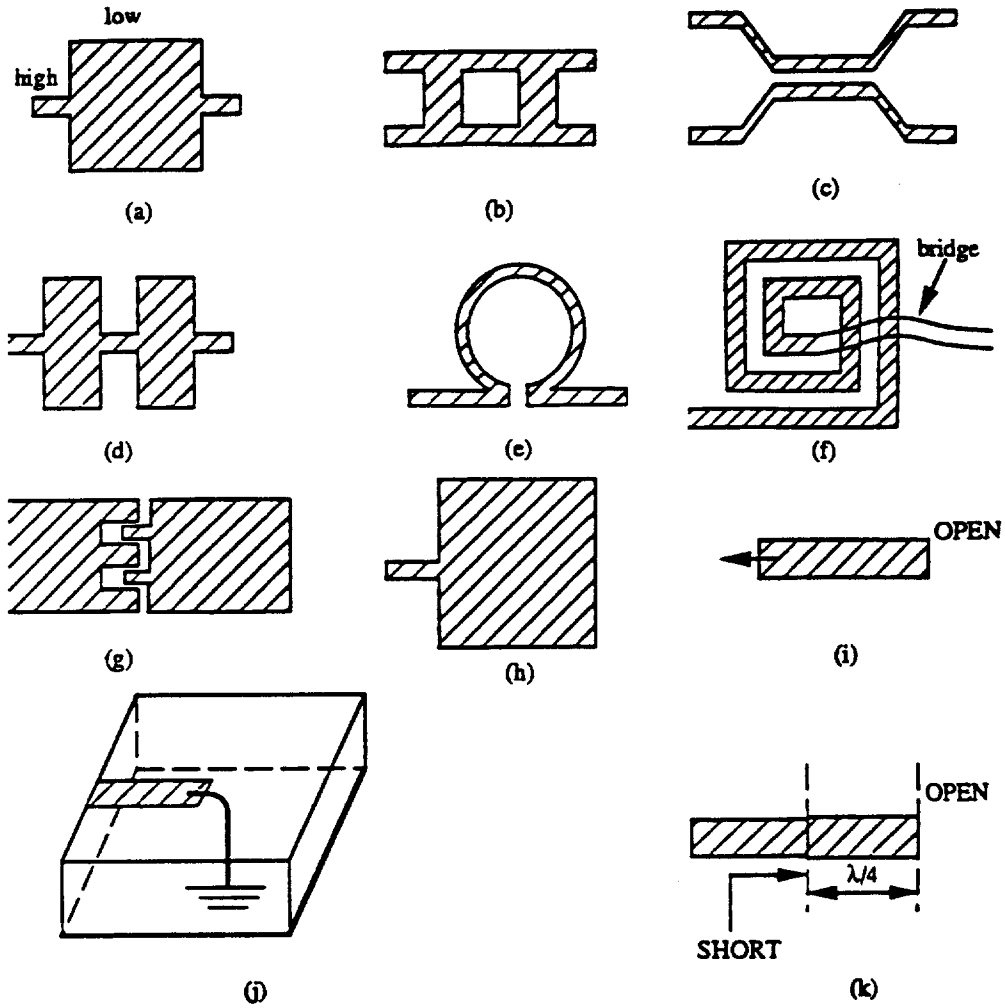



Figure 1.17: A microstrip impedance step (a), directional couplers (b and c), filter (d), inductors (e and f), capacitor (g), patch antenna (h), open circuit (i) and two ways of making a short circuit (j and k).

## 1.9 Waveguides

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

$$\vec{E}(x,y,z) = e^{-\gamma z}\vec{\mathcal{E}}(x,y); \quad \vec{H}(x,y,z) = e^{-\gamma z}\vec{\mathcal{H}}(x,y) \tag{1.70}$$

where

$$\gamma = \alpha + j\beta \tag{1.71}$$

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 $\mu$ and $\epsilon$ 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

$$\gamma = \alpha = 2\pi f_c\sqrt{\mu\epsilon}\sqrt{1-\frac{f^2}{f_c^2}}; \qquad f < f_c \tag{1.72}$$

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

$$\gamma = j\beta = j2\pi f\sqrt{\mu\epsilon}\sqrt{1-\frac{f_c^2}{f^2}}; \qquad f > f_c \tag{1.73}$$

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* $\lambda_g$, defined by:

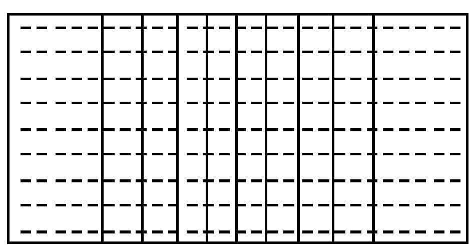
$$\lambda_g = \frac{2\pi}{\beta} = \frac{1}{f\sqrt{\mu\epsilon}\sqrt{1-\frac{f_c^2}{f^2}}} \tag{1.74}$$

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

$$f_c = \frac{c}{2a} \tag{1.75}$$

and whose electric field spans the shorter dimension of the waveguide as shown in Fig. 1.18. The electric



(1) $TE_{10}$

Figure 1.18: Electric (—) and magnetic (- - -) transverse field distributions of the $TE_{10}$ mode of a metallic rectangular waveguide.

field is largest in the center of the waveguide cross section, going to zero at the side walls as

$$\sin\left(\frac{\pi x}{a}\right) \tag{1.76}$$

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

| 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.

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 $f_c$.

## 1.10 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 total 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 $\alpha$ 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 an SWR?

15. 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?

16. 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.

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

18. How do the thickness and permittivity of the substrate affect the way a 50-$\Omega$ microstrip line looks?

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

20. Why, based on qualitative physical arguments, is a series gap in a microstrip line capacitive, and a loop (such as the one in Fig. 1.17(e) inductive?

21. What are the main differences, from the electrical point of view, between a physical short circuit as shown in Fig. 1.17(j), and a "virtual" short circuit made using a $\lambda/4$ long section of open microstrip line as shown in Fig. 1.17(k)?

22. What standard circuit element does the quarter wavelength long line remind you of? Explain.

## 1.11 Homework Problems

**1.** Use microwave design software 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. 1.19. [SPICE

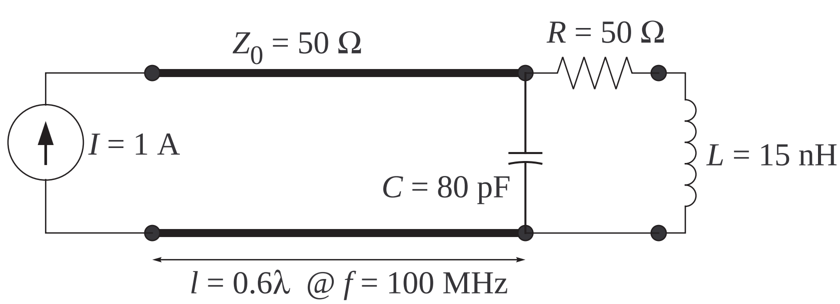



Figure 1.19: Transmission line terminated in lumped element load.

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. 1.20. Find an expression for the reflection coefficient of each lumped element for a

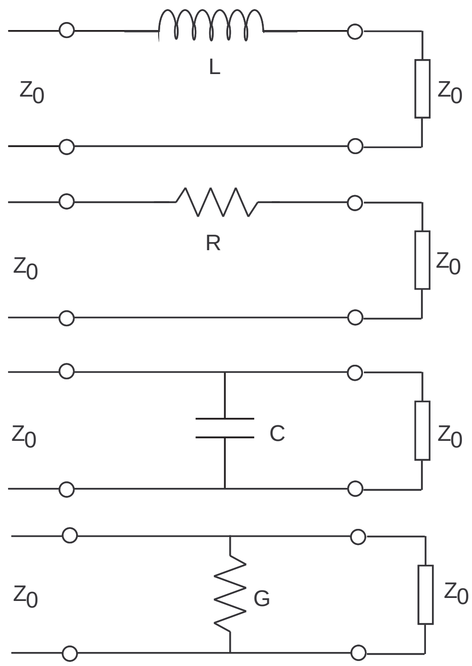



Figure 1.20: Lumped elements in a transmission line.

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 $Z_L = 100$ Ω connected at $z = 0$.

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

**7.** Consider an unbalanced artificial transmission line of the type shown in Fig. 1.4 having twenty (20) sections, with $L_\Delta = 1$ mH and $C_\Delta = 10$ nF. Use shunt capacitors at the beginning and end of the line of value $C_\Delta/2$ as in Fig. 1.12 to ensure that the unit cells are symmetric.

(a) What is the cutoff frequency of this line?

(b) Use microwave design software to do an AC analysis of the behavior of the line. Use an operating frequency of $f = 10\,\text{kHz}$: what is the electrical length $B = \text{Im}(\Gamma)$ at this frequency? Choose the source to be a voltage generator of strength 1 V, and load impedances of $Z_L = 316\ \Omega$, $Z_L = 0\ \Omega$, and $Z_L =$ (a 50 $\Omega$ resistor in series with a $1\mu$F capacitor) in turn. 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: To obtain such a plot, run an AC simulation at a single frequency, then use the methods described in Appendix C to extract the data and plot it using a suitable external program.*

(c) Next, replace the artificial transmission line with 20 sections of an actual lossless transmission line, each with a characteristic impedance $Z_0 = 316\ \Omega$ and terminated with these same loads. The electrical length $\beta\Delta z$ of each section should be chosen to be the same as the value of $B$ found in the first part of the problem. Use microwave design software to calculate the magnitude of the load voltage for each case, and compare with the results for the artificial line.

**8.** Use microwave design software to analyze the behavior of the voltage and current waves along a long 50 $\Omega$ 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 modeling software output.]

**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 $\Omega$ coaxial transmission line is connected to a load impedance of 50 k$\Omega$ (this would be typical of the input impedance to an audio amplifier). Consider an audio frequency of $f = 10\,\text{kHz}$.

(a) If the dielectric filling the transmission line has a relative permittivity of 2.25, calculate $\beta$ 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 k$\Omega$ to a generator connected to that end as frequency is varied through the audio range (20 Hz to 20 kHz) for the line lengths considered in (b) and (c)? 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 $\Omega$, 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 $\Omega$ line connected to such a load at $f = 20$ kHz?

**11.** A lab receiver measures an RF signal to have a strength of 35 dB$\mu$V. What is the signal level in volts? In watts? In dBm? Assume the instrument has an input impedance of 50 $\Omega$.

**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?

**13.** Can information about the magnitude only of $V(z)$ determine whether $\beta$ is positive or negative? From (1.37), and expressing the reflection coefficient at $z = 0$ as

$$\rho(0) = |\rho(0)|e^{j\phi_0}$$

obtain an expression for $|V(z)|^2$ in terms of $|V_+(0)|^2$, $|\rho(0)|$, $\phi_0$ and $\beta z$ only. To what does this expression reduce if $\rho(0) = -j$ (a capacitive load)? As we move from $z = 0$ towards more negative values of $z$, does $|V(z)|$ increase or decrease if $\beta > 0$? If $\beta < 0$?

**14.** Repeat Problem 7(a) and (b), but for the backward-wave artificial line shown in Fig. 1.14, and using an operating frequency of 100 kHz. Make the line have symmetric cells by choosing the first and last shunt inductors to have twice the values of all the others.

**15.** Use microwave design software 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. 1.21. Simulation

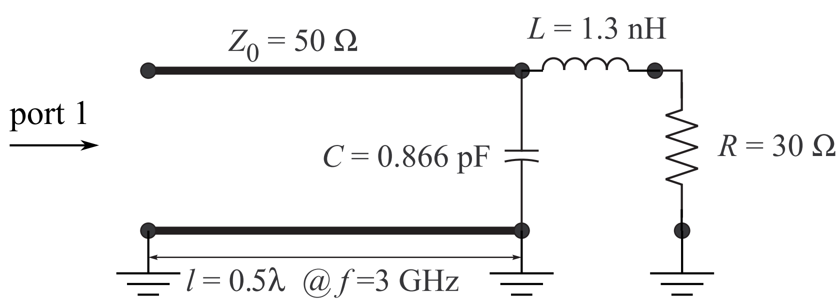



Figure 1.21: Transmission line terminated in lumped element load.

should cover the frequency range from 2 to 4 GHz. Plot the reflection coefficient $|S_{11}|$ (in dB) and $\angle S_{11}$ (in degrees) versus frequency in rectangular plots, and plot the input impedance $Z_{11}$ at port 1 versus frequency on a Smith chart, place markers at 2, 3 and 4 GHz in all plots, and include the schematic and plots with your solution. Explain why the plotted quantities behave the way they do.

# Lab 1: 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 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 $\Delta z$ 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. 1.22. Each section of it consists of an

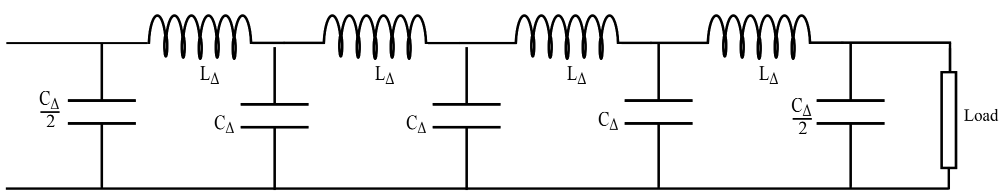



Figure 1.22: The artificial transmission line.

*L*-*C* section as shown in Fig. 1.4 that models a short section of transmission line. Artificial lines of this type are said to be unbalanced, because one conductor is a common ground, while the series impedances (here, the inductors) appear only in the other conductor of the line. 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, for which neither conductor acts as a "preferred" ground. 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 artificial line used in this experiment has series and shunt elements mounted on banana-plug connectors that are plugged into a ladder-network whose nodes are connected to a computer-controlled data-acquisition network. Resistances can readily be connected in parallel either with the series impedances or the shunt admittances. They can also be connected in series, although this is somewhat more awkward to accomplish. In this way losses corresponding to those in the metal conductors or the dielectric of an actual line can be incorporated into the artificial line.

The inductance per unit length of the transmission line of which a length $\Delta z$ is represented by one unit cell of the artificial line is

$$L = \frac{L_\Delta}{\Delta z};$$

and the capacitance per unit length is

$$C = \frac{C_\Delta}{\Delta z}.$$

Similar relations hold when resistances are included in the line. For the artificial line you will use in this experiment, $L_\Delta = 1\,\mathrm{mH}$ and $C_\Delta = 10\,\mathrm{nF}$.

Connect the experimental setup as shown in Fig. 1.23. Although the artificial lines have 36 sections, due to computer limitations we will only use 32 of them, and components are not inserted in the unused

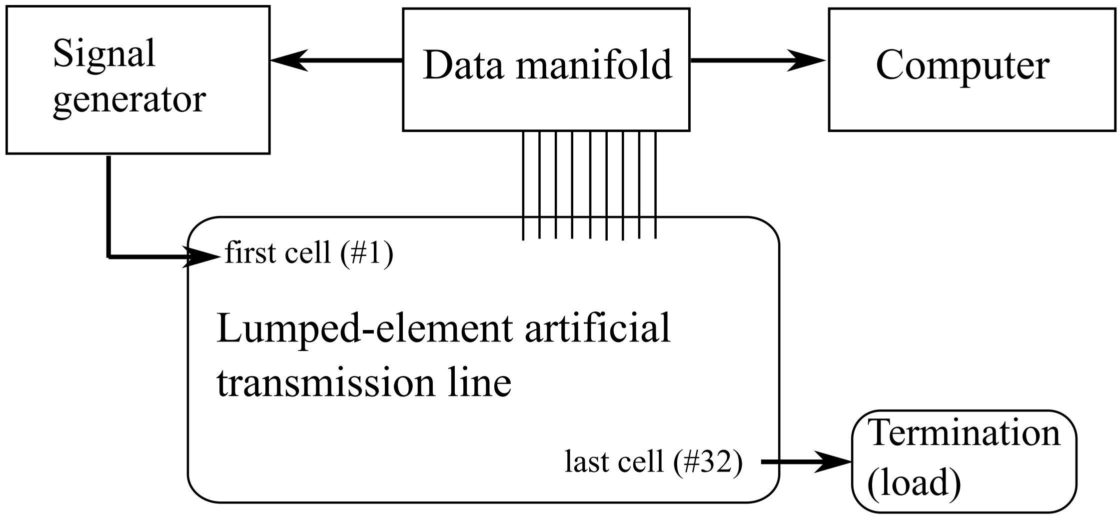



Figure 1.23: The setup for the artificial transmission line experiment.

sections. The signal generator (contained in the manifold attached to the computer) is connected at section 1, and the load at section 32. We will use only the inductors and capacitors at this stage. Note that, in order to have a line with symmetrical unit cells as shown in Fig. 1.12, we must use capacitors of value $C_\Delta/2 = 5$ nF at each end instead of $C_\Delta$.

## Part I.

The signal generator is contained in the same computer interface as the data acquisition manifold. In the LabView interface program, set the frequency on the signal generator to 10 kHz.

**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 $f = 10\,\text{kHz}$? What is the value of the ratio $\Delta z/\lambda$? How do these answers change if each segment represented a length $\Delta z = 1\,\text{km}$ of actual transmission line?

**Q2:** What would the wavelength of the wave be on an actual, air-filled transmission line at this frequency? How long would it have to be to have the same length in wavelengths 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 $\Delta z$ selected for the actual transmission line it represents?

Using the LabView-based data acquisition program, measure and plot the line voltage as a function of position along the line. The results acquired can be saved to a file and plotted as a so-called *normalized graph*, which means that the maximal value is made equal to unity. For the horizontal axis of all your plots, use the coordinate

$$\frac{z}{\lambda} = \frac{n\Delta z}{\lambda}$$

where $n$ is the section number (position) along the line.

**Q4:** Why is the output voltage (the voltage at the load) smaller than the input voltage? Is the line really lossless? If you think not, where do you think the losses come from?

**Q5:** Was your line well matched? Explain. Try to reduce the VSWR by adjusting the value of the load resistance and thereby find the best match.

Increase the oscillator frequency in steps of 10 kHz until you reach 150 kHz, and observe the plot of voltage versus position for each frequency. If at some frequency you observe strange behavior of the voltage vs. position plot, compare the oscillator frequency to the sampling rate shown at the lower left of the LabView window. What happens if one is an integer multiple of the other? To eliminate this problem, change the sampling rate a little (e. g., from 100 kHz to 110 kHz), but in no case increase it above 200 kHz.

**Q6:** Based on these observations, what do you estimate the cutoff frequency of this artificial line to be? How does this compare to the theoretically calculated value? Include the most important such voltage plots with your answer (but not all of them).

# Part II.

Next, reset the oscillator frequency to 10 kHz. Connect a 150 Ω resistor in parallel with each inductor of the artificial line.

**Q7:** At the operating frequency, what equivalent *series* resistance $R_{2\Delta}$ and inductance $L_{2\Delta}$ give the same impedance as the parallel combination of $R_{1\Delta}$ and $L_{1\Delta}$ that are actually present (see Fig. 1.24)?

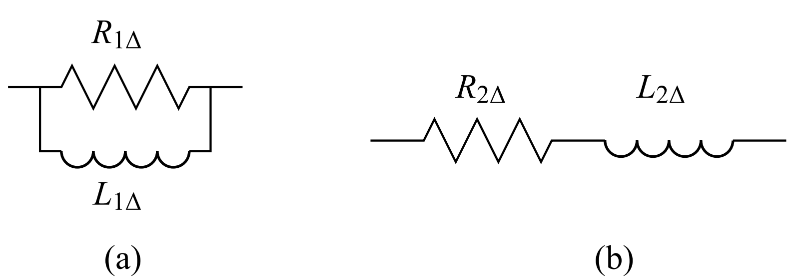



Figure 1.24: Equivalent series $RL$ circuit for a parallel $RL$ circuit.

**Q8:** Use the values of $R_{2\Delta}$ and $L_{2\Delta}$ determined above to find the characteristic impedance of the line now, assuming that the shunt conductance per section of the artificial line is $G_\Delta = 0\,\text{S}$.

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 same graph as in Part I.

**Q9:** Calculate the lattice attenuation coefficient $A = \alpha\Delta z$ from your measurement and from theory. How do they compare?

**Q10:** From your measurements, infer the magnitude of the current as a function of position along the line.

## Part III.

We will now look at what happens when the line is not well matched. Remove the resistors to obtain a nominally lossless line once again, 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.

**Q11:** 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 (1) a short circuit, and then (2) an open circuit, and plot the voltage versus position along the line for each case.

**Q12:** For these two terminations, what would the voltage as a function of position look like if the $\lambda/4$ transmission line was a perfect one?

**Q13:** 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?

**Q14:** Reset the generator frequency to 10 kHz, and set the decade capacitor to a value for which $\omega C Z_0 \simeq 1$ (it does not have to be exact) and use this as the load impedance. Obtain and save a plot of the voltage magnitude versus position.

## Part IV. – Optional

Now set the signal generator to a frequency of 80 kHz. Interchange the inductors and capacitors in the line, so that the capacitors are in series and the inductors are in parallel. Note that at each end you should now have shunt inductors of value $2L_\Delta = 2$ mH in order that the unit cells are symmetric (why twice the value instead of one half the value as was done for the shunt capacitors before?). NOTE: Hold the banana-plug connectors, NOT the inductors or capacitors themselves when removing them and inserting them in the artificial line. Stressing the component wires may break them off.

**Q15:** Can this configuration still function as a transmission line? Obtain formulas for the characteristic impedance and propagation constant of this line. Show that if we require the characteristic impedance to be positive, the lattice propagation constant $B = \beta\Delta z$ must be negative.

**Q16:** Connect the decade resistor to the load of this line. Obtain a plot of the voltage as a function of position along this line, and adjust the value of the load until reflection is minimized. What is the value of this resistance? How does it compare with the value of characteristic impedance calculated in **Q15**?

**Q17:** Can you tell just by looking at the voltage plot that the propagation constant is negative? Change the load to a pure capacitance, with the decade capacitor set so that $\omega C Z_0 \simeq 1$ (remember that the frequency is now 80 kHz), and plot the standing wave voltage pattern. How is it different from what you obtained in **Q14**? Is $\beta > 0$ or $< 0$ (refer to problem 13 of chapter 1)?

Decrease the oscillator frequency in steps of 10 kHz until you reach 10 kHz, observing the voltage plot for each frequency.

**Q18:** Based on these observations, what do you estimate the cutoff frequency of this artificial line to be? How does this compare to the theoretically calculated value? Include the most important such voltage plots with your answer (but not all of them).

Return the configuration of the artificial line to its original one before leaving the lab.

# Chapter 2

# Scattering Parameters, the Smith Chart and Impedance Matching

## 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 access points of transmission lines or waveguides connected to a circuit. These transmission lines support waves traveling 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 into 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 $i$ (it is thus assumed hereafter that the externally connected transmission lines have real characteristic impedances; some of the results derived under this assumption are not true if $Z_{0i}$ is complex). 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

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

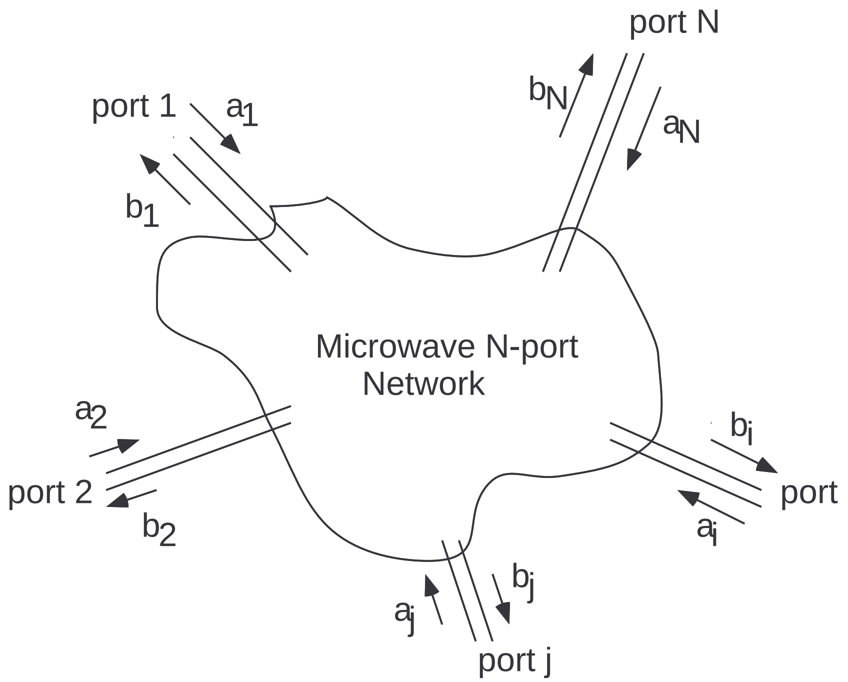



Figure 2.1: An $N$-port used for defining $S$-parameters.

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

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

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 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 $2 \times 2$ 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

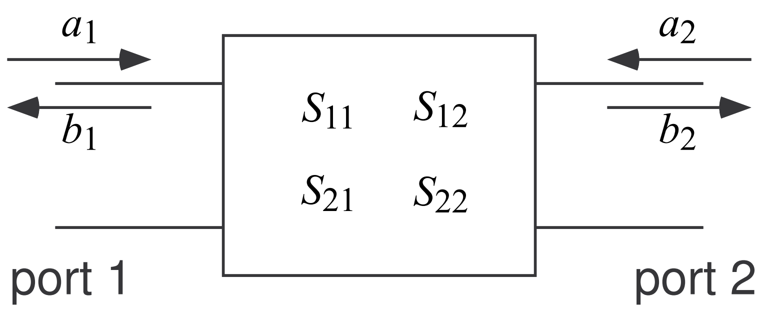



Figure 2.2: A two-port microwave network. The waves $a_1$ and $a_2$ are the input waves, the waves $b_1$ and $b_2$ are the output waves, and $\mathbf{S}$ is the scattering matrix for this network. One example of a two port network is just a section of transmission line.

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

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 modeling $S$-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.

## 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 $4 \times 4$ scattering matrix, and in this case the network is determined

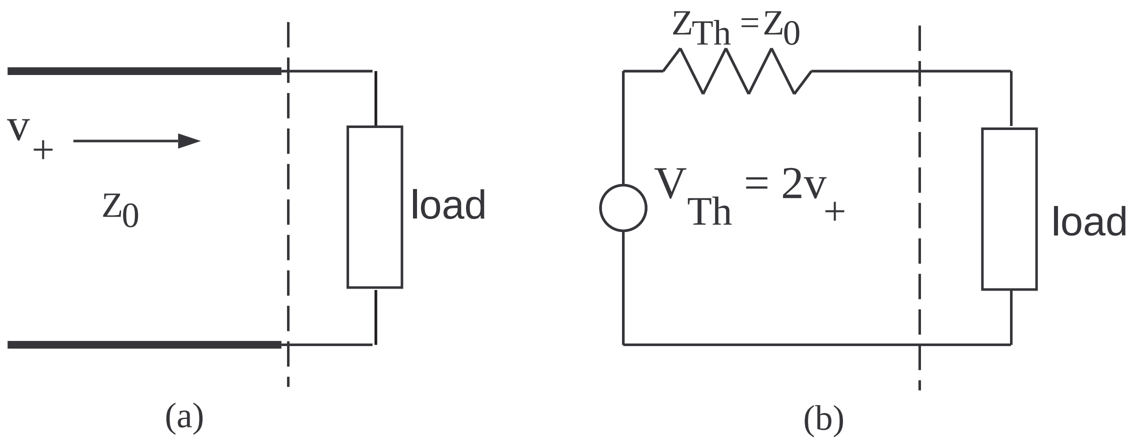



Figure 2.3: (a) Transmission line with an incident wave $v_+$, terminated by a load. (b) Equivalent circuit with the transmission line replaced by its Thévenin equivalent circuit.

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

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

Now consider a network terminated in matched loads with a single incident wave $a_j^{(1)}$ at port $j$, Fig. 2.4. The output at port $i$ is a scattered wave $b_i^{(1)}$ given by

$$b_i^{(1)} = S_{ij} a_j^{(1)}. \tag{2.12}$$

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

$$b_j^{(2)} = S_{ji} a_i^{(2)}. \tag{2.13}$$

If the system is reciprocal, the scattering matrix is symmetric, so that

$$a_j^{(1)} b_j^{(2)} = a_i^{(2)} b_i^{(1)}. \tag{2.14}$$

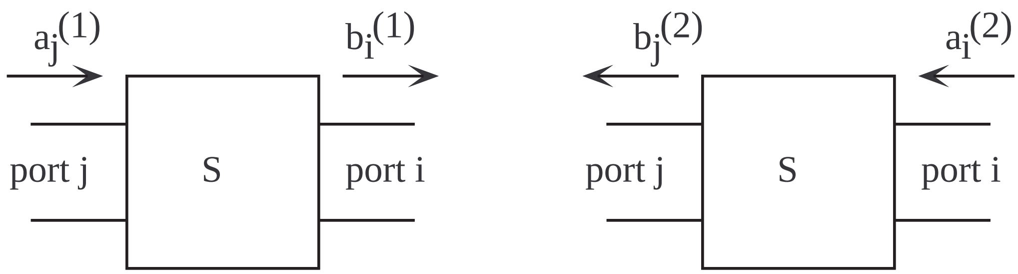



Figure 2.4: Reciprocity applied to a network.

By (2.4) and (2.5), this means that if we interchange which port is the input and which is the output, the power delivered through the network is the same. 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.

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

$$\mathbf{b}^\dagger \mathbf{b} = \mathbf{a}^\dagger \mathbf{a}. \tag{2.15}$$

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

$$\mathbf{S}^\dagger \mathbf{S} = \mathbf{I}, \tag{2.16}$$

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. In other words, the columns of the scattering matrix form an orthonormal set, which cuts down the number of independent $S$-parameters by a factor of two. In the case of a lossless reciprocal two-port, (2.16) leads to the three independent scalar identities:

$$|S_{11}|^2 + |S_{21}|^2 = 1 \quad ; \qquad |S_{22}|^2 + |S_{21}|^2 = 1 \quad ; \qquad S_{11}^* S_{21} + S_{21}^* S_{22} = 0 \quad . \tag{2.17}$$

which serve to further reduce the number of independent quantities in the scattering matrix.

## 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. Networks with three or more ports will be discussed in Chapter 6 of these lecture notes.

## 2.3 The Smith Chart

The Smith chart is a certain kind of plot of the complex reflection coefficient $\rho$ (or $S_{11}$, $S_{22}$, etc.; it is also used, though less commonly, to plot other $S$-parameters such as $S_{21}$). It provides 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 be 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.

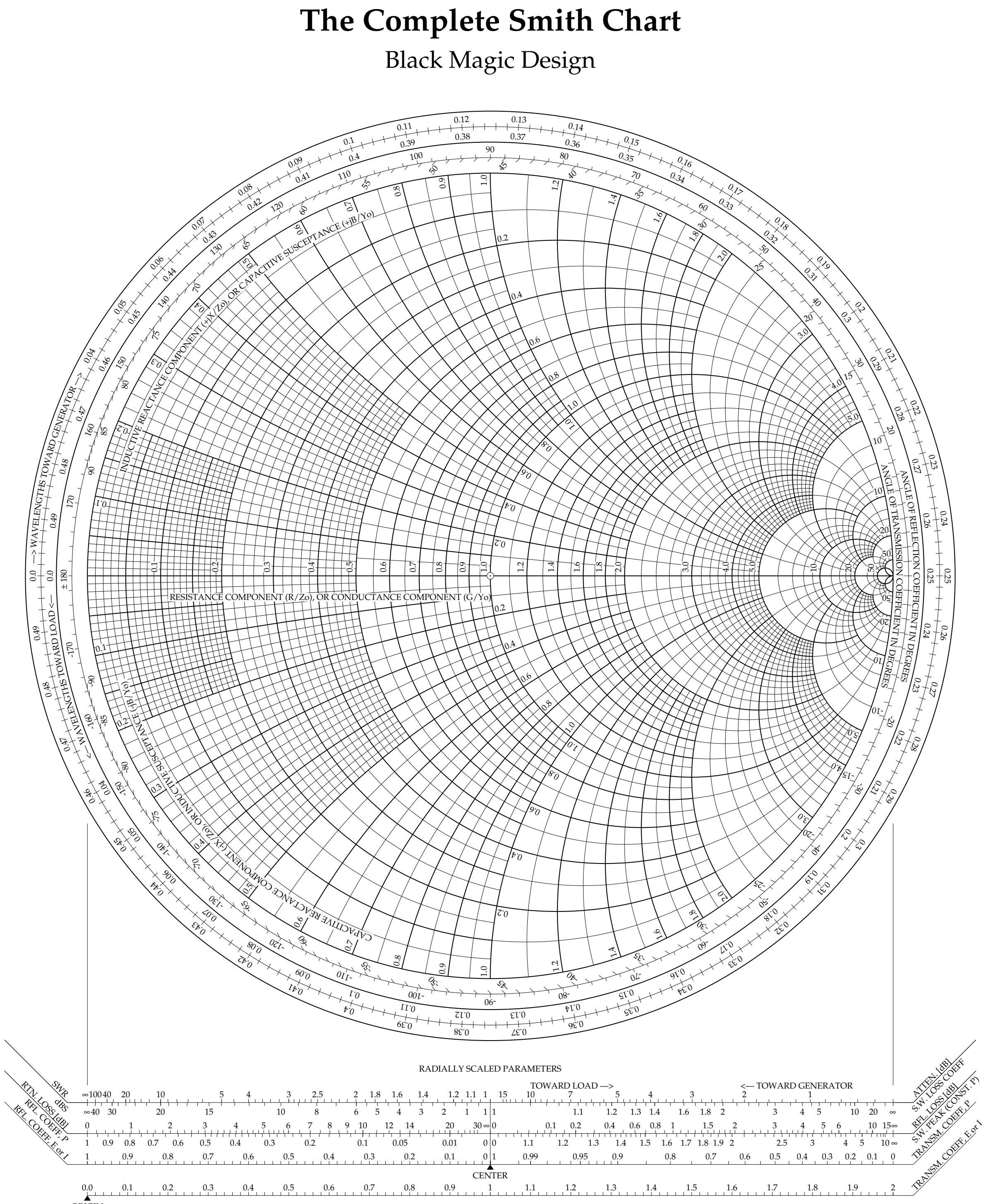



Figure 2.5: The Smith chart, with associated scales.

Here are a few simple rules to guide you in the use of the Smith chart:

1. Circles concentric to the origin are circles of constant reflection coefficient. The largest such circle represents a reflection coefficient magnitude equal to $|\rho| = 1$.

2. The Smith chart is a normalized graph: impedances $Z$ must be converted to normalized values $z = r + jx = Z/Z_0$ in order to be entered onto the chart using the impedance grids. Usually normalization is done with respect to the characteristic impedance of the transmission line (usually $50\,\Omega$), but other reference impedances may also be used.

3. Mathematically, equations (1.30) that relate $\rho$ and $z$ are called *bilinear* relationships. This means that circles in the complex $z$-plane map into circles in the complex $\rho$-plane, and vice versa. In this context, straight lines are considered as limiting cases of circles.

4. Circles of constant resistance are not concentric. The centers of these circles lie on the horizontal diameter of the chart. Point 1 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 ($r \to \infty$). 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.

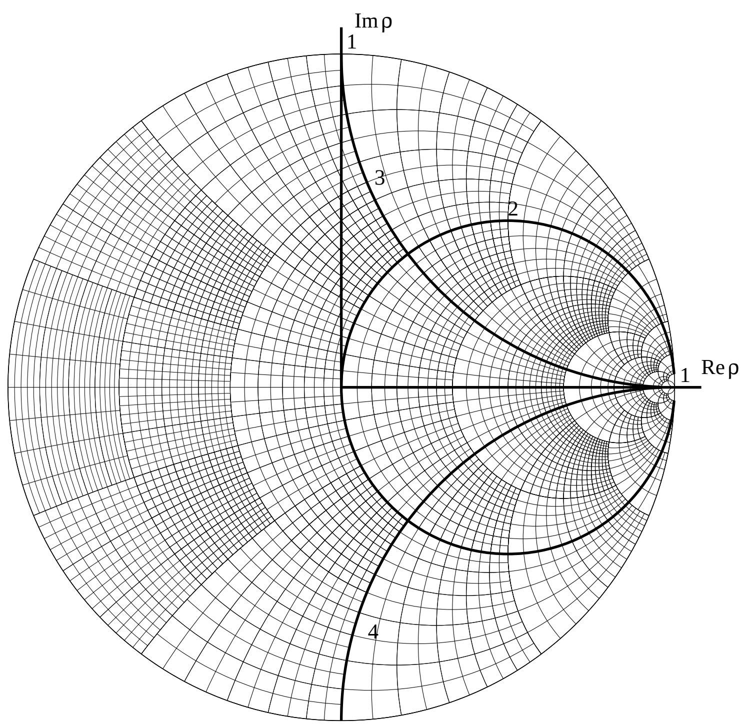



Figure 2.6: The Smith chart. Circle 2 is one of constant $r = 1$, and circles 3 and 4 correspond to lines of constant reactance $x = \pm 1$ respectively.

5. Circles of constant normalized reactance $x = X/Z_0$ have centers on the vertical axis along the line $\text{Re}(\rho) = 1$. The top half of the chart corresponds to inductances, and the bottom half 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

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

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

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

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

6. 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 $0.5\lambda$, which means that it includes the periodicity of transmission lines with half a wavelength. When you move around the chart by $2\pi$, 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.

7. 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.

### 2.3.1 Example 1

Consider the same example as in Section 1.6.1 above. Find the return loss in dB, SWR on the line, and the reflection coefficient at the load, using the Smith chart. Since all impedances on a Smith chart are

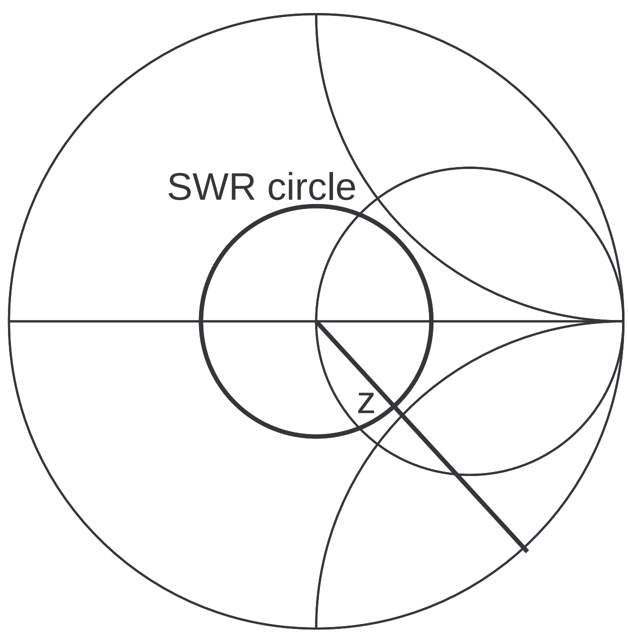



Figure 2.7: Solution to Example 2.3.1.

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 $|\rho| = 0.37$ (if plotting by hand, we might make a slight error and obtain, for example, $|\rho| = 0.36$). 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 that 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 (centered at the origin where $\rho = 0$) through the load impedance point, 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.

Note that our numerical values obtained using the Smith chart do not have the same precision as those found earlier using analytical methods.

### 2.3.2 Example 2

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 $Z = Z_0 z = 260 + j180\,\Omega$.

## 2.4 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 $0.2\lambda$ 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. 2.8, 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 $0.2\lambda$, 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 $0.2\lambda$ and reading off the impedance grid.

## 2.5 Parasitics

Real circuit elements have a geometries that induce behavior beyond that of the "intended" operation of the component. Wires, for example, have resistance due to finite conductivity of the metal, and this resistance increases with frequency due to the skin effect. Changes in dimension of the conductors of a

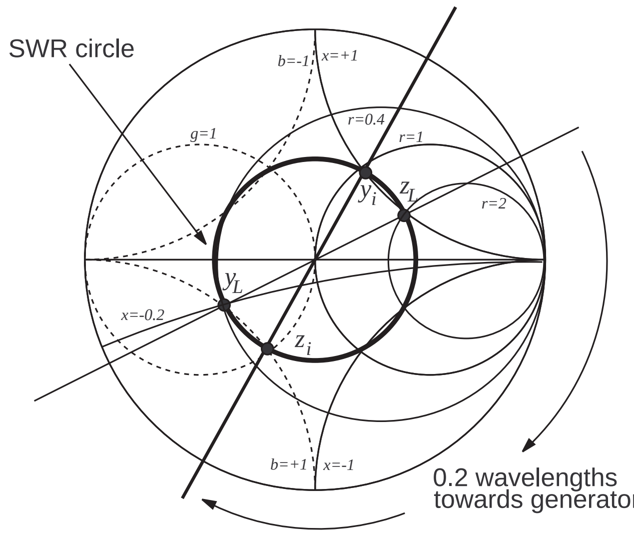



Figure 2.8: Using the Smith chart as an admittance chart. The impedance scale is shown as solid lines, and the admittance scale as dashed lines. The SWR circle is shown as a heavy line.

transmission line can cause excess charge accumulation at corners, and bunching of the current density as the cross section of a conductor is reduced. This causes extra capacitance and inductance beyond what would be predicted from a uniform transmission line theory. Such phenomena are called *parasitic*: they may be due to the interconnection wiring, the effects of a physical package of a device, junction effects in semiconductor devices, radiation or interaction with the environment. At microwave frequencies and above, parasitic effects can be very important, and must be accounted for in the design of a system. For example, most RF and microwave software design tools include models for parasitic effects when microstrip lines of different conductor strip widths are joined.

As an explicit example, consider a straight wire of radius $r_0$ and length $l$. The inductance of this wire at high frequencies (skin depth small compared to $r_0$) can be shown to be

$$L = \frac{\mu_0 l}{2\pi}\left[\ln\left(\frac{2l}{r_0}\right) - 1\right] \tag{2.18}$$

if $l \gg r_0$. For $l = 1$ cm and $r_0 = 0.5$ mm, this inductance is $L = 5.38$ nH. At $f = 2$ GHz, this presents an inductive reactance of $X = \omega L = 67.6\ \Omega$, which is enough to significantly alter the performance of a circuit based on transmission lines with a characteristic impedance of $Z_0 = 50\ \Omega$.

A second example is that of a gap of length $d$ in a wire of radius $r_0$ as shown in Fig. 2.9. A lumped

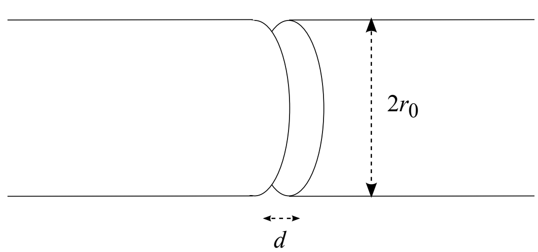



Figure 2.9: A gap in a conducting wire.

element device such as a diode or packaged capacitor might be connected in this gap. In addition to whatever intrinsic impedance the connected element has, the gap itself has a capacitance of

$$C = \epsilon \left[ \frac{\pi r_0^2}{d} + 2r_0 \left( \ln \frac{r_0}{d} + 0.4228 \right) \right] \tag{2.19}$$

to a good approximation if $r_0/d > 1$. The first term in square brackets is of course the parallel-plate contribution to the capacitance; the second term is the extra capacitance due to fringing fields outside the radius of the wire. If an element with a given capacitance is connected in this gap, its capacitance should be substituted for the parallel-plate term in (2.19). If $r_0 = 0.5$ mm, $d = 0.1$ mm and $\epsilon_r = 2.2$, then $C = 0.19$ pF, and

$$X_C = -\frac{1}{\omega C} = -413\,\Omega \qquad \text{at } f = 2 \text{ GHz}$$

which is a value that can significantly alter circuit performance.

## 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. 2.10. A matching network is placed between a load and a transmission line.

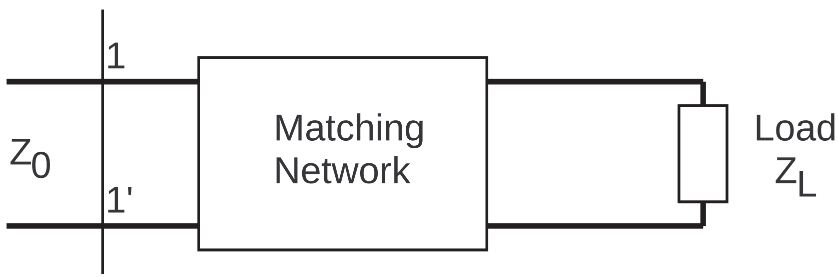



Figure 2.10: A matching network matches a load impedance to a transmission line within a certain frequency bandwidth.

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:

- 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 parasitic 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. The series element may be placed either first or second in the circuit, and each element may be either a capacitor or an inductor, making for eight possible versions of this type of matching network. We will examine one version in detail in the following example.

### Example

Consider two lossless transmission lines with different real characteristic impedances $Z_{01} = R_1$ and $Z_{02} = R_2$ as shown in Fig. 2.11. We wish to place a capacitor in parallel with the second line to change

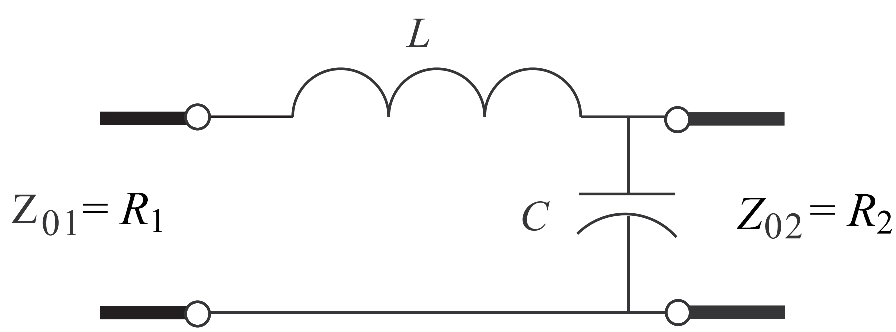



Figure 2.11: Lumped element matching example.

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 $R_1$, and therefore the reflected wave on the first line is eliminated.

It is straightforward to show that this requires:

$$R_1 = j\omega L + \frac{R_2}{1 + j\omega C R_2}$$

We multiply both sides by $1 + j\omega C R_2$, and force the real and imaginary parts of the equation to hold separately. The result is:

$$\frac{L}{C} = R_1 R_2$$

and

$$1 - \omega^2 LC = \frac{R_1}{R_2}$$

We can see from the second of these that this circuit can work only if $R_1 < R_2$ (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:

$$L = \frac{\sqrt{R_1(R_2 - R_1)}}{\omega}; \quad C = \frac{1}{\omega R_2}\sqrt{\frac{R_2 - R_1}{R_1}} \tag{2.20}$$

Evidently, the match only works for a single frequency, if $R_1$, $R_2$, $L$ and $C$ do not change with $\omega$. 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. 2.12. The distance $d$ between the stub and load

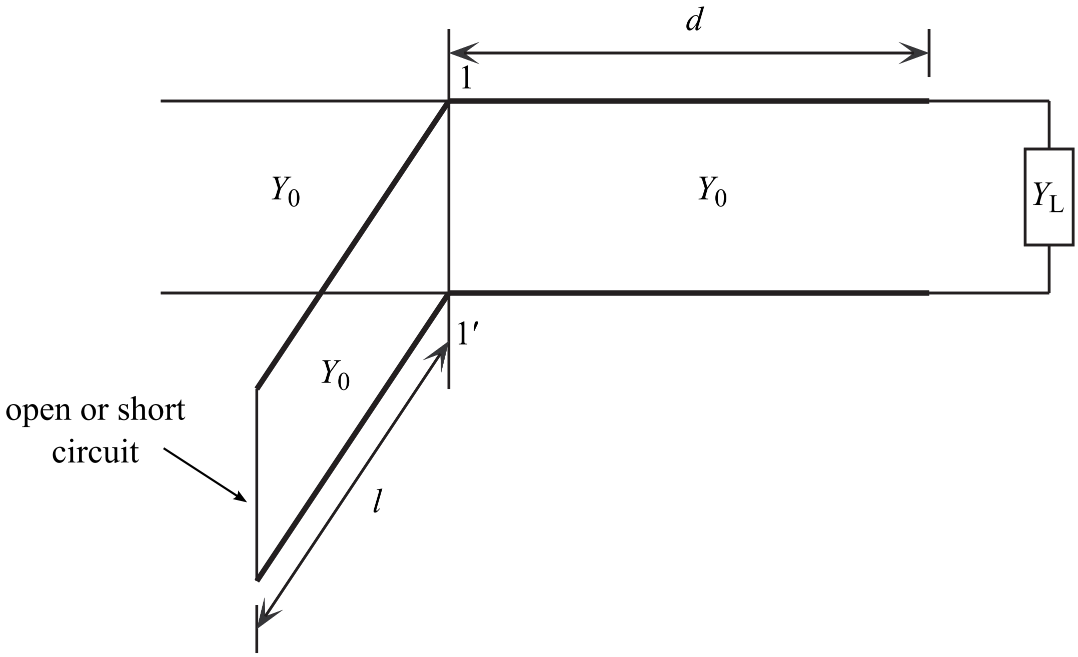



Figure 2.12: Shunt single-stub matching.

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.

### 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, enter the normalized load impedance $z_L = 0.3 + j0.2$, then rotate its locus 180° about the center to obtain the locus of the normalized admittance $y_L = 1/z_L = 2.31 - j1.54$ and the corresponding SWR circle, as shown in Fig. 2.13. The two points of intersection of the SWR circle and the $g = 1$ circle give two possible admittance values $y_1 = 1 - j1.33$ (which we will denote as case 1) and $y_2 = 1 + j1.33$ (which we denote case 2) as shown. 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, respectively.

The stub itself can be either open-circuited or short-circuited at its end, as is the more convenient for a particular kind of transmission line. 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 $y = 0$

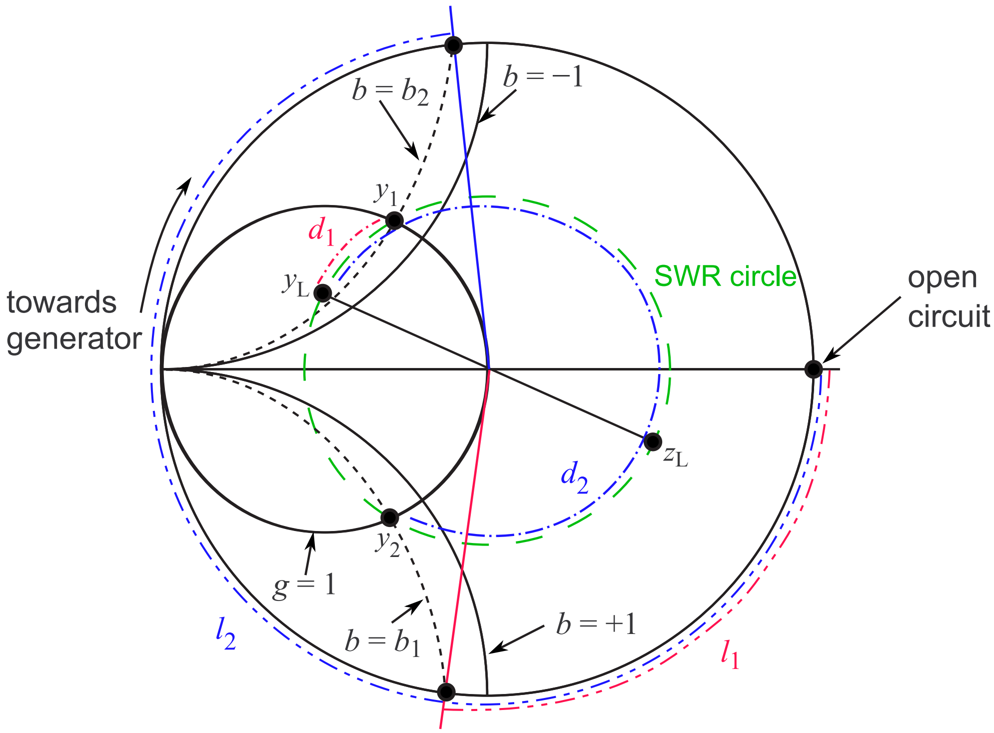



Figure 2.13: Normalized load admittance $y_L$ and stub match calculations on a Smith chart (case 1 calculations shown in red, case 2 in blue).

along the outer edge of the Smith chart ($g = 0$) to the circle corresponding to $b_1$ by $l_1 = 0.147\lambda$ in case 1. Similarly, the solution for case 2 gives $l_2 = 0.353\lambda$.

Do these two matching networks behave in the same way? In order to find that out, let us look at how the two matched circuits described above function as the frequency changes. First we need to say at what frequency we matched the load and what the load is (specifically, how it varies with frequency). 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, which is shown in Fig. 2.14. Case 1 has broader bandwidth than does case 2, and this makes sense since the lengths of the transmission line sections are shorter, so we expect there to be less dependence on wavelength.

### 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

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

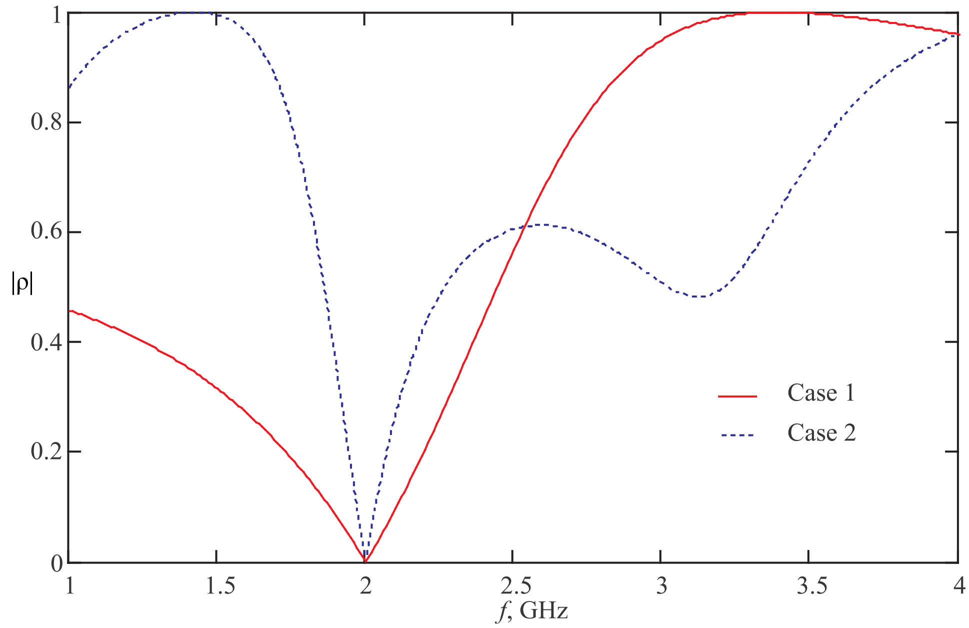



Figure 2.14: Reflection coefficient of single-stub matched circuit as a function of frequency for two different stub matching solutions.

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

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

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

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

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

or

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

If $R_L = Z_0$, the solution degenerates to the single value $x = -X_L/2Z_0$. 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, the length must be

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

while for a short-circuited stub

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

Once again, if the calculated 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 $Z_L =$

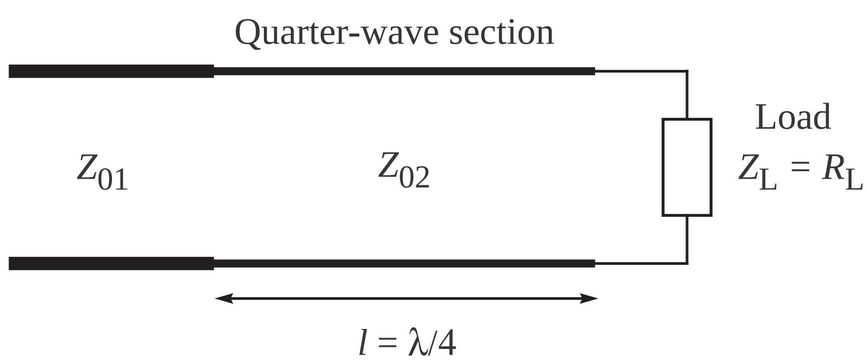



Figure 2.15: Matching with a single quarter-wave section of transmission line.

$R_L$ needs to be matched using a quarter-wave long section of characteristic impedance $Z_{02}$ to a line impedance $Z_0$ at some frequency $f$, as in Fig. 2.15, the length of the matching section will be $l = \lambda/4 = c/4f$ long, and its characteristic impedance will be equal to

$$Z_{02} = \sqrt{Z_{01}Z_L}. \tag{2.30}$$

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 Vector Network Analyzers (VNAs)

$S$-parameters are such a fundamental part of RF and microwave analysis and design that it is important to have a simple but accurate way of measuring them. The vector network analyzer (VNA) is an instrument that is capable of making such measurements quickly (most often for two-port networks, but analyzers capable of handling more than two ports are available). A VNA can measure complex quantities in the frequency domain (analyzers that can measure only magnitudes are called scalar network analyzers or SNAs), and so can determine not only reflection and transmission coefficients but also impedances and admittances. Using the Fast Fourier Transform (FFT), the time domain response of microwave networks can be measured as well. The VNA is a part of every modern microwave lab. Major manufacturers of VNAs are Agilent (formerly Hewlett-Packard), Anritsu and Rohde & Schwarz. 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.

We will present a more detailed description of the operation of a VNA in sections 6.5-6.6 of chapter 6, but we wish here to focus only on a very important aspect of VNA operation—its *calibration*. A real, two-port VNA can be represented as an ideal (error-free) two-port VNA to which an "error box" has been connected at each port, and one between the ports, as shown in Fig. 2.16. Of course, this is

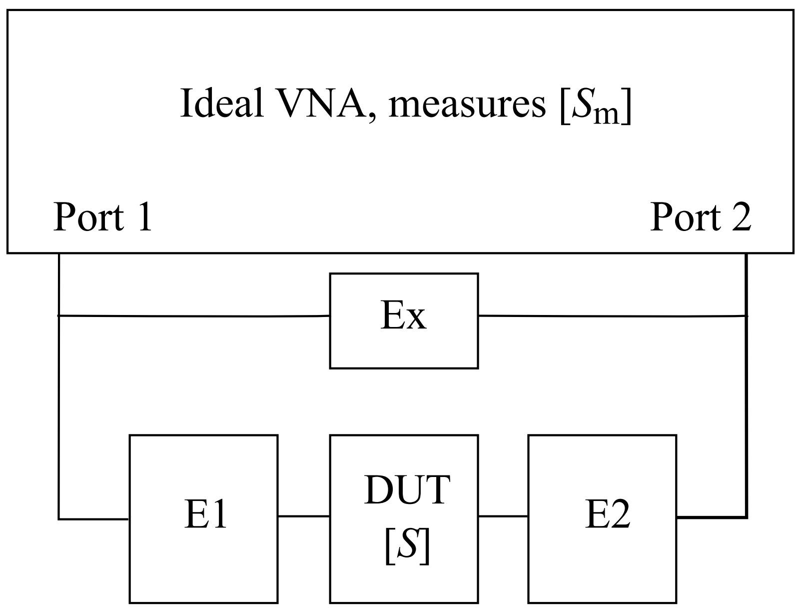



Figure 2.16: Model of a two-port VNA as an ideal VNA with an error box at each port connected to the device under test (DUT), and an error box between port 1 and port 2.

not what the VNA actually looks like inside, but it is an equivalent circuit for it in the same spirit as a Thévenin equivalent circuit is for an ordinary two-terminal network. The ideal VNA measures not the $S$-parameters of the device under test (DUT), but the parameters $[S_m]$ of the composite two-port formed by cascading error box E1 with the DUT and the second error box E2, and the result connected in parallel with an error box Ex representing imperfect isolation effects (leakage between ports 1 and 2) that are important in transmission calibration.

The error boxes account for the many ways in which inaccuracies may enter into the "raw" measured $S$-parameters $[S_m]$: the effect of the test port cables and coaxial connectors, tolerances on components used in the VNA circuits, variations caused by changes in temperature, pressure, humidity, etc. At lower frequencies, the errors of measurement instruments such as multimeters are generally small and can be compensated for on a long term basis (over a number of years), but at microwave frequencies their effect is more serious, and we must perform a calibration of the VNA before every use in order to determine the $S$-parameters of the error boxes and remove their influence on the raw measured data.

Calibration is accomplished by connecting known (or at least partially known) devices known as *calibration standards* (or *cal standards* for short) to the measurement ports of the VNA, and using the measured raw $S$-parameters to deduce information about the error boxes. Once we have this information, we can measure unknown DUTs on the VNA, which will use an internal microprocessor to eliminate the effect of the error boxes from the raw data and display the actual $S$-parameters of the DUT. As an example, if a one-port termination whose reflection coefficient is $S_{11}$ is attached to port 1, the raw reflection coefficient measured at the ideal VNA can be shown to be

$$S_{m11} = r_{1d} + t_1^2 \frac{S_{11}}{1 - r_{1s} S_{11}} \tag{2.31}$$

where the $[S]$ matrix of error box E1 (assumed to be reciprocal) is given by

$$[S_{E1}] = \begin{bmatrix} r_{1d} & t_1 \\ t_1 & r_{1s} \end{bmatrix} \tag{2.32}$$

In the following subsections we will indicate two ways of finding the $S$-parameters of the error boxes using measurements of cal standards. Once the elements of $[S_{E1}]$ have been found, we can determine $S_{11}$ from the "raw" or "uncalibrated" measured value $S_{m11}$ as:

$$S_{11} = \frac{S_{m11} - r_{1d}}{t_1^2 + r_{1s}\left(S_{m11} - r_{1d}\right)} \tag{2.33}$$

### 2.10.1 SOLT Calibration

A widely used calibration technique employs calibration standards of high precision. The most common such method is SOLT (short-open-load-thru) calibration, wherein a short circuit, an open circuit and a matched load are successively connected to one of the ports and measured on the VNA to accomplish the reflection calibration. After this, a thru standard is connected between ports 1 and 2, and transmission measurements carried out. Let us look in detail at the reflection calibration process.

Suppose our cal standards (denoted A, B and C) have $S_{A11} = -1$, $S_{B11} = +1$ and $S_{C11} = 0$ (it is not necessary for the cal standards to be ideal ones like this; the same general procedure will apply with somewhat more complicated equations to solve). Then we have the three equations

$$S_{mA11} = r_{1d} + t_1^2 \frac{-1}{1 + r_{1s}} \tag{2.34}$$

$$S_{mB11} = r_{1d} + t_1^2 \frac{1}{1 - r_{1s}} \tag{2.35}$$

$$S_{mC11} = r_{1d} \tag{2.36}$$

where $S_{mA11}$, $S_{mB11}$ and $S_{mC11}$ denote the raw measured values of $S_{m11}$ at the ideal VNA with cal standards A, B or C connected to port 1, respectively. These equations are easily solved to give

$$t_1^2 = 2\frac{\left(S_{mA11} - S_{mC11}\right)\left(S_{mB11} - S_{mC11}\right)}{S_{mA11} - S_{mB11}} \tag{2.37}$$

$$r_{1s} = \frac{2S_{mC11} - S_{mA11} - S_{mB11}}{S_{mA11} - S_{mB11}} \tag{2.38}$$

$$r_{1d} = S_{mC11} \tag{2.39}$$

An exactly similar procedure applies to the determination of the $S$-parameters of error box E2 at port 2. Calculations related to the transmission calibration to determine Ex are analogous, though algebraically more complicated, and will not be presented here.

### 2.10.2 TRL Calibration

Calibration kits (sets of high precision cal standards) are expensive and are not always available for nonstandard or outdated types of transmission line. Fortunately, there are alternative calibration methods that do not require high precision cal standards and make only a modest sacrifice in accuracy. The most popular of these methods is TRL (thru-reflect-line), in which the "thru" standard is simply a direct connection between ports 1 and 2 (the connectors must be such as to permit this), the "reflect" standard is some load with a high reflection coefficient (its phase can be known only approximately) and the "line" standard is a length of transmission line whose electrical length is known very approximately (e. g., it is closer to $\lambda/4$ than to $3\lambda/4$). Similar algebra to that used for SOLT calibration shows that TRL calibration can not only determine the $S$-parameters of the error boxes, but also can determine precisely the reflection coefficient $\rho_{\rm cal}$ of the reflect standard and the transmission factor $e^{-\gamma l}$ of the line standard (see D. M. Pozar, *Microwave Engineering*, 3rd edition, pp. 193-196 for details).

At a frequency where the length of the line standard is at or near an integer number of half wavelengths, measurement of this cal standard will give the same values (or close to them) as measurement of the thru standard. Because of this, there are frequency limitations imposed on TRL calibrations that are not present in SOLT calibrations. As an example, if the thru standard has zero electrical length, and the length of the line standard in cm is $l_{\rm cm}$, the largest contiguous frequency range in GHz that can be calibrated by the TRL method is

$$\frac{1.67}{l_{\rm cm}\sqrt{\epsilon_e}} < f_{\rm GHz} < \frac{13.33}{l_{\rm cm}\sqrt{\epsilon_e}} \tag{2.40}$$

which is a maximum frequency ratio of 8:1 (there are also higher frequency ranges such as $\frac{16.67}{l_{\rm cm}\sqrt{\epsilon_e}} < f_{\rm GHz} < \frac{28.33}{l_{\rm cm}\sqrt{\epsilon_e}}$, etc., but these have a much smaller relative bandwidth). Modern network analyzers have built-in capabilities for both SOLT and TRL calibrations, and often other calibration methods as well.

## 2.11 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. What is the Smith chart? Which quantities can you plot on it?

4. What do concentric circles centered at the middle of the chart represent?

5. What do circles of constant resistance and those of constant reactance look like?

6. 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?

7. Derive the equations for the constant resistance and reactance circles. Start by writing $\rho = u + jv$, $z = r + jx$ and $\rho = (z-1)/(z+1)$.

8. Why does a full rotation around the chart correspond to half a wavelength?

9. 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.

10. 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.

11. Why is impedance matching so important at high frequencies? Why did you not have to worry about it in your circuits classes?

12. When would you consider using lumped elements for impedance matching?

13. When is it convenient to use a quarter-wave transmission line section for matching?

14. 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 $Z_0$.

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

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

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

17. Without deriving the detailed mathematical equations, explain why there must be a limitation on the frequency range of a TRL calibration. [*Hint*: What would happen if the length of your LINE calibration standard was a multiple of a half-wavelength?]

## 2.12 Homework Problems

**1.** Find the $S$-parameters of the networks (a) and (b) of Fig. 2.17. Are these networks lossless?

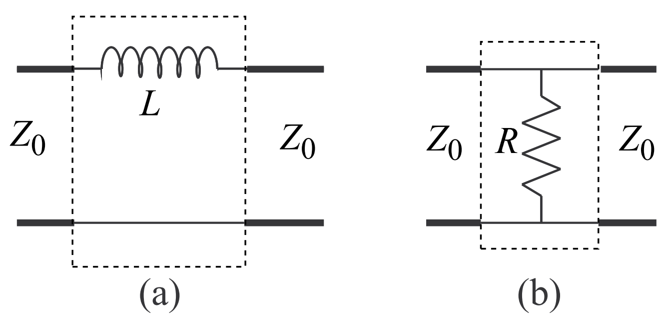



Figure 2.17: Determine the $S$-parameters of the two-port networks shown in (a) and (b).

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 the network of Fig. 2.18, which is a length $l$ of transmission line with characteristic impedance $Z$ connected between two lines of characteristic impedance $Z_0$ as shown. The "electrical length" $\theta$ of the middle section is equal to $\beta l$.

**3.** Repeat problem 1 for the case of the T-network shown in Fig. 2.19.

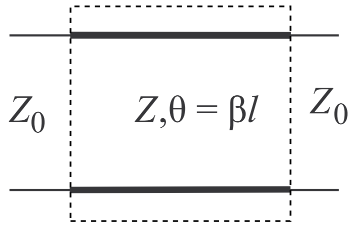



Figure 2.18: Determine the $S$-parameters of the two-port network shown above.

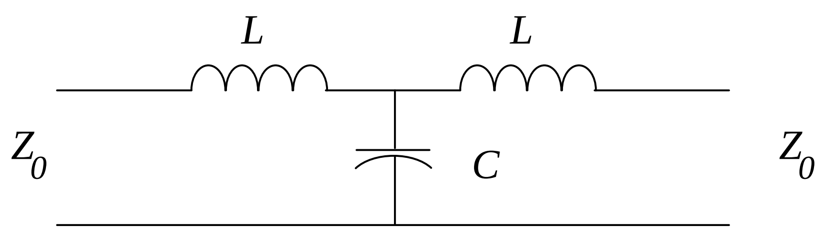



Figure 2.19: Determine the $S$-parameters of the two-port T-network shown.

**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 $\theta \equiv \beta l = 90°$ long ($l = \lambda/4$) at 2 GHz.

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

**6.** 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)?

**7.** 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.

**8.** Repeat problem 4, using a load impedance of $Z_L = 30 - j80\ \Omega$ instead.

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

**10.** 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 (in this case, choose values for the characteristic impedances $Z_{01}$ and $Z_{02}$ of these two sections so as to maximize the bandwidth of the match). Model these circuits in microwave design software, and plot the magnitude of the reflection coefficient for each from 1 GHz to 3 GHz.

**11.** Use an open-circuited single shunt stub to match a $20 - j15\ \Omega$ 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 microwave design software, 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?

**12.** Use lumped elements (capacitors and/or inductors) to match an impedance of $Z_L = 100 + j75\,\Omega$ to a $50\,\Omega$ transmission line using microwave design software. 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.

**13.** Verify eqn. (2.31) and eqns. (2.37)-(2.39).

**14.** Design three different matching circuits to match a 30 Ω load to a 50 Ω transmission line at 2 GHz and again at 10 GHz. Compare your results in terms of bandwidth, power delivered to the load and circuit size when it is implemented in microstrip using a Duroid substrate with relative permittivity $\epsilon_r = 10.2$ and thickness $h = 0.01$ in. You can do this analytically, or using microwave design software.

**15.** Prove that the matching circuit in Fig. 2.11 also works in the reverse direction. In other words, show that the input impedance looking left in Fig. 2.20 is $R_2$, if $L$ and $C$ are given by (2.20).

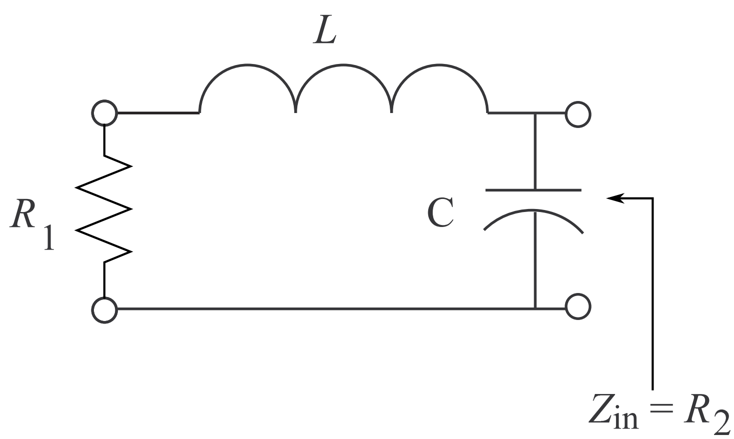



Figure 2.20: Lumped element matching in the reverse direction.

**16.** Consider the lossless, reciprocal 2-port matching network shown in Fig. 2.21(a), which matches the resistance $R_2$ to a different value $R_1$. Show that this same matching network

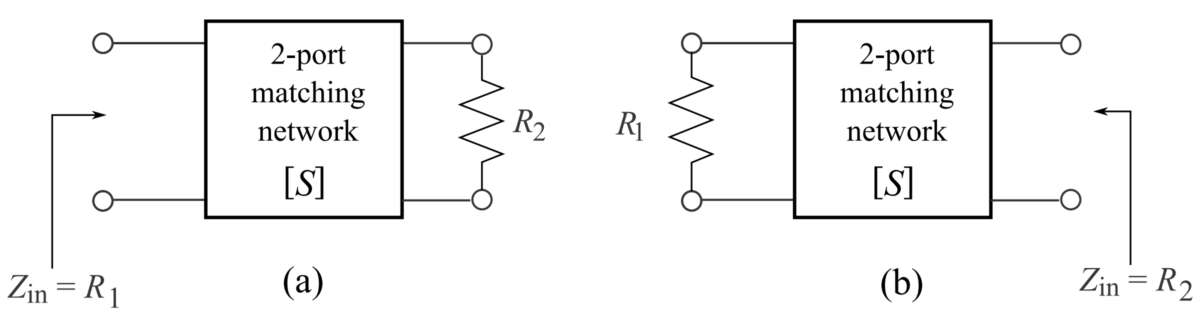



Figure 2.21: General matching network (a) matching $R_2$ to $R_1$; (b) matching $R_1$ to $R_2$.

also matches $R_1$ to $R_2$ as shown in Fig. 2.21(b). In other words, the impedance match works in both directions. Give counterexamples to show that this is not necessarily true if:

(a) the 2-port is lossy,

(b) the 2-port is nonreciprocal, or

(c) either of the loads is a complex impedance ($R_1 \to Z_1$ or $R_2 \to Z_2$).

## Lab 2: Network Analyzer Operation

In this lab, you will learn how to use a modern network analyzer (the Agilent 8753ES). 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 will only need to do an SOL calibration. This type of calibration uses three standards: a short circuit, an open circuit and a matched (50 Ω) 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, etc.), and can even be made of microstrip, waveguide or any other transmission medium.

**Before you use the network analyzer, read the following section on safety (to the instrument).** This is a very expensive piece of equipment, which costs a lot of money even to repair, 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 Ω 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 for most purposes in these experiments).

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 frequency points to 201, set the start frequency to 300 kHz 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)? [Hint: check the settings under SWEEP and AVG.]

Do a full $S_{11}$ one-port calibration using an HP/Agilent 3.5 mm calibration kit (short, open, matched load). Refer to the manual. Make sure the selected cal kit matches the one you are using (not all cal kits are the same, even for the same size connectors). After any calibration, it is a good idea to save the state of the instrument (including the calibration) using the SAVE/RECALL button on the front panel. You will be replacing someone else's previous saved state, so choose a register with a date at least one year old (or at any rate, as old as possible) so you are not erasing what someone else needs. If the analyzer freezes or acts unusually, you can reboot the instrument using the green PRESET button, after which you can recall your saved state.

**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 $S_{11}$ (these may be captured using Agilent's Excel add-in software). 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. If it has not already been done for you, 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 without changing the calibration from previous steps measure the $S_{11}$ of the capacitor and inductor using the General Radio (GR) component mount with adapters to SMA connectors. Then measure the component mount with no component connected, just the empty cavity formed when the cover shield is on. Include the plots of $S_{11}$.

**Q5:** What are the factors that affect the accuracy of these measurements?

The SMA to GR874 adapters connected between the component mount and the network analyzer port 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 include 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 (here you should be sure to use a torque wrench on the 3.5 mm connectors, since the much larger GR874 lines and connectors tend to unscrew the smaller connectors when you manipulate them). Additionally attach a 90° GR874 bend to port 2 to allow for easier connection of devices between port 1 and port 2. The GR874 connector at port 1 should be rotated so that it can readily be connected to port 2. Since no commercial cal kits are made with GR874 connectors, we will have to use a TRL calibration, which only needs approximate information about the cal standards and does not require them to be made with high precision.

A TRL calibration is a two-port calibration, even if you only intend to do one-port measurements. Select the "USER KIT" as your cal kit; it should be designated "GR874TRL" or something similar. The three standards you will use are a "THRU" (no physical standard, just a direct connection between ports 1 and 2), a "REFLECT" (a model WN short circuit) and a "LINE" (a 10 cm section of GR874 air line). You must set the start frequency to 170 MHz and the stop frequency to 1.1 GHz, because the TRL "LINE" type cal standard determines a limited frequency range over which it can be used; multiple standards are necessary if more than an 8:1 bandwidth is desired. Perform a TRL calibration using these standards to eliminate the effect of the adapters. The calibration information for these standards has already been stored in the network analyzer by the instructor or TA.

**Q6:** Where is the reference plane located now?

**Q7:** The LINE "standard" is not really a uniform section of transmission line, because of the dielectric supports at each end. If all such dielectric supports were the same, would this affect the accuracy of the calibration or not, and why? What other possible sources of error in this TRL calibration can you think of?

Make an independent check of your calibration by connecting a 50 Ω coax load to port 1 and measuring $S_{11}$. What is the maximum magnitude of $S_{11}$? Next, 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 $S_{11}$.

**Q8:** What differences do you see between the measurements made using the different calibrations?

Connect a 50 Ω coaxial load to the end of a 10 cm long GR air line marked "not constant $Z_0$" (i. e., *not* the one you used in the TRL calibration), and measure $S_{11}$. Include the plot in your lab report.

**Q9:** Assuming that the characteristic impedance of this 10 cm line is uniform, but not equal to 50 Ω, what is the value of its $Z_0$? [Hint: plot the value of $S_{11}$ on a Smith chart for various values of $Z_0$ and choose the value that most closely matches your measurement.] Can you suggest some reasons why the measured plot differs from what is predicted theoretically?

## 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 again perform a TRL calibration using the same GR874 calibration standards as before. 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.

**Q10:** 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? Give a reason for the difference. What is the measured SWR of the load equal to?

Now place the connecting tee and the shorted stub and adjustable line length in the system as shown in Fig. 2.22. Use the GR coaxial 100 Ω load at the right end of the adjustable length line. Since the

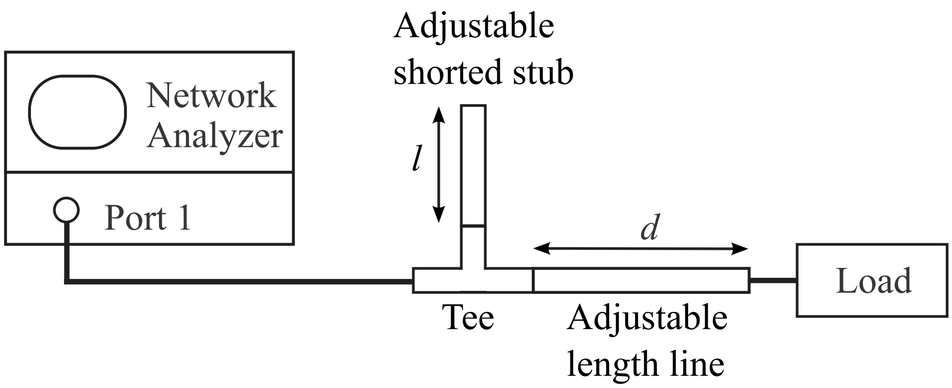



Figure 2.22: Coaxial stub matching: the quality of the match is measured with a network analyzer.

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 for the 8753ES network analyzer). Push the Marker Fctn 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 press FORMAT → SMITH CHART to see the admittance chart display.

With the load placed at the right end of the line, we will now 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 (if time is short, you may skip one of them).

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 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 **Q10**. 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.

**Q11:** 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. Make sure there is *some* resistance in the load. 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, and move a line length $d$ towards the load.

**Q12:** What is the impedance of the load? If you connect the load directly at the reference plane, do you measure approximately the same impedance?

# 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, 1m̂W – 10 W and larger than 10 W, respectively. Power is most commonly expressed in decibels 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* (the rate of variation is so rapid that it is difficult or impossible to directly measure instantaneous power or phase at RF frequencies or higher). 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

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

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 $T_d$:

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

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):

$$P_{\mathrm{pulse}} = \frac{P_{\mathrm{av}}}{\mathrm{Duty\ Cycle}}. \tag{3.5}$$

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, the basic idea of which is shown in Fig. 3.2. The thermistor ($R_T$) is DC-biased through the bridge to a known value of resistance. When no RF power is incident on the thermistor, its resistance is determined solely by the DC power $P_{T0}$ dissipated in it, which can be determined from a knowledge of the DC voltage $V_0$ across the bridge, and the value of the resistance $R$ in the upper arm of the bridge:

$$P_{T0} = \frac{V_T^2}{R_T} = \frac{V_0^2 R_T}{(R+R_T)^2} \tag{3.6}$$

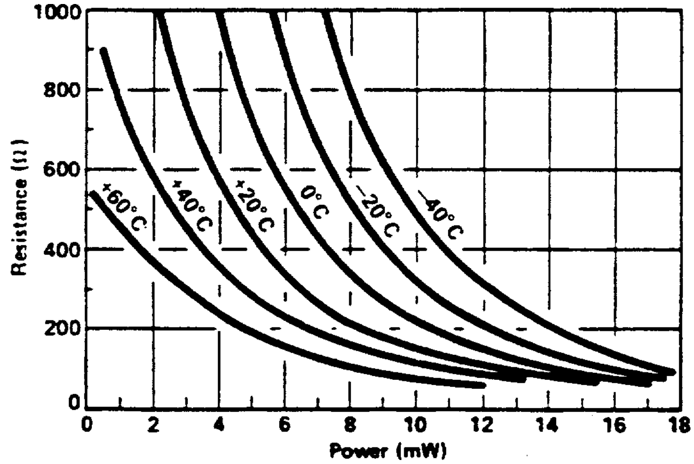



Figure 3.1: Characteristic curves of a typical thermistor bead.

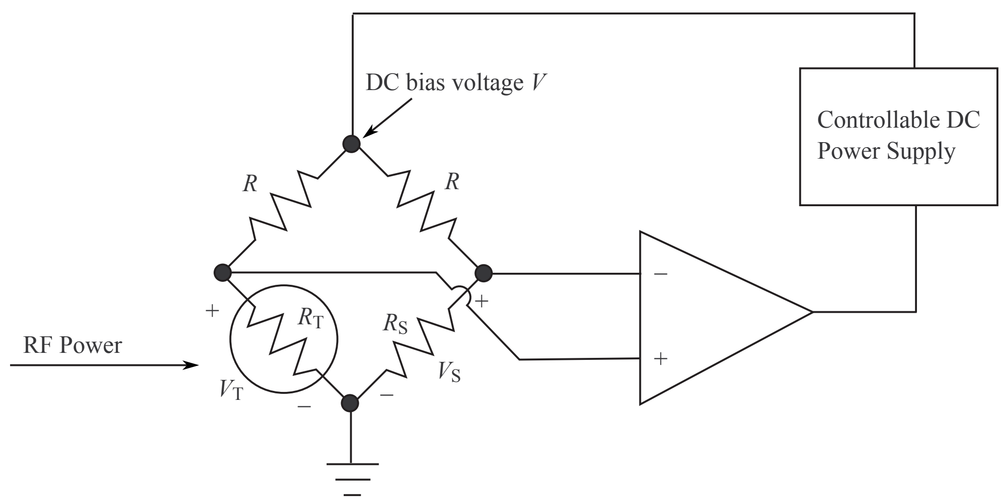



Figure 3.2: Diagram of a self-balanced Wheatstone bridge.

A particular desired value of $R_T$ can be assured if we select the resistance $R_S$ in the opposite arm of the bridge to be equal to this desired value, and then adjust the bias voltage $V$ to the value $V_0$ that balances the bridge (so that $V_T = V_S$ in Fig. 3.2).

When there is incident RF power applied to the thermistor, this power added to the DC power already present will tend to decrease $R_T$, 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 to a new value $V$ in such a way as to re-balance the bridge. The new DC bias voltage will be lower than before, 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. The RF power is readily calculated from knowledge of the old and new DC voltages across the bridge:

$$P_{\rm RF} = P_{T0} - P_T = \frac{V_0^2 R_T}{(R+R_T)^2} - \frac{V^2 R_T}{(R+R_T)^2} = \frac{(V_0^2 - V^2)R_T}{(R+R_T)^2} \tag{3.7}$$

and is translated into a power 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.

This process is automated in commercial thermistor-based power meters such as the HP 432A, a diagram of which is shown in Fig. 3.3. The main parts are two self-balanced bridges (an "RF" bridge

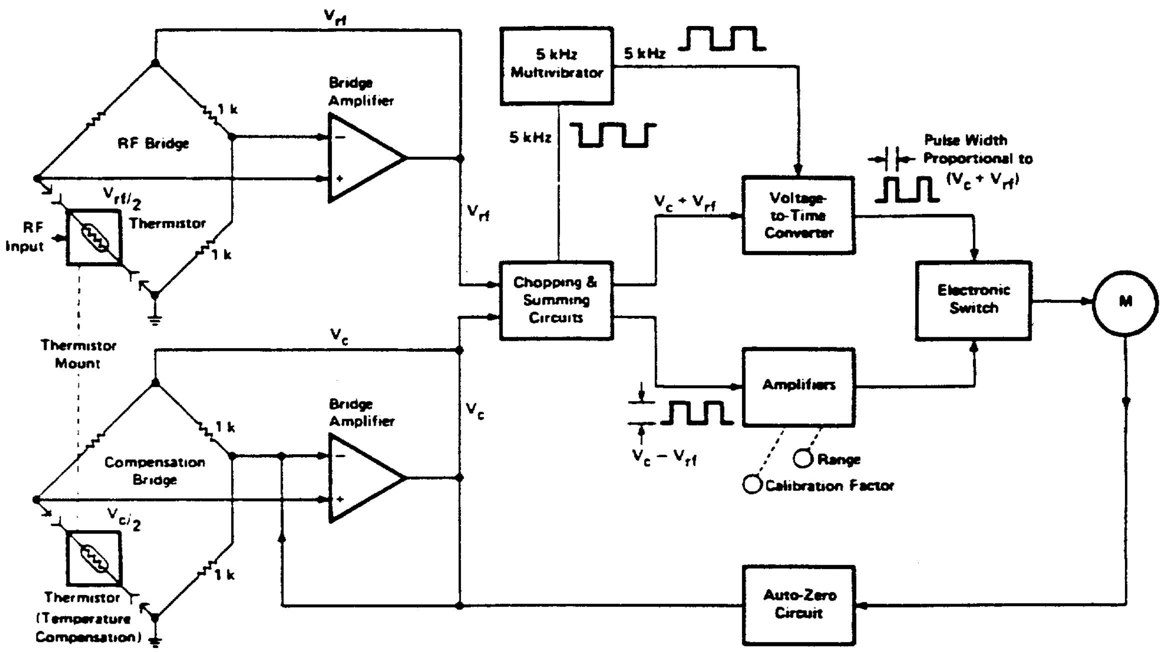



Figure 3.3: Diagram of the HP 432A power meter.

and a compensating bridge), a logic section, and an auto-zeroing 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 $V_{\rm rf0}$, which is the value of $V_{\rm rf}$ in the absence of RF input power. When RF power is applied to the detecting thermistor, $V_{rf}$ decreases, so that

$$P_{\rm rf} = \frac{{V_{\rm rf0}}^2}{4R} - \frac{{V_{\rm rf}}^2}{4R} = \frac{{V_c}^2 - {V_{\rm rf}}^2}{4R} = \frac{1}{4R}\left(V_c - V_{\rm rf}\right)\left(V_c + V_{\rm rf}\right). \tag{3.8}$$

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 (3.8). Ambient temperature changes cause changes in both $V_c$ and $V_{\rm rf}$, such there is no change in $P_{\rm rf}$. In the lab experiment, you will be balancing 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

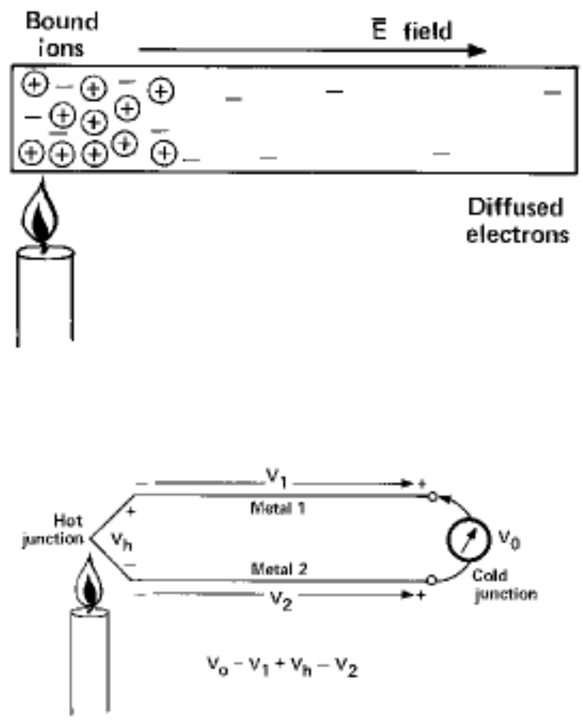



Figure 3.4: Simplified explanation of thermocouple operation. The total measured EMF is a resultant of several thermoelectric voltages generated along the circuit ($V_1$ and $V_2$ are Thomson EMFs, while $V_h$ is a Peltier EMF).

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. A photograph of a thermocouple silicon chip is shown in Fig. 3.5.

The thermocouple used in the HP8481A sensor head is shown in Fig. 3.6. 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

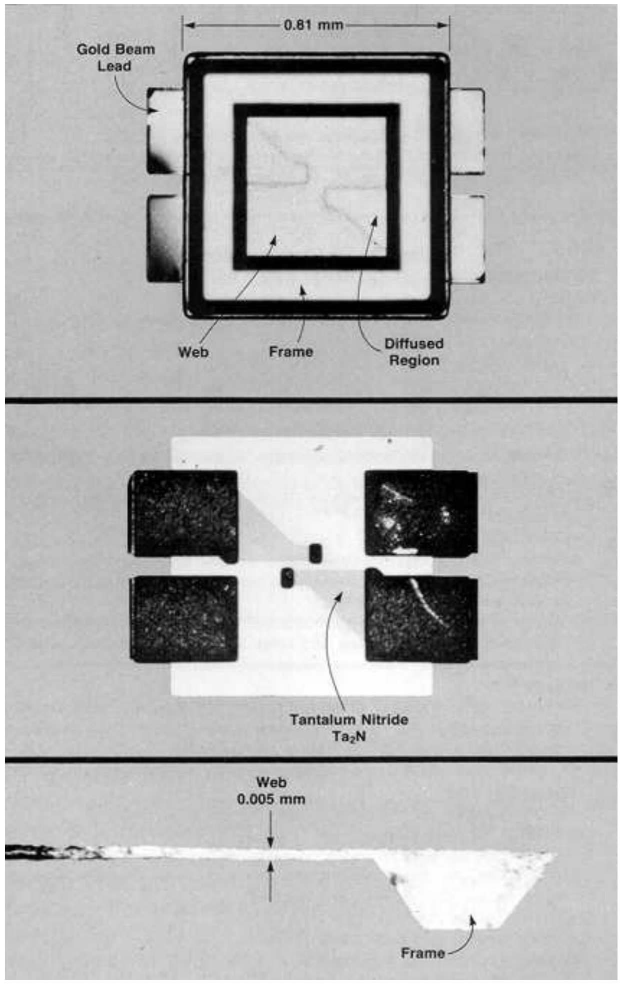



Figure 3.5: Photo-micrograph of the structure of the HP 8481A thermocouple chip on a thin silicon web.

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. 3.7.

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.

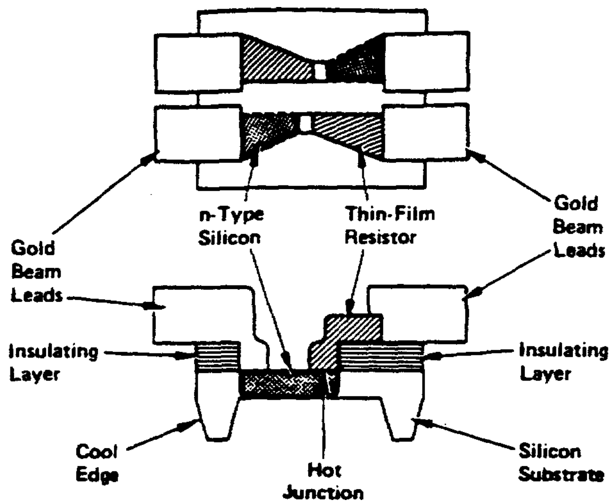



Figure 3.6: Structure of the thin-film thermocouple used in the HP8481A sensor.

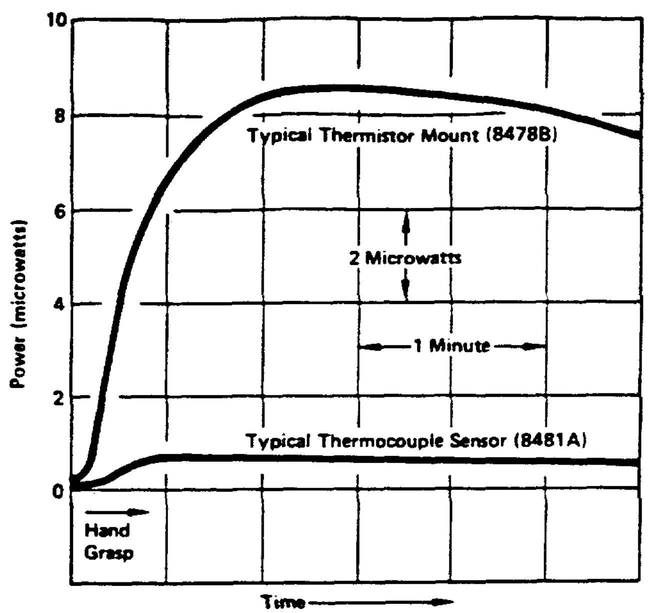



Figure 3.7: Difference in behavior of thermistor and thermocouple when grasped by the hand. (From HP Application Note 64-1.)

## 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. The schematic symbol for a Schottky diode is shown in Fig. 3.8, but often an ordinary diode symbol is used instead.

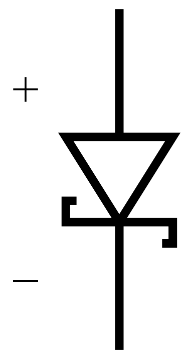



Figure 3.8: Schematic symbol and reference directions for a Schottky diode.

At sufficiently low frequencies, a Schottky diode behaves as a nonlinear resistor for which the current-voltage dependence obeys approximately the following equation, shown graphically in Fig. 3.9:

$$I(V) = I_S\left(e^{\alpha V} - 1\right), \tag{3.9}$$

where $I_S$ is the so called reverse saturation current (or leakage current) whose value is small (between about $10^{-5}$ A and $10^{-15}$ A), and $\alpha = n/(25\,\text{mV})$ at room temperature. $n$ is called the ideality factor, and it depends on the diode structure. For a typical Schottky diode, $n = 1.2$, while 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

$$V = V_0 + v, \tag{3.10}$$

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:

$$\begin{aligned} I(V) &= I_0 + v\left.\frac{dI}{dV}\right|_{V_0} + \frac{1}{2}v^2\left.\frac{d^2I}{dV^2}\right|_{V_0} + \ldots \\ \left.\frac{dI}{dV}\right|_{V_0} &= \alpha I_S e^{\alpha V_0} = \alpha(I_0 + I_S) = G_d = \frac{1}{R_j}, \\ \left.\frac{d^2I}{dV^2}\right|_{V_0} &= \alpha^2(I_S + I_0) = \alpha G_d = {G_d}'. \end{aligned} \tag{3.11}$$

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

$$I(V) = I_0 + i = I_0 + vG_d + \frac{v^2}{2}{G_d}' + \ldots \tag{3.12}$$

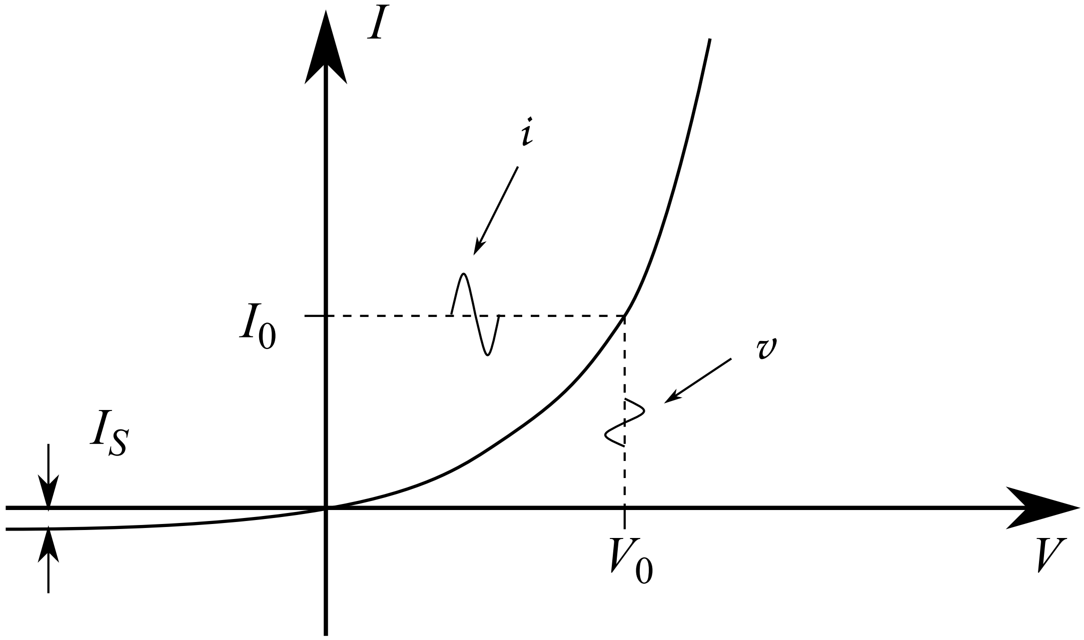



Figure 3.9: Current-voltage characteristic of a diode.

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

$$v = \sqrt{2} V_{\mathrm{RF}} \cos \omega t \tag{3.13}$$

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

$$v^2 = 2V_{\mathrm{RF}}^2 \cos^2 \omega t = V_{\mathrm{RF}}^2 (1 + \cos 2\omega t) \tag{3.14}$$

we can gather the terms from (3.12) which are constant in time:

$$I_{DC} = I_0 + \frac{1}{2} G_d' V_{\mathrm{RF}}^2 \tag{3.15}$$

All other terms from (3.12) 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, and typically the diode response is low-pass filtered before connection to the voltmeter. This assures that only (3.15) will register on the instrument, and the RF behavior of the circuit under test is not modified (see the discussion on bias networks in section 3.5 below).

Now, the RF power delivered to the diode junction is

$$P_{j,\mathrm{RF}} = \frac{V_{\mathrm{RF}}^2}{R_j} \tag{3.16}$$

so by (3.15) and (3.11),

$$I_{DC} = I_0 + \frac{\alpha}{2} P_{j,\mathrm{RF}} \tag{3.17}$$

By (3.9), the DC voltage at the junction is then

$$V_{j,DC} = \frac{1}{\alpha} \ln \left( 1 + \frac{I_{DC}}{I_S} \right) = \frac{1}{\alpha} \ln \left( 1 + \frac{I_0}{I_S} + \frac{\alpha P_{j,\mathrm{RF}}}{2I_S} \right) \tag{3.18}$$

Since typical values of the part of $V_{j,DC}$ due to $P_{j,\mathrm{RF}}$ are tens or perhaps a few hundreds of millivolts, it is usual to set the DC bias voltage $V_0 = 0$ so that the small DC voltage proportional to the RF power is not lost in noise. If $I_0 = 0$, then

$$V_{j,DC} = \frac{1}{\alpha} \ln \left( 1 + \frac{\alpha P_{j,\mathrm{RF}}}{2I_S} \right) \tag{3.19}$$

Two limiting cases of (3.19) are of interest. If $P_{j,\mathrm{RF}} \ll P_{r0}$, where

$$P_{r0} = \frac{2I_S}{\alpha} \qquad \text{(typical values would be between -20 dBm and 0 dBm)}$$

then

$$V_{j,DC} \simeq \frac{P_{j,\mathrm{RF}}}{2I_S} = \frac{1}{\alpha} \frac{P_{j,\mathrm{RF}}}{P_{r0}} \tag{3.20}$$

This is the so-called square-law approximation, valid when the RF voltage $v$ is a small enough signal, meaning that the input power at the diode is not too large. In this range, $V_{j,DC}$ is proportional to $P_{j,\mathrm{RF}}$ in watts. On the other hand, if $P_{j,\mathrm{RF}} \gg P_{r0}$ we have

$$V_{j,DC} \simeq \frac{1}{\alpha} \ln \left( \frac{\alpha P_{j,\mathrm{RF}}}{2I_S} \right) = \frac{1}{\alpha} \ln \left( \frac{P_{j,\mathrm{RF}}}{P_{r0}} \right) \tag{3.21}$$

Thus, in this regime $V_{j,DC}$ is proportional to $P_{j,\mathrm{RF}}$ in dBm plus a constant, and for such large input power levels the diode is said to be saturated, the dependence of $V_{j,DC}$ on $P_{j,\mathrm{RF}}$ being a fundamentally nonlinear one. The different regimes of a typical diode detector with respect to the input power level are shown in Fig. 3.10. The quantity $P_{r0}$ thus has the meaning of a reference power level representing the transition between square-law behavior and saturation of the diode. If it is desired to measure

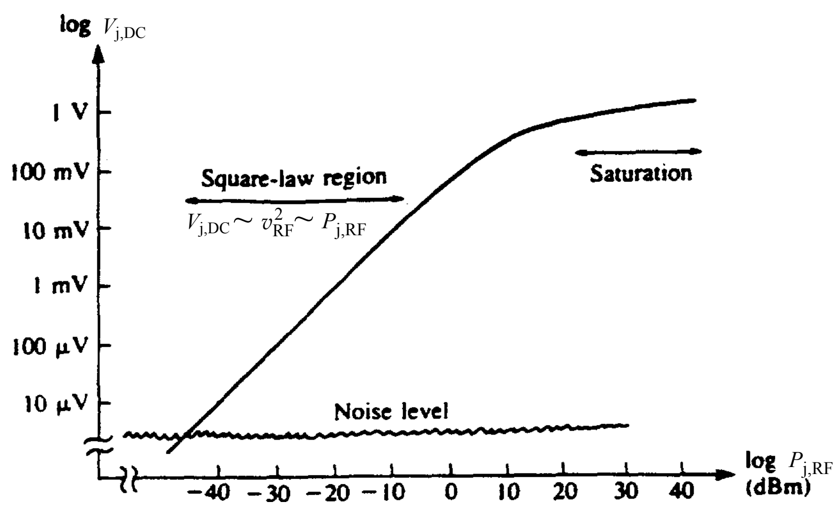



Figure 3.10: Diode detector regimes of operation.

higher levels of RF power, a resistive attenuator can be placed before the diode in order to make sure of operation in the square-law region and prevent damage to the diode.

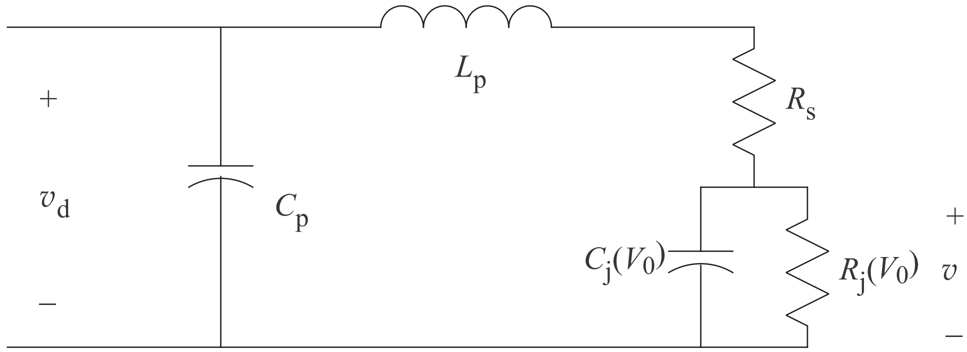



Figure 3.11: Equivalent small-signal AC circuit for a Schottky diode.

As operating frequency increases, other effects must be accounted for when modeling the behavior of a diode. A typical *AC* small-signal equivalent circuit valid at microwave frequencies is shown in Fig. 3.11. $L_p$ and $C_p$ are due to the diode package—$L_p$ is a series inductance due to the wire leads, and $C_p$ is a shunt capacitance of the contacts. The resistance $R_s$ is that of the wire leads of the diode and the spreading of the current at the connection of the leads.

The capacitance $C_j$ due to the diode's *pn* junction is dependent on the total voltage $V_j$ across the junction, which can be approximated by the DC bias voltage $V_0$. The value of the resulting capacitance is typically a fraction of a pF. The RF resistance of the diode is obtained from the slope of the *I*-*V* curve as $R_j = 1/(\alpha I_s)$ (if $V_0 = 0$), and is usually a few k$\Omega$, for leakage currents $I_S$ on the order of $10\,\mu$A. 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.

It should be noted that the passive impedances in a diode's equivalent circuit will generally be different at higher harmonic frequencies than at the fundamental, and furthermore the DC equivalent circuit will be very different than at RF. In the latter case, a Thévenin or Norton equivalent circuit for the diode junction will apply, as shown in Fig. 3.12. The DC junction resistance of the diode is given by

$$R_{DC} = \frac{V_{j,DC}}{I_{DC}} \tag{3.22}$$

using (3.17) and (3.18).

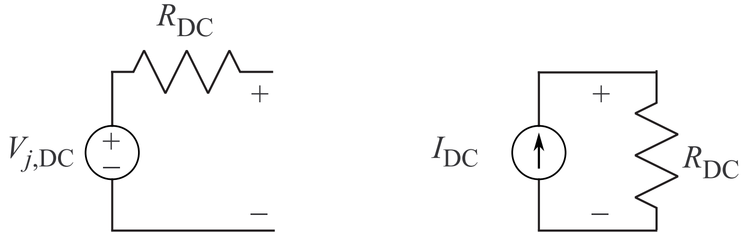



Figure 3.12: DC equivalent circuits for a diode junction.

## 3.5 DC Biasing Networks

Providing a DC bias voltage to a diode must be done in such a way so as not to disrupt the RF operation of the circuit. The following general principles must be followed in designing a network to provide this bias voltage.

1) The bias network needs to be "invisible" to the RF signals, i. e. be as close to an open circuit as possible. We do not wish to lose any of the RF power to the biasing circuit and power supply.

2) Conversely, 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 to an amplifier or other sensitive device.

3) These constraints must hold over all frequencies that might be important to operation of the circuit, including those outside the primary design frequencies of interest. This is because, as we will see in Lectures 5 and 8, unwanted oscillations may occur when amplification is present in the circuit.

In order to satisfy the first criterion, the DC bias lines need to have inductance in series with the power supply as shown in Fig. 3.13. It is difficult to make an inductor at microwave frequencies due to

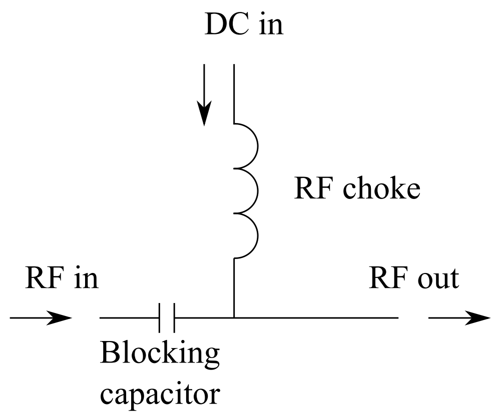



Figure 3.13: A bias Tee equivalent circuit.

parasitics, so another option is to design the bias lines with the characteristics of a low-pass filter, as will be discussed in section 8.6. In order to satisfy the second requirement, a DC blocking capacitor needs to be added to the circuit in series with the RF signal path. A bias tee is similar to a single unit cell of an artificial transmission line of the type discussed in section 1.7, so performance could in principle be improved by adding more sections to this circuit. Capacitors are not ideal shorts, nor are inductors ideal opens at microwave frequencies (they have parasitics), so they need to be taken into account in microwave frequency designs. More details about biasing will be given in Lecture 8.

## 3.6 Measuring Reflection Coefficients: The Slotted Line

We have seen that time-average power is virtually the only thing about a microwave signal that can be easily measured. In this section, we will see one way that such power measurements can be used indirectly to measure voltages, impedances and reflection coefficients. This technique uses a slotted-line configuration, which enables direct sampling of the electric field amplitude (via diode detection) of a standing wave. In modern labs, network analyzers are used to measure impedances and $S$-parameters. 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 can help you get a better feeling for some quantities than staring at the network analyzer screen.

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 complex reflection coefficient

$$\rho = |\rho| e^{j\theta}. \tag{3.23}$$

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

$$|\rho| = \frac{SWR-1}{SWR+1}. \tag{3.24}$$

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

$$\theta = \pi + 2\beta l_{min}. \tag{3.25}$$

(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 $\theta$, and therefore find $\rho$ from (3.23). From this, we can find the unknown complex load impedance to be

$$Z = Z_0 \frac{1+\rho}{1-\rho}. \tag{3.26}$$

We will illustrate the procedure by means of an example.

### 3.6.1 Example

We will consider the example of a load connected to a $50\,\Omega$ air-filled coaxial slotted line. Any slotted line measurement must be preceded by a calibration step: the location of the *proxy planes* that are an integer number of half wavelengths from the load position. We do this by 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. 3.14(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, which are the locations of the proxy planes. [Question: What are the wavelength and the frequency equal to?]

The short is next replaced by the unknown load, for which the SWR is measured to be 1.5 and the voltage minima (not as sharp as when the short was connected) are found at $z = 0.72$, 2.72, and 4.72 cm, as shown in Fig. 3.14(b). From these voltage minima, knowing that they repeat every $\lambda/2$, we can find the distance $l_{min} = (4.2 - 2.72)\,\text{cm} = 1.48\,\text{cm} = 0.37\lambda$. Then

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

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

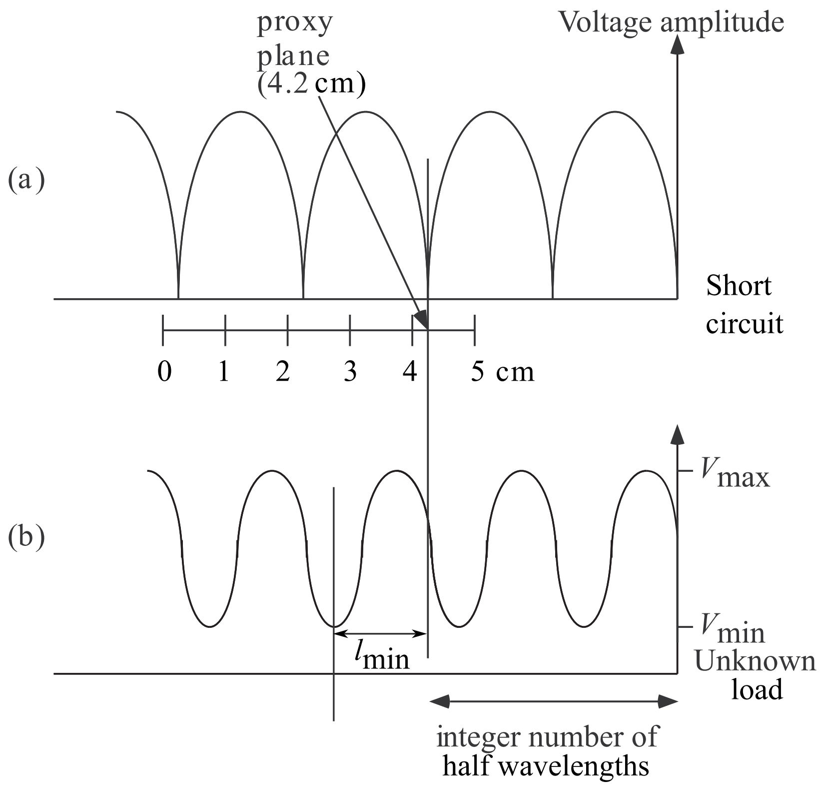



Figure 3.14: Voltage standing waves on a line terminated in a short (a) and an unknown load (b).

## 3.7 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?

13. Do a Google search to obtain a copy of Agilent Application Note 1449-3 on uncertainty in power measurement. Read it and give a general summary (in one paragraph) of the factors that affect the accuracy of power measurement at RF and microwave frequencies. If a power measurement is made in a system for which the generator and load reflection coefficients may be as high as 0.2, what is the maximum uncertainty (in dB) of this power measurement?

## 3.8 Homework Problems

**1.** Express low ($\ll 1\,\text{mW}$), medium ($1\,\text{mW}$) and high power levels in dBm and dBW (referenced to $1\,\text{W}$).

**2.** In the lab, you will characterize a thermistor power detector using a bridge circuit as shown in Fig. 3.15. Assume that the resistors $R$ and $R_S$ are known. For a measured DC voltage

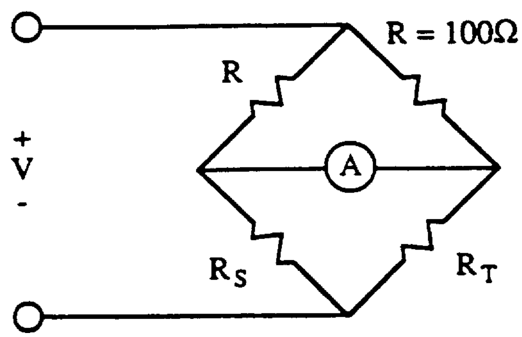



Figure 3.15: Bridge circuit for characterizing the thermistor $R_T$

$V$ 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. 3.16. 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 total resistance of $100\,\Omega$. 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 by a DC signal, and by an RF signal? 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?

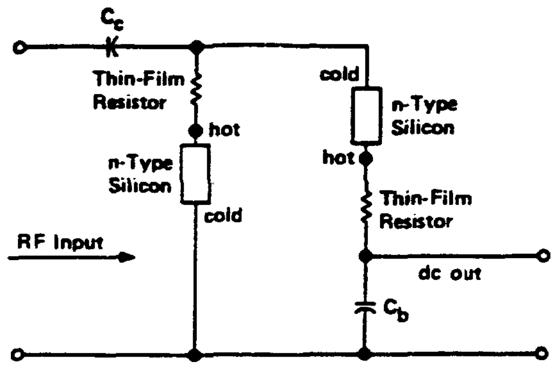



Figure 3.16: Diagram of the HP8481A dual-thermocouple coaxial mount.

**4.** The equivalent circuit parameters of a packaged diode, Fig. 3.11, 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 microwave design software to calculate and plot the complex RF impedance and magnitude of the reflection coefficient 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 (3.9)-(3.12) are those across and into $R_j$ in Figure 3.11). Is the diode well matched to 50 Ω? If so, over what frequency range?

**5.** Assuming the diode equivalent circuit of problem 4, suppose an incident wave $a_1 = 0.001$ W$^{1/2}$ is incident on the diode from the left, from a transmission line with $Z_0 = 50$ Ω. 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 4? To determine this, first use microwave design software 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.

**6.** Refer to problem 8 of Lecture 1. Use 18 sections of transmission line whose electrical length is 20° each. Perform a microwave design software 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? In terms of the slotted-line method of section 3.6, is its value consistent with the value of $Z_L$ that you used?

**7.** 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.*

**8.** Suppose you do not have a short circuit to use in the calibration step of the slotted-line measurement procedure, but you do have an accurately known capacitor whose capacitance is $C_0$. How should you change the measurement process in order to make use of this capacitor?

**9.** Find the latest version of Agilent's Application Note 1449-3, "Fundamentals of RF and Microwave Power Measurements (Part 3): Power Measurement Uncertainty per International Guides" online, and read it. Give a general summary (in one paragraph) of the factors that affect the accuracy of power measurement at RF and microwave frequencies. If a power measurement is made in a system for which the generator and load reflection coefficients may be as high as 0.15, what is the maximum uncertainty (in dB) of this power measurement?

## Lab 3: 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.

In this part of the lab you will measure the reflection coefficients of the various waveguide-mounted devices used to measure power. This must be done at the Agilent 8719 network analyzer because the other VNAs in the lab only operate up to 6 GHz. Since there is only one model 8719 VNA in the lab, your group may have to wait to perform this part until another group is finished. This part may be done at any time during the lab period.

First, bring the power meter sensor (attached to a coax-to-waveguide adapter) to the network analyzer station. The VNA has already been calibrated for you with a waveguide cal kit for the frequency range 8-12 GHz. Connect the waveguide adapter to port 1 of the VNA and measure $S_{11}$, including a plot of $|S_{11}|$ in your report.

**Q1:** In view of your measured values of $S_{11}$, how accurate should power measurements with the power meter be?

Next, take the thermistor mounted in an X-band waveguide (HP X487B) and attach it to port 1 of the VNA. Take the bridge network and decade resistor box and connect these to the thermistor BNC connector, DC power supply and multimeter as shown shown in Figure 3.17 and discussed in the lecture. Whenever connecting waveguide-mounted thermistors or diodes, it is important that the ground (shield)

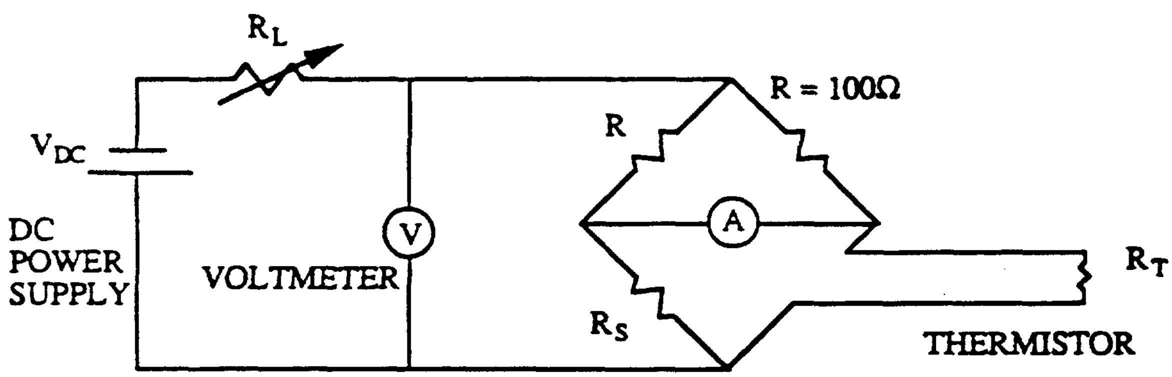



Figure 3.17: Bridge for measuring the DC characteristics of a thermistor.

of all cables be connected to the ground side of the voltmeter or any other device in order to avoid ground loops and erroneous readings. You will measure the DC properties of the thermistor and observe how these affect the value of $S_{11}$ measured in the waveguide.

**Q2:** If there is no current flowing through the ammeter, what does that tell you about the values of the resistors in the bridge? Explain.

**Q3:** For a known value of $R_S$ in a balanced bridge, you will measure the voltage $V$ across the bridge 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 $V$ and the value of $S_{11}$ on the VNA. Decrease $R_S$ by $100\,\Omega$ and repeat the measurement, recording the resistance of the thermistor, the voltage $V$ and the value of $S_{11}$ at 9 GHz. Continue until you reach $R_S = 100\ \Omega$. Make a table with columns $R_T$, $V$, $P_{DC}$ (calculated from your answer to **Q2**) and $|S_{11}|$. 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.

**Q4:** Why does $S_{11}$ change as the value of $P_{DC}$ is changed? To the extent that $S_{11} \neq 0$, what effect will this have on the accuracy of power measurements made using the thermistor? Based on this criterion, what is the best DC operating point $P_{DC}$ at which to use the thermistor (i. e., at what value of $R_T$ is the thermistor best matched to the waveguide)? Is there any other advantage to choosing this operating point?

Finally, while you are still at the VNA station, take the waveguide-mounted diode (the "crystal detector" that will be used in parts III and IV of this lab) to the network analyzer station, connect it to port 1 and measure $S_{11}$, including a plot of $|S_{11}|$ in your report.

**Q5:** What effect will the fact that $S_{11} \neq 0$ have on the accuracy of diode power measurements?

## 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 now connect the thermistor, bridge, etc. in the setup shown in Figure 3.18. Connect the thermistor to the bridge, DC power supply and multimeter next to

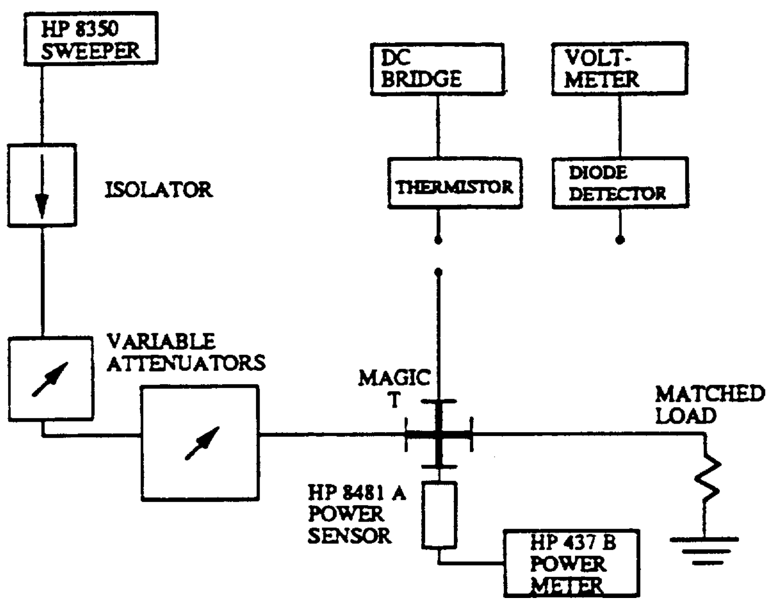



Figure 3.18: Setup for measuring microwave power with a waveguide mounted thermistor or diode detector.

the sweep oscillator/attenuator setup. Set the sweeper at 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 300 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. We will use the equal power splitting property of a magic T as described in chapter 6 to monitor the power going into the thermistor connected to one port, with the HP 437B power meter and sensor connected to the other.

Connect the setup shown in Figure 3.18. Reset both attenuators to 0 dB. Adjust the bridge resistors for a thermistor resistance of 300 Ω with no RF power input (RF power off on sweeper). Now turn the oscillator RF power on adjust its output so that the power meter reads 5 dBm, and balance the bridge by changing $R_L$. Record the DC voltage and the power measured by the HP437B power meter.

**Q6:** 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, changing the incident RF power levels by setting one of the attenuators to 1 dB, 2 dB, . . . , 10 dB in succession. Fill in a table as shown in Table 3.1. Plot, on a linear scale, the

| Power from Power Meter | | DC Voltage $V$ Across Bridge | Calculated RF Power |
|---|---|---|---|
| (dBm) | (mW) | (V) | (mW) |
| | | | |

Table 3.1: Thermistor RF power measurements.

power measured by the power meter versus the RF power measured by the thermistor bridge. Include your graph.

# Part III.

Before performing power measurements with the diode detector, you will determine some of its nonideal characteristics. To determine its DC Thévenin equivalent circuit, connect the diode BNC connector to the resistor decade box as shown in Fig. 3.19. Connect the waveguide port of the diode detector to the output of the variable attenuators. Connect a DC voltmeter across the decade resistor to measure $V_{DC}$ as shown.

First, set the decade resistance $R_{\text{decade}} = \infty$ (i. e., disconnect the decade resistor). Set the variable attenuators to a total attenuation of 30 dB, and set the output power of the sweep oscillator so that

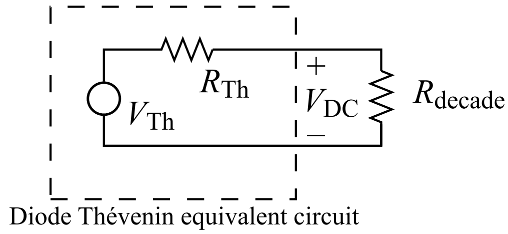



Figure 3.19: Resistor decade box connected to the diode Thévenin equivalent circuit.

the reading on the power meter is -30 dBm. Measure the open-circuit DC voltage $V_{DC} = V_{\text{oc}}$ on the voltmeter. Now connect the decade resistor, set initially to $R_{\text{decade}} = 1\ \text{k}\Omega$, and observe the DC voltage $V_{DC} = V_1$ on the voltmeter. Adjust the decade resistance until $V_1 = V_{\text{oc}}/2$, and note this value of $R_{\text{decade}}$. Repeat these measurements with the variable attenuators set to total attenuations of 20 dB, 10 dB and 0 dB.

**Q7:** For each setting of the variable attenuators, use the measured voltage $V_{\text{oc}}$ and the decade resistance $R_{\text{decade}}$ determined above to calculate the Thévenin equivalent circuit parameters $V_{\text{Th}}$ and $R_{\text{Th}}$. Why are they different for different RF power levels? Are the nonzero values of $R_{\text{Th}}$ likely to degrade the accuracy of your power measurements?

**Q8:** At each RF power level, what is the maximum DC power that can be extracted from the diode? What is the RF to DC conversion efficiency $\eta = P_{\text{DC}}/P_{\text{RF}}$ equal to at each RF power level?

## Part IV.

In this part of the experiment, you will use the 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. At low enough power levels, 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 3.2. Start at

| Power from Power Meter | | DC Voltage $V$ Across Diode | DC Diode Voltage |
|---|---|---|---|
| (dBm) | (mW) | (V) | (dBV) |
| | | | |

Table 3.2: Diode detector RF power measurements.

9 dBm and go down in 3 dB 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.

**Q9:** 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? Use this to find the constant $K$ in the relation

$$V_{DC} = KP_{\mathrm{RF}}$$

## Tables

The following tables may be useful as you collect data:

| Rs (Rt) [Ω] | V [mV] | \|S11\| [dB] | PDC |
|---|---|---|---|
| 1000 | | | |
| 900 | | | |
| 800 | | | |
| 700 | | | |
| 600 | | | |
| 500 | | | |
| 400 | | | |
| 300 | | | |
| 200 | | | |
| 100 | | | |

| Att. [dB] | V [mV] | Ppm [dBm] |
|---|---|---|
| 0 | | |
| 1 | | |
| 2 | | |
| 3 | | |
| 4 | | |
| 5 | | |
| 6 | | |
| 7 | | |
| 8 | | |
| 9 | | |
| 10 | | |
| V (no RF) = | | |

| -30 dBm | | |
|---|---|---|
| Rdec [Ω] | V1 [mV] | |
| ∞ | | Voc |
| 1000 | | |
| | | Voc/2 |

| -20 dBm | | |
|---|---|---|
| Rdec [Ω] | V1 [mV] | |
| ∞ | | Voc |
| 1000 | | |
| | | Voc/2 |

| -10 dBm | | |
|---|---|---|
| Rdec [Ω] | V1 [mV] | |
| ∞ | | Voc |
| 1000 | | |
| | | Voc/2 |

| 0 dBm | | |
|---|---|---|
| Rdec [Ω] | V1 [mV] | |
| ∞ | | Voc |
| 1000 | | |
| | | Voc/2 |

| P [dBm] | V [mV] |
|---|---|
| 9 | |
| 6 | |
| 3 | |
| 0 | |
| -3 | |
| -6 | |
| -9 | |
| -12 | |
| -15 | |
| -18 | |
| -21 | |
| -24 | |
| -27 | |
| -30 | |
| -33 | |
| -36 | |
| -39 | |
| -42 | |
| -45 | |

# Chapter 4

# Time-Domain Reflectometry

## 4.1 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.

## 4.2 Reflection from Simple Loads

### 4.2.1 Example: 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. 4.1(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

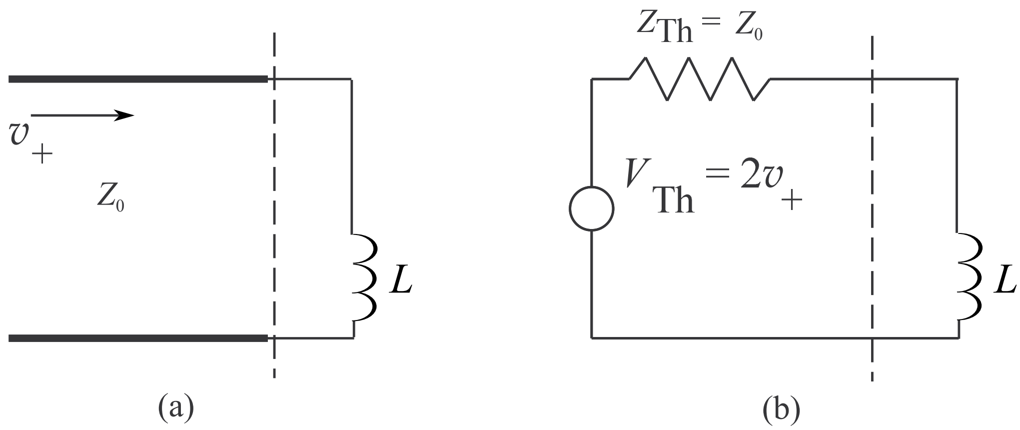



Figure 4.1: (a) Transmission line with an incident wave $v_+$, terminated in an inductive load. (b) Lossless transmission line is replaced by its Thévenin equivalent circuit.

theorem 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 shown in Fig. 4.1(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

$$V(s) = \frac{2}{s}\frac{sL}{sL + Z_0} = \frac{2}{s + Z_0/L}. \tag{4.1}$$

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

$$v(t) = 2e^{-t/t_L} \quad , \quad t > 0. \tag{4.2}$$

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:

$$i(t) = \frac{1}{L}\int_0^t v\,dt = \frac{2}{Z_0}\left(1 - e^{-t/t_L}\right) \quad , \quad t > 0. \tag{4.3}$$

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:

$$v_-(t) = v(t) - v_+(t) = 2e^{-t/t_L} - 1 \quad , \quad t > 0. \tag{4.4}$$

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 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. 4.2.

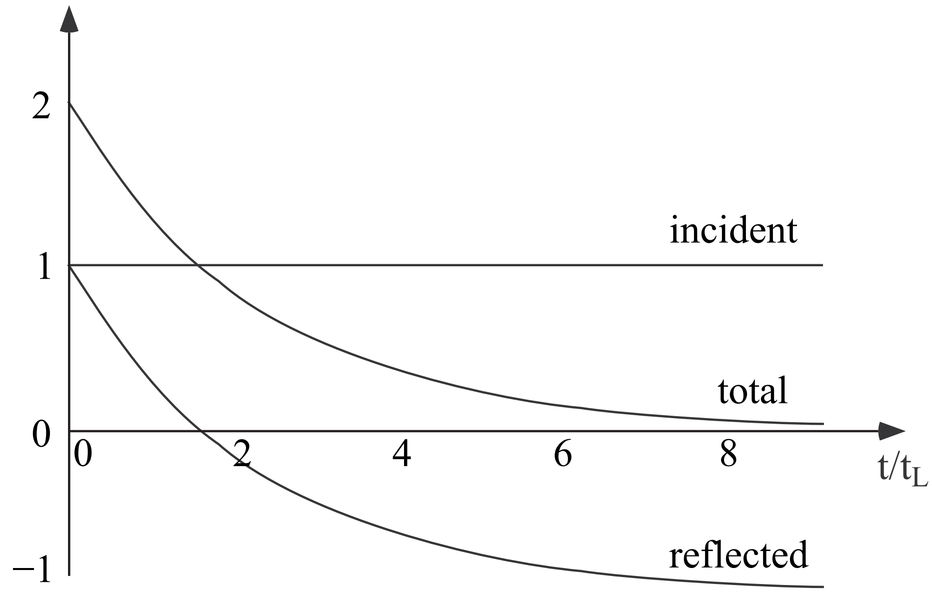



Figure 4.2: The incident, reflected and total voltages for an inductive load.

### 4.2.2 Apparent Impedance

It is common practice in TDR to display not the time-domain reflection coefficient $\rho(t) = v_-(t)$, but instead the "apparent impedance" (always a real quantity) defined by

$$Z_{\text{app}}(t) = \frac{v(t)}{i(t)} = Z_0 \frac{v_+ + v_-}{v_+ - v_-} = Z_0 \frac{1 + \rho(t)}{1 - \rho(t)} \tag{4.5}$$

where $Z_0$ is the reference impedance of the system (the characteristic impedance of the transmission line carrying the incident wave), normally equal to 50 Ω. The meaning of $Z_{\text{app}}$ is that at any given instant of time it is equal to the resistance that would cause a reflected voltage equal in value to what is caused by the actual load at that time. Note that under some circumstances (e. g., for resonant loads) we may have $|v_-(t)| > 1$ for some ranges of $t$. This means that $Z_{\text{app}}$ can be negative for some values of $t$. The apparent impedance is undefined for $t < 0$, since both incident and reflected voltages are zero.

Thus, for a purely resistive load, a TDR instrument would display the value of that load resistance, delayed by the round trip travel time to the load. Loads that have a reactive component will exhibit a more complicated time dependence of $Z_{\text{app}}(t)$ that is at first not obviously related to the nature of the load, but after some experience is gained you will be able to interpret such display traces readily. For example, the inductive load studied in section 4.2.1 has a time-domain reflection coefficient given by $v_-(t)$ in (4.4), so by (4.5) we have for this load

$$Z_{\text{app}}(t) = Z_0 \frac{e^{-t/t_L}}{1 - e^{-t/t_L}} \tag{4.6}$$

This impedance is plotted in Fig. 4.3, and has a clear physical interpretation. At early times, the inductor does not allow an abrupt change in current from its initial value of zero. In other words, it initially acts as an open circuit, and this is the value of $Z_{\text{app}}(0) = \infty$. At very late times, the inductor approaches the DC steady state, at which point it behaves like a short circuit, and $Z_{\text{app}}(\infty) = 0$ in agreement with that.

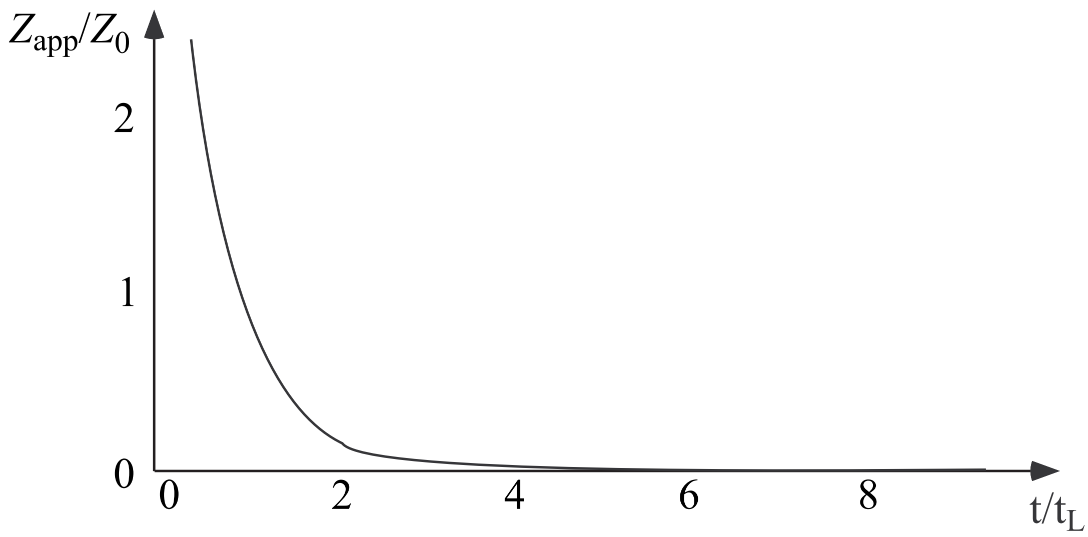



Figure 4.3: Normalized value of $Z_{\text{app}}$ as a function of time for an inductive load.

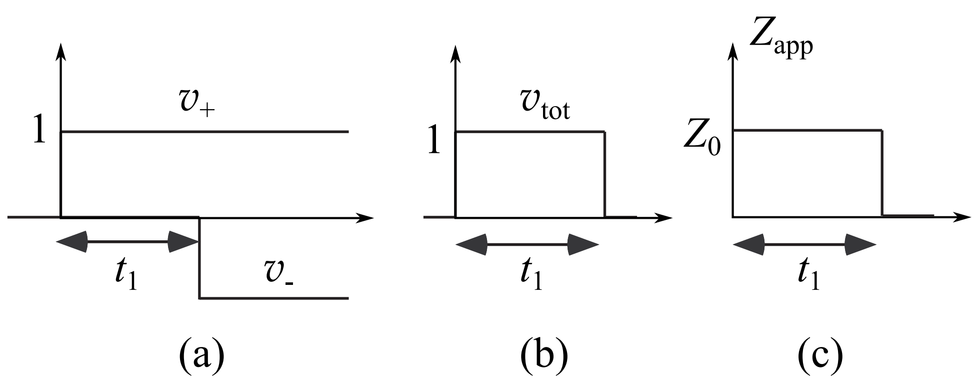



Figure 4.4: (a) Reflected voltage wave off of a short-circuited transmission line with an incident unity step function. (b) A standard TDR instrument display shows the reflected wave added on to the incident step function. (c) Apparent impedance of the short circuit.

### 4.2.3 Example: Short Circuit

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. 4.4(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. 4.4(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). Fig. 4.4(c) shows a plot of the apparent impedance of the short-circuit terminated line. Before the arrival of the reflected pulse, $Z_{\text{app}}$ is equal to $Z_0$, which would be observed on an infinitely long line with no reflections.

### 4.2.4 Example: Series $RL$ Load

A third, slightly more complicated example, is that of a series combination of an inductor $L$ and a resistor $R$, shown in Fig. 4.5. The incident voltage is again a step with unit amplitude. The voltage

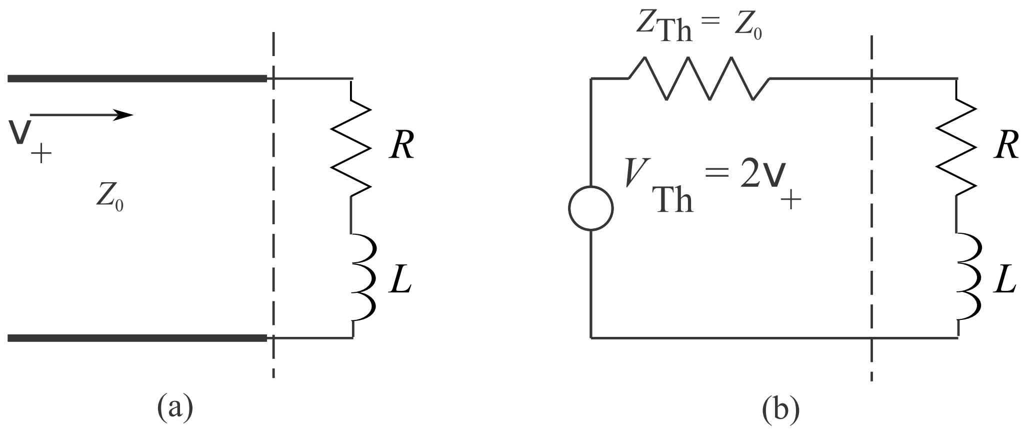



Figure 4.5: (a) Transmission line with an incident wave $v_+$, terminated in an $RL$ load. (b) Lossless transmission line is replaced by its Thévenin equivalent circuit.

across the inductor is, as before,

$$v_L(t) = 2e^{-t/t_L}, \quad t > 0, \tag{4.7}$$

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

$$i_L(t) = \frac{2}{Z_0 + R}(1 - e^{-t/t_L}), \quad t > 0, \tag{4.8}$$

and the load voltage becomes

$$v(t) = v_L(t) + Ri_L(t) = 2\left[\frac{R}{R + Z_0} + \frac{Z_0}{R + Z_0}e^{-t/t_L}\right], \quad t > 0. \tag{4.9}$$

The reflected voltage wave is now

$$v_-(t) = v(t) - v_+(t) = \left[\frac{R - Z_0}{R + Z_0} + 2\frac{Z_0}{R + Z_0}e^{-t/t_L}\right], \quad t > 0. \tag{4.10}$$

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 complete reflected waveform is shown in Fig. 4.6(a). The resulting plot of the total voltage as would be shown on a TDR (incident step plus reflected wave) is shown in Fig. 4.6(b).

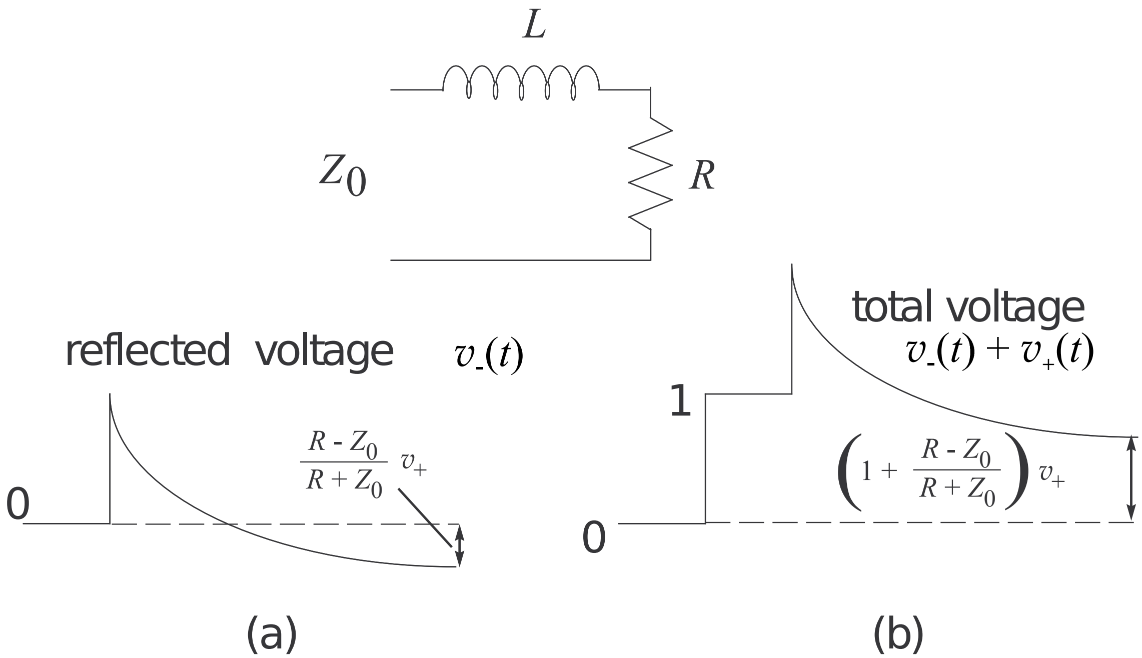



Figure 4.6: (a) Voltage wave reflected from a series $RL$ combination with an incident unit step function. (b) A standard TDR instrument display shows the reflected wave added on to the incident step function.

### 4.2.5 Example: Lossy Lines

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 match-terminated) line can be written as:

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

where the per-unit-length line parameters are here denoted by $R'$, $L'$ and $C'$ to distinguish them from the lumped elements and

$$C_l = \frac{2L'}{Z_0 R'} = \frac{2}{R' c_0}$$

This means that a line with series loss will give the same TDR response as a matched resistor in series with an equivalent series capacitance, the value of which depends on the series resistance per unit length and the wave velocity of the line.

### 4.2.6 Determination of the 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 $t_1 = 0.69 t_L$, where $t_L$ is the time constant we used for an inductive load, but the same principle also holds for a capacitive load. This procedure is shown qualitatively in Fig. 4.7.

However, in practice the reflected waveform may be noisy, due either to experimental errors or to the presence of parasitics in the lumped elements. In that case, a more accurate way to determine the time

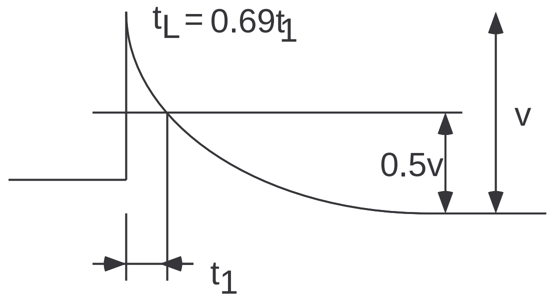



Figure 4.7: Determining the time constant of an exponential TDR response.

constant is by computing the area under the time response curve, which tends to reduce the effect of such errors. Consider the example of an inductance $L$ connected in series with two sections of transmission line of characteristic impedance $Z_0$. The reflected wave can be obtained as a special case of the $RL$ response found in section 4.2.4. Setting $R = Z_0$ in (4.10), we have

$$v_-(t) = e^{-t/t_L} \tag{4.11}$$

where $t_L = L/2Z_0$. This time constant can be found as the area under the curve $v_-(t)$ as shown in Fig. 4.8:

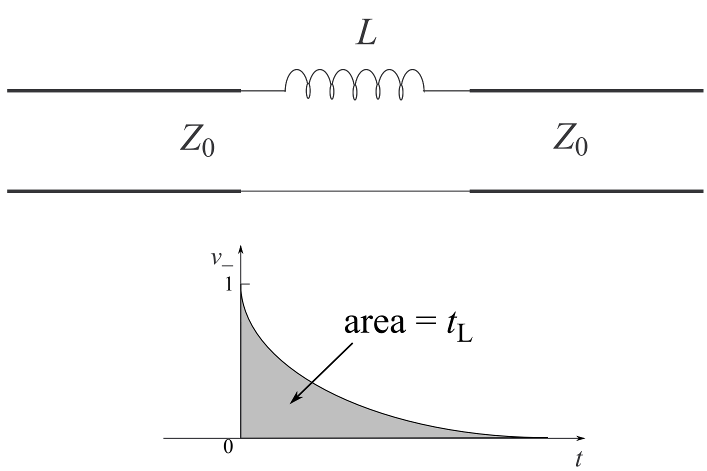



Figure 4.8: Area under reflected wave from a series inductor in a transmission line.

$$t_L = \int_0^\infty v_-(t)\, dt \tag{4.12}$$

Note that formula (4.12) is not universal, and has to be derived individually for each particular kind of circuit. When measured data is available in numerical form, it can be numerically integrated (in Matlab, for example) to compute $t_L$ with higher accuracy than by the method of Fig. 4.7.

## 4.3 TDR Measurements in Time and Frequency Domains

A TDR instrument is really just an oscilloscope together with a source which can produce an almost step-function incident wave with a very fast rise time. The fast rise time of the TDR source 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 adapter. You have already noticed that we have not used BNC connectors in the microwave frequency range, and Agilent Application Note 1304-7 illustrates why. Typically, reflection performance changes with edge speed 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 reflected back from a DUT increases.) Notice in Fig. 4.9 (a measurement of a 50-Ω SMA to BNC adapter) that

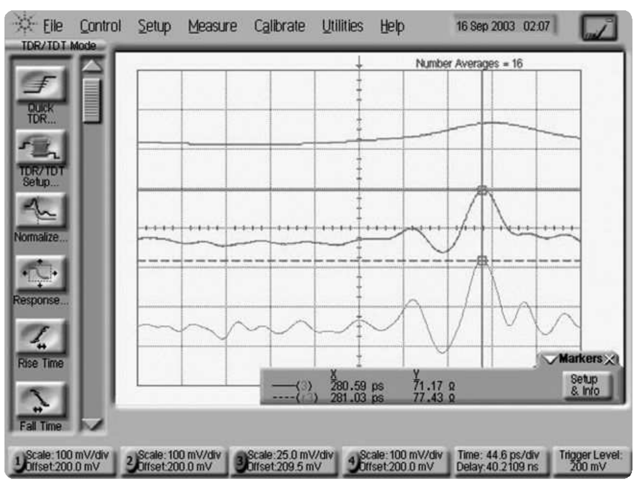



Figure 4.9: Effect of excitation rise time on TDR response of a SMA to BNC adapter. The rise times shown are 100 ps, 35 ps and 20 ps, from top to bottom.

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 an apparent impedance of about 56 Ω. When the edge risetime is increased to 35 ps, more reflections are observed, with the dominant one corresponding to $Z_{\text{app}} = 71\ \Omega$. At a step speed of 20 ps, the impedance peak 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 an oscilloscope)? Consider the time domain response of a circuit with multiple discontinuities, as shown in Fig. 4.10. The two mismatches produce reflections that can be analyzed separately. The mismatch at

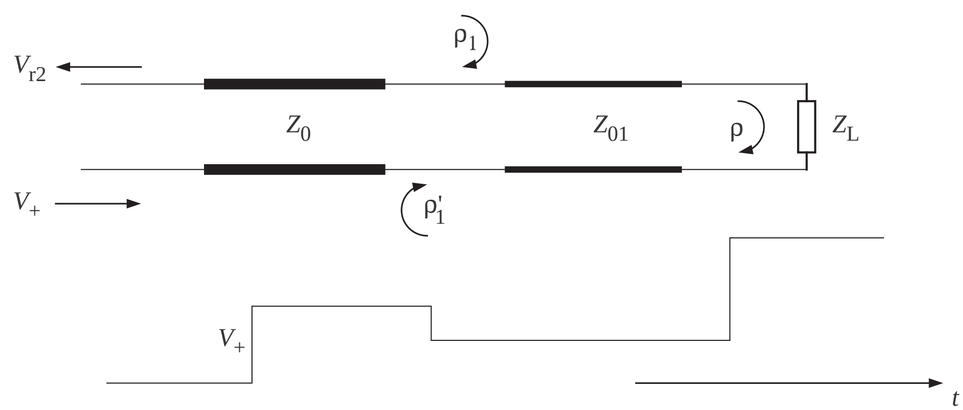



Figure 4.10: Time domain response of a circuit with multiple discontinuities, where $Z_{01} < Z_0$ and $Z_{01} < Z_L$.

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 $\rho$. First, the voltage step incident on $Z_L$ is $(1 + \rho_1)V_+$ (not just $V_+$). Second, the reflected voltage from the load is $\rho(1 + \rho_1)V_+ = V_{rL}$, but this is not equal to $V_{r2}$ since a re-reflection 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:

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

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).

## 4.4 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. 4.11. The designer wants to make sure that a 1

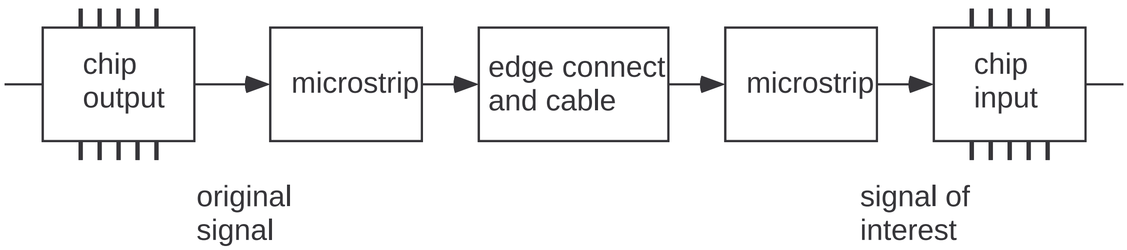



Figure 4.11: A typical digital circuit transmission path.

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 forgiving. 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. 4.12 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. 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. Yesterday's 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 today's 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.

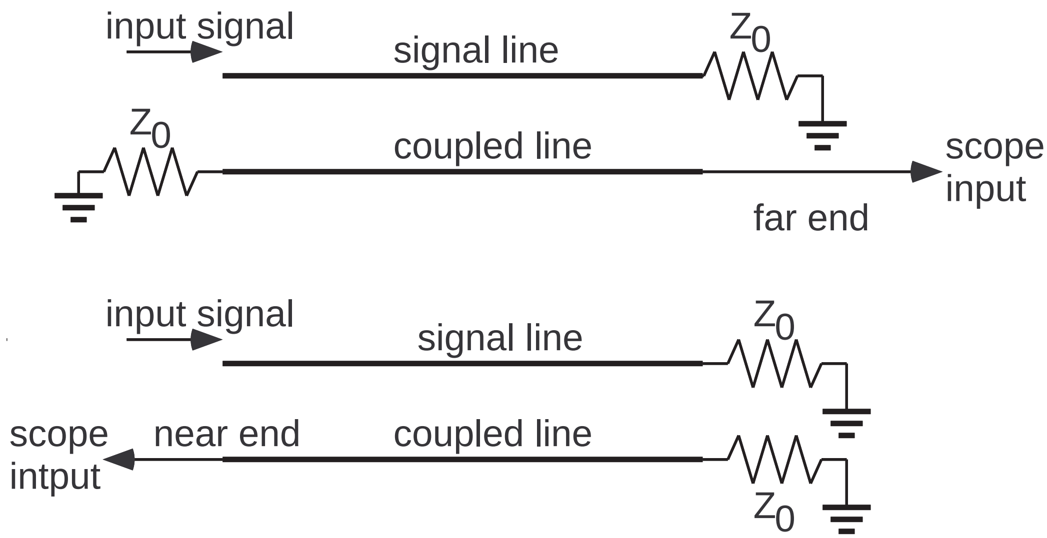
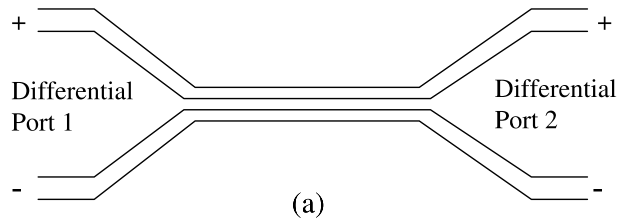



Figure 4.12: Near and far-end cross talk measurements in the case of two adjacent PC board traces.

In order to limit unwanted effects of radiative coupling (electromagnetic interference - EMI), differential circuits are common practice, Fig. 4.13(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.

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. 4.13(a). The two lines of this basic differential circuit initially have a single ended impedance of $50\,\Omega$. 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-$\Omega$ line, then a 70-$\Omega$ section, and a 50-$\Omega$ section before being terminated in a $50\,\Omega$ 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 $100\,\Omega$, which was the intent of the transmission line design. The odd-mode impedance (one trace of the transmission line grounded when




(a)

| | | | |
|---|---|---|---|
| $s_{11}$ | $s_{12}$ | $s_{13}$ | $s_{14}$ |
| $s_{21}$ | $s_{22}$ | $s_{23}$ | $s_{24}$ |
| $s_{31}$ | $s_{32}$ | $s_{33}$ | $s_{34}$ |
| $s_{41}$ | $s_{42}$ | $s_{43}$ | $s_{44}$ |

Single-ended measurements:
$S_{11}$ – return loss, single ended
$S_{21}$ – insertion loss
$S_{31}$ – near-end cross talk
$S_{41}$ – far-end cross talk

| Response | | Excitation: Differential Signal Port 1 | Differential Signal Port 2 | Common signal Port 1 | Common signal Port 2 |
|---|---|---|---|---|---|
| Differ. signal | Port 2 | $S_{DD11}$ | $S_{DD12}$ | $S_{DC11}$ | $S_{DC12}$ |
| | Port 1 | $S_{DD21}$ | $S_{DD21}$ | $S_{DC21}$ | $S_{DC22}$ |
| Common signal | Port 2 | $S_{CD11}$ | $S_{CD12}$ | $S_{CC11}$ | $S_{CC12}$ |
| | Port 1 | $S_{CD21}$ | $S_{CD22}$ | $S_{CC21}$ | $S_{CC22}$ |

(b)

Figure 4.13: (a) Differential passive interconnect circuit (implemented in microstrip). (b) Single-ended and differential-common mode $S$-parameters.

driven differentially) is close to $50\,\Omega$. 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. 4.13(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. 4.13. 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.

## 4.5 Practice questions

1. Why could we not use simple transmission-line analysis when calculating the step response of the inductor in Figure 4.1?

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 4.2.6. 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 6 is related to differential and common mode analysis in this lecture.

8. Give an example of a load for which the apparent impedance $Z_{\mathrm{app}}(t)$ can be negative for some values of $t$.

## 4.6 Homework Problems

**1.** Determine the Thévenin equivalent 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 each of the cases shown in Fig. 4.14. In addition, sketch the apparent impedance $Z_{\mathrm{app}}(t)$ for each case. [*Hint*: Follow the examples of the short circuit termination and the

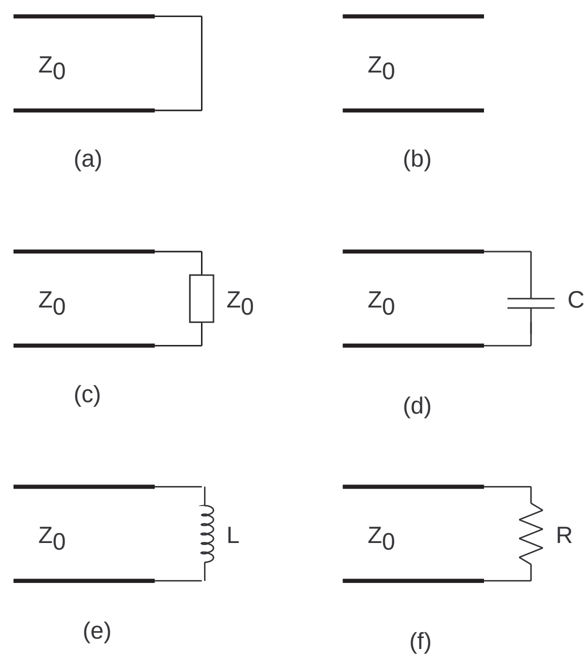



Figure 4.14: Find the waveforms of a TDR display for a short circuited line (a), an open circuit (b), a matched load (c), a capacitive load (d), an inductive load (e), and a resistive load (f).

inductive load that were done in the lecture 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. 4.15. 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 and the apparent impedance $Z_{\mathrm{app}}(t)$.

**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, the apparent impedance

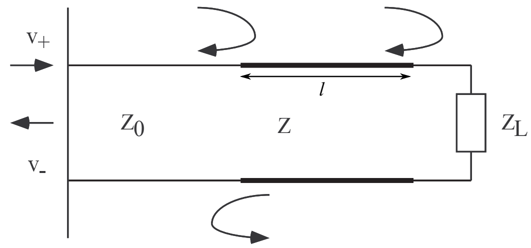



Figure 4.15: A cascade of transmission lines terminated in a load.

and the sum of the reflected and incident voltage at the load, for a unit step incident voltage wave. Explain your result.

**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 when the load is a series combination of an inductance $L$ with a capacitance $C$.

**8.** Repeat problem 4 for the case of a load consisting of a series combination of an inductor $L$, a capacitor $C$ and a resistor $R$, using the following sets of element values: $Z_0 = 10\ \Omega$, $L = 1$ nH, $C = 10$ pF, and (a) $R = 5\ \Omega$, or (b) $R = 50\ \Omega$.

**9.** Using microwave design software, simulate the time domain responses of circuits from problems 4-7.

**10.** Using microwave design software, simulate in the time domain a 50-ohm source feeding a matched line of some electrical length $L$ with a 25 $\Omega$ 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 from the load?

**11.** Do an internet search to obtain a copy of Agilent Application Note 1304-7 on High-Precision TDR and read it (you may also want to locate and read the article "Time-domain reflectometry," by B. M. Oliver, in the *Hewlett-Packard Journal*, vol. 15, no. 6, pp. 1-8, 1964). On page 4, it is stated that a BNC to SMA adapter shows a reflection equivalent to "77 $\Omega$" for an incident signal with a rise time of 20 ps on a 50 $\Omega$ transmission line. Does this response correspond to an equivalent parasitic capacitance in parallel with the transmission line, or to an equivalent parasitic inductance in series with the line? Indicate the assumptions you have made (you can use the approximate step function

$$\begin{aligned} u_R(t) &= 0 \quad (t < 0) \\ &= \frac{t}{T_R} \quad (0 < t < T_R) \\ &= 1 \quad (t > T_R) \end{aligned}$$

to include the effect of a rise time $T_R$). What is the approximate value of this capacitance or inductance?

**12.** Derive an integral expression similar to (4.12) for the time constant of the reflected voltage response of the load in problem 5.

**13.** Derive an integral expression similar to (4.12) for the time constant of the reflected voltage response of the load in problem 6.

# Lab 4: 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 8753ES Network analyzer, use the following procedure (similar procedures apply for the other network analyzers):

(0) USE THE GROUNDING STRAP!

(1) Choose the SYSTEM key in the INSTRUMENT STATE group.

(2) Choose TRANSFORM MENU

(3) Choose LOW PASS STEP softkey

(4) Press TRANSFORM ON

(5) Choose a START time of -20 ns and a STOP time of +20 ns.

(6) In the CALIBRATE MENU, pick FULL 2-PORT

(7) Perform the (two-port) calibration using 3.5 mm cal standards (be sure to save the state of the VNA after calibration, since these instruments seem to be more prone to freezing when the time-domain option is used)

(8) Go to LOCAL MENU

(9) GP-DIAG ON

(10) Go to FORMAT, MORE, REAL

Observe the START and STOP time settings on the network analyzer. Using the TRANSFORM MENU, toggle between TRANSFORM ON and TRANSFORM OFF. With TRANSFORM OFF, what are the START and STOP frequencies? With TRANSFORM ON, change the START and STOP times. Does this have any effect on the START and STOP frequencies when you turn TRANSFORM OFF?

**Q2:** On some VNAs, the choice of START and STOP times affects the allowed frequency interval. Why would this be so? Does this happen for the VNA you are using?

## Part I.

In this part of the experiment, you will measure the length (time delay multiplied by an assumed propagation velocity) of a GR cable. What velocity does the network analyzer assume for this purpose? You should try both a short and an open at the other end of the cable. 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).

**Q3:** 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 (remove the GR cable for this measurement). This is just a piece of air coax terminated with a short that can slide along the coax to shrink it or stretch it.

**Q4:** 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?

**Q5:** 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, GR 100 Ω, and GR 50 Ω loads, using the GR component mount if necessary.

**Q6:** Calculate the reflection coefficient from a 50 Ω transmission line terminated in a 100 Ω 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.

**Q7:** Calculate the reflection coefficient from a 50 Ω transmission line terminated in a 20 Ω 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 **Q6** and **Q7**? 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 (select its capacitance so that the time constant will be less than 10 nsec). Sketch the display and label it carefully.

**Q8:** From the display, and your homework results, calculate the capacitance of the capacitor. The characteristic impedance of the coax is $50\,\Omega$.

(C) Obtain three inductors (winding your own on a piece of straw or a pen if necessary) as tight solenoids with different numbers of windings.

**Q9:** What is the inductance of each solenoid proportional to? Why? [*Hint:* Remember the flux definition of the inductance $L = \Phi/I$.]

The dimensions and number of windings for each inductor should be chosen so that the expected time constants will all again be less than 10 nsec. Do not use ferrite-core inductors such as toroids, because these often have a complicated frequency dependence of the inductance that does not correspond to that of an ordinary inductor. Moreover, for such inductors $L$ is usually so large that the resulting time constant will too long to be readily observable in a $\pm 20$ nsec interval.

(D) Connect your inductors one by one to the jig, and sketch the resulting displays.

**Q10:** 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.

**Q11:** 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?

Finally, connect the ends of the 10 cm air line marked "non-constant $Z_0$" to ports 1 and 2 of the network analyzer. Display and include plots of $S_{11}$ and $S_{21}$ in the time domain.

**Q12:** Does this line have a uniform characteristic impedance along its length? Make a rough sketch of what you think $Z_0$ looks like as a function of position along the line.

# Chapter 5

# Nonlinear Microwave Circuits

## 5.1 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 or line frequency AC power is supplied to it, and it produces AC power at a desired frequency. 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 (traveling 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 $\mu$W 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. 5.1. 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

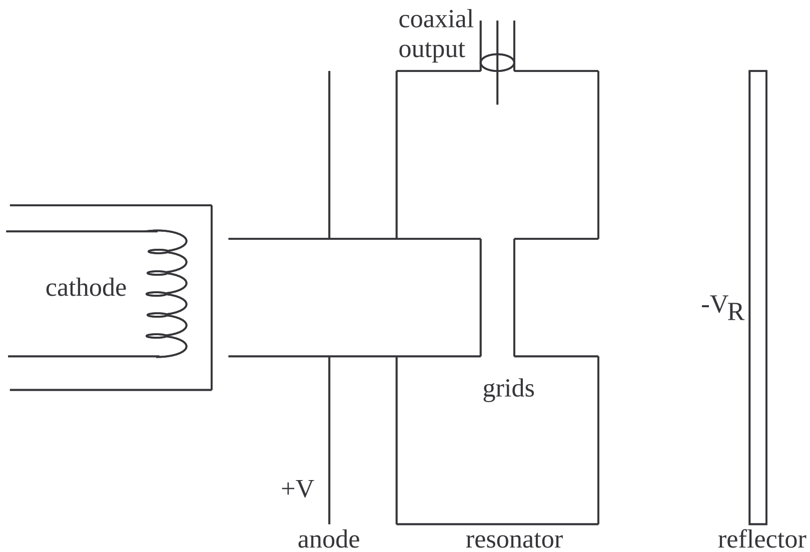



Figure 5.1: The reflex klystron.

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 (determined by the condition that its dimension is an integer 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 $V_a$, a simplified version of which is shown in Fig. 5.2. This voltage produces a radial electric field $E$ between the electrodes; an electron

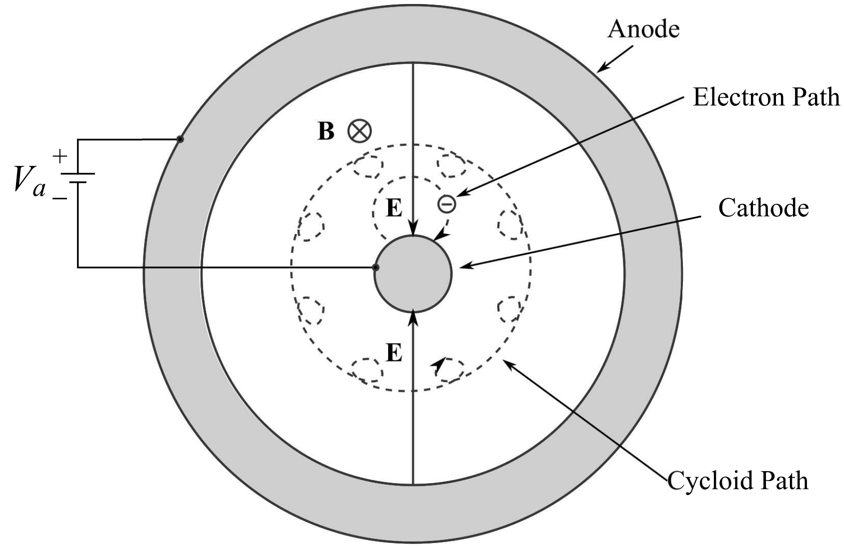



Figure 5.2: Electron trajectory in a simplified magnetron cavity.

placed in this field would tend to move along a straight radial path from the cathode towards the anode. If additionally a DC magnetic field **B** is imposed parallel to the tube axis, the electron under the $e\mathbf{v} \times \mathbf{B}$ magnetic force would tend to move back to the cathode in a circular path, if $V_a$ is not too large. For larger $V_a$, the electrons move in cycloidal paths to form a cloud between the anode and cathode. If an external circuit resonant at the cyclotron frequency of this electron motion is connected between the anode and cathode, then the radial portion of the electron motion in the cloud is reinforced and creates a strong RF current in the external circuit at this frequency.

The magnetron can be modified by corrugating the anode as shown in Fig. 5.3, forming of an even number of resonant cavities in this wall. A high frequency electric field is established between these cavities, such that the field polarities alternate between them. In this way the external resonant circuit has been integrated into the magnetron itself, and the RF signal is then extracted from the resonant cavities. A half period later in time than what is shown in the figure, 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 overall effect is similar to a negative conductance device, in which a slight increase in anode voltage produces a slight decrease in current. We will study he behavior of a negative-conductance device in more detail when we consider the Gunn diode in section 5.2 below.

The anode of the magnetron 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.

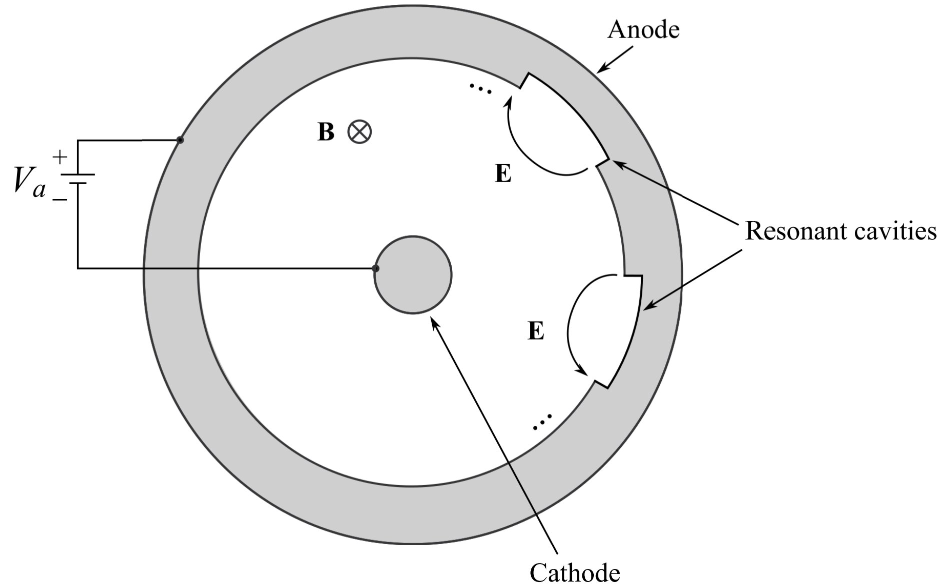



Figure 5.3: A magnetron – the center electrode is the cathode surrounded by a coaxial anode formed of an even number of resonant circuits.

The *TWT* (Traveling Wave Tube) is a tube amplifier that falls in the category of *slow-wave devices*. As shown in Fig. 5.4(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

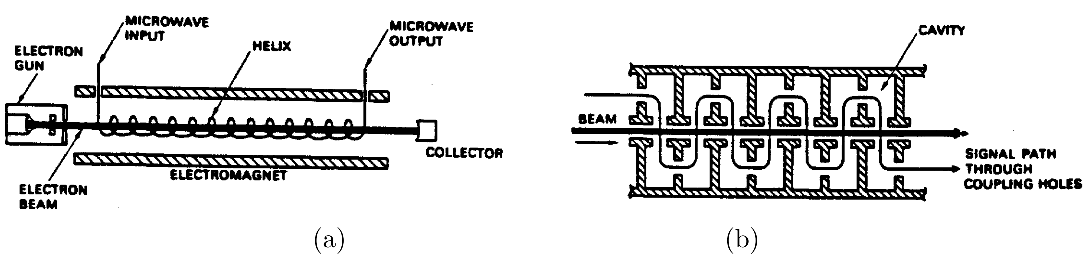



Figure 5.4: The TWT with a helix slow wave structure (a) and periodically loaded waveguide slow-wave structure (b).

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. 5.4(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 (GaAs), indium phosphide (InP), silicon carbide (SiC), 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. 5.5(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. 5.5(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 $\mu$m 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. 5.6. 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

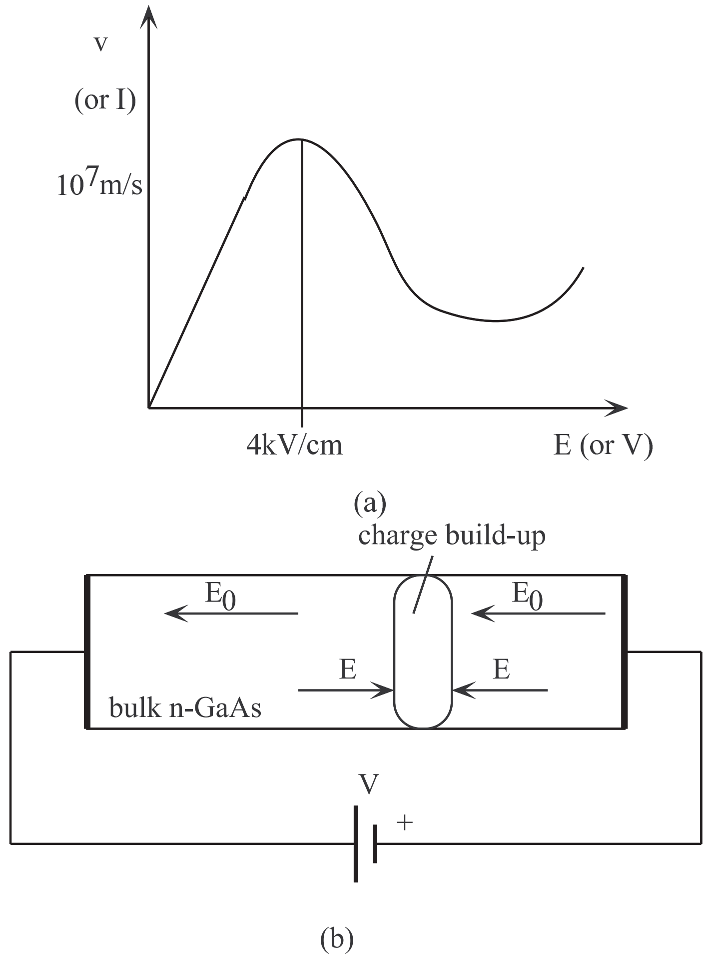
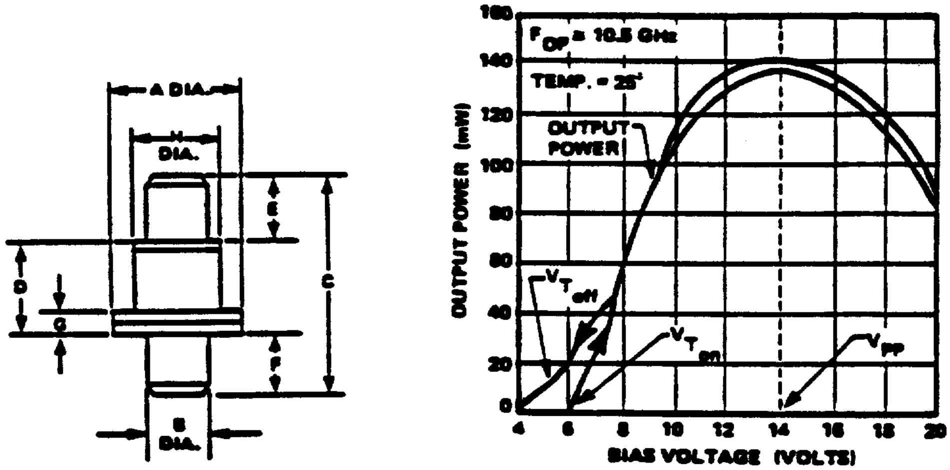



Figure 5.5: Electron velocity versus electric field magnitude for GaAs (a) and oscillation buildup in a Gunn diode (b).

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 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.




(a) (b)

Figure 5.6: A commercial MaCom Gunn diode package (a) and specifications (b).

## 5.2 Oscillators

In the lab you will look at a Gunn-diode version of a two-port negative-conductance oscillator as shown in Fig. 5.7. Let us examine the operation of this circuit.

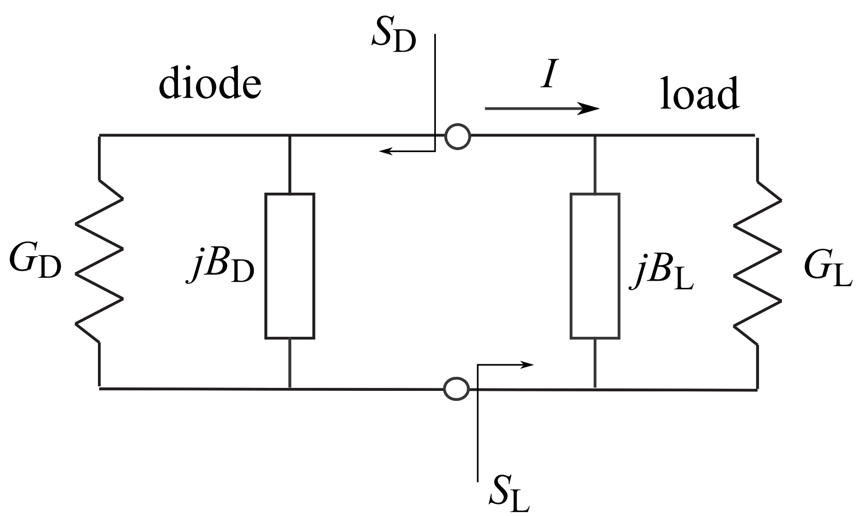



Figure 5.7: A one-port RF oscillator circuit.

### Quasiharmonic Description of Negative-Conductance Diodes

As we have already learned in chapter 3, a diode is a nonlinear device whose current-voltage ($I$-$V$) characteristic can be written in the form of a Taylor series:

$$I(V) = I_0 + i = I(V_0) + v_d G_d + \frac{v_d^2}{2} G_d' + \frac{v_d^3}{6} G_d'' + \dots \tag{5.1}$$

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. The DC bias must be supplied through a biasing network, as described in section 3.5. 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 (5.1) 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:

$$v_d \simeq \sqrt{2} V_1 \cos \omega t \tag{5.2}$$

where $V_1$ is nonnegative and real, the nonlinear nature of (5.1) would create in the current $i$ not only a current at this *fundamental* frequency $\omega$, but also at the second harmonic $2\omega$, the third harmonic $3\omega$ and so on. Indeed, because of the trigonometric identities

$$\cos^2 \omega t = \frac{1}{2}(1 + \cos 2\omega t), \qquad \cos^3 \omega t = \frac{1}{4}(3\cos\omega t + \cos 3\omega t) \tag{5.3}$$

etc., we can write (5.1) as

$$\begin{aligned} I(V) \quad = \quad & \left[ I(V_0) + \frac{1}{2} G_d' V_1^2 + \dots \right] \\ & + \sqrt{2} \left[ G_d V_1 + \frac{1}{4} G_d'' V_1^3 + \dots \right] \cos \omega t \\ & + \sqrt{2} \left[ \frac{1}{2\sqrt{2}} G_d' V_1^2 + \dots \right] \cos 2\omega t \\ & + \sqrt{2} \left[ \frac{1}{12} G_d'' V_1^3 + \dots \right] \cos 3\omega t + \dots \\ = \quad & I_0 + \sqrt{2} I_1 \cos \omega t + \sqrt{2} I_2 \cos 2\omega t + \sqrt{2} I_3 \cos 3\omega t + \dots \end{aligned} \tag{5.4}$$

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 (5.4), 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:

$$I_1 = G_D(V_1) V_1 \tag{5.5}$$

where

$$G_D(V_1) = G_d + \frac{1}{4} G_d'' V_1^2 + \dots \tag{5.6}$$

It is assumed that $|I_2|, |I_3|, \ldots \ll |I_1|$. 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 $V_1$.

## Oscillator Resonance Condition

The diode in Fig. 5.7 is connected to a passive load admittance $Y_L(\omega) = G_L(\omega) + jB_L(\omega)$. Kirchhoff's voltage and current laws give us

$$[Y_D(V_1) + Y_L(\omega)] V_1 = 0. \tag{5.7}$$

or, if the RF voltage $V_1$ is not zero:

$$Y_D(V_1) + Y_L(\omega) = 0 \tag{5.8}$$

This can also be written in $S$-parameter form as

$$S_L = \frac{1}{S_D} \tag{5.9}$$

where

$$S_L = \frac{Y_0 - Y_L}{Y_0 + Y_L}; \quad \text{and} \quad S_D = \frac{Y_0 - Y_D}{Y_0 + Y_D} \tag{5.10}$$

are the reflection coefficients of the load and diode respectively, $Y_0 = 1/Z_0$ being the characteristic admittance of the system. For a given RF voltage amplitude $V_1$, the solutions $\omega_r$ 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:

$$G_D(V_1) + G_L(\omega) = 0 \qquad \text{and} \qquad B_L(\omega) = 0. \tag{5.11}$$

Since we have neglected the susceptance $B_D$ of the diode, the second of eqns. (5.11) 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 (5.11) can be represented graphically on the Smith chart. First, we plot a line corresponding to the variation of $Y_L$ with $\omega$. 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 (5.11) (see Fig. 5.8). This graphical approach can also

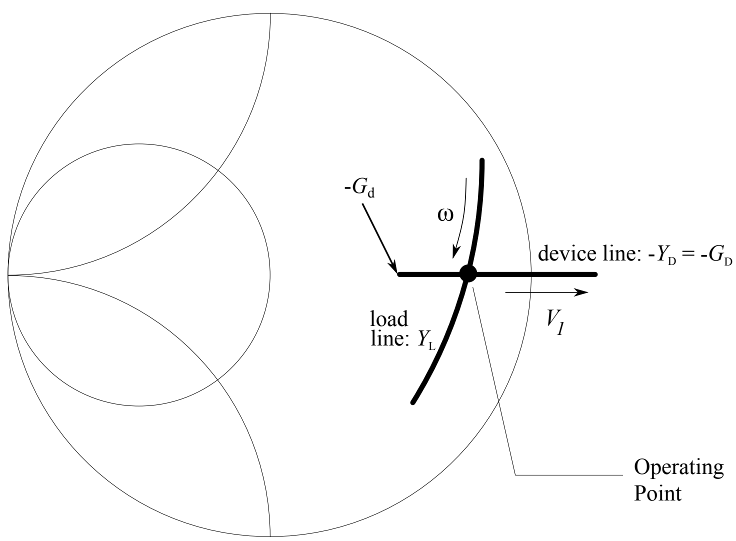



Figure 5.8: Admittance Smith chart showing intersection of device line and load line for oscillation condition.

be used when the susceptance $B_D$ of the device needs to be taken into account.

For typical negative-conductance diodes, the zero-amplitude conductance $G_D(0) = G_d$ in (5.6) is negative. In fact, since the coefficient $G_d''$ of the $I$-$V$ curve is positive, $G_d$ represents the most negative value that $G_D(V_1)$ can take (the small-signal limit, see Fig. 5.8). For oscillation to occur, therefore, $G_L$ cannot be too large, or equivalently $R_L$ cannot be too small. Another way of saying this is that the quality factor $Q$ of the resonant load must be sufficiently large for oscillation to occur. The way in which oscillations build up with time in the circuit is also of some interest. From the signs of $G_d$ and $G_d''$ noted above, the typical device conductance vs. $V_1$ dependence is as shown in Fig. 5.9. If

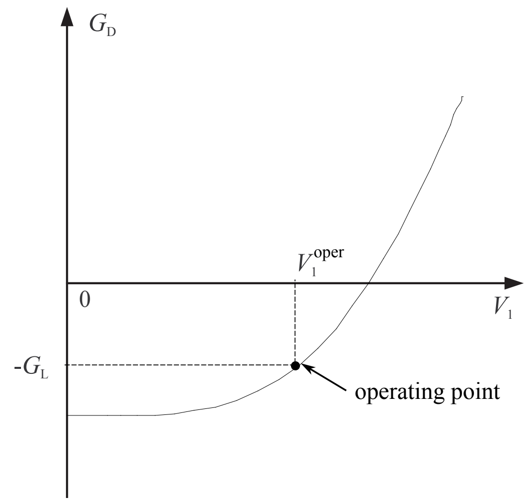



Figure 5.9: Typical $G_D$ vs. $V_1$ curve for a negative-conductance diode.

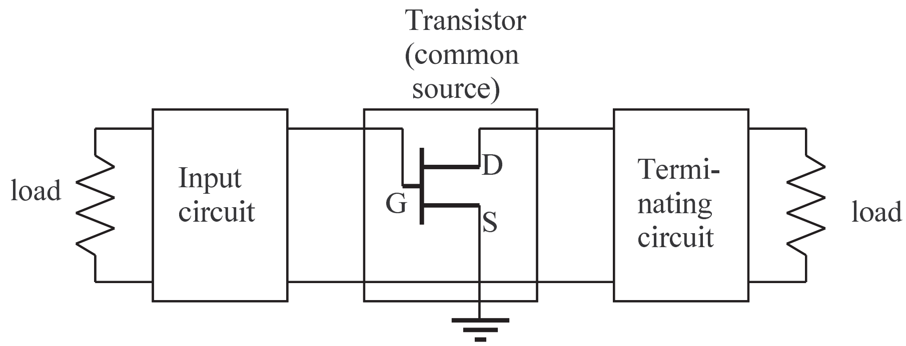
there is initially no signal in the circuit, but a small amount of noise (corresponding to a small value of $V_1$) is present, then $G_D$ is more negative than it will be at the oscillation point, $G_D + G_L$ is negative, and the natural frequency of the network will have a negative imaginary part (or in Laplace transform terminology, is in the right half-plane). This means that the circuit response is a growing exponential, 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 operating-point value $V_1^{\text{oper}}$ 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

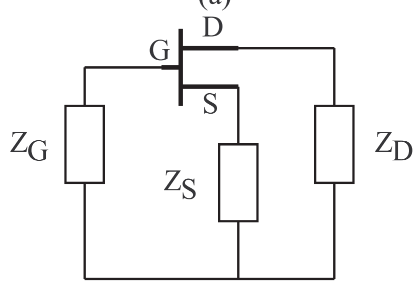
An oscillator can be coaxed into oscillating at a slightly different frequency than the free-running oscillation frequency determined by the load admittance as in (5.11). This is done by injecting a small RF signal at the desired frequency (sufficiently close to the free-running frequency) into the oscillator 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. 5.10. The analysis of two-port oscillators is carried out by the




(a)




(b)

Figure 5.10: A two-port transistor oscillator RF circuit (a), and a three-port transistor circuit (b), where each of the leads is connected to transmission lines with a common ground lead.

same methods that are used in Lecture 8 to analyze microwave amplifiers, but will not be carried out in detail here.

## 5.3 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 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. The resonator looks like the one in Fig. 5.11. 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.)

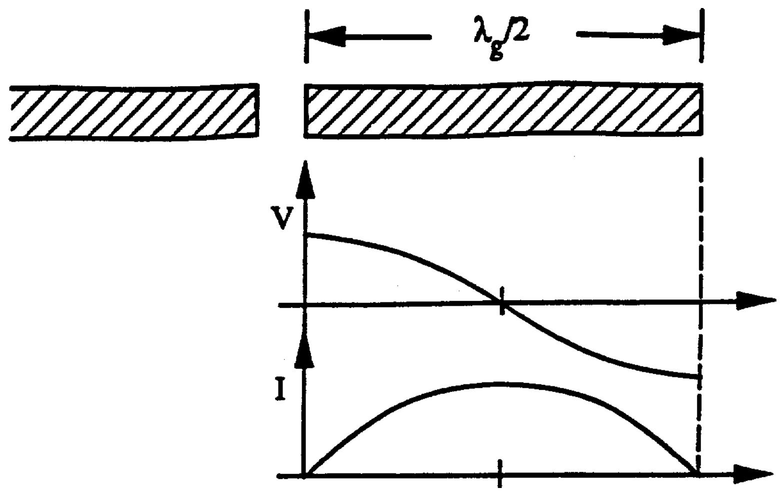



Figure 5.11: Microstrip resonator with associated voltage and current waveforms. The energy is usually coupled in through a gap (capacitive coupling).

Waveguide cavity resonators are usually $\lambda/2$ shorted stubs. The output signal can be extracted with a short loop (magnetic coupling), a short monopole (electric coupling) or via a small aperture (both electric and magnetic 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. 5.12(a).

The $(l, m, n)$ subscripts usually used in waveguide resonators are modified to $(l, m, \delta)$, where the $\delta$ 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. 5.12(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 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

$$Q = \frac{1}{\omega RC} \qquad \text{and} \qquad f = \frac{1}{2\pi RC}. \tag{5.12}$$

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.

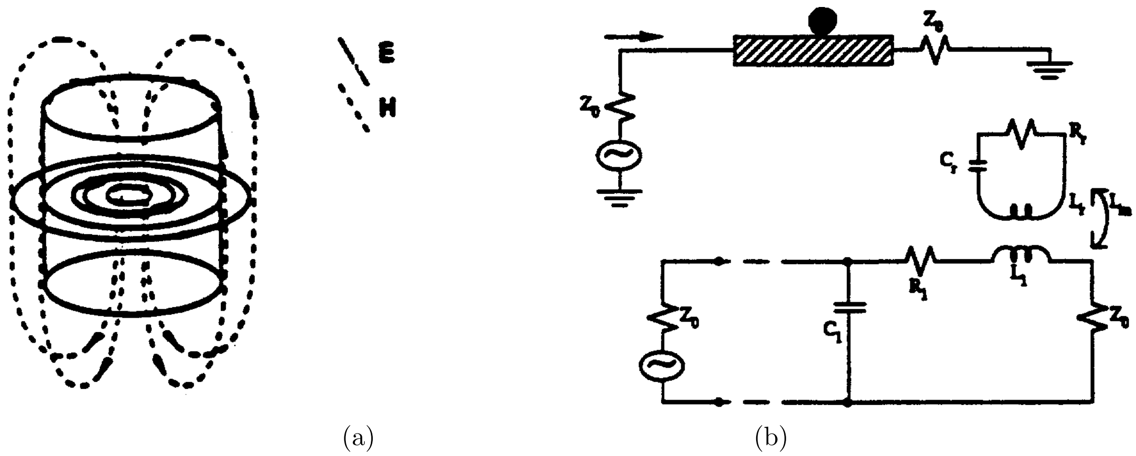



Figure 5.12: The $TE_{01\delta}$ mode in a dielectric resonator, and the equivalent circuit for a dielectric resonator coupled to a microstrip line (b).

## 5.4 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. 5.13 shows a simplified block diagram of a spectrum analyzer. Microwave spectrum analyzers cover frequencies from several hundred MHz to

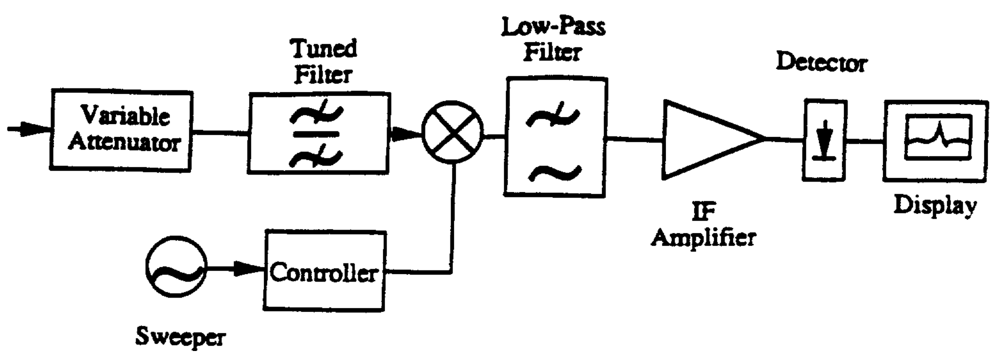



Figure 5.13: Block diagram of a spectrum analyzer.

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 publication about Spectrum Analyzer basics is Application Note 150.

## 5.5 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

$$f(t) = f(t + nT), \quad n \text{ is an integer.} \tag{5.13}$$

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 $\phi_k$'s, we can represent the function $f$ as:

$$f(t) = \sum_{k=-\infty}^{+\infty} a_k e^{jk\omega t}, \tag{5.14}$$

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

$$a_k = \frac{1}{T} \int_0^T f(t) e^{-jk\omega t}\, \mathrm{d}t. \tag{5.15}$$

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 $|a_k|$ and a phase $\angle a_k$. When you use the spectrum analyzer, you can get the amplitude of the different harmonics, but no information about the phase, and as a consequence, you cannot define the Fourier series. As an example, Fig. 5.14 shows two time waveforms that would have the same spectra shown on a spectrum analyzer. The Transition Analyzer

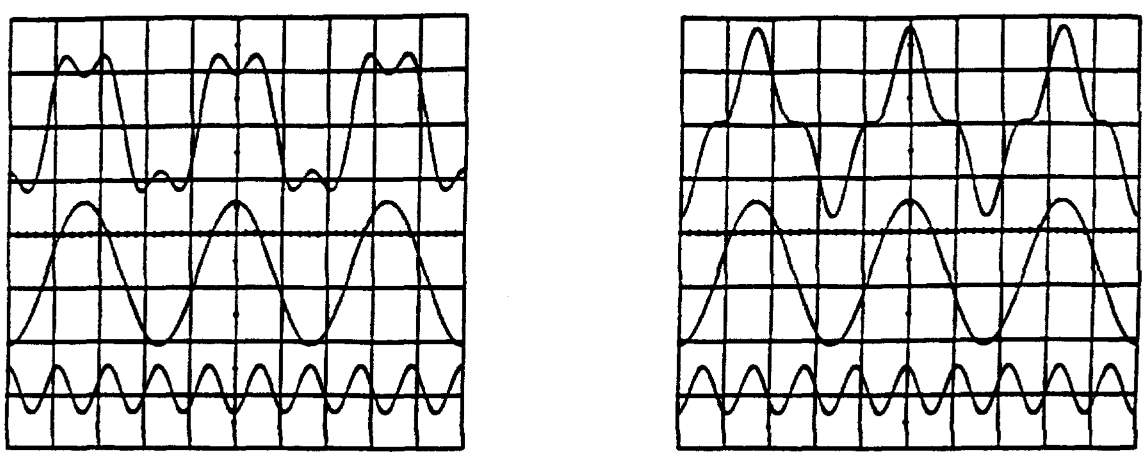


Figure 5.14: Two waveforms that give the same output on a spectrum analyzer.

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 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.

## 5.6 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?

## 5.7 Homework Problems

**1.** Find the Fourier series for a train of pulses of unit amplitude and 50% duty cycle at 1 GHz, Fig. 5.15. Which are the three most important frequencies that describe the square wave?

**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 $R_0 > 0$) as shown in Fig. 5.16. You will examine the circuit behavior dependence on the value of the resistor $R$ (it can be positive or negative).

   (a) What is the input impedance of the resonant circuit equal to? What is the resonant frequency $\omega_0$ equal to?

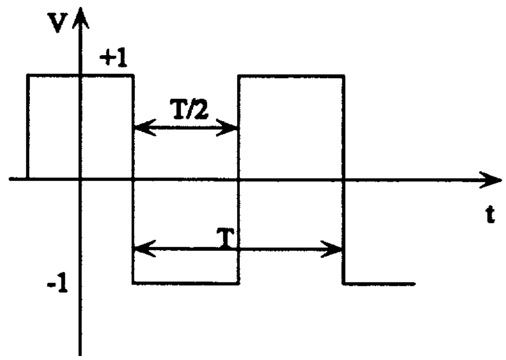



Figure 5.15: A square wave with unit amplitude. Find the Fourier series.

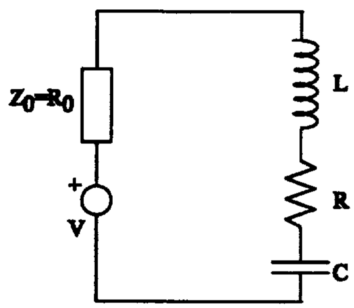



Figure 5.16: A series resonant circuit fed from a one-port generator of characteristic impedance $R_0$.

(b) Derive the expression for the reflection coefficient $\rho$ 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 $\rho$ (whether it is larger or smaller than 1) and the phase of $\rho$ at three frequencies around resonance, $\omega_1 = \omega_0 - \delta$, $\omega_0$ and $\omega_2 = \omega_0 + \delta$, where $\delta \ll \omega_0$ for the following four cases:

(i) $R \geq R_0$

(ii) $R_0 \geq R \geq 0$

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

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

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

(e) Use microwave design software 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. Show a Smith chart for each case.

**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. 5.17. Use a substrate

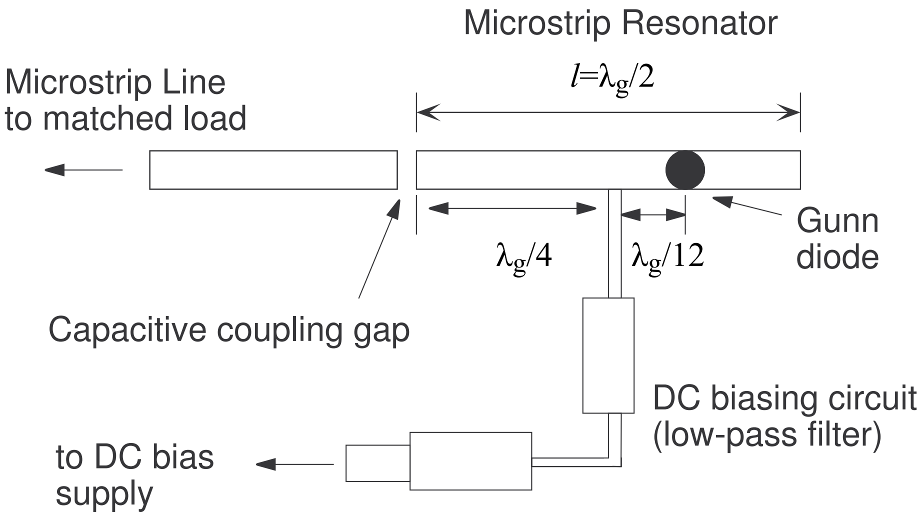



Figure 5.17: A microstrip Gunn diode oscillator.

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. 5.17, and assume the input microstrip line is terminated on the left 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 microwave design software 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 $80\Omega$ and $25\Omega$ 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 $0.5\Omega$ resistor (connected from the end of the bias line to ground), use microwave design software 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 microwave design software 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 (alternatively, you may use any other microwave design software, if it is capable of time-domain simulation and analog behavioral modeling). The Gunn diode is modeled here by a (somewhat oversimplified) equivalent circuit consisting of a voltage-dependent current source with 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 $K = 0.8$, $V_{\rm thr} = 4\,$V, and $I_{\rm thr} = 0.1\,$A, along with some parasitic elements that make the model more realistic and help make time-domain simulations more stable. The complete Gunn diode equivalent circuit is shown in Fig. 5.18. In PSPICE (OrCAD/Cadence),

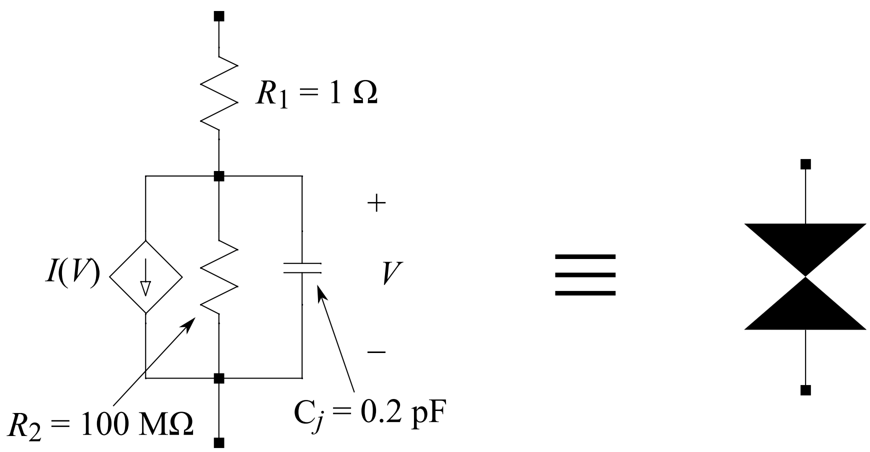



Figure 5.18: Equivalent circuit for a Gunn diode, with parasitics.

the nonlinear voltage-dependent current source in the Gunn diode equivalent circuit can be modeled using the SPICE element G1, a voltage-controlled current source. 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 (this is also sometimes called a behavioral element). In LTspice, this nonlinear dependent current source is realized using the BI behavioral element.

The complete Gunn diode oscillator circuit is shown in the schematic in Fig. 5.19. A DC bias

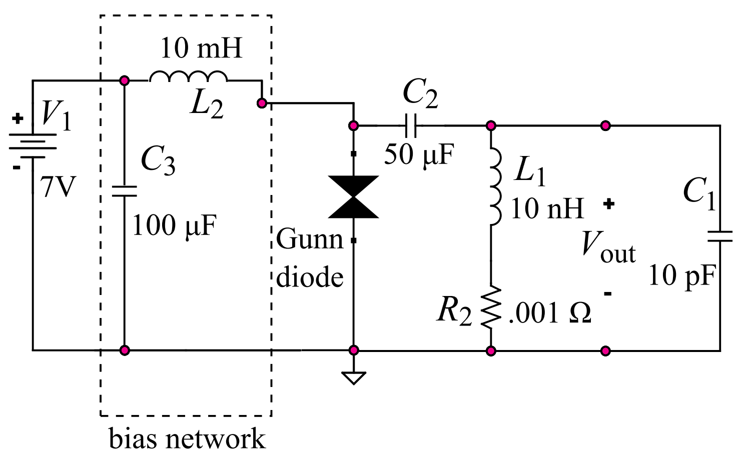



Figure 5.19: Circuit for modeling a lumped-element Gunn diode oscillator in SPICE.

voltage of 7 V is provided through an RF filter (bias network) as shown, and the Gunn diode is connected to a resonant parallel RLC network through the large capacitor $C_2$ (inserted to remove the DC voltage) to achieve oscillation. An oscillator in the real world would start the oscillations from naturally occurring noise in the circuit elements, while in the problem, "numerical noise" (floating-point roundoff error) achieves the same effect.

(a) Using a DC sweep, plot $I$ vs. $V$ in the voltage range from 0 to 10 V. For what voltages is negative RF conductance achieved?

(b) Do a *transient* analysis of this circuit (a `.TRAN` command in SPICE) covering the time interval $0 < t < 1$ $\mu$s. You may need to adjust the value of the maximum time-step used by the SPICE simulator to achieve satisfactory results: if the time step is too large, the simulation will not be able to follow the oscillation of the signal; if it is too small, it may either take too long to run the simulation or fail to generate enough numerical noise to start the oscillation in the time interval specified. Plot the voltage $V_{\mathrm{out}}$ indicated in the figure over this time interval. Explain what is happening. How long does it take for the oscillation to build up to its steady-state amplitude?

(c) Perform a Fourier transform of your data to display the frequency content of the waveform you obtained in part (b). In PSPICE, this is obtained in the PROBE display via the suitable menu item, while in LTspice, you use the FFT command in the View menu when the time-domain plot is highlighted. At what frequency is the spectrum largest? What is the half-power bandwidth of this spectrum (i. e., find the frequencies at which the voltage 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.

(d) Replace the lumped-element resonant circuit by a quarter-wave section of transmission line terminated by a small resistance, as shown in Fig. 5.20, and repeat parts (b) and (c) above. What are the major differences between the Fourier transform spectra in this

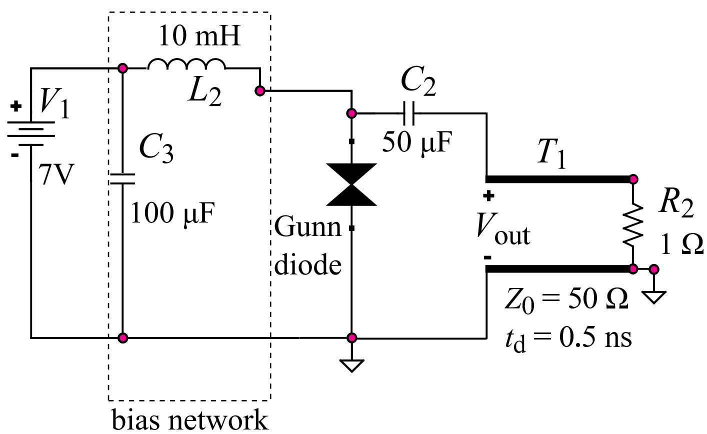



Figure 5.20: Circuit for modeling a transmission-line Gunn diode oscillator in SPICE.

case compared to that for the lumped-element oscillator circuit? What is the reason for these differences? In the time-domain simulation results, expand the time scale on the horizontal axis to show only a few cycles of oscillation in order to help explain your results.

(e) Change the bias voltage to 5 V, and repeat parts (b) and (c) for the lumped-element oscillator in Fig. 5.19. You may need to change the time step in order to get oscillation to occur in the simulation. What is the effect of the change in bias voltage, and what is the reason for the difference?

**5.** Using microwave design software, simulate a half-wave microstrip resonator at 4 GHz capacitively coupled to a 50 Ω input port. Use a realistic (lossy) transmission line for the resonator (copper). For the first part, couple the resonator to the feed with a 0.05 pF capacitor, and observe the magnitude of the input reflection coefficient: plot $|S_{11}|$ for frequencies between 3

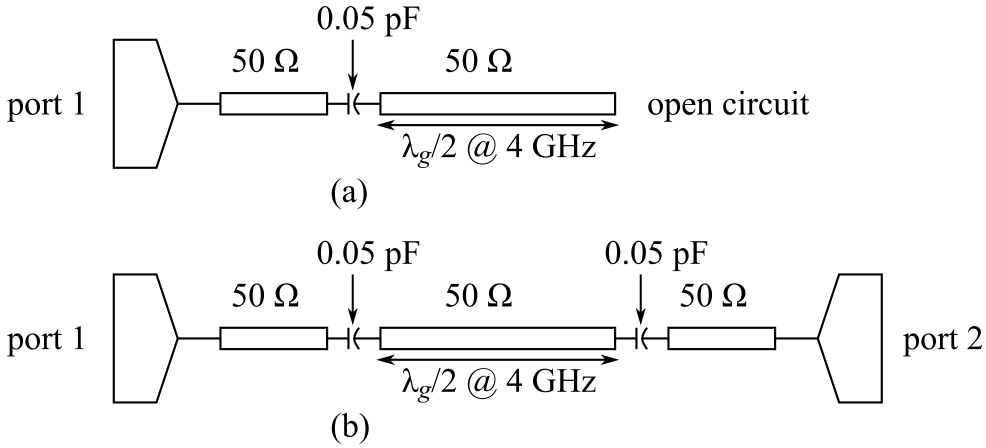



Figure 5.21: Capacitively coupled oscillators: (a) one-port; (b) two-port.

and 5 GHz. Change the following parameters and comment on how they affect the resonator behavior:

(1) impedance of resonator line (is the line exactly 180 degrees long?);

(2) Q factor (higher loss);

(3) value of coupling capacitor.

Repeat this analysis using a coupling capacitor at both ends, observing the magnitudes of both $S_{11}$ and $S_{12}$. Compare with the measurements for $S_{11}$ alone that you would do for the one-port case—which case would be easier to get precise results from? What values of coupling capacitors are reasonable to use? How would you implement them?

**6.** From the first of eqns. (5.11), derive an expression for the time-average RF power $P_L$ delivered to the real part of the load impedance $Y_L$, in terms of $G_L$, $G_d$ and $G_d''$. What is the maximum possible value of $P_L$?

# Lab 5: The Gunn-Diode Waveguide-Mounted Oscillator

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 5.22. The Gunn diode cathode

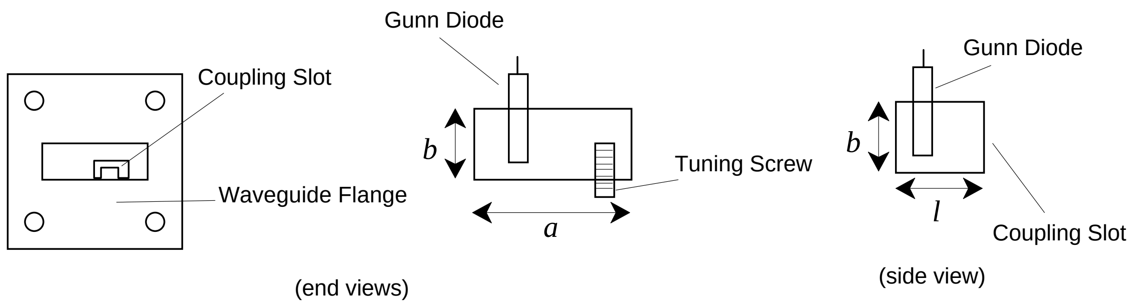



Figure 5.22: A waveguide cavity Gunn diode oscillator.

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. 5.23.

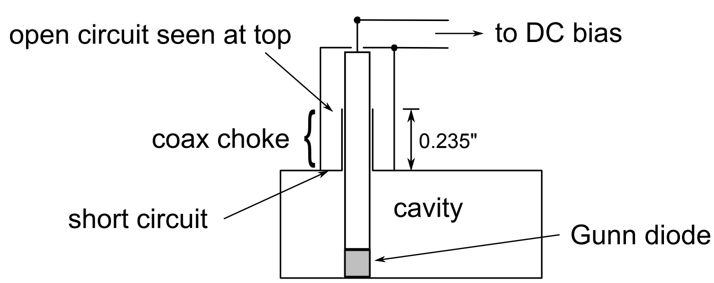



Figure 5.23: Waveguide Gunn diode housing showing coax choke.

**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 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). The diode would get very hot when DC bias voltage is applied were it not for the heat sink effect of the waveguide cavity walls. 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:** What significant feature of the *I-V* plot is present at the point where oscillations appear on the spectrum analyzer?

## 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. The tuning screw in the side of the Gunn diode resonator inserts or retracts a dielectric rod in the cavity, allowing the resonant frequency to be changed. With a small screwdriver, *carefully* adjust the tuning screw (it is soft, and easily damaged), and observe the effect on the resonant frequency. Note: if the dielectric rod has fallen out of the tuning screw holder, you may not observe any frequency change, in which case this part can be skipped.

**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 may see more than one frequency on the spectrum analyzer in this case. If so, this should occur at around 4 V.

**Q5:** What are the oscillation frequencies? Does any of these correspond to the length of the resonator? Capture the data from the spectrum analyzer to the computer to include in your lab report, and indicate the relative amplitudes of all frequencies you observe in your plot. 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 5 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 $f_{\text{osc}}$ vs. $V$. Does $f_{\text{osc}}$ depend strongly on the bias voltage?

Reset the DC bias to obtain a stable, single-frequency oscillation; about 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 10 MHz span of frequencies centered around the oscillator frequency. Connect a signal generator, set to produce a 1 MHz 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 MHz signal until you observe a change in the spectrum analyzer display. You may need to reduce the resolution bandwidth of the spectrum analyzer for this part as well as for Part V.

**Q10:** Capture the display of the spectrum analyzer to the computer. 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 MHz sine wave. What differences do you see?

**Q11:** Repeat **Q10** using modulation frequencies of 10 MHz and 100 MHz. Report and comment on your observations.

## Part V.

In this part, you will attempt what is known as *injection locking* of the oscillator. Remove the isolator and connect a 10 dB directional coupler between the Gunn diode resonator and the HP 8350 sweep oscillator as shown in Fig. 5.24. Set the bias on the Gunn oscillator to a low level as in Part III. Set the

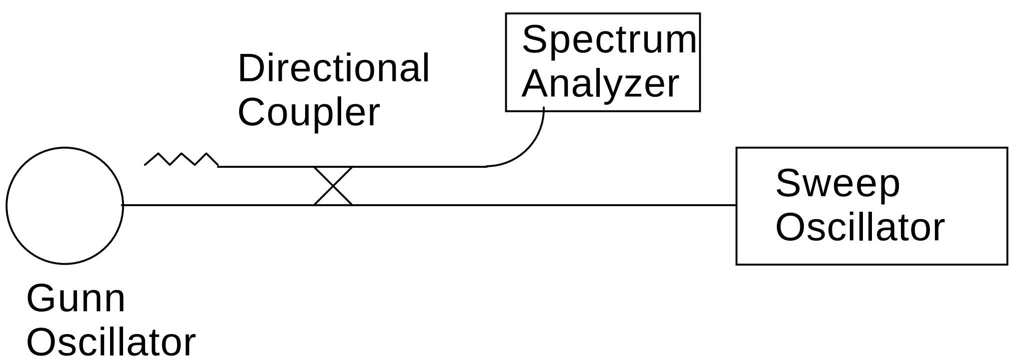



Figure 5.24: Injection locking of a waveguide Gunn diode oscillator.

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 RF output of 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. Include the spectrum analyzer plots in your lab report.

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.

**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?

# Chapter 6

# 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. $\mathbf{S}^{\dagger}\mathbf{S} = \mathbf{I}$;
- 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.

## 6.1 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 tuner is a device capable of presenting a wide range of impedances at both its ports. A common measurement technique that uses tuners is a so called *load-pull* system. It is desirable for a tuner to cover the entire passive Smith chart, and thus 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 the presence of loss prevents a tuner from reaching 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. 6.1, and are referred to as "slug tuners". Focus Microwaves, Inc. and Maury Microwave produce tuners capable of realizing VSWRs higher than 40:1 (see, e. g. http://www.focus-microwaves.com – look under manual and automatic tuners).

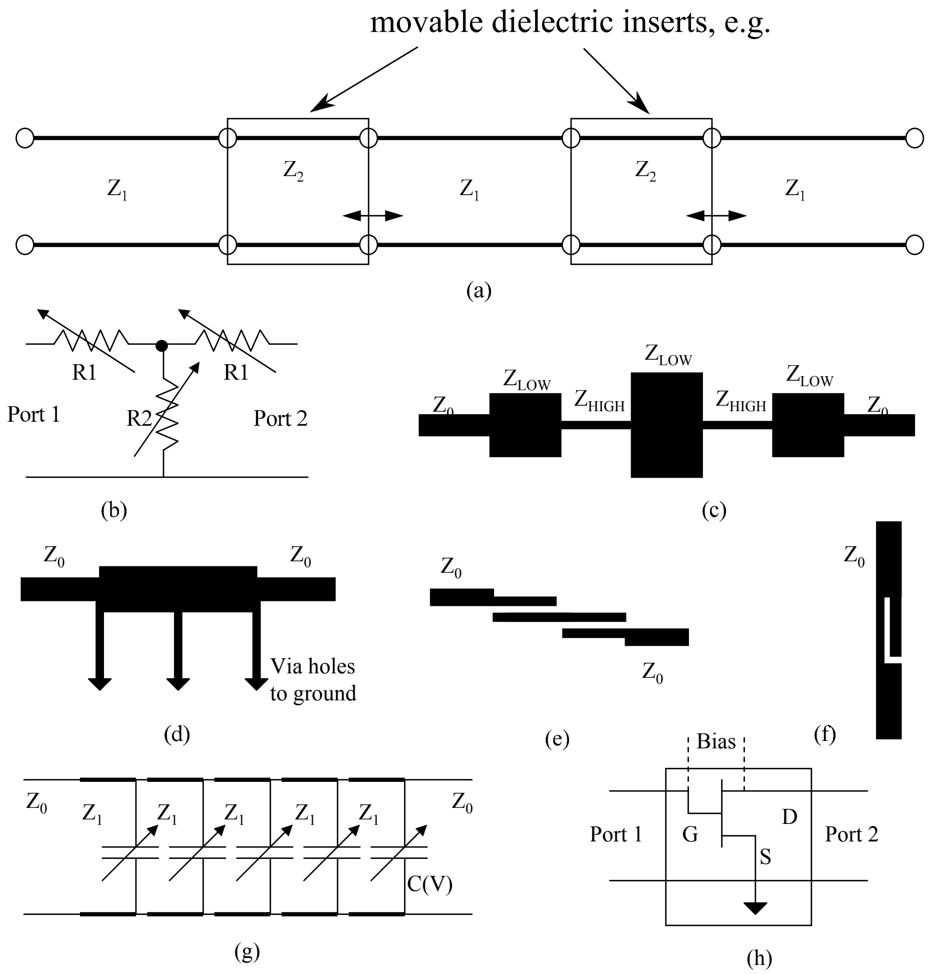



Figure 6.1: (a) Electrical model of a slug tuner. (b) Variable attenuator. (c) Low-pass filter implemented in microstrip, (d) Microstrip high-pass filter, (e) Microstrip coupled-line bandpass filter, (f) Microstrip band-stop filter, (g) Block diagram of a loaded-line phase shifter with varactor diodes, 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. 6.1(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.

Two different quantities are commonly used for characterizing the loss of an attenuator (or indeed, any two-port network). Defined in dBs, they are the *insertion loss*:

$$\mathrm{IL} = -10\log\frac{P_{\mathrm{out}}}{P_{\mathrm{avail}}} = -10\log|S_{21}|^2,$$

and the *transmission loss*:

$$\mathrm{TL} = -10\log\frac{P_{\mathrm{out}}}{P_{\mathrm{in}}} = -10\log\frac{|S_{21}|^2}{1-|S_{11}|^2}.$$

Here, $P_{\mathrm{out}}$ refers to the power emanating from port 2, $P_{\mathrm{avail}}$ refers to the available power—that which could enter port 1 if there were no reflection at the input, and $P_{\mathrm{in}}$ refers to the power that actually enters port 1.

(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, as with the artificial transmission line of section 1.7. They can also be realized as cascades of with high-impedance line sections (which behave like inductors) and low-impedance line sections (which behave like capacitors). An example of such a microstrip filter layout is shown in Fig. 6.1(c). A high-pass microstrip filter is shown in Fig. 6.1(d). What is the low-frequency circuit equivalent of this filter? In both cases, a desired cutoff frequency of the filter can be obtained using artificial transmission line theory.

The main difference between the lumped-element and transmission-line section implementations is in their overall frequency response: in transmission-line filters, as frequency increases, the properties will eventually repeat. For example, when the line lengths become a half-wavelength, the filter is electrically not present. Bandpass filters can be made using coupled-line sections (as shown, for example, in Fig. 6.1(e)), which effectively behave as resonant circuits. A stop-band microstrip filter, shown in Fig. 6.1(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:

$$\mathbf{S} = \begin{bmatrix} 0 & 0 \\ e^{j\theta} & 0 \end{bmatrix}. \tag{6.1}$$

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

$$\mathbf{S} = \begin{bmatrix} 0 & e^{j\theta} \\ e^{j\theta} & 0 \end{bmatrix}. \tag{6.2}$$

In transmission, the phase shifter adds a phase $\theta$ 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. 6.1(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 the DC bias voltage $V_{\text{dc}}$. 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.

(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. 6.1(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 Lecture 8 and lab 8.8.

## 6.2 Three-Port Networks

Networks with three or more ports are frequently used in practical microwave circuits. In connection with Lab 6.8 (multiport networks), we will next study some three-ports and four-ports in more detail.

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 reflection coefficients at the three ports are zero, so the scattering matrix of such a device has to look like:

$$\mathbf{S} = \begin{bmatrix} 0 & S_{12} & S_{13} \\ S_{12} & 0 & S_{23} \\ S_{13} & S_{23} & 0 \end{bmatrix}. \tag{6.3}$$

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

$$\begin{aligned} |S_{12}|^2 + |S_{13}|^2 &= 1 & \quad S_{13}{}^* S_{23} = 0 \\ |S_{12}|^2 + |S_{23}|^2 &= 1 & \quad S_{23}{}^* S_{12} = 0 \\ |S_{13}|^2 + |S_{23}|^2 &= 1 & \quad S_{12}{}^* S_{13} = 0. \end{aligned} \tag{6.4}$$

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 three-port network that is simultaneously lossless, reciprocal and matched is not possible (It should be noted that sometimes one port of a 4-port network is terminated in a matched load and is then sealed and not accessible for external connections. Thus, a 4-port device can be effectively turned into a matched, reciprocal three-port. This three-port is a lossy one, however, because of the matched load termination at the fourth port).

If one of these three conditions is dropped, the device becomes feasible, and devices useful for certain purposes can be made. For example, if we assume the device is not reciprocal, but is matched and lossless, the following scattering matrix results:

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

A device that has such scattering parameters is called a *circulator*, and its symbol is shown in Fig. 6.2(a). It has the property that all power coming into port 1 will emerge from port 2, and none will go out of

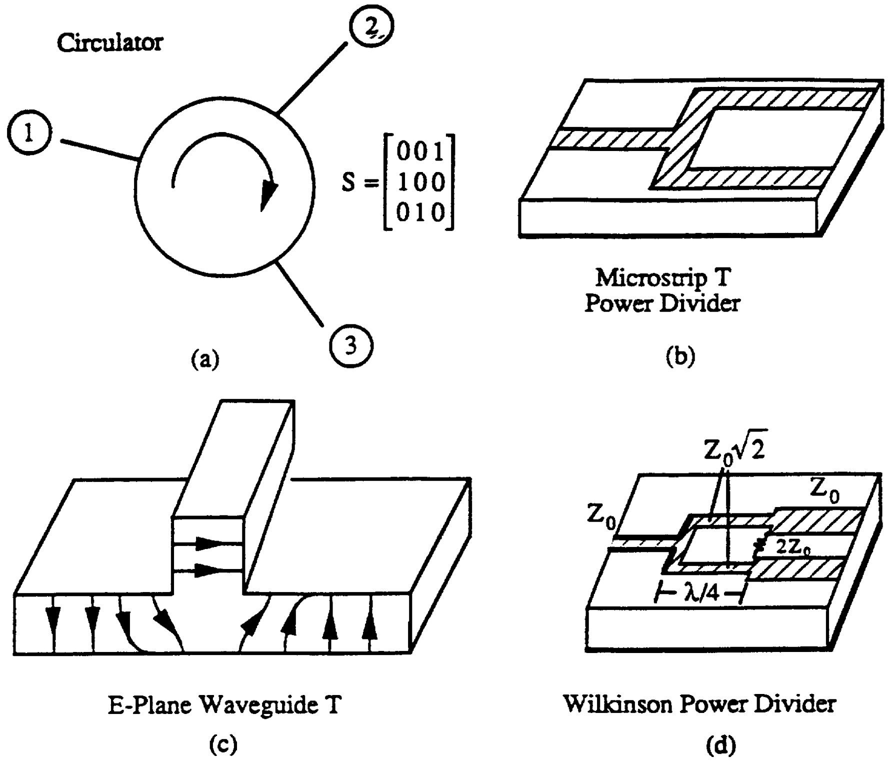



Figure 6.2: The symbol and scattering matrix for a circulator (a), and examples of lossless (b), (c) and resistive (d), power dividers.

port 3, power going into port 2 will only come out of port 3, and power going into port 3 will only come out of 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. 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. 6.2(b),(c). Finally, if we relax the lossless condition, we get a resistive power divider, an example of which is shown in Fig. 6.2(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.

Sometimes we need a three-port device to separate frequencies from different bands into separate outputs; these are called *diplexers*. These circuits can also be designed to have more than two outputs. All of these components are often used in microwave circuits.

Finally, some three-port networks are nonlinear active circuits, such as amplifiers, mixers and oscillators. They cannot exactly be represented with conventional scattering parameters, because conversion occurs between different frequency components. However, we shall see in Lecture 8 that S-parameters can still be used in the linear (small-signal) approximation in the design of amplifiers.

## 6.3 Four Port Networks—The Directional Coupler

Four-port microwave devices are also used very often, especially ones called *directional couplers* or *hybrid networks*. A hybrid network has 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*, meaning that no signal emerges from it. The terminology came from the *hybrid coil*, a multi-winding transformer that was used to isolate signals traveling in different directions in the early days of telephony. A special kind of hybrid often used in waveguide circuits is the magic T. Different schematic symbols used for directional couplers are shown in Fig. 6.3. Three parameters are defined for a directional coupler, and these are what are usually

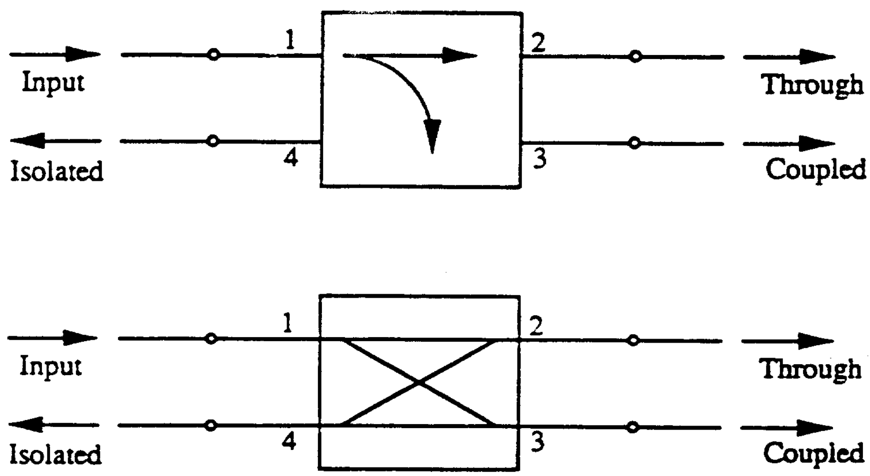



Figure 6.3: Symbols for directional couplers.

given in spec sheets when you buy one:

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

The coupling factor tells us what fraction of the input power emerges from the through or output port. The directivity and isolation are different measures of the ideality of the device—how well the coupler isolates forward from backward traveling waves. An ideal directional coupler has an infinite directivity and isolation ($S_{41} = 0$).

By writing down equations similar to the ones we used for three-ports, we can find out that a lossless, reciprocal and matched four-port with perfect isolation between ports 1 and 4, and between ports 2 and

3 can indeed exist. Its scattering matrix will have the form

$$\mathbf{S} = \begin{bmatrix} 0 & S_{21} & S_{31} & 0 \\ S_{21} & 0 & 0 & S_{42} \\ S_{31} & 0 & 0 & S_{43} \\ 0 & S_{42} & S_{43} & 0 \end{bmatrix} \tag{6.5}$$

The property of reciprocity gives the conditions $S_{ik} = S_{ki}$, while from (2.16) the fact that the network is lossless gives

$$S_{21}S_{42}^* + S_{31}S_{43}^* = 0; \qquad S_{21}S_{31}^* + S_{42}S_{43}^* = 0 \tag{6.6}$$

$$|S_{21}|^2 + |S_{31}|^2 = 1; \qquad |S_{31}|^2 + |S_{43}|^2 = 1; \qquad |S_{42}|^2 + |S_{43}|^2 = 1 \tag{6.7}$$

which also imply that

$$|S_{42}| = |S_{31}|; \qquad |S_{43}| = |S_{21}| \tag{6.8}$$

All other consequences of (2.16) are redundant or trivial.

The positions of the reference planes at the ports can still be adjusted, leaving some degrees of freedom in the phases of these S-parameters. If we assume symmetry of the junction (that is, we can interchange ports $1 \leftrightarrow 2$ and $3 \leftrightarrow 4$ or similarly switch the top and bottom ports without changing the network) then suitable choices of reference planes can be made so that $S_{21} = S_{43} = A$, a real number. Then $|S_{31}| = |S_{42}| = \sqrt{1 - A^2}$, and (6.6) implies that the phase angles of $S_{31}$ and $S_{42}$ must be related (within a multiple of $2\pi$) by:

$$\phi_{31} + \phi_{42} = \pi \tag{6.9}$$

Abbreviating $\phi_{31} = \phi$, we see that

$$S_{31} = Be^{j\phi}; \qquad S_{42} = -Be^{-j\phi} \tag{6.10}$$

where $B = \pm\sqrt{1 - A^2}$ is a real number and $\phi$ is a free parameter.

The two most common choices of $\phi$ are $\pm 90°$, which results in

$$\mathbf{S} = \begin{bmatrix} 0 & A & jB & 0 \\ A & 0 & 0 & jB \\ jB & 0 & 0 & A \\ 0 & jB & A & 0 \end{bmatrix}, \tag{6.11}$$

and 0° or 180°, which results in

$$\mathbf{S} = \begin{bmatrix} 0 & A & B & 0 \\ A & 0 & 0 & -B \\ B & 0 & 0 & A \\ 0 & -B & A & 0 \end{bmatrix}, \tag{6.12}$$

where $A$ and $B$ are real numbers such that $A^2 + B^2 = 1$. 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.

An important design parameter is how the power splits between the two non-isolated ports—that is, the choice of $A$. A special, and very often used case, is when the power splits equally. This is called a 3 dB coupler. In this case, $A = \pm 1/\sqrt{2}$ and $B = \pm 1/\sqrt{2}$, so the two matrices from above become:

$$\mathbf{S} = \pm\frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & j & 0 \\ 1 & 0 & 0 & j \\ j & 0 & 0 & 1 \\ 0 & j & 1 & 0 \end{bmatrix} \quad , \quad \mathbf{S} = \pm\frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{bmatrix}. \tag{6.13}$$

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. 6.4. If port 1 is the input, then port 2 is the through port. A small amount of

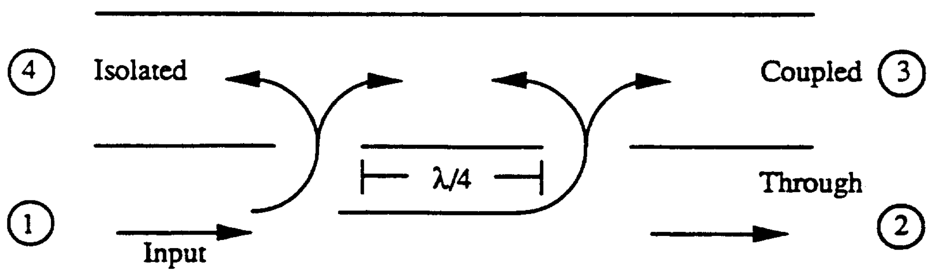



Figure 6.4: A waveguide directional coupler with two coupling holes $\lambda/4$ apart.

the input wave leaks out through the two coupling holes into the second waveguide, and both of these 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 traveled 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.

The magic T (or magic tee) shown in Figure 6.5(a) is a special kind of so-called hybrid waveguide junction with four ports. If one of the ports is the input, and a certain other one is terminated with a matched load, the power will split equally between the two remaining ports. To see how this works, let a $TE_{10}$ mode be incident at port 1. There is an odd symmetry about the waveguide at port 4, and the field lines of $E_y$ behave as shown in Figure 6.5(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 ports 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 will split equally between ports 2 and 3, but with a 180° phase shift. The field lines for this case are sketched in Figure 6.5(c).

## 6.4 Odd and Even Mode Analysis of a Branch Line Directional Coupler

A directional coupler made with sections of transmission line is shown in Fig. 6.6. 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. 6.7. We will assume that the circuit is linear, so that superposition holds. 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, for each of which the circuit essentially reduces to two identical two-port networks. Let us see how this works on the example of a branch line coupler.

Fig. 6.8(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. 6.8(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

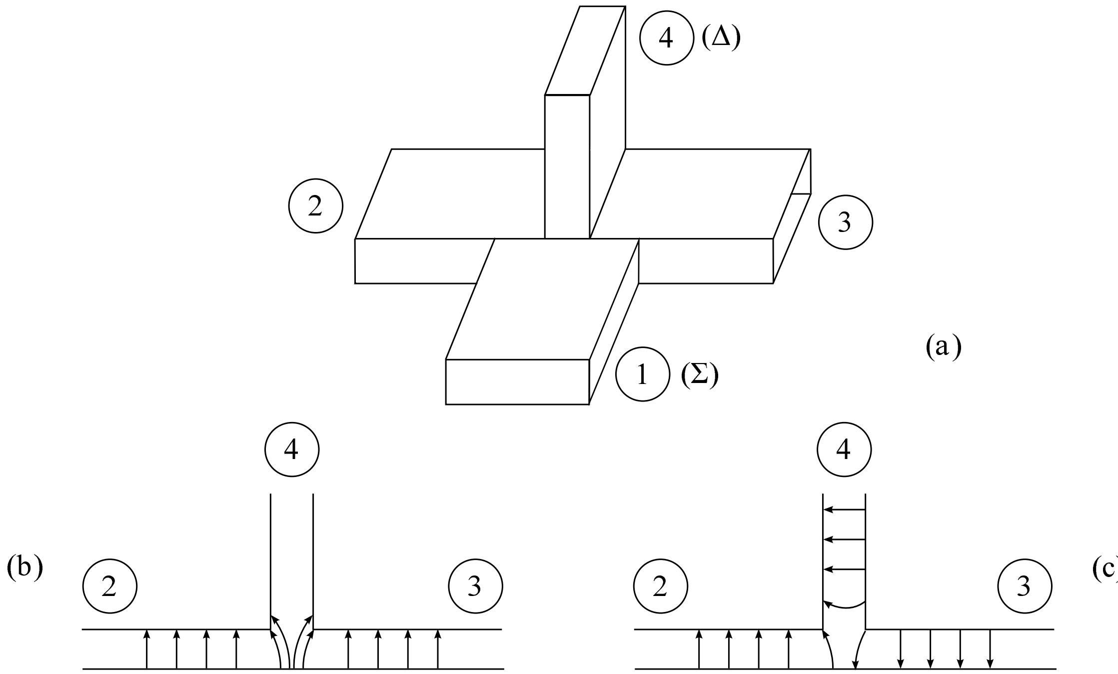



Figure 6.5: The magic T (a) and electric field lines for $TE_{10}$ mode excitation from port 1 (b) or port 4 (c).

added up to get the response of the four-port coupler. If we denote the odd and even mode reflection and transmission coefficients by $\rho_o$, $\rho_e$, $\tau_o$ and $\tau_e$, the following equations can be written for the amplitudes of the waves coming out of the four ports:

$$\begin{aligned} b_1 &= \frac{1}{2}\rho_e + \frac{1}{2}\rho_0 \\ b_2 &= \frac{1}{2}\tau_e + \frac{1}{2}\tau_0 \\ b_3 &= \frac{1}{2}\tau_e - \frac{1}{2}\tau_0 \\ b_4 &= \frac{1}{2}\rho_e - \frac{1}{2}\rho_0, \end{aligned} \tag{6.14}$$

assuming there is an incident wave $a_1 = 1$ at port 1 only ($a_2 = a_3 = a_4 = 0$), as shown in Fig. 6.8.

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

$$\begin{array}{ll} \rho_e = 0 & \rho_0 = 0 \\ \tau_e = -\frac{1}{\sqrt{2}}(1+j) & \tau_0 = \frac{1}{\sqrt{2}}(1-j). \end{array} \tag{6.15}$$

From these, we get

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

This result agrees with the first row and column of the scattering matrix [given in the first of (6.13)] for a 90° hybrid symmetrical directional coupler.

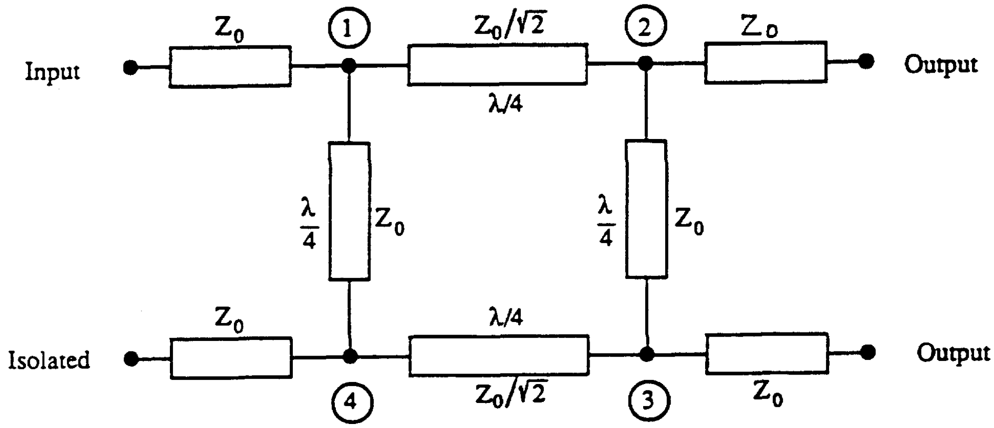



Figure 6.6: A branch line directional coupler.

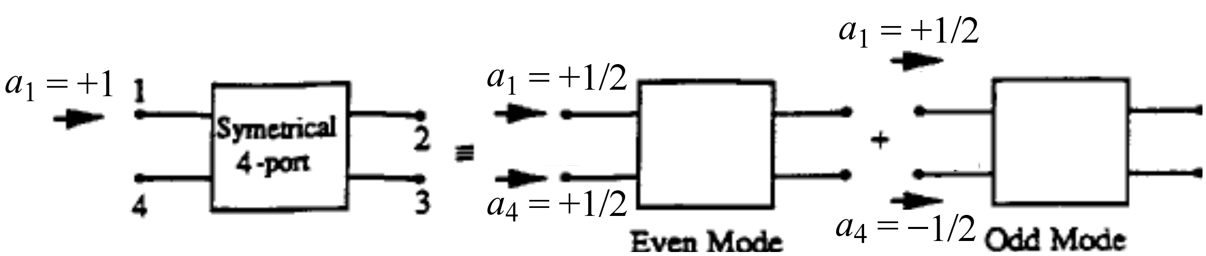
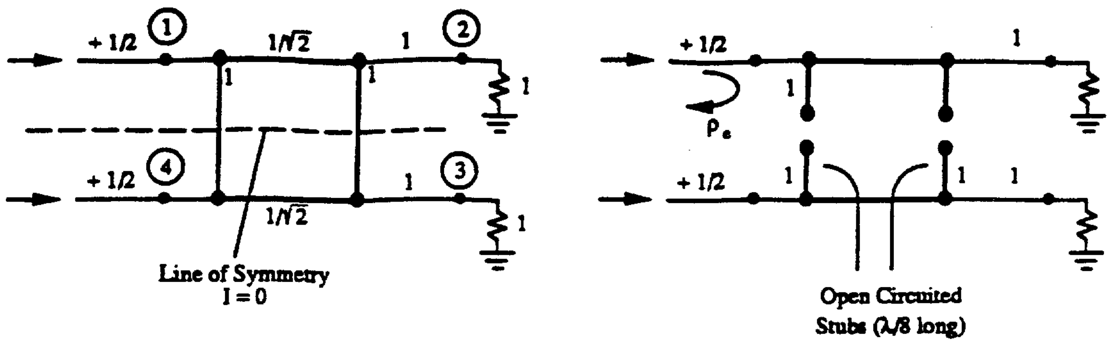
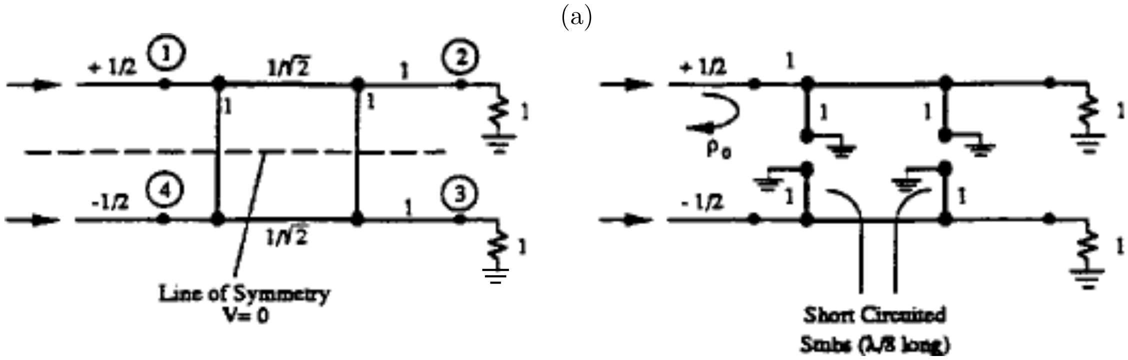



Figure 6.7: Odd and even mode decomposition for a symmetric four-port.




(a)




(b)

Figure 6.8: Normalized branch line coupler decomposed into even (a) and odd (b) mode excitation.

## 6.5 Operating Principles of the VNA

In section 2.10, we modeled a VNA by a hypothetical ideal VNA to which error boxes were connected to account for all the nonideal aspects of a real VNA. This allowed us to present the principles behind calibration of a VNA, but did not discuss how the ideal VNA could be realized, even approximately. As we saw in the previous chapter, the rapid variations with which the phase of a microwave signal is associated are virtually impossible for most instrumentation to follow accurately, and only time-average power (or amplitude) is directly measurable. In this section and the next one, we will show how such information can be used to accomplish the task of measuring ratios of wave amplitudes in RF and microwave networks.

The phase of a microwave signal can usually only be measured indirectly: for example, a reference signal and a signal to be measured can be sent to a mixer (see section 9.2) 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.

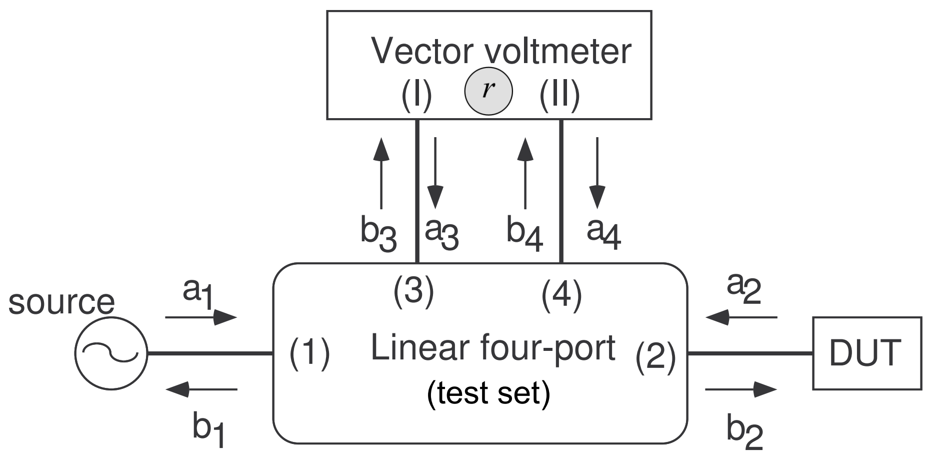



Figure 6.9: The vector voltmeter based network analyzer.

A common way to perform power measurements, which is used in most commercial instruments, is to combine a linear four-port (called a *test set*) with a vector voltmeter. The vector voltmeter used in commercial network analyzers is a two-port device that is linear with respect to the waves at its two RF ports ($I$ and $II$). It generates a complex response that can be written in the form

$$r = K_0 + K_r \frac{a_I}{a_{II}} \tag{6.16}$$

where $a_I$ and $a_{II}$ are the incident wave amplitudes at its ports, 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 ports 3 and 4 of a test set. The other two ports are connected to a source and the device under test (DUT), Fig. 6.9. The configuration described here is usually called a *reflectometer*, and the test set typically consists of two directional couplers along with some auxiliary circuitry. The principle of operation of this circuit 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 port 1 of the test set, 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 via ports 3 and 4 of the test set and get partially reflected there as well. The vector voltmeter samples different combinations of the waves that are incident and reflected at the DUT. Since the test set and the vector voltmeter are both linear multiports, we can write the outgoing waves from the test set in terms of $a_1$ and $a_2$ only:

$$\begin{aligned} b_2 &= M_1 a_1 + M_2 a_2 \\ b_3 &= L_1 a_1 + L_2 a_2 \\ b_4 &= K_1 a_1 + K_2 a_2 \end{aligned} \tag{6.17}$$

where the $K$'s, $M$'s and $L$'s are complex constants that depend on the $S$-parameters of the test set and those of the vector voltmeter.

By definition, $a_2 = \rho b_2$, where $\rho$ 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 $b_2$:

$$\begin{aligned} b_2 &= \frac{M_1}{1-\rho M_2} a_1 \\ b_3 &= L_1 a_1 + L_2 \rho b_2 \\ b_4 &= K_1 a_1 + K_2 \rho b_2 \end{aligned} \tag{6.18}$$

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 $\rho$ in the numerator, you should get the following expression:

$$r = \frac{\rho + A}{B\rho + C}. \tag{6.19}$$

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 $r$, 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. The three unknown complex constants in (6.19) 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, as we saw in section 2.10, three known cal standards can be measured to determine these constants. From the formula for $r$,

$$-A + \rho r B + rC = \rho. \tag{6.20}$$

which is a linear equation in the unknown constants $A$, $B$ and $C$. Therefore, by observing the three values of $r$ that result from measuring three known values of $\rho$, the constants $A$, $B$ and $C$ can be determined. Just as we saw from the error-box model in section 2.10, 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. 6.10. The part of the test set connected to port 1 of the DUT is the reflectometer from above (it is in the dashed box in Fig. 6.10), 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

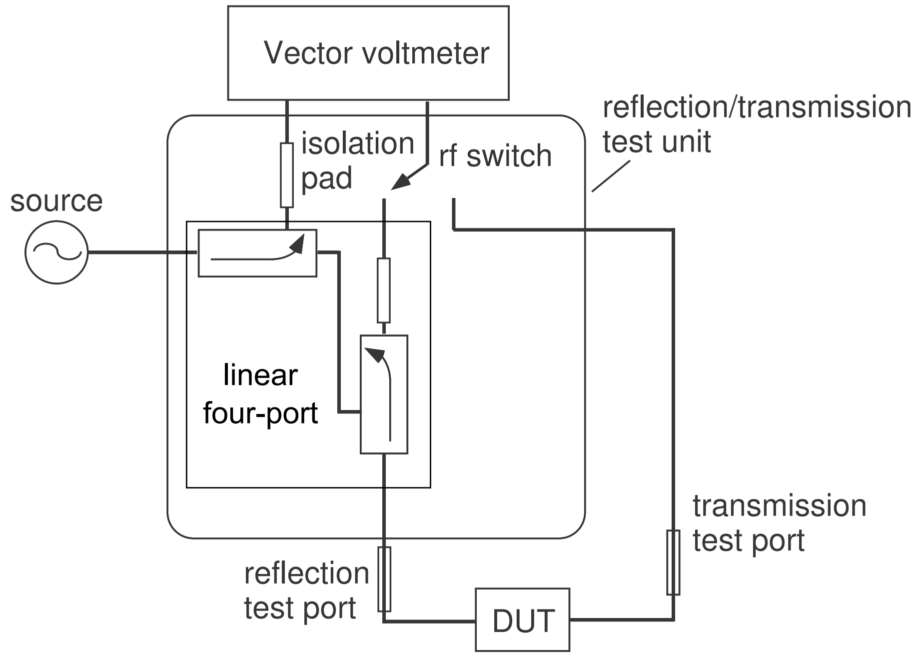



Figure 6.10: A reflection-transmission test set for full $S$-parameter measurements using a 4-port network analyzer.

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?

## 6.6 Multiport Reflectometry

As we saw in section 6.5, the principle behind most commercial microwave VNAs is that of multiport reflectometry. There, we postulated the existence of a vector voltmeter without inquiring how such a device could be realized in practice. In this section we will discuss in detail another related modern method for reflection coefficient measurement that is also based on the use of interconnected multiports, but uses time-average power measurements directly rather than through the intermediary of a vector voltmeter. In the process, we will gain deeper insight into the various nonideal factors that are accounted for by the calibration process.

Suppose that a microwave oscillator is connected to two directional couplers with fairly small coupling coefficients $C_3$ and $C_4$, and then to an unknown load whose reflection coefficient is $\rho$, as shown in Fig. 6.11. Power detectors, such as diode detectors, are connected to the coupled ports of the directional

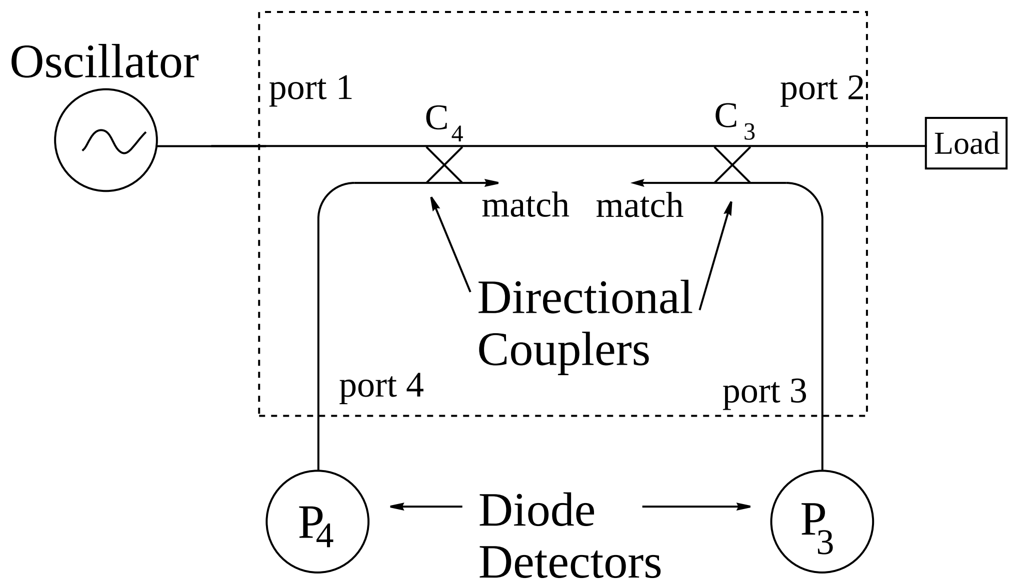



Figure 6.11: Four-port reflectometer (the test set is inside the dashed box).

couplers. The power detectors, oscillators and directional couplers will be assumed to have small reflection coefficients, so that $P_3 = |C_3|^2|a_1|^2$ and $P_4 = |C_4|^2|b_1|^2$. Under these conditions, if the detectors are operated in their square-law regimes the ratio of the measured powers will be

$$\frac{P_4}{P_3} = q_{43}|\rho|^2 \tag{6.21}$$

where $q_{43} = |C_4/C_3|^2$ is independent of the unknown reflection coefficient $\rho$. If diode detectors are used to measure power, (6.21) can be replaced by

$$\frac{V_4}{V_3} = Q_4|\rho|^2 \tag{6.22}$$

where $V_3$ and $V_4$ are the DC voltages measured at the diode (under zero bias voltage conditions). The constant $Q_4$ now depends on the properties of the directional couplers as well as those of the diodes, but still not on the unknown $\rho$; it is called a calibration constant and can be determined by a calibration procedure.

For the present case, it is sufficient to connect a short circuit calibration standard to the load port, for which we know that $\rho = -1$. Thus, the measurements of $V_3$ and $V_4$ would give:

$$Q_4 = \left.\frac{V_4}{V_3}\right|_{\rho=-1} \tag{6.23}$$

We do not need to have the directional couplers or diodes to be made to precise specifications, or even to be identical to each other. As long as the calibration measurement can be made, the constant $Q_4$ can be computed. Once this is done, the actual load can be connected to the load port, and the diode voltages measured. From (6.22), the magnitude of $\rho$ can be determined.

Unfortunately, this procedure gives no information about the phase of $\rho$ (we have made only a scalar network analyzer so far). To obtain phase information, we must modify the configuration of our reflectometer. Suppose now that we replace the circuit shown in Fig. 6.11 with the one shown in Fig. 6.12. Instead of a match, the fourth port of the second coupler is terminated in a short circuit whose

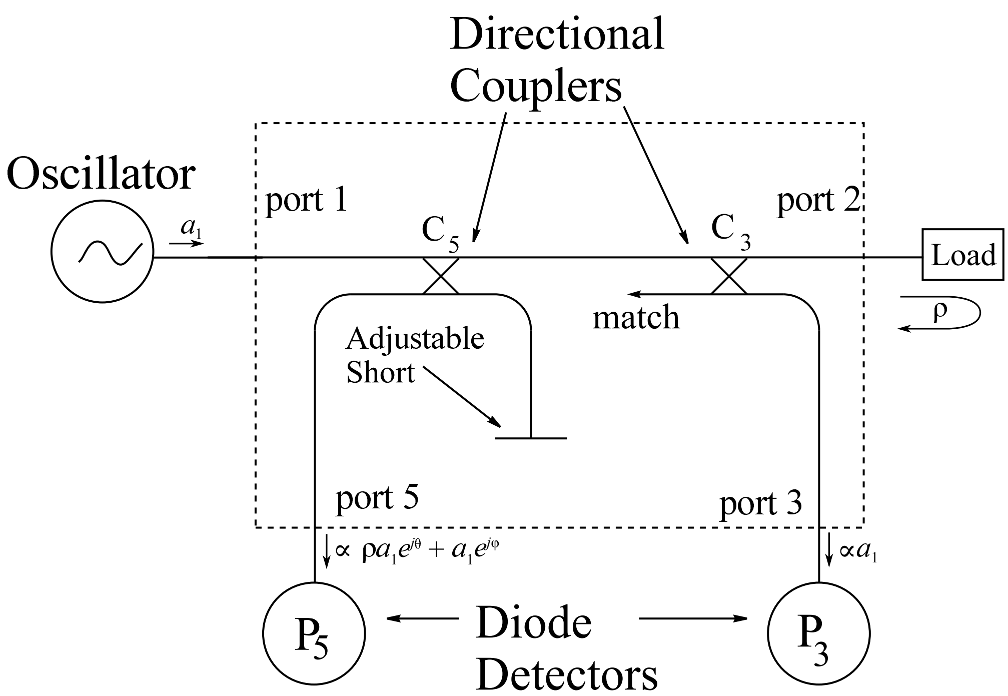



Figure 6.12: Modified four-port reflectometer.

position is adjustable. We rename the port where power is measured on this coupler to be port 5, to avoid confusion with the previous configuration. If the coupling coefficients of the directional couplers are again assumed to be small (and furthermore to be identical), then the ratio of the two diode voltages will have the form

$$\frac{V_5}{V_3} = Q_5|\rho - e^{j\phi_5}|^2 \tag{6.24}$$

where $Q_5$ and the angle $\phi_5$ are calibration constants. The phase angle $\phi_5$ is dependent on the path length difference between $a_1$ and $b_1$ due to the routes they travel on the main line and on the coupled line. All that really matters is that all calibration constants are independent of the load reflection coefficient, and can be determined by calibration measurements (the details of which we will postpone until a little later).

To simplify our equations somewhat, we introduce the notation

$$\Lambda_n = \frac{V_n}{V_3}; \qquad n = 4, 5, \dots \tag{6.25}$$

for ratios of measured voltages (these are real and $\geq 0$). If the calibration constants $Q_5$ and $\phi_5$ are known, then (6.24) says that the unknown $\rho$ lies on a circle in the complex plane, centered at $\rho_5 = e^{j\phi_5}$ and having a radius

$$\sqrt{\frac{\Lambda_5}{Q_5}} \equiv R_5 \tag{6.26}$$

If this information is combined with that from (6.22), which says that $\rho$ lies on a circle centered at the origin with a radius

$$\sqrt{\frac{\Lambda_4}{Q_4}} \equiv R_4 \tag{6.27}$$

we conclude that $\rho$ must be located at one of the intersection points of these two circles. To resolve the question of which intersection is the correct one, we must take a further measurement, replacing the circuit of Fig. 6.12 with a similar one, wherein the short circuit is adjusted so as to give a different phase angle $\phi_6$. Since the diode and directional coupler properties may also be different, we have

$$\Lambda_6 = Q_6|\rho - e^{j\phi_6}|^2 \tag{6.28}$$

meaning that $\rho$ is located on yet a third circle, centered at $\rho_6 = e^{j\phi_6}$ and with a radius

$$\sqrt{\frac{\Lambda_6}{Q_6}} \equiv R_6 \tag{6.29}$$

The intersection of the three circles thus determines both the magnitude and phase of $\rho$ uniquely, as shown in Fig. 6.13.

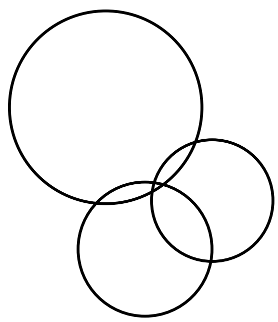


Figure 6.13: Intersection of three circles to determine complex $\rho$.

We do not need to use three different experimental configurations to make our measurements. All three can be combined into one circuit, as shown in Fig. 6.14. If this is done, there can arise significant

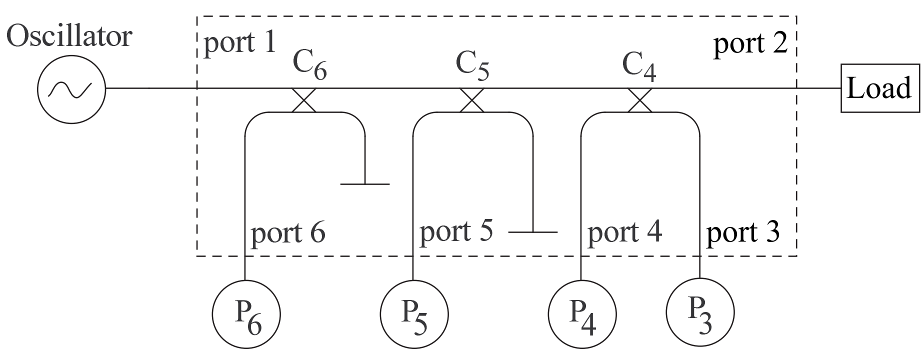



Figure 6.14: Six-port reflectometer.

deviations from the assumptions we made in deriving our equations because the couplers do not have perfect match, isolation, etc. and the diodes are not perfectly matched. For example, the directional couplers can reintroduce some of the sampled waves back into the main line, thus modifying what we are trying to measure. For this reason, we have to modify the equations (6.22) (6.24) and (6.28) to allow the centers of these circles in the complex $\rho$-plane to be at arbitrary locations $\rho_4$, $\rho_5$ and $\rho_6$, so that for the reflectometer in Fig. 6.14 they take the form

$$|\rho - \rho_n|^2 = \frac{\Lambda_n}{Q_n} = R_n^2; \qquad n = 4, 5, 6 \tag{6.30}$$

for some positive real calibration constants $Q_4$, $Q_5$ and $Q_6$, and some complex calibration constants $\rho_4$, $\rho_5$ and $\rho_6$, all independent of the load reflection coefficient $\rho$.

If the calibration constants in (6.30) are known, we have three real equations for two real unknowns (the real and imaginary parts of $\rho = u + jv$). Experimental error will always result in the three circles not intersecting exactly, but merely coming close to doing so. The extent to which they do not all intersect at the same point is one measure of the error in determining the reflection coefficient. Indeed, drawing the three circles on a graph (a Smith chart will do nicely, but is not necessary) and locating their near intersection point is one way to compute $\rho$ from the measured data. This process can be automated using an optimization technique such as "least squares" solution of (6.30). However, the determination of $\rho$ in this way is somewhat arbitrary, since the choice of error criterion that gives a "best fit" solution of (6.30) can be made in many different ways. A more systematic approach is to find the *radical center* of the three circles, as we will now describe. Once this is done it is a good idea to substitute the calculated value of $\rho$ back into the original equations to check that a good solution has indeed been found, and to estimate the experimental error.

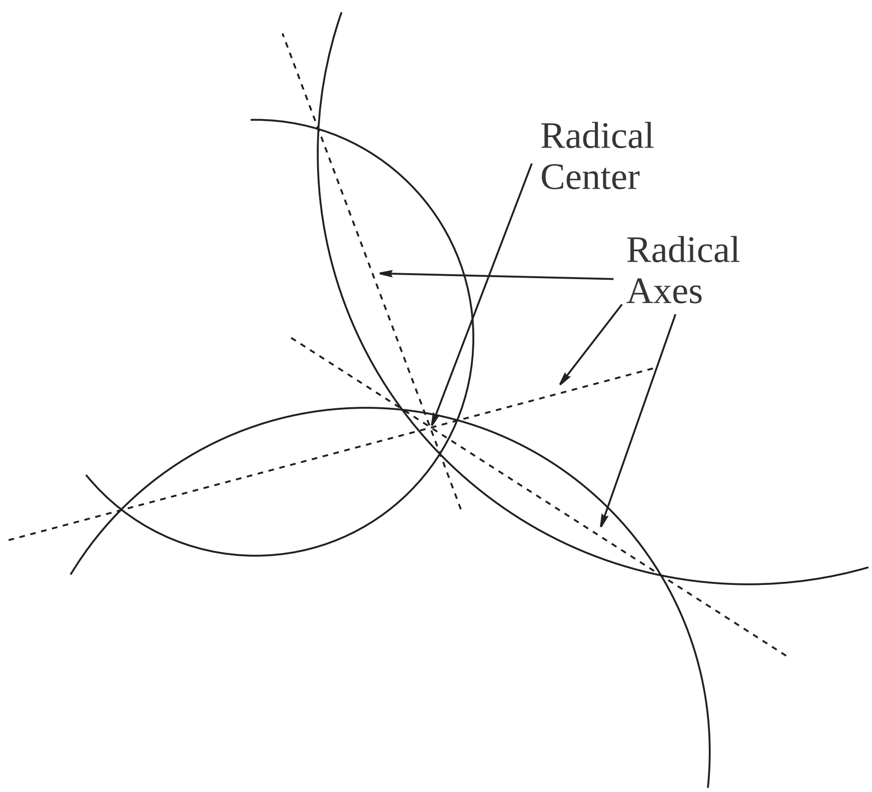



Figure 6.15: Radical axes and radical center of three circles in the complex $\rho$-plane.

If we take a closer look at the near intersection of the three circles, we have the situation shown in

Fig. 6.15. A straight line drawn through the intersection points of any pair of circles is called the *radical axis* of that pair of circles. A theorem of geometry tells us that the three radical axes generated by our three circles must all intersect in a single point, if the centers $\rho_4$, $\rho_5$ and $\rho_6$ of the circles are all different. This point is called the radical center of the three circles, and from Fig. 6.15 we can see that it is in a certain sense the best fit solution of equations (6.30). It can be calculated by first deriving the equation for a typical radical axis, say the one for the circles centered at $\rho_4$ and $\rho_5$. We have

$$(u-u_4)^2+(v-v_4)^2=R_4^2; \qquad (u-u_5)^2+(v-v_5)^2=R_5^2 \tag{6.31}$$

Subtracting these two equations and rearranging gives an equation for the radical axis:

$$u\,(u_4-u_5)+v\,(v_4-v_5)=K_4-K_5 \tag{6.32}$$

where

$$K_n=\frac{u_n^2+v_n^2-R_n^2}{2}=\frac{|\rho_n|^2-R_n^2}{2}; \qquad n=4,5,6 \tag{6.33}$$

In a similar way, the equation for the radical axis of the circles centered at $\rho_5$ and $\rho_6$ is

$$u\,(u_5-u_6)+v\,(v_5-v_6)=K_5-K_6 \tag{6.34}$$

and the radical center is found by solving (6.32) and (6.34) using determinants:

$$\boxed{u=\frac{N_u}{D}; \qquad v=\frac{N_v}{D}} \tag{6.35}$$

where $N_u$, $N_v$ and $D$ are the matrix determinants

$$N_u=\det\begin{bmatrix}(K_4-K_5) & (v_4-v_5)\\ (K_5-K_6) & (v_5-v_6)\end{bmatrix} \qquad N_v=\det\begin{bmatrix}(u_4-u_5) & (K_4-K_5)\\ (u_5-u_6) & (K_5-K_6)\end{bmatrix} \tag{6.36}$$

and

$$D=\det\begin{bmatrix}(u_4-u_5) & (v_4-v_5)\\ (u_5-u_6) & (v_5-v_6)\end{bmatrix} \tag{6.37}$$

Equation (6.35) is ready to use once the calibration constants are found: $u_n$, $v_n$ and $Q_n$ are known from calibration, and $\Lambda_n$ (and therefore $R_n$ and $K_n$) are known from measurement. The resulting value of $\rho=u+jv$ should be substituted back into (6.30) to check the accuracy of the solution. Note that the centers $\rho_4$, $\rho_5$ and $\rho_6$ of the circles (6.30) should be well separated in the complex plane in order to avoid having two approximately identical circles, which would degrade the accuracy of the measurement. The separation can be accomplished by appropriate settings of the adjustable short circuits in Fig. 6.14.

The calibration constants have to be determined by a calibration procedure, just as with any network analyzer. Because there are more calibration constants to determine than when we assumed $\rho_4=0$ and that $\rho_5$ and $\rho_6$ lie on the unit circle, more calibration measurements and calculations will have to be done than in that case, but the resulting value found for an unknown $\rho$ will be more accurate. Because there are three real quantities (one real $Q_n$ and one complex $\rho_n$) for each $n$ to be determined in (6.30), it would seem that three calibration standards will have to be measured, giving three conditions to be satisfied for $Q_n$ and $\rho_n$ for each $n=4,5,6$. However, we encounter a situation similar to the one above when finding the unknown load reflection coefficient $\rho$, in that three cal standards will give two possible solutions for the calibration constants, and an extra cal standard is needed to resolve this ambiguity.

Let our first three cal standards have known values of reflection coefficient $\rho_a=u_a+jv_a$, $\rho_b=u_b+jv_b$ and $\rho_c=u_c+jv_c$ (chosen in some convenient way), and let the final cal standard be a matched load ($\rho_d=0$). If we denote the corresponding measured values of $\Lambda_n$ (for each $n=4,5,6$) by $\Lambda_{na}$, $\Lambda_{nb}$, $\Lambda_{nc}$ and $\Lambda_{nd}$, then at each port $n$, (6.30) gives four separate equations

$$|\rho_a-\rho_n|^2=\frac{\Lambda_{na}}{Q_n}; \qquad n=4,5,6 \tag{6.38}$$

$$|\rho_b - \rho_n|^2 = \frac{\Lambda_{nb}}{Q_n}; \qquad n = 4, 5, 6 \tag{6.39}$$

$$|\rho_c - \rho_n|^2 = \frac{\Lambda_{nc}}{Q_n}; \qquad n = 4, 5, 6 \tag{6.40}$$

$$|\rho_n|^2 = \frac{\Lambda_{nd}}{Q_n}; \qquad n = 4, 5, 6 \tag{6.41}$$

for determining $Q_n$ and $\rho_n$, subject to the additional constraint that $Q_n > 0$. We can temporarily eliminate $Q_n$ from these equations by dividing each of (6.38)-(6.40) by (6.41), giving:

$$|\rho_a - \rho_n|^2 = \frac{\Lambda_{na}}{\Lambda_{nd}} |\rho_n|^2; \qquad n = 4, 5, 6 \tag{6.42}$$

$$|\rho_b - \rho_n|^2 = \frac{\Lambda_{nb}}{\Lambda_{nd}} |\rho_n|^2; \qquad n = 4, 5, 6 \tag{6.43}$$

$$|\rho_c - \rho_n|^2 = \frac{\Lambda_{nc}}{\Lambda_{nd}} |\rho_n|^2; \qquad n = 4, 5, 6 \tag{6.44}$$

Some algebra will show that (6.42)-(6.44) can be rewritten as

$$\left|\rho_n - \frac{\Lambda_{nd}}{\Lambda_{nd} - \Lambda_{na}}\rho_a\right|^2 = \frac{\Lambda_{na}\Lambda_{nd}}{(\Lambda_{nd} - \Lambda_{na})^2} |\rho_a|^2; \qquad n = 4, 5, 6 \tag{6.45}$$

$$\left|\rho_n - \frac{\Lambda_{nd}}{\Lambda_{nd} - \Lambda_{nb}}\rho_b\right|^2 = \frac{\Lambda_{nb}\Lambda_{nd}}{(\Lambda_{nd} - \Lambda_{nb})^2} |\rho_b|^2; \qquad n = 4, 5, 6 \tag{6.46}$$

$$\left|\rho_n - \frac{\Lambda_{nd}}{\Lambda_{nd} - \Lambda_{nc}}\rho_c\right|^2 = \frac{\Lambda_{nc}\Lambda_{nd}}{(\Lambda_{nd} - \Lambda_{nc})^2} |\rho_c|^2; \qquad n = 4, 5, 6 \tag{6.47}$$

It will be seen that these equations are circles in the complex $\rho_n$-plane in the same form as (6.30), although now the unknown is $\rho_n$, while $\rho_a$, $\rho_b$, $\rho_c$, $\Lambda_{na}$, $\Lambda_{nb}$, $\Lambda_{nc}$ and $\Lambda_{nd}$ are known. To ensure accurate calibration, the centers of the circles (6.45)-(6.47) should be well separated. Although the settings of the adjustable short circuits will have some effect on this separation as noted above, the choices of phase for $\rho_a$, $\rho_b$ and $\rho_c$ will be even more important in achieving the desired positioning of the circle centers.

We may thus use the same method as before to find the $\rho_n$. After some algebra we arrive at

$$\boxed{u_n = \frac{N_{un}}{D_n}; \qquad v_n = \frac{N_{vn}}{D_n}} \tag{6.48}$$

where $N_{un}$, $N_{vn}$ and $D_n$ are the matrix determinants

$$N_{un} = \frac{1}{2}\det\begin{bmatrix} \left(\frac{|\rho_a|^2}{\Lambda_{nd} - \Lambda_{na}} - \frac{|\rho_b|^2}{\Lambda_{nd} - \Lambda_{nb}}\right) & \left(\frac{v_a}{\Lambda_{nd} - \Lambda_{na}} - \frac{v_b}{\Lambda_{nd} - \Lambda_{nb}}\right) \\ \left(\frac{|\rho_b|^2}{\Lambda_{nd} - \Lambda_{nb}} - \frac{|\rho_c|^2}{\Lambda_{nd} - \Lambda_{nc}}\right) & \left(\frac{v_b}{\Lambda_{nd} - \Lambda_{nb}} - \frac{v_c}{\Lambda_{nd} - \Lambda_{nc}}\right) \end{bmatrix} \tag{6.49}$$

$$N_{vn} = \frac{1}{2}\det\begin{bmatrix} \left(\frac{u_a}{\Lambda_{nd} - \Lambda_{na}} - \frac{u_b}{\Lambda_{nd} - \Lambda_{nb}}\right) & \left(\frac{|\rho_a|^2}{\Lambda_{nd} - \Lambda_{na}} - \frac{|\rho_b|^2}{\Lambda_{nd} - \Lambda_{nb}}\right) \\ \left(\frac{u_b}{\Lambda_{nd} - \Lambda_{nb}} - \frac{u_c}{\Lambda_{nd} - \Lambda_{nc}}\right) & \left(\frac{|\rho_b|^2}{\Lambda_{nd} - \Lambda_{nb}} - \frac{|\rho_c|^2}{\Lambda_{nd} - \Lambda_{nc}}\right) \end{bmatrix} \tag{6.50}$$

and

$$D_n = \det\begin{bmatrix} \left(\frac{u_a}{\Lambda_{nd} - \Lambda_{na}} - \frac{u_b}{\Lambda_{nd} - \Lambda_{nb}}\right) & \left(\frac{v_a}{\Lambda_{nd} - \Lambda_{na}} - \frac{v_b}{\Lambda_{nd} - \Lambda_{nb}}\right) \\ \left(\frac{u_b}{\Lambda_{nd} - \Lambda_{nb}} - \frac{u_c}{\Lambda_{nd} - \Lambda_{nc}}\right) & \left(\frac{v_b}{\Lambda_{nd} - \Lambda_{nb}} - \frac{v_c}{\Lambda_{nd} - \Lambda_{nc}}\right) \end{bmatrix} \tag{6.51}$$

With $\rho_n = u_n + jv_n$ now determined, we obtain $Q_n$ from (6.41):

$$\boxed{Q_n = \frac{\Lambda_{nd}}{|\rho_n|^2}} \tag{6.52}$$

Once again, it is important to verify that the original equations (6.38)-(6.40) are accurately satisfied once the computation of the calibration constants has been completed. It may be necessary to adjust the positions of the sliding shorts in the six-port network in order to be able to obtain a more accurate calibration.

To summarize, we must first perform calibration measurements and use equations (6.48) and (6.52) to evaluate the calibration constants. Then the DUT is connected and measured, from which we determine the real and imaginary parts of the unknown $\rho$ using (6.35).

## 6.7 Practice questions

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

2. Show that the ideal isolator is lossy.

3. Show that the ideal phase shifter is lossless.

4. 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. 6.1(c).

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

6. 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?

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

8. Sketch the odd and even mode circuits of a Wilkinson power divider, shown in Figure 6.2(d).

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

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

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

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

13. 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?

14. 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?

15. For three mutually intersecting circles as shown in Fig. 6.15, each pair of circles intersects at two points. Show that if the three radical axes are not parallel to each other, they intersect at a single point, which can be used as a best fit to the equations described by the three circles.

16. What constraints on directivity and isolation of the two terminated directional couplers in Fig. 6.11 are needed for proper operation of this reflectometer? Could the two couplers be replaced by a single four-port coupler? What constraints on directivity and isolation of this coupler would be needed?

## 6.8 Homework Problems

**1.** Find the $S$-parameters of an attenuator that has a 6 dB attenuation (in power). Assume that ports 1 and 2 are matched.

**2.** Find the $S$-parameters of a matched isolator that has 0.2 dB insertion loss and a -25 dB backward attenuation at the design frequency.

**3.** Using microwave design software, 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?

**4.** 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 microwave design software. (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.)

**5.** Using microwave design software, design a transmission-line Tee network in microstrip with no loss. Use FR-4 substrate ($\epsilon_r = 4.5$, $h = 0.508$ mm). Assume the network is to be designed to feed two antennas in phase from a common 50 Ω feedpoint. You want as much power to get to the antennas as possible for the frequency range 1.8 GHz $< f <$ 2.2 GHz. The antennas have a 50 Ω input impedance at and around 2 GHz. Can you use this circuit as a basis for networks that feed 4 and 8 antenna elements, and how would you accomplish that?

**6.** A 50-Ω resistor is placed in series in a 50-Ω transmission line as shown in Fig. 6.16, with the continuation of the line terminated in a matched load. If 1 W of power is incident from the generator, determine

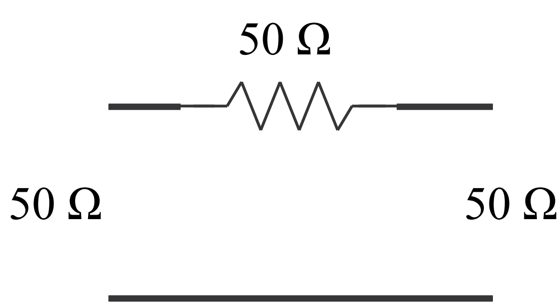



Figure 6.16: Series resistor in a transmission line.

(a) how much power is reflected,

(b) how much power is delivered to the matched load, and

(c) how much power is lost in the series resistor.

What is the insertion loss of the series resistor, viewed as a two-port network?

**7.** Use microwave design software to design a branch-line directional coupler, Fig. 6.17, at 2 GHz. Each of the four ports is assumed to be terminated in 50 Ω. From formulas (1.41)-(1.44) of

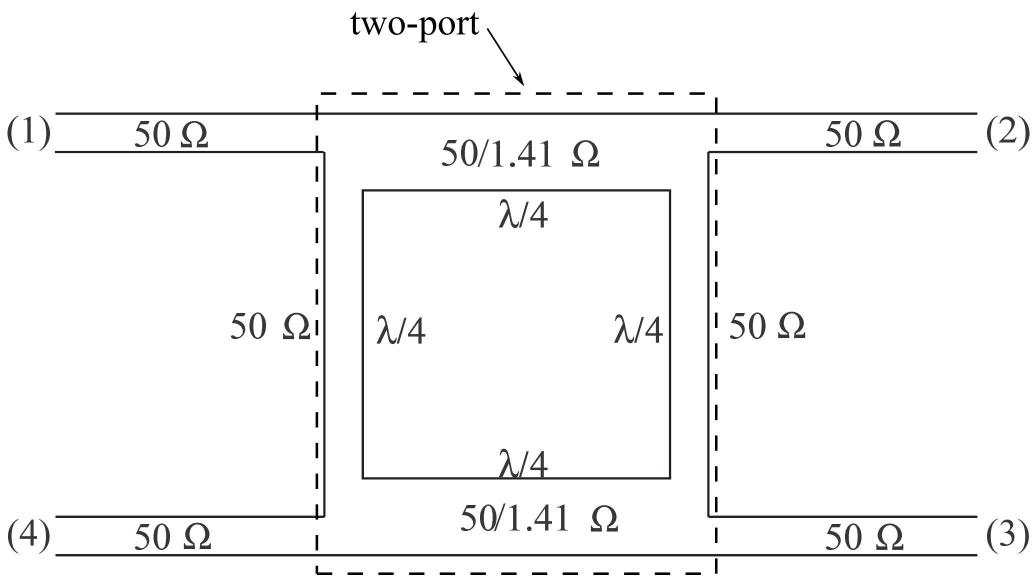



Figure 6.17: A branch-line microstrip directional coupler.

the lecture notes (or from suitable software), 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)?

**8.** Repeat problem 7, 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.

**9.** Use microwave design software to design a stop-band filter using alternating low ($Z_{0l} < Z_0$) and high-impedance ($Z_{0h} > Z_0$) quarter-wave line sections ($l = \lambda/4$) as shown in Fig. 6.18 centered at $f_0 = 2$ GHz. How does the number of lines (and whether that number is even

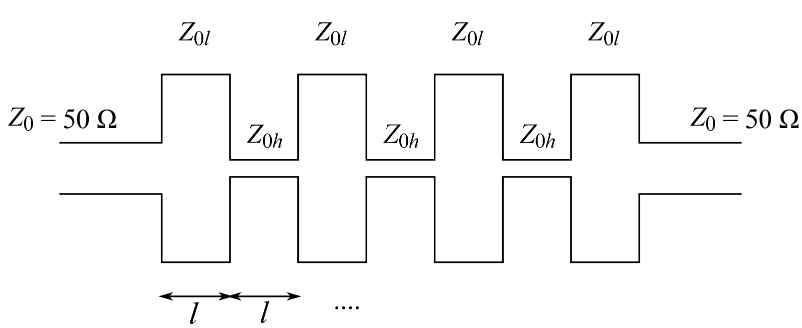



Figure 6.18: Stepped-impedance transmission-line filter.

or odd) affect the frequency response in the pass and stop bands? Adjust the values of $Z_{0l}$ and $Z_{0h}$ to try to improve the stop-band performance. What is the 3 dB bandwidth in each case? 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.

**10.** By choosing the values of $Z_{0l}$ and $Z_{0h}$ in problem 9 in a different way, show that this structure can also function as a low-pass filter when $l \ll \lambda/4$, with an upper frequency of the passband equal to some $f_u$ that is less than the $f_0$ given in problem 9. Try to improve the low-pass behavior for flattest response with your choice of impedances.

**11.** Use microwave design software to design a microstrip high pass filter, with the same substrate parameters as in problem 9. Are there any parasitic elements you need to take into account (e.g. the inductance of a via to ground)?

**12.** Provide details of the derivation of (6.15).

**13.** Simulate a quarter-wave transformer matching circuit in microstrip using microwave design software, both with and without parasitic effects included. Use FR-4 substrate, matching a 50 Ω line to a 100 Ω line at 5 GHz, and plot $S_{11}$ between 4 GHz and 6 GHz.

**14.** A three-port reciprocal lossless network is matched at port 3 (i. e., $S_{33} = 0$), and has identical magnitudes of transmission from port 1 to port 3 and from port 2 to port 3 ($|S_{13}| = |S_{23}|$). Evaluate the magnitudes of all the $S$-parameters $|S_{11}|$, $|S_{12}|$, $|S_{13}|$, etc.

**15.** From the formulas (6.18) of the lecture notes, derive an expression for the ratio $b_3/b_4$. Then use expression (6.16) for $r$ to derive (6.19). What would you pick as known standards to do the calibration? For those standards, determine the expressions for the unknown constants.

**16.** Suppose that it is known that diode voltages measured at ports 3 to 6 of a six-port reflectometer are related to the reflection coefficient $\rho$ of the device under test connected to the load port by eqns. (6.30). Four calibration standards are connected in turn to the load port, with reflection coefficients $\rho_a = -1$ (a short circuit), $\rho_b = j$, $\rho_c = 1$ and $\rho_d = 0$. The corresponding diode voltages are $V_3^a$, $V_4^a$, …, $V_3^b$, …, etc. If the measured data are

$$V_3^a = 5.21 \text{ mV}; \quad V_3^b = 5.07 \text{ mV}; \quad V_3^c = 5.16 \text{ mV}; \quad V_3^c = 8.33 \text{ mV}$$

$$V_4^a = 28.9 \text{ mV}; \quad V_4^b = 57.1 \text{ mV}; \quad V_4^c = 29.3 \text{ mV}; \quad V_4^d = 22.4 \text{ mV}$$

$$V_5^a = 21.1 \text{ mV}; \quad V_5^b = 6.8 \text{ mV}; \quad V_5^c = 0.84 \text{ mV}; \quad V_5^d = 7.3 \text{ mV}$$

$$V_6^a = 3.1 \text{ mV}; \quad V_6^b = 26.2 \text{ mV}; \quad V_6^c = 78.7 \text{ mV}; \quad V_6^d = 28.0 \text{ mV}$$

then determine the calibration constants $\rho_4$, $\rho_5$, $\rho_6$, $Q_4$, $Q_5$ and $Q_6$, using a mathematical software program of your choice. Estimate the accuracy of this calibration by plotting the three circles (6.45)-(6.47), and evaluating the error your solution exhibits in these equations.

**17.** Suppose the calibration constants of a six-port reflectometer have been determined to be

$$Q_4 = 3.5; \quad Q_5 = .34; \quad Q_6 = 10.1$$

$$\rho_4 = .01 - j(.9); \quad \rho_5 = -.98 + j(.1); \quad \rho_6 = .44 + j(.76)$$

If, with an unknown load (DUT) connected to port 2, the measured DC voltages at ports 3-6 are

$$V_3 = 3.85 \text{ mV}; \quad V_4 = 10.1 \text{ mV}; \quad V_5 = 2.55 \text{ mV}; \quad V_6 = 30.4 \text{ mV}$$

determine (using a mathematical software program of your choice) the best approximate value of the complex reflection coefficient $\rho$ of the DUT. Plot the three circles (6.30) and provide an estimate of the accuracy of this result.

**18.** A six-port reflectometer is made from an arbitrary linear five-port and an ideal (i. e., perfectly isolated and matched) directional coupler as shown in Fig. 6.19. Let diodes be used at ports 3 to 6 to measure power, and assume they are matched RF loads. Show that eqns. (6.30)

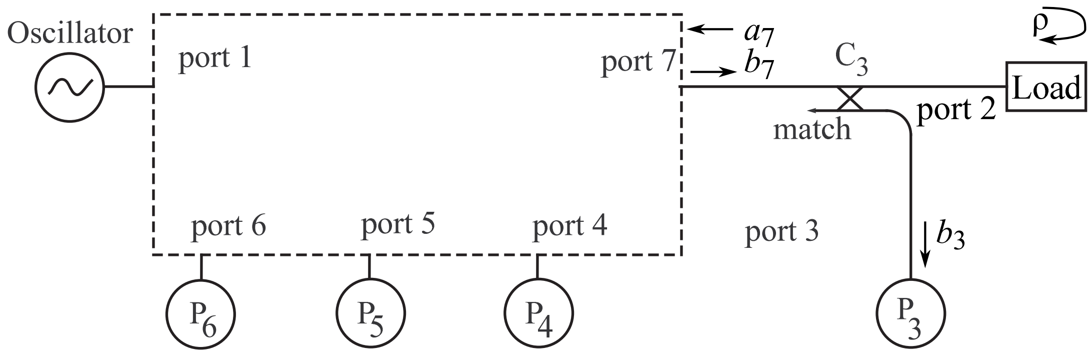



Figure 6.19: Six-port reflectometer consisting of a five-port and an ideal directional coupler.

hold.

**19.** Derive equations (6.45)-(6.47) from (6.42)-(6.44), and show that equations (6.48)-(6.51) give the radical center of the circles (6.45)-(6.47).

# Lab 6-1: 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 evaluate the quality of your measurement.

## Part I.

We begin by performing a two-port 3.5mm SOLT (short-open-load-thru) calibration of the Agilent 8753ES 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. In this case, a “thru” is just a direct connection between the (male) connector at port 1 and the (female) connector at port 2. 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 compare 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 microstrip circuits using the network analyzer. Make plots of the relevant S-parameters in a frequency range you determine to be relevant, somewhere between 1 GHz and 6 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 very simple two-port circuit. At what frequencies do the amplitudes of the S-parameters have dips? Explain why: how long is the open/shorted stub at those frequencies? Measure the stub and, assuming the dielectric is 30 mils = 30 x 24.5 $\mu$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 2.4 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?

The following circuits are various types of filters. For all of the circuits, when you discuss performance and comparisons, please pay attention to the following parameters:

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?

You may need to state how you define the different parameters.

**Q2:** Circuit #2 is a low-pass filter. Explain how the circuit works based on physical principles and draw a simple equivalent lumped element circuit. Can you check if the circuit is lossless? Try calculating $|S11|^2 + |S21|^2$ at a frequency point 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.

Compare uncalibrated and calibrated measurements for this circuit.

**Q3:** Turn calibration back on and measure Circuit #3. What is the functionality of Circuit #3, and how is it different from Circuit #2? Quantify in terms of (1) to (4) above and comment on the pass-band ripple. Draw the lumped-element low-frequency analogue of this circuit as well and compare to that of Circuit #2.

**Q4:** Measure Circuit #4 and explain its functionality. Save relevant S-parameters and evaluate criteria (1) to (4) from above. What is the lumped-element equivalent circuit and how would you obtain the equivalent lumped-element values? How is the circuit behavior different from that of a lumped-element equivalent?

**Q5:** Quantify the performance of Circuit #5 and compare it to Circuit #4.

**Q6:** Measure Circuit #6, explain its functionality and quantify all relevant parameters.

**Q7:** Circuit #7 is a bandpass filter. Explain how it is designed based on your knowledge of the previous circuits. Quantify all relevant parameters such as pass-band insertion loss, stop-band attenuation, bandwidth.

**Q8:** Circuit #8 is a coupled-line bandpass filter. Quantify all relevant parameters and compare it to Circuit #7.

## Part III – Three-Port Networks

**Q9:** What is the function of Circuit #9? 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 (3.5mm) adapter, measure the impedances of the matched loads to be 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?

Attach the matched loads to ports 2 and 3, and make a plot of S11 from 1 GHz to 6 GHz. What frequency is the matching circuit designed for? How big is the 10-dB (about 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).

**Q10:** 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 S23 (the isolation), S22 and S33 (input match at each port) on the display. What should these S-parameters be if the circuit is to function well as a power combiner?

**Q11:** Connect ports 1 and 2 of the network analyzer to ports 1 and 2 of Circuit #9, and terminate port 3 of this circuit with SMA coaxial matched loads. Measure the transmission coefficient S21 from port 1 to port 2. Based on your answers to **Q9** and **Q11**, 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?

**Q12:** Circuit #10 is a Wilkinson power divider. Measure this circuit as a three-port and quantify:

1. Power division (coupling coefficient);
2. Return loss (how well the circuit is matched);
3. Isolation between ports 2 and 3;
4. Insertion loss between input and coupled ports;
5. Bandwidth;
6. Relative phase between the two coupled ports.

# Part IV – Four-Port networks

**Q13:** Use the network analyzer to characterize Circuit #11, the four-port microstrip hybrid directional coupler (as studied in your prelab homework). Measure the S-parameters for this circuit between 1 GHz and 6 GHz. Make use of all the symmetries you can to reduce the number of plots. Include plots of the amplitudes in your notebook, and look at the phases on the network analyzer. What is the phase difference between S21 and S31 at the center frequency? Compare un-calibrated and calibrated measurements.

**Q14:** 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? If you had no matched loads, could you still characterize the circuit, and if yes, what would you do?

**Q15:** At the operating frequency, how large is the through, coupled, reflected and isolated-port power assuming 1mW (0 dBm) incident at port 1?

**Q16:** Sketch the circuit and label the ports as they are labeled on the circuit. What are the relative phases of S11, S21, S31 and S41 at the design frequency? Does this agree with the analysis from your prelab?

**Q17:** How large is the 3-dB bandwidth of the coupler? What is the 10-dB bandwidth equal to?

**Q18:** Characterize Circuit #12 between 1 GHz and 6 GHz. Make use of all the symmetries you can to reduce the number of plots. Compare the performance to Circuit #11:

- What are the relative phases of S11, S21, S31 and S41 at the design frequency?
- How large is the 3-dB bandwidth of the coupler? What is the 10-dB bandwidth equal to?

# Part V – Coaxial Components

In this part of the lab, you will characterize a few coaxial components. Set the network analyzer for a frequency range of 100MHz to 6 GHz. Perform a full two-port calibration. Circuit #13 is a coaxial directional coupler. Circuits #14 and #15 are unspecified components whose functions you will determine using the network analyzer.

**Q19:** What range of frequencies is circuit #13 designed for? What are its coupling factor and isolation?

**Q20:** What does circuit #14 do, and why do you think so?

**Q21:** What does circuit #15 do, and why do you think so? [Hint: Compare S21 with S12.]

## Lab 6-2: The Slotted Line

Connect the slotted line setup as shown in Fig. 6.20. The HP 8654B RF signal generator goes up to

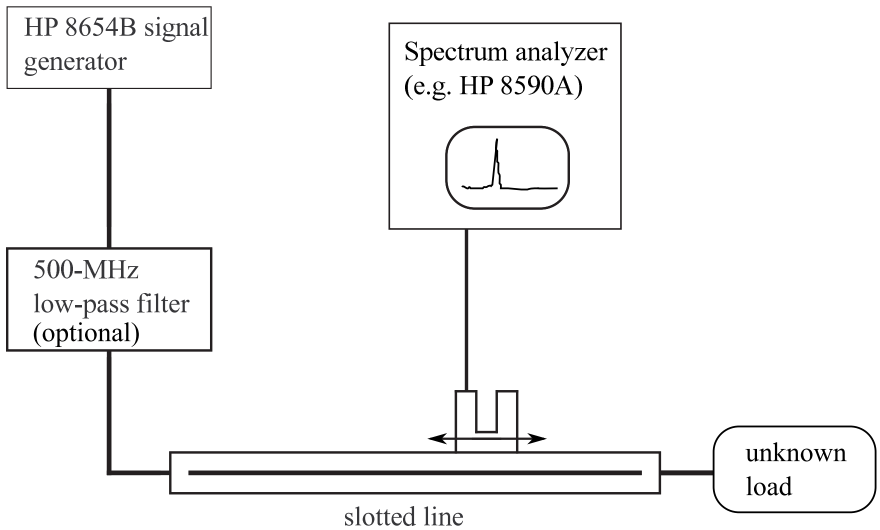



Figure 6.20: Setup for slotted line measurements.

about 550 MHz in frequency. It has the possibility of being amplitude modulated with a 1 kHz signal (although we will not use this feature here). The output power you will use is 10 dBm. The output from the generator is analyzed for harmonic content to determine sinusoidal purity, then again after passing through an optional 500 MHz low pass filter. The purpose of this filter is to get rid of all the higher-order harmonics from the generator in order to maintain as pure a sine wave output as possible. The output of the filter is the input to a coaxial slotted line, which is about 60 cm long. The other end of the slotted line is terminated with the unknown load we wish to measure. All connectors in this system are GR874 connectors.

The slotted line has a probe with a diode detector built into the sliding unit, along with a connector to which an adjustable length shorted stub can be connected to help impedance match the diode for maximum signal output. The diode is designed to detect the 1 kHz modulating signal from the modulated RF signal, and this detected signal is then fed to a special device called an SWR meter (basically just a narrow band, high gain amplifier for the detected signal), which reads the SWR directly on its front panel. The high gain ensures good measurement sensitivity.

However, we will not be using this detector system. Rather, we will measure the small amount of RF power that leaks from the slider probe in the slot in the line through the connector normally used to attach a matching stub using an HP 8590A spectrum analyzer. To set up the generator and spectrum analyzer, first turn power on to the spectrum analyzer and generator with no connections between them. Before pushing the "RF On" button on the generator, check to see that the spectrum analyzer input has a reference attenuation of 10 dB, and that the generator is set to maximum output power (10 dBm). Connect the generator to the input of the spectrum analyzer, and adjust the center frequency and span to observe the generator output near 500 MHz. Narrow the frequency span to about 5 MHz. Adjust the generator frequency to $500.0 \pm 0.1$ MHz (use peak track to help keep track of the signal maximum, which can drift over time). Now, change the center frequency to 1 GHz and 1.5 GHz to try to observe signal generator harmonics.

**Q1:** Why would or wouldn't these harmonics be important for slotted line measurements? Would the use of a spectrum analyzer rather than a diode detector with an SWR meter make a difference in whether harmonics were important?

Next, connect the generator to the one end (the left, say) of the slotted line, and the spectrum analyzer to the left connector on the slider.

**Q2:** Does it make a difference whether we connect the spectrum analyzer to the diode port rather than directly to the probe in the slotted line? Is it be better to leave the diode in the slotted line carriage, or to remove it? Why?

When measuring the slider signal we will remove the default value of 10 dB of attenuation from the input of the analyzer, lower the reference level to -50 dBm, and reduce the resolution bandwidth to 1 kHz or the video bandwidth to 3 kHz to obtain maximum sensitivity. In doing so we will also reduce the span to 10 kHz to be able to sweep often enough for practical measurements and to accurately determine the power of the signal.

# Part I.

In this first part of the experiment you will use a short circuit termination. We use the short to establish the location of reference (or proxy) planes, thus effectively calibrating the length scale on the slotted line. This use of a short to determine the reference plane location is analogous to the calibration procedure of a network analyzer. Re-adjust the frequency to $f = 500.00 \pm 0.01$ MHz. Find and record the locations of the nulls of the standing wave pattern. These nulls correspond to short circuits ($z = 0$) on the Smith chart.

**Q3:** Why are the minima of a standing wave pattern on a short-circuited line sharp, and the maxima broad and flat? What is the measured wavelength on the slotted line? How does this compare to the free-space wavelength at 500 MHz? Why should one choose a null and not a maximum for the reference plane?

The sharp nulls will sometimes be at the level of noise of the spectrum analyzer. A good way to find their position precisely is to move towards the null from both sides, locate points A and B from Fig. 6.21, and take the null to be midway in between them. From the location of these nulls determine the location of the proxy (short-circuit) reference plane from the scale on the sliding short.

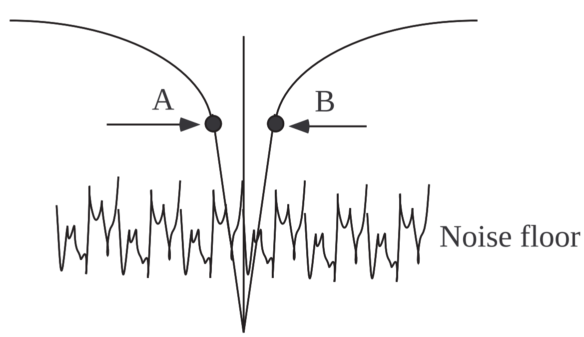



Figure 6.21: Determining the position of the standing wave minimum for a shorted line.

**Q4:** What are the dynamic range and sensitivity of your measurements (in dB)? How accurately can you determine the location of a null (in cm)? How should you set up the spectrum analyzer for maximum dynamic range?

# Part II.

In this part of the experiment, you will measure the complex impedance of several test loads at 500 MHz (a second frequency of 250 MHz is optional). Each load consists of one or more lumped elements mounted in the metal GR 874 component mount from experiment 2.12.

Connect the load, and measure the distance between the proxy plane and the first dip (minimum) next to the proxy plane. You can either move towards or away from the load, but be sure to note this direction for later use with the Smith chart. Note also the ratio of the spectrum analyzer power at the dip and the broad maximum. This is the standing wave ratio (SWR).

**Q5:** Fill in a table like the one shown in Table 6.1. The distance $l$ in the sixth column is the electrical distance between columns 4 and 5 (negative, or, away from the generator, if appropriate). Use this distance on the Smith chart to find the normalized load impedance $z = Z/Z_0$ from the measured SWR. Then find the load impedance $Z$ by multiplying by the characteristic impedance of the coax. Repeat for each unknown load. Include your Smith charts.

| Load | $f$ (MHz) | $\lambda$ (cm) | Proxy plane position (cm) | Nearest minimum (cm TG or AG) | $l/\lambda$ (TG or AG) | $z = Z/Z_0$ | $Z$ ($\Omega$) | SWR |
|---|---|---|---|---|---|---|---|---|
| A | $f_1 =$ | | | | | | | |
| B | $f_2 =$ | | | | | | | |
| C | | | | | | | | |
| D | | | | | | | | |
| E | | | | | | | | |

Table 6.1: Measurements and calculations of unknown loads at various frequencies.

# Part III.

Decrease the generator power by 10 dB and repeat the measurements.

**Q6:** Why are these measurements not as good as the previous ones? What gets worse: the minimum or the maximum of the voltage?

# Part IV.

Return to the original full power level of the generator. At 500 MHz, measure the impedance of a coaxial 50 Ω termination, repeating the procedure from Part II. Include a Smith chart.

**Q7:** Why is the 50 Ω termination harder to measure than a badly matched load?

# Part V.

In the last part of this experiment, you will measure the impedance of a monopole antenna above a metal plane at 500 MHz. The antenna is 15 cm long, which is about a quarter wavelength at this frequency. This is called a quarter-wave monopole. The impedances of this monopole with and without a ground plane are very different.

**Q8:** What basic equations of electromagnetics would you use to find the electric and magnetic fields of such a system? Illustrate your answer with a simple drawing (you do not have to explicitly find the fields).

**Q9:** The impedance of an antenna is defined as the voltage across its terminals divided by the current flowing through the antenna feed terminals. Compare a monopole above a ground plane and a dipole constructed by imaging the same monopole with respect to the ground plane, and then removing the ground plane. Considering the dipole and monopole to be fed by the same *current*, find the feed point voltage for each. What then is the relationship between their impedances?

Repeat the same measurement as in Part II, now covering the frequency range from 420 to 520 MHz in 20 MHz steps. Remember that the wavelength changes when the frequency is changed, and a fresh determination of the proxy plane position must be made for each frequency. Also, the frequency span can be temporarily increased to make finding the oscillator signal on the spectrum analyzer screen easier. Since we are using a sensitive spectrum analyzer the antenna will need to be placed a few meters away from the spectrum analyzer using a cable. Does placing the antenna near the spectrum analyzer affect your measurements? For each frequency, you will get a point on the Smith chart for the normalized antenna impedance. If you could connect these points you would have a graph of the antenna impedance versus frequency. Include your Smith chart plot.

## Lab 6-3: Four-Port and Six-Port Network Analysis

In this lab, you will perform $S$-parameter measurements "by hand", duplicating the same types of calculations made by the computer which is inside an automatic network analyzer. This will give you insight into how the information obtained in the calibration process is used to establish a value for the complex reflection coefficient of an unknown load when only power measurements can be made.

## Part I.

First, we will show how a four-port network can be used to measure the *magnitude* of the reflection coefficient of some unknown WR-90 waveguide components.

**Q1:** Why can't you use an HP 8753ES network analyzer to characterize these components?

Take a pair of 10 or 20 dB directional couplers connected as shown in Figure 6.22. Connect a

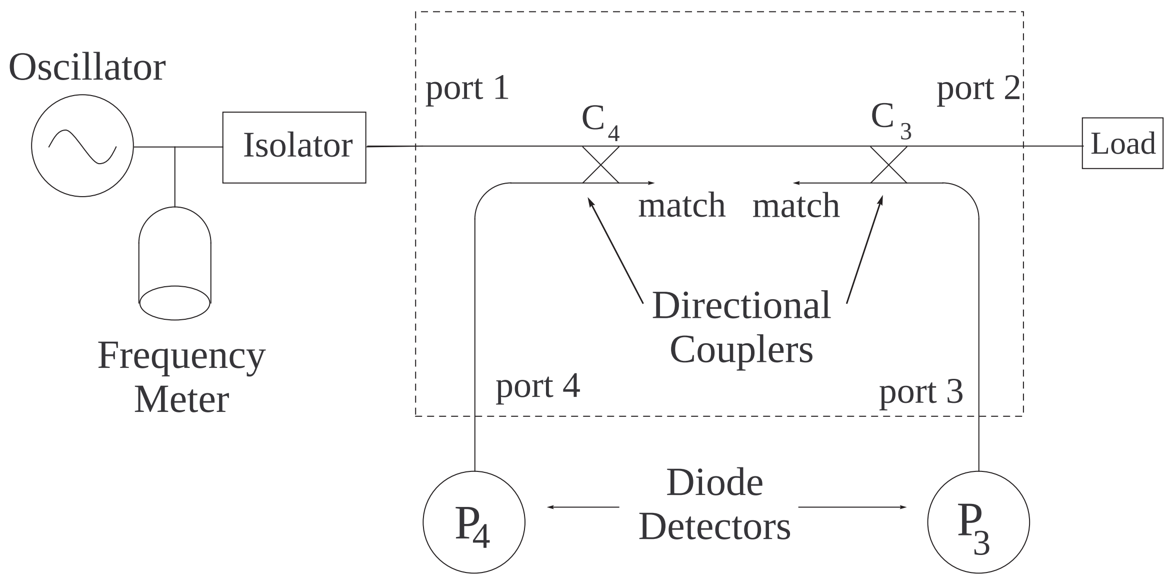



Figure 6.22: Directional coupler reflectometer.

microwave generator (the HP 8350 sweep oscillator set to CW operation) to port 1 of the directional coupler combination (through a coax-to-waveguide adapter and an isolator) as shown. Choose an operating frequency as follows. If your bench is supplied with .282″ and .476″ offset waveguide short circuits for use later in this lab, use a frequency of 10.5 GHz. If you have the .181″ and .376″ offset shorts, use a frequency of 12 GHz. An unknown load impedance is to be connected to port 2 (the far right end of the coupler combination). We will measure the power emerging from the remaining ports (3 and 4) of the coupler combination, using waveguide-mounted diode detectors as in Experiment 3.8. Attach a DC voltmeter to each of the diode detectors. When connecting waveguide-mounted diodes, be sure that the ground (shield) of all cables is consistently connected to the ground sides of the voltmeters in order to avoid ground loops and erroneous readings. Because of the polarity of the diodes in the mounts, each meter will show a negative DC voltage when a microwave signal is incident at the diode—this is normal. Verify the operating frequency using the frequency meter: adjust the meter until you observe a dip in power measured at the diode detectors, and read the frequency from the dial.

The side arm of a 20 dB (for example) directional coupler takes a small amount (−20 dB) of the power going in the forward direction in the main waveguide, and virtually none from the wave in the backward direction (the direction towards the match—a phantom fourth port of the coupler not accessible to you). Thus, $V_3$ should measure the power in the forward wave, and $V_4$ the power in the reflected wave with minimal disturbance to the waves in the main waveguide section. We expect a more accurate measurement of $|\rho|$ by this method.

In section 6.6 of chapter 6, it was shown that the magnitude squared of the unknown reflection coefficient $\rho$ of the load connected to port 2 can be expressed as

$$Q_4|\rho|^2 = \frac{V_4}{V_3} \equiv \Lambda_4$$

where $Q_4$ is some positive real constant.

**Q2:** Connect the matched waveguide load to the end (port 2) of the pair of directional couplers and measure $V_3$ and $V_4$. What do you measure for the DC voltage $V_4$ from the diode at port 4? Is this what you should expect? Record also the voltage $V_3$ from the diode at port 3.

Now connect each of the three waveguide short circuits (one is flush, or $0''$ offset; the other two are marked with the length of their offsets in inches) in turn to port 2, and record the voltages $V_3$ and $V_4$ in each case.

**Q3:** What is the value of $|\rho|$ that you expect for these short circuits? Use the measured voltages for the case of the short circuit with $0''$ offset to determine the value of the constant $Q_4$ from the equation above. This is the *calibration step* for the measurement of $|\rho|$.

**Q4:** Do the same with the various offset waveguide short circuits attached. Use the zero-offset short to calibrate the system, and then compute a measured value of $|\rho|$ for each of the other waveguide shorts from your measurements of $V_3$ and $V_4$. Are your answers consistent?

**Q5:** Use this setup to measure $|\rho|$ for the two unknown loads (a horn antenna and one other load made from a post tuner and a matched termination).

## Part II.

This same (or a similar) technique can also be employed to measure the magnitudes of $S$-parameters of multiport networks, which we will do in this part of the lab. In fact, the values of the diagonal elements of the $S$-matrix—$|S_{11}|$, $|S_{22}|$ and so on—can be measured using the setup of Fig. 6.22, if all ports but the one being measured are terminated with matched loads. There are waveguide to coax adapters, waveguide directional couplers, waveguide matched loads, waveguide-mounted diode detectors, a sweep oscillator and a waveguide attenuator at your bench, some of which were used in the first part of this experiment. Carry out the measurements of the diagonal $S$-parameters for each of the WR-90 multiport components on your bench that are to be characterized.

Next, connect the components supplied in a different configuration, suitable for measuring the magnitudes $|S_{ij}|$ of the off-diagonal elements ($i \neq j$) for the multiport waveguide components. You must figure out this configuration on your own. *Do not turn the sweeper RF power on before the TA or instructor checks out your circuit.* First, attach calibration standards to the ports of the directional couplers as appropriate, and measure the diode DC voltages to calibrate your measurement setup. Then connect the directional couplers to each of the multiports as appropriate.

**Q6:** How is the directional coupler that samples an incident wave $|a_j|$ connected? How is the one that samples $|b_i|$ connected? What do you have to connect to the remaining ports of the multiport?

Measure the diode DC voltages for each of the multiport waveguide components. You will need to connect the directional couplers in different ways in order to get a full set of values for $|S_{ij}|$. You may omit measurements of any $S$-parameters which, by symmetry or reciprocity are obviously identical to ones you have already measured (justify this omission in each case). The components are numbered, so make sure you record the number of the component you are measuring, in order that the grader will know which ones you used. For each component answer the following questions:

**Q7:** What is this component? (Name it, and describe in one sentence what it does and how you figured it out.)

**Q8:** Given the available equipment, what parameters of the component can you quantitatively determine? List them and indicate the measured values.

# Part III.

Unfortunately, the foregoing method is not capable of giving us the values of the phases of the $S$-parameters. We need more measured data to obtain this additional information. A popular method is to use a six-port instead of a four-port between the source and load to provide this data. In this part of the lab, we will show how this is done for a one-port load. Measurement of the complex $S$-parameters of multiport networks by six-port methods requires even more complicated data processing, and we will not attempt that here.

Now between the source and the first directional coupler as shown, connect two more directional couplers with diodes and adjustable short circuits installed on the coupled ports as shown in Fig. 6.23. If desired, the directional couplers leading to ports 3 and 4 may be combined into a single one as shown.

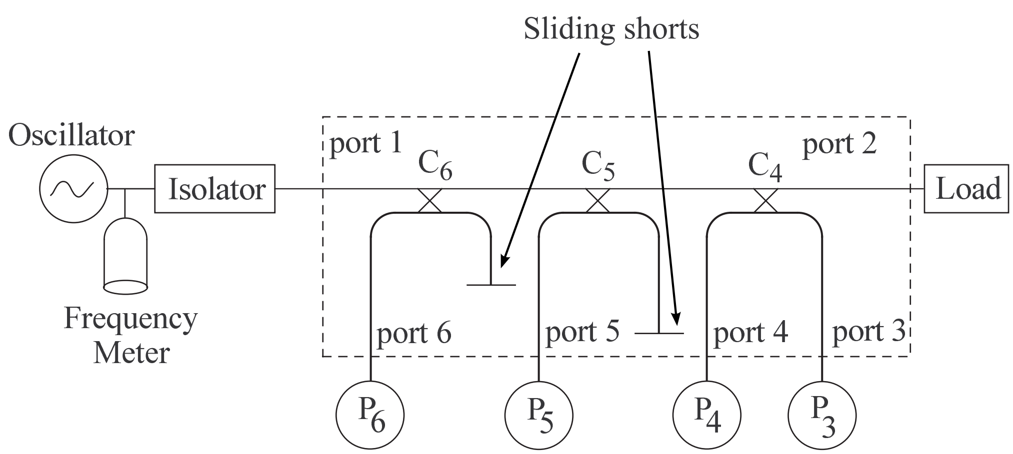



Figure 6.23: Six-port reflectometer configuration.

We will first carry out a kind of "pre-calibration"; the sliding short circuits will be adjusted to have optimal locations for the centers $\rho_n$ ($n = 4, 5, 6$) of our three $\rho$-circles (this may already have been done for you by the instructor or TA; in that case do not adjust the sliding shorts at all). We would like them to be approximately 120° apart from each other on the unit circle in the complex plane—only rough settings of the sliding short circuits will be needed in this step. Only $\rho_5$ and $\rho_6$ can be changed very much by adjusting the sliding shorts; in fact, adjusting the sliding short connected to coupler $C_5$ affects mostly the DC diode voltage $V_5$, while adjusting the sliding short connected to coupler $C_6$ affects mostly $V_6$. Proceed first by connecting the various offset short circuits ($\rho_a$, $\rho_b$ and $\rho_c$) to the load port (port 2) and note the measured DC diode voltages $V_{3a,b,c}$, $V_{4a,b,c}$, $V_{5a,b,c}$ and $V_{6a,b,c}$. It should be the case that adjusting either of the sliding shorts will not significantly affect the values of $V_{3a,b,c}$ and $V_{4a,b,c}$ (verify this). By (6.30), the one of $\rho_a$, $\rho_b$ and $\rho_c$ that lies closest to $\rho_4$ in the complex plane should be the one that gives the smallest value of $V_4/V_3$. Let us call this offset short $\rho_a$. Ideally, the reflection coefficients of the other two offset shorts should differ by approximately $+120°$ and $-120°$ from $\rho_a$, but the parts available may not allow for this.

**Q9:** At your chosen operating frequency (10.5 or 12 GHz), what is the guide wavelength $\lambda_g$ of the fundamental mode of this waveguide? What change in length of an offset short would cause a $120°$ change in its reflection coefficient?

Connect either of the other offset shorts (call this one $\rho_b$) to port 2, and adjust sliding short #5 so as to minimize the value of $V_{5b}/V_{3b}$. This should result in a value of $\rho_5$ that is closer to $\rho_b$ than either of the other two offset shorts. Next, connect the remaining offset short (call this one $\rho_c$) to port 2, and adjust sliding short #6 so as to minimize the value of $V_{6c}/V_{3c}$. This should result in a value of $\rho_6$ that is closer to $\rho_c$ than either of the other two offset shorts. You should now have a system in which the values of $\rho_4$, $\rho_5$ and $\rho_6$ are reasonably well separated, providing a stable and accurate means to measure the reflection coefficient of an unknown load. Finally, considering equations (6.45)-(6.47) for the calibration constants, we need to verify that the three centers of these circles

$$\frac{\Lambda_{nd}\rho_{(a,b,c)}}{\Lambda_{nd} - \Lambda_{n(a,b,c)}} = \frac{V_{nd}\rho_{(a,b,c)}}{V_{nd} - V_{n(a,b,c)}}$$

for $n = 4, 5, 6$ are well separated in the complex plane. Since we have already ensured that $\rho_4$, $\rho_5$ and $\rho_6$ are well-separated, we only need to make sure that all of the ratios

$$\frac{V_{nd}}{V_{nd} - V_{n(a,b,c)}}$$

have the same sign ($> 0$ or $< 0$) for each $n = 4, 5, 6$. If this condition is not met for one or more of the offset shorts, try adjusting one of the sliding shorts somewhat and check again until all conditions are met. The pre-calibration is now complete: do not change the settings of the adjustable short circuits after this.

Now for our actual calibration step we will measure four sets of the three DC diode voltage ratios

$$\Lambda_4 = \frac{V_4}{V_3}; \qquad \Lambda_5 = \frac{V_5}{V_3}; \qquad \Lambda_6 = \frac{V_6}{V_3}$$

for each of four different *calibration standards* (loads with known reflection coefficients) connected to the load port (port 2) of the six-port. The standards we will use are a zero-offset short and two different offset shorts (lengths 0.282″ and 0.476″, for example—these three standards are $\rho_a$, $\rho_b$ and $\rho_c$) and a matched load ($\rho_d$).

**Q10:** What is the theoretical value of $\rho$ for each of your calibration standards at the chosen operating frequency?

**Q11:** Measure $V_3$, $V_4$, $V_5$ and $V_6$ for each of the calibration standards mentioned above. For your lab write-up, use suitable mathematical software to obtain the calibration constants that appear in eqns. (6.30). Obtain the complex constants $\rho_4$, $\rho_5$ and $\rho_6$ from (6.48), and the real constants $Q_4$, $Q_5$ and $Q_6$ from (6.52) using your measured data. Check your computations by substituting the known reflection coefficients and calculated constants into the equations and checking how good the agreement is with equations (6.38)-(6.40).

Finally, we will measure the reflection coefficients of some unknown loads using our six-port reflectometer.

**Q12:** Measure the two unknown loads by this method. You will again need to post-process the data using mathematical software. What do you calculate for the complex reflection coefficients of these loads? How do their magnitudes compare with those obtained in Part I?

**Q13:** Finally, take your two unknown loads to a network analyzer capable of X-band measurements (an Agilent 8719, for example), which has already been calibrated for you. Observe all the usual precautions for the operation of any network analyzer. How do the measured values of $S_{11}$ compare with what you obtained in **Q11**?

# 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.

Some of the vocabulary used in antenna engineering is as follows:

*Azimuth ($\phi$) and Elevation ($\theta$)* – Antennas radiate or receive at different power levels depending on the direction relative to the antenna. A spherical coordinate system is usually used to describe this direction in terms of the azimuth angle $\phi$ and elevation angle $\theta$ (see Fig. 7.1). In the standard mathematical definition, $0° < \theta < 180°$ (you can think of $\theta = 0°$ as the "north pole" and $\theta = 180°$ as the "south pole") and $0° < \phi < 360°$. However, in antenna engineering we often use $(\theta, \phi) = (0°, 0°)$ as the direction of the main beam (maximum radiation), in which case we usually employ the ranges $-90° < \theta < 90°$ and $-180° < \phi < 180°$. In relation to the earth, we often think of the elevation angle as describing the vertical direction, and the azimuth the horizontal direction.

*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 ($\eta$)* – 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.

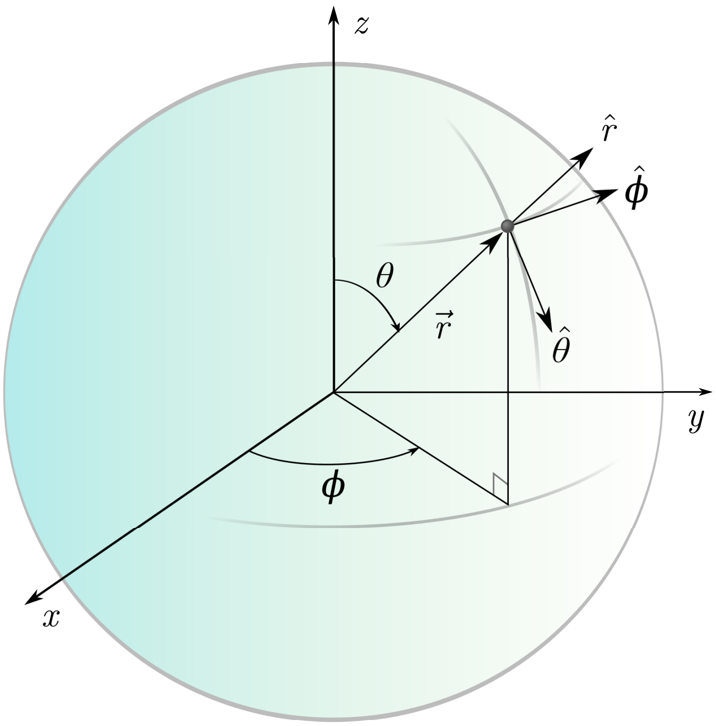



Figure 7.1: Spherical coordinate system used for antenna patterns.

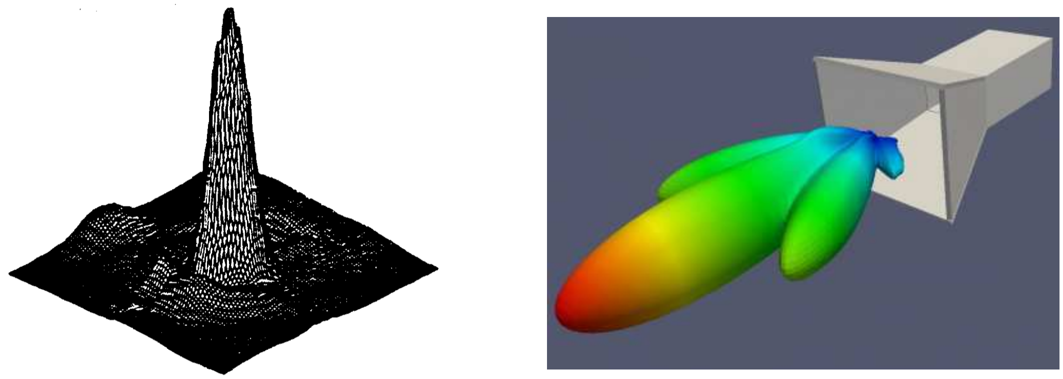
For example, AM broadcasting antennas have almost linear vertical polarization, while FM antennas have linear horizontal polarization, and this determines how the receiving antennas are oriented. Some antennas are designed to radiate and receive two polarizations (used in atmospheric sensing). Circularly 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: 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 accurately. However, for most purposes 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. The far field begins at a distance of at least one wavelength ($\lambda$) away from the antenna. If the largest linear dimension $d$ of the antenna is larger than $\lambda$, we must also impose the commonly used rule of thumb that requires the observation point to be a distance of at least $2d^2/\lambda$ from the antenna.

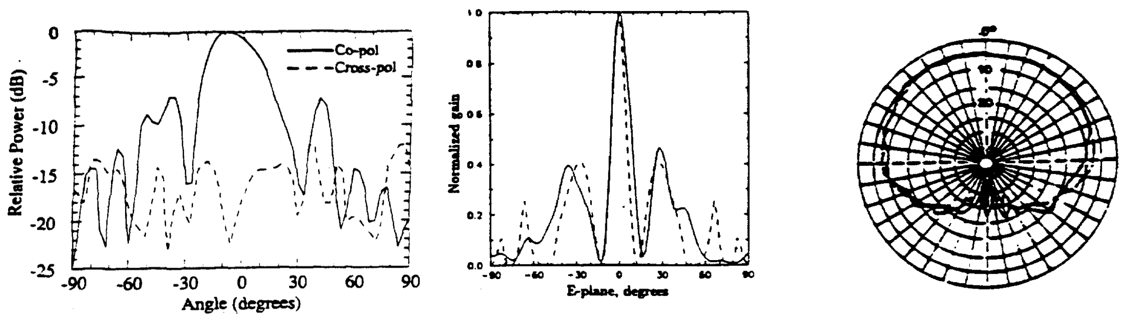
*Radiation pattern* – An antenna will radiate better in some directions, and worse in others. A graphical description of the radiation is the radiation pattern, which is usually the plot of radiated power (on a linear or dB scale) versus spatial angle in the far-field. Different ways of plotting a full three-dimensional radiation pattern are shown in Fig. 7.2. We can also plot certain planar cuts or cross-sections of the full pattern. This can be done in cartesian form, as in Fig. 7.3(a) or (b), but very often such planar cuts are plotted in a polar plot, like the one shown in Fig. 7.3(c).

The most commonly used cuts for linearly polarized antennas are the *E-plane* and *H-plane* patterns.



(a) (b)

Figure 7.2: Plots of three-dimensional patterns: (a) cartesian form; (b) solid angle form.




(a) (b) (c)

Figure 7.3: Ways of plotting radiation patterns: (a) a normalized power pattern plotted in dB or (b) on a linear scale, (c) a polar pattern plot.

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. For antennas that are not linearly polarized, the E-plane and H-plane have no meaning, and in general one presents azimuth plots (constant $\theta$ as $\phi$ is varied) or elevation plots (constant $\phi$ as $\theta$ is varied). The constant values of $\theta$ or $\phi$ in such plots is usually chosen so as to contain the direction of maximum radiation.

Various features of a typical radiation pattern are shown in Fig. 7.4. The pattern can have *nulls* in various directions, at which the radiated power is zero. Between these nulls, radiation increases to a local maximum—these portions of the pattern are called *lobes*. The *main lobe* is the lobe that has the largest power level of all the lobes. Next to the main lobe we often find *side lobes* of considerably lower power than the main one. Many antennas have a *back lobe*, lower in power than the main lobe, but considerably stronger than the side lobes.

The radiation pattern is usually plotted in terms of the power contained in one polarization of the electric 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

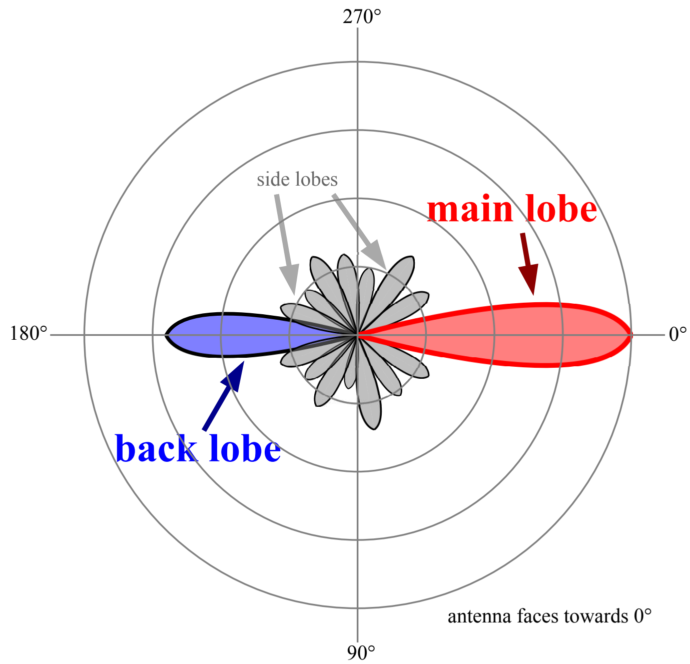



Figure 7.4: Typical features of a radiation pattern.

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. 7.5. The antennas shown in Fig. 7.5(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. The antenna shown in Fig. 7.5(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. 7.5(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. 7.5(g) and (h) are of the reflector type, and are designed for very high gain, and are usually very large (measured in wavelengths).

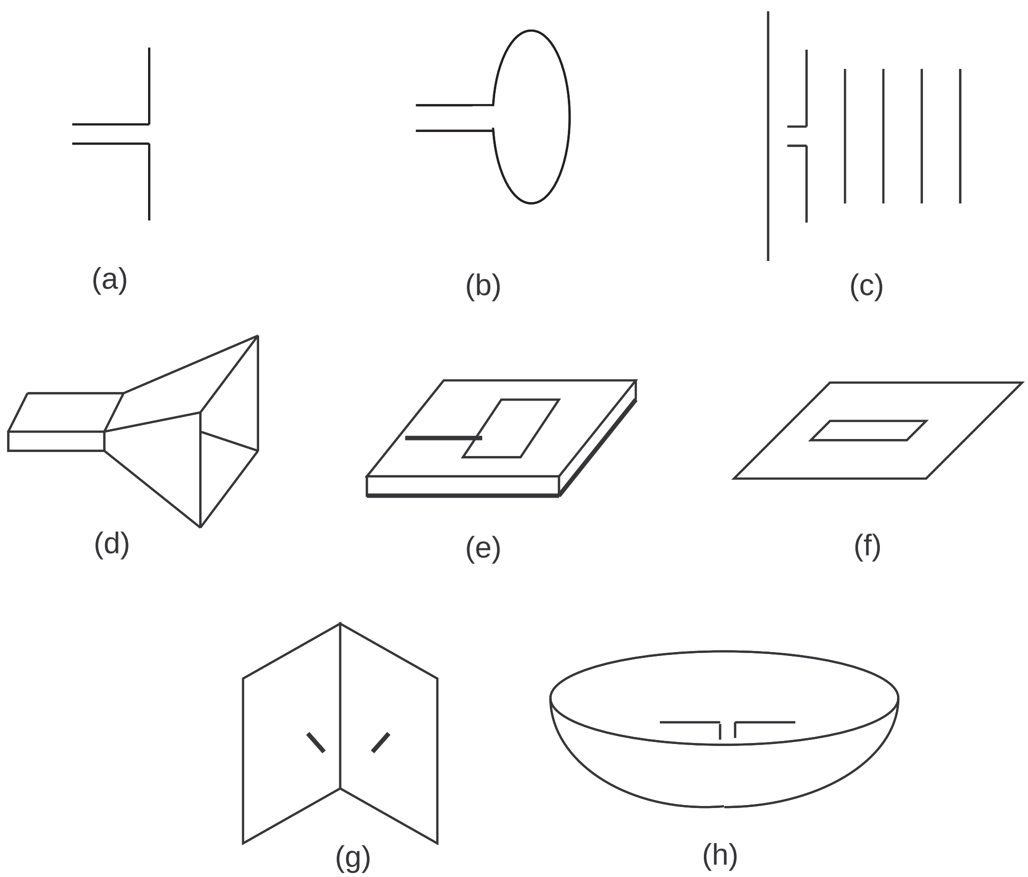



Figure 7.5: Most commonly used antenna types. A wire dipole antenna (a), a wire loop antenna (b), a Yagi-Uda array (c), a horn antenna (d), a microstrip patch antenna (e), a printed slot (f), a corner reflector (g) and a parabolic reflector antenna (h).

## 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. 7.6(a). From the generator's point of view, the

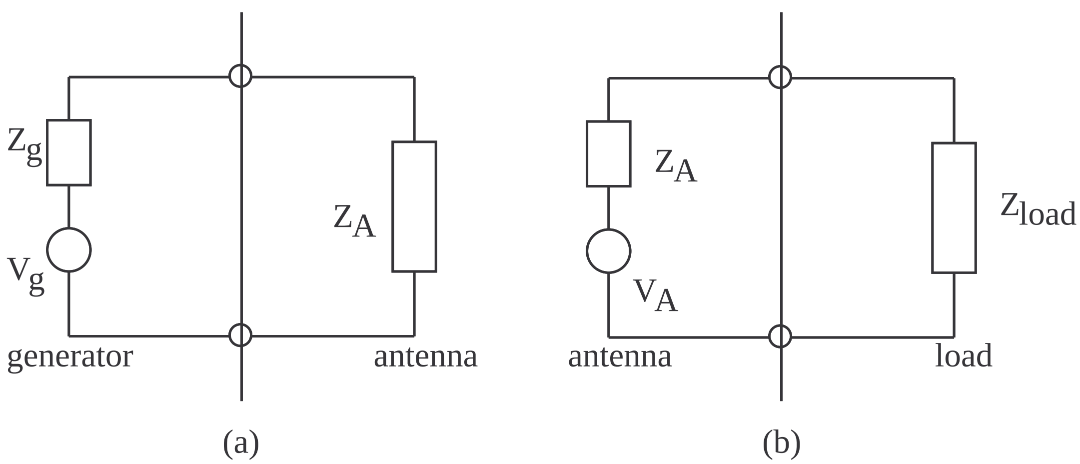



Figure 7.6: The equivalent circuit for (a) a transmitting and (b) a receiving antenna.

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 $R_{\text{loss}}$ 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 $\eta$ 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 $(\theta,\phi)$, 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:

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

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 $\theta$ 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

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

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 $\Omega$ through its differential element

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

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 $\theta$ and $\phi$:

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

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 $P_{\text{loss}}$ 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 (valid for sufficiently small beamwidths) formula for obtaining the absolute 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 $\theta_E$ and $\theta_H$ are the 3-dB beamwidths in the E and H-planes in degrees, respectively. The beamwidth in a given plane depends on the size of the antenna aperture and the uniformity of the field distribution in the aperture. As a rough estimate, you can assume that the beamwidth in a given plane is inversely proportional to the dimension of the aperture in that plane, provided that dimension is large compared to a wavelength and the illumination of the aperture is approximately uniform.

## 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 $I_A$. We can use superposition to study the fields of more complicated antennas by adding the fields of collections of electric current elements. If the current flow is directed 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

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

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

As in usual practice, we have defined the angle $\theta$ so that it ranges from 0 to $\pi$, and the equatorial plane is at $\theta = \pi/2$. In antenna engineering, it is often standard practice to define $\theta$ 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

$$S(\theta, \phi) = E_\theta H_\phi^* = \frac{\zeta_0}{4} \frac{|I_A \Delta l|^2}{\lambda^2} \frac{\sin^2\theta}{r^2} \tag{7.20}$$

By (7.7), the total radiated power from this element is

$$P_{tr} = \frac{2\pi\zeta_0}{3} \frac{|I_A \Delta l|^2}{\lambda^2}$$

and thus from (7.2), 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 (7.3) and (7.12) 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 $(\theta_{tr}, \phi_{tr})$. 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 $\vec{l}_{\rm eff}$ in the direction of this dipole moment whose length $l_{\rm eff} = \Delta l$ is what the electric current element would need to have to produce the same (far-field) electric field in the direction $(\theta_{tr}, \phi_{tr})$ as the original antenna. That is,

$$\vec{l}_{\rm eff}(\theta_{tr}, \phi_{tr}) = \frac{2\lambda r\vec{E}(\theta_{tr}, \phi_{tr})e^{jk_0r}}{jI_A\zeta_0} \tag{7.23}$$

Note that $\vec{l}_{\rm eff}$ may be a complex number in the general case. It can be easily related to the radiated power density by using (7.20):

$$S(\theta_{tr}, \phi_{tr}) = \frac{\zeta_0|I_A|^2|\vec{l}_{\rm eff}|^2}{4\lambda^2r^2} \tag{7.24}$$

## Example – An Electrically Short Dipole

An electrically short dipole $l$ long is shown in Fig. 7.7. 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 $(\theta_{tr}, \phi_{tr})$ 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

$$l_{\rm eff} = \frac{\sin\theta_{tr}}{I_A}\int_{-l/2}^{l/2} I(z)\,dz = \frac{l\sin\theta_{tr}}{2} \tag{7.25}$$

and that the direction of $\vec{l}_{\rm eff}$ is parallel to the unit vector $\hat{\theta}_{tr}$, calculated at the angle $\theta_{tr}$. We may thus write

$$\vec{l}_{\rm eff} = \frac{l\sin\theta_{tr}}{2}\hat{\theta}_{tr} \tag{7.26}$$

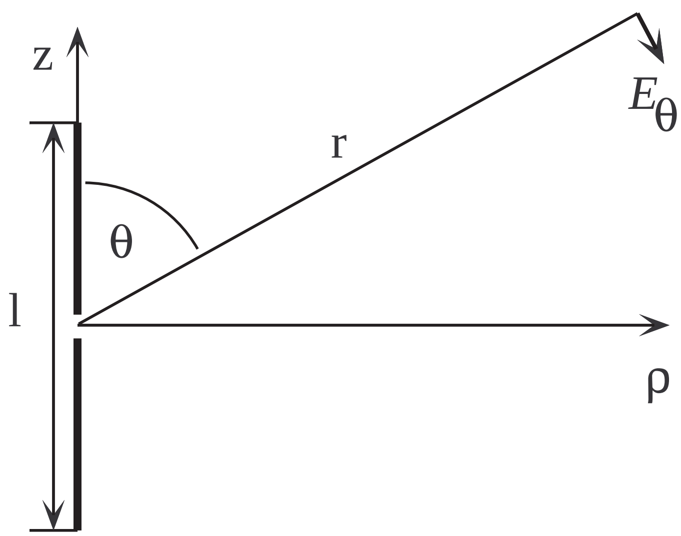



Figure 7.7: An electrically short dipole.

### 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. 7.6(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 $(\theta, \phi)$:

$$A(\theta, \phi) = \frac{P_r}{S(\theta, \phi)}. \tag{7.28}$$

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 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:

$$P_r = \frac{V^2}{4R_A} = \frac{(l_{\text{eff}} E)^2}{4R_A}. \tag{7.29}$$

From the definition of the effective area, we now find

$$\frac{A(\theta,\phi)}{\zeta_0} = \frac{|l_{\text{eff}}(\theta,\phi)|^2}{4R_A}, \tag{7.30}$$

since the power density magnitude is just $S_{inc} = E_{inc}H_{inc}$ (from Poynting's theorem) and $E_{inc}/H_{inc} = \zeta_0$.

The effective area of an antenna may be very different from its physical or geometrical area $A_{\text{geom}}$. For example, from (7.21), (7.26) and (7.29), the maximum effective area of an electrically short dipole is

$$A_{\max} \simeq \frac{3\lambda^2}{8\pi} \tag{7.31}$$

assuming that $R_A = R_r$ and taking $\Delta l = l/2$. This is much larger than the actual area of the antenna. On the other hand for an electrically large aperture antenna (aperture dimensions large compared to wavelength) whose aperture field is uniform (we say the aperture is "perfectly illuminated" in this case), we have $A_{\max} = A_{\text{geom}}$. So if the physical size of the projected area of the antenna in the direction of propagation is large compared to a wavelength in both directions, then the effective area will be approximately equal to this projected area. This fact can be used to provide quick estimates of the properties of such antennas.

## 7.3 The Friis Transmission Formula

A communication system, Fig. 7.8, 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

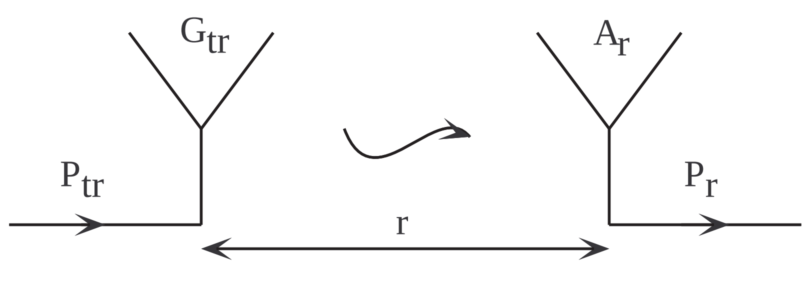



Figure 7.8: A communication system.

gain $G$, and the receiving antenna has an effective area $A$. The power density $S$ incident at the receiving antenna is given by

$$S(\theta_{tr},\phi_{tr}) = \frac{\text{ERP}}{4\pi r^2}, \tag{7.32}$$

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 $(\theta_{tr},\phi_{tr})$ along which the wave travels from the transmitting antenna to reach the receiving antenna. Physically, (7.32) 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*:

$$P_r = \frac{P_{tr}G_{tr}A_r}{4\pi r^2}. \tag{7.33}$$

where $A_r$ is implicitly a function of the direction $(\theta_r, \phi_r)$ from which the wave incident from the transmitting antenna arrives at the receiving antenna.

The Friis formula (7.33) is valid if the polarizations of the transmitting and receiving antennas are the same. If they are not, we must additionally multiply (7.33) by the *polarization loss factor* $PLF$, which for linearly polarized antennas is equal to $\cos^2\theta_p$, where $\theta_p$ is the angle between the polarization directions of the transmitting and receiving antennas.

## 7.4 Reciprocity and Antennas

The property of reciprocity discussed in chapter 2 can also be applied to antennas, since they are, generally speaking, made from linear and isotropic materials. Consider again two possible excitations, 1 and 2, of the network shown in Fig. 2.4. It is easy to show from the relationship between the wave amplitudes and the port voltages and currents that reciprocity also implies the formula

$$V_j^{(1)} I_j^{(2)} = V_i^{(2)} I_i^{(1)} \tag{7.34}$$

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 $dz$ 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 (7.34),

$$V dI_{sc}(z) = E_{z,inc}(z) dz I(z) \tag{7.35}$$

so that the *total* short-circuit current at the feed-point due to the incident wave is

$$I_{sc} = \int dI_{sc}(z) = \frac{1}{V}\int E_{z,inc}(z) I(z)\, dz \tag{7.36}$$

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,

$$V_A = \frac{Z_A}{V} E_{z,inc} \int I(z)\, dz = E_{z,inc} l_{\text{eff}} \tag{7.37}$$

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. 7.9. 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 7.9 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—excitation 1—let us use the antenna on the

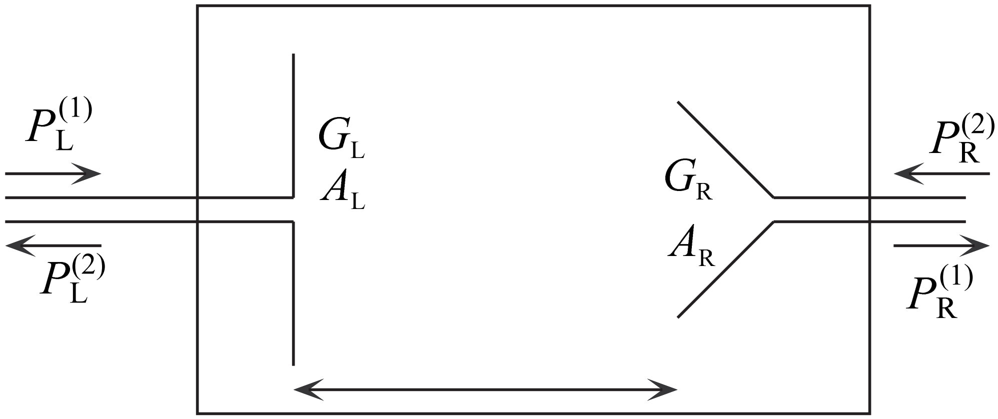



Figure 7.9: Reciprocity applied to a system consisting of two antennas.

left (subscript $L$) as the transmitter, and the right one (subscript $R$) as the receiver, and then the other way round—excitation 2. The reciprocity condition (2.14) in terms of transmitted and receiving powers now becomes

$$P_L^{(1)} P_L^{(2)} = P_R^{(2)} P_R^{(1)}. \tag{7.38}$$

The Friis transmission formula (7.33) allows us to rewrite the received powers $P_{R1}$ and $P_{L2}$ in terms of the gains and effective areas:

$$P_{L1}\left(P_{R1}\frac{G_R A_L}{4\pi r^2}\right) = P_{R1}\left(P_{L1}\frac{G_L A_R}{4\pi r^2}\right), \tag{7.39}$$

which after eliminating the common terms gives us the following:

$$\frac{G_L}{A_L} = \frac{G_R}{A_R}. \tag{7.40}$$

This means that the gain and effective area of an antenna are related by a universal constant (units are $\text{m}^{-2}$) 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

$$\frac{G}{A} = \frac{4\pi}{\lambda^2} \tag{7.41}$$

Eqn. (7.41) allows us to rewrite the Friis formula (7.33) in terms of antenna gains only:

$$P_r = P_{tr} G_{tr} G_r \left(\frac{\lambda}{4\pi r}\right)^2 \tag{7.42}$$

or in terms of effective areas only:

$$P_r = P_{tr}\frac{A_{tr} A_r}{\lambda^2 r^2} \tag{7.43}$$

This formula assumes that the antenna has no return loss ($\rho = 0$).

## 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

$$G(\theta, \phi) = \frac{4\pi r^2 S(r, \theta, \phi)}{P_{tr}}. \tag{7.44}$$

where $S(\theta, \phi)$ is the power density in W/m$^2$ transmitted in the direction $(\theta, \phi)$.

When (7.44) is combined with (7.11) and integrated over all angles, the resulting integral looks like

$$\oint G \mathrm{d}\Omega = \frac{4\pi}{P_{tr}} \oint \frac{\mathrm{d}P_{tr}}{\mathrm{d}\Omega} \mathrm{d}\Omega, \tag{7.45}$$

which in turn gives us the value of the gain integral:

$$\oint G \, \mathrm{d}\Omega = 4\pi. \tag{7.46}$$

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

$$\oint A \, \mathrm{d}\Omega = \lambda^2. \tag{7.47}$$

One proof of (7.47) depends only on properties which follow from Maxwell's field equations and is quite simple, making use of (7.41) and (7.46). 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.

## 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. 7.10. 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 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.

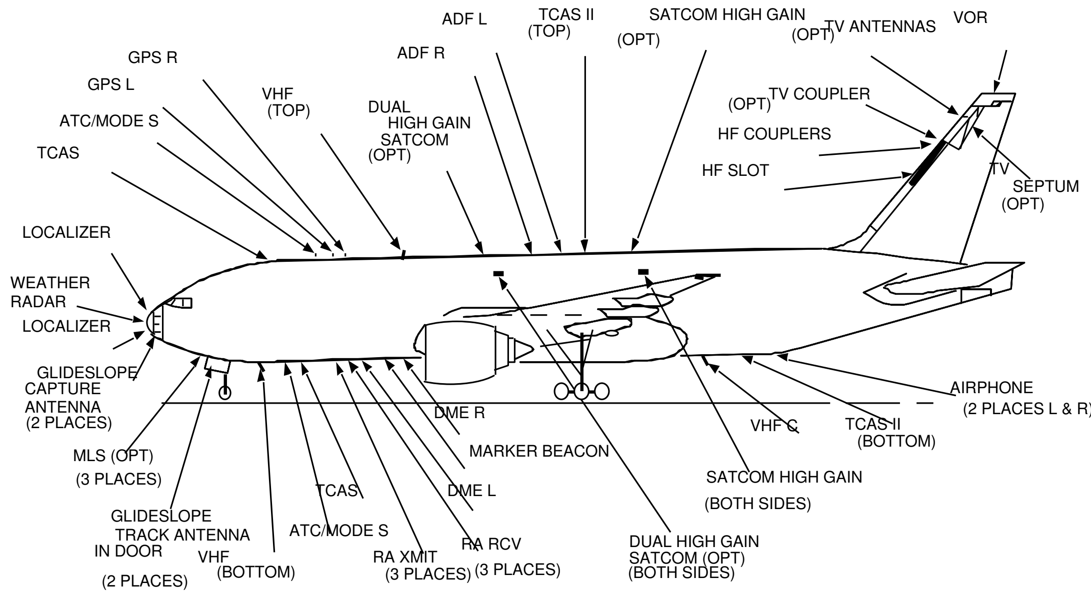



Figure 7.10: The 54 different antennas on a jet airliner used for navigation, communication and weather prediction.

- 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 lightning detector antenna measures ratios of power levels at different frequencies to detect nearby lightning.

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. 7.11 shows a sketch of approximate location and shape of the 14 antennas on a Mercedes Class S car.

Antennas vary drastically in size, both physical and electrical. For example, a number of manufacturers currently offer “chip” antennas, Fig. 7.12. 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.

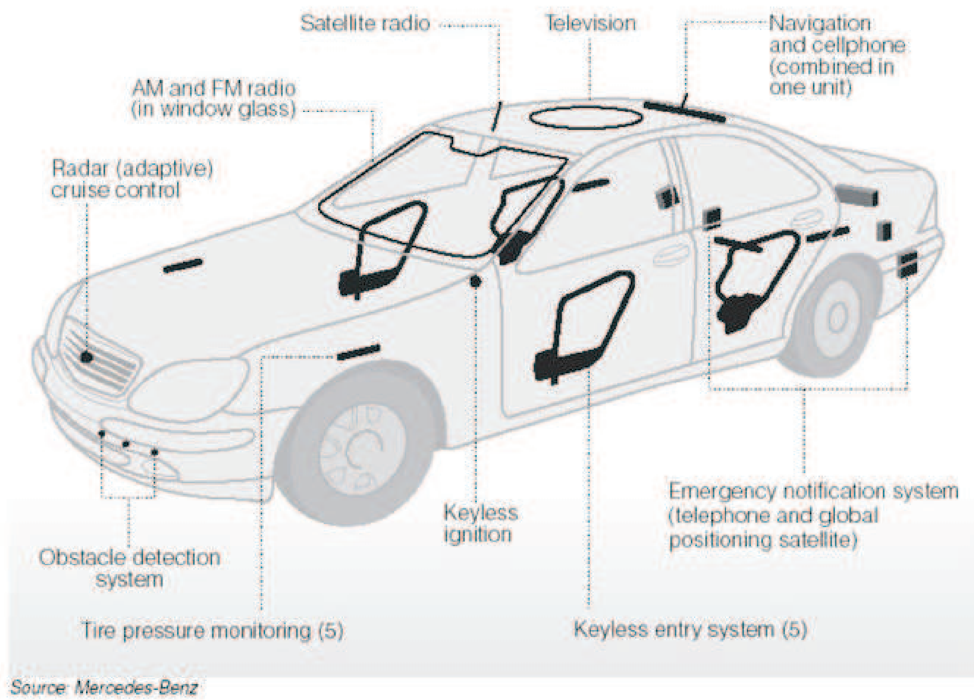



Figure 7.11: The 14 different antennas on a Mercedes Class S automobile (2005).

## 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)?
7. In a horn antenna, the mouth of the horn can be viewed approximately 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. How would you make a horn with a high gain?
8. Define the E- and H-planes for a rectangular horn antenna such as the one shown in Fig. 7.5(d). Show the geometry of these planes relative to the antenna.

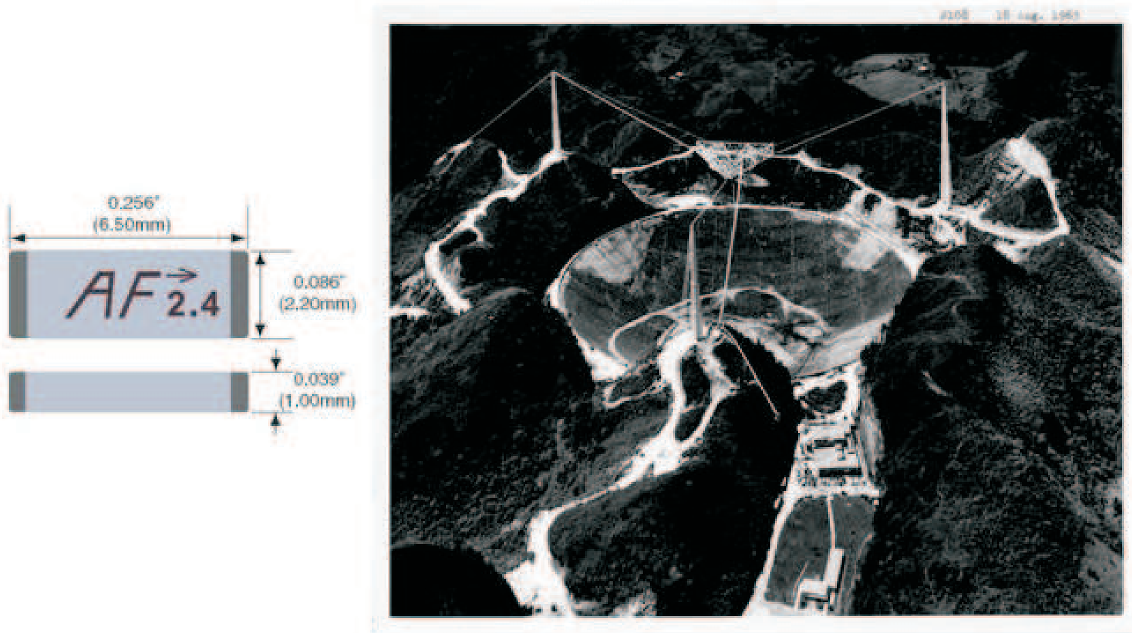



Figure 7.12: Currently marketed (2005) surface-mount "chip" antennas with dimensions indicated (left) and the largest single antenna reflector in the world during its construction in 1963 in Arecibo, Puerto Rico.

## 7.8 Homework Problems

**1.** Find formulas for the radiation resistance and effective area of the short dipole shown in Fig. 7.7.

**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 and are separated by a distance of $r = 100$ km. If the required received power is $P_r = 1$ 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. 7.13. 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 normal to 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

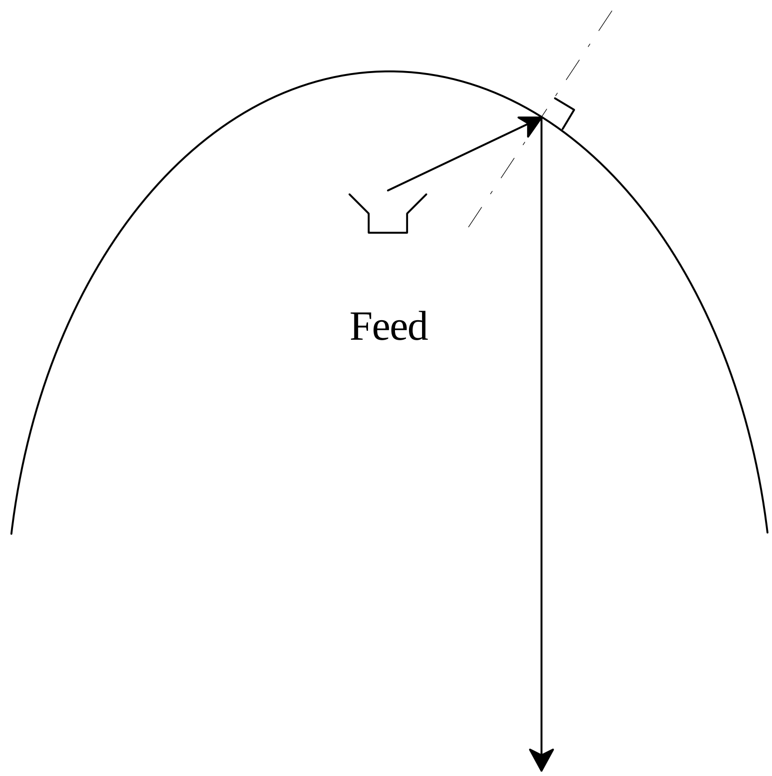



Figure 7.13: Incident and reflected rays at parabolic cylinder reflector.

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.** Place the antenna from problem 5 inside a conducting surface bent at an angle of 90° as shown in Fig. 7.14 (a so-called *corner reflector*). Repeat the exercise of problem 5. Which reflector do you think will provide a larger increase (or at least a smaller decrease) in the gain of the original antenna?

**7.** The "chip" antenna in Fig. 7.12 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 no 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 maximum effective gain, assuming a worst-case VSWR of 2, that the feed will not radiate and that 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 shown in Fig. 7.12, 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 is rarely used that low). Arecibo is used mostly for receiving very low noise powers from distant objects, but also sometimes as a planetary radar.

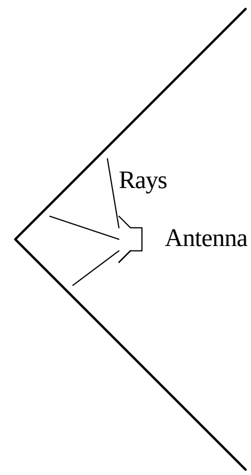



Figure 7.14: Antenna radiating towards a 90° corner reflector.

At 1 GHz, how much power would get to Jupiter if 1 GW was transmitted from Arecibo when the beam points exactly in the direction of Jupiter?

**9.** Calculate the minimal far-field distance of (1) a 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).

**10.** Draw a horn antenna as an extension of a $TE_{10}$ rectangular waveguide. Label the E and H planes of the co-polarized radiation pattern. Sketch two sets of patterns that you would expect from two differently sized horns: one flared only in the E-plane to a size of $2\lambda$ on a side, and one flared in both planes to $10\lambda$ on a side.

## Lab 7: 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 chamber antenna measurements in the second half of the lab.

## Part I. Pattern Measurements in the Anechoic Chamber

At microwave frequencies, antenna measurements are usually done in a special room called an *anechoic chamber*, as depicted in Figs. 7.15(a,b). The purpose of the anechoic chamber is to absorb any reflected

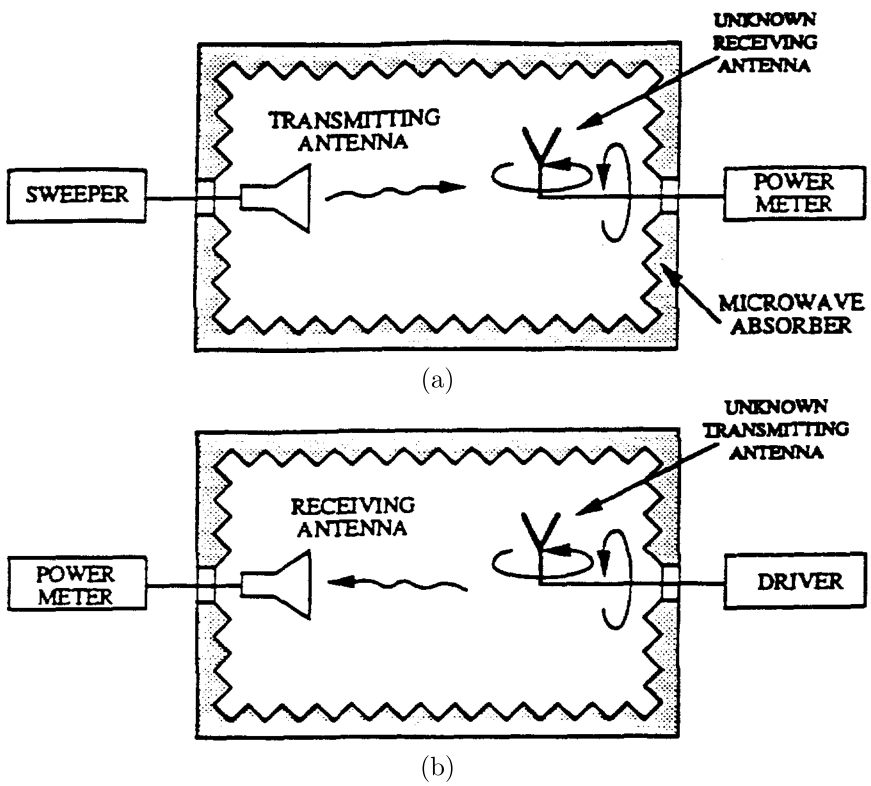



Figure 7.15: Antenna pattern measurement for a receiving (a) and transmitting (b) antenna in an anechoic chamber.

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.

**Q1:** The chamber in the lab is about 4 m long from front wall to back wall, and the two antennas are about 3 m apart. The transmitting horn at the back end is a wideband ridged horn that can be used from 1 to 18 GHz. How large 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?

**Q2:** We will use an HP 437B power meter with the HP 8141D 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 11 dB, and that the effective area of the antenna under test (AUT) is that determined in **Q1**. State what your S/N criterion is.

In this part of the lab you will first measure the radiation of a microwave horn antenna at X-band. Antenna pattern measurements will be made using the Diamond Antenna Measurement Studio (DAMS). The AUT is mounted on a special positioner platform attached to a tripod, both made of mostly nonmetallic materials to minimize unwanted scattering effects on the measurements. The positioner can rotate in both the horizontal and vertical planes, and so is capable of making E-plane and H-plane measurements without re-mounting the AUT. The motion of the positioner is controlled by the benchtop 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 ridged horn antenna at the back of the chamber, fed by a sweep oscillator.

Carefully remove the tripod from the chamber. Mount the AUT on the positioner such that it is centered above the center of horizontal rotation of the platform. Make sure the AUT is oriented so that its main beam is directed horizontally, and the vertical arm of the positioner is on the right side of the AUT as you look at the back of the AUT. Connect the RF cable from the positioner mount to the AUT. Replace the tripod in the anechoic chamber so that the label "This side towards other antenna" on the *tripod base* is facing the ridged horn at the back of the chamber. The center of the AUT should be located halfway between the left and right walls of the chamber. Make sure the control cables and RF cables are neatly arranged away from the signal path from the AUT to the ridged horn, and that the RF cable connected to the AUT is free to move easily when the positioner moves, and does not block the signal path.

Make sure AC power is on to the stepper motor drive, the DAMS Platform Controller, the power meter and the sweep oscillator. Check that the Enable switch on the stepper motor drive is set to On. Before beginning your antenna pattern measurements, you must choose the polarization of the incident wave. This is controlled by a separate LabView program "Polarization Selector" on the PC that drives a stepper motor rotating the transmitting ridged horn. To exit this program when finished, you must first click on the "STOP" button in its program window, and then close the window.

Now open the DAMS software to prepare to measure the radiation pattern. A more detailed description of the use of the DAMS software is given in Appendix E, but basic instructions will be provided in what follows. Configure the angular extents as follows:

- Azimuth set to Beam Cut ($-180°$ to $+180°$)
- Azimuth Start: $-90°$; Azimuth Stop: $+90°$
- Elevation set to $-90°$ to $+90°$
- Elevation Start: $-60°$; Elevation Stop: $+60°$

Set the angular step to 7.5°.

Next, use the Jog controls in the DAMS software to adjust the position of the AUT until it directly faces the transmitting horn. To do this, use the drop-down menus to select a number of degrees to move the positioner up, down, right or left, even if the box is already showing the number of degrees you want to move it. When you have finished, make this position $(0°, 0°)$ by clicking on "Zero Positioner", then OK. Notice that the current angular position of the AUT is indicated in the DAMS software window as

"AzPos" and "ElPos". For measuring the horn antenna, select sweep oscillator frequencies of 8 and 10 GHz and a power level of 20 dBm in the Source configuration menu.

Measure the 3D radiation patterns for incident waves of both vertical and horizontal polarization (use the "Scan AzEl" type of measurement in DAMS). The horizontal and vertical angle ranges of the AUT in the DAMS software should be as you chose above unless otherwise specified. If for some reason you need to interrupt a measurement at any time, click on the "Stop Measurement" button to do so gracefully. The measured results for received power show up in the DAMS program window, and can be exported to a file for plotting in your lab report. This is done with the "Data Processing" and "Data Visualization" features of the DAMS software. Results can be saved as a text file, a spreadsheet file, or as a graphics file.

**Q3:** Sketch the location of the AUT on the positioner, and indicate the E-plane and H-plane in your sketch.

**Q4:** What part of the measured data is needed to plot the E-plane patterns, and what part to plot the H-plane patterns? Include these plots in your lab report.

**Q5:** How wide are the 3 dB beamwidths of the patterns?

**Q6:** Calculate approximately the directivity of the waveguide horn from its E and H-plane co-polarized patterns using (7.17). Compare this to the value obtained using the effective area, if you assume $A \simeq A_{\rm geom}$.

**Q7:** 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 maximum co-polarized power to the maximum cross-polarized power, and is usually quoted in dB.)

Repeat **Q3** to **Q7** for the other antenna(s) provided, but measure only 2D patterns (for both azimuth and elevation) with an angular step size of 5° instead of a full 3D pattern to save time. Be sure to reset the Zero Position if necessary after mounting a new antenna, as described above. Measure E-plane and H-plane patterns using the "Azimuth Cut" and "Elevation Cut" buttons in the software, making sure that the angle not being varied is set to 0°. In the case of a resonant antenna such as a microstrip patch, first measure the feed-point reflection coefficient of the antenna (see Part II below) and choose two (or more) frequencies at which the antenna is best matched to make the chamber measurements, rather than 8 and 10 GHz. Are the definitions of E-plane and H-plane obvious now? What about co-polarization and cross-polarization? How much higher or lower is the gain of the other antenna(s) than that of the waveguide horn? Is the difference consistent with the ratio of their geometrical areas?

## Part II. $S_{11}$ of the Antennas

In this portion of the lab, you will look at the reflection coefficient of the AUTs on the Agilent 8719 network analyzer, since the antenna is designed for frequencies above 6 GHz, and thus we cannot use the HP 8753ES machines. This time, an SMA-to-waveguide transition is connected to the horn, so it may have a different reflection coefficient than when the horn itself was measured at the end of a waveguide in an earlier lab. The network analyzer will have been calibrated for you using an SOLT cal kit for 3.5 mm connectors.

**Q8:** Plot the magnitude of the $S_{11}$ coefficient for the waveguide horn from $f = 2$ GHz to 13 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.

**Q9:** Repeat the above for the other antenna(s) provided.

## Part III. Friis Formula and Polarization Effects

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.

**Q10:** 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.

**Q11:** 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? Below 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 Plexiglas mount 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.

**Q12:** 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.

**Q13:** 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 **Q5** for the 3-dB beamwidths?

## Part IV. Parabolic Reflector

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 Lecture 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.

**Q14:** How does the received signal strength compare with that obtained directly between the antennas without using a reflector, if the distance between the antennas is left unchanged?

**Q15:** What is the effect of adjusting the upwards angle of the feed horn (remembering to re-orient the reflector position for maximum signal strength)?

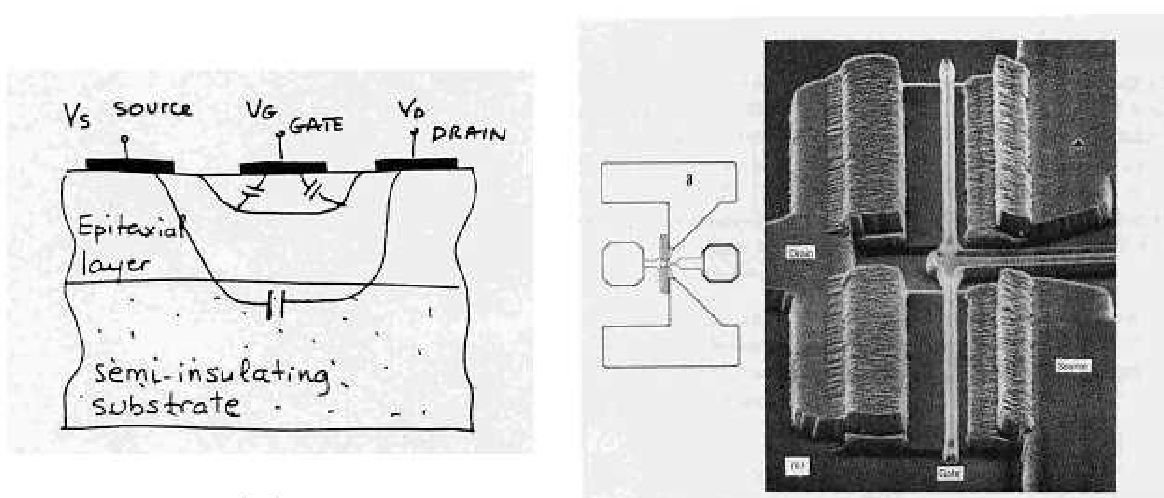
**Q16:** Can you think of other possible enhancements to increase the gain of the horn antenna?

# 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. 8.1(a), and a typical electrode layout is shown in Fig. 8.1(b). 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.




(a) (b)

Figure 8.1: A cross section of a MESFET (a) and photograph and electrode layout (b) (courtesy Tom Midford, Hughes).

When you purchase a MESFET, it can come either in a package or in chip form. You will be furnished with 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. In particular, the $S_{21}$ parameter corresponds to the gain of the device in common-source configuration. The amplitudes and phases of all four parameters are given at many discrete frequencies.

Another way to represent the transistor is with an equivalent circuit, as 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 than scattering parameters. Admittance parameters for a two-port network are defined by the relation

$$\mathbf{i} = \mathbf{Y}\mathbf{v} \tag{8.1}$$

where $\mathbf{i}$ and $\mathbf{v}$ are column vectors made up of the currents and voltages at the ports, analogous to the wave amplitude column vectors $\mathbf{a}$ and $\mathbf{b}$ used to define the scattering matrix. The admittance matrix $\mathbf{Y}$ is a $2 \times 2$ matrix, often given in a form normalized to the characteristic admittance of an appropriate transmission line:

$$y_{ik} = \frac{Y_{ik}}{Y_0} = Y_{ik} Z_0 \tag{8.2}$$

If we wish to find the values of the elements in the equivalent circuit, we could first solve for the $S$-parameters of the equivalent circuit in terms of the unknown circuit 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 first, because they are rather simply related to the equivalent circuit parameters. These are subsequently transformed to $S$-parameters using the formulas given, e. g., in *Microwave Engineering* (third edition) by David Pozar, page 187.

What is the order of magnitude of the S-parameters of a MESFET? As an example, let us choose an Agilent MESFET in chip form similar to the package you will use in the lab, and at 500 MHz with $V_{DS} = 3\,\text{V}$ and $I_{DS} = 20\,\text{mA}$, 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,

$$S_{21} = 5.0\angle 166^\circ \approx -5,$$

which represents the gain of a common-source (inverting) amplifier. In the reverse direction, ideally a transistor does not transmit a signal, so

$$S_{12} = 0.029\angle 77^\circ \approx 0.$$

(if we assume $S_{12} = 0$ exactly, this is called the *unilateral* model of the device). Finally, the output reflection coefficient of the transistor reduces to

$$S_{22} = 0.52\angle -11^\circ \approx 0.5.$$

We see that at such low frequencies, say below 1 GHz, the capacitances and inductances associated with a MESFET are quite small, and we can assume they are negligible. The same is true for the resistive losses, and in this case we arrive at a simple low-frequency model by approximating $S_{11} = 1$ and $S_{12} = 0$. These values are achieved from the equivalent circuit shown in Fig. 8.2(a). For this equivalent circuit, the normalized $Y$-parameters are

$$\mathbf{y} = \begin{bmatrix} 0 & 0 \\ g_m & g_{ds} \end{bmatrix}, \tag{8.3}$$

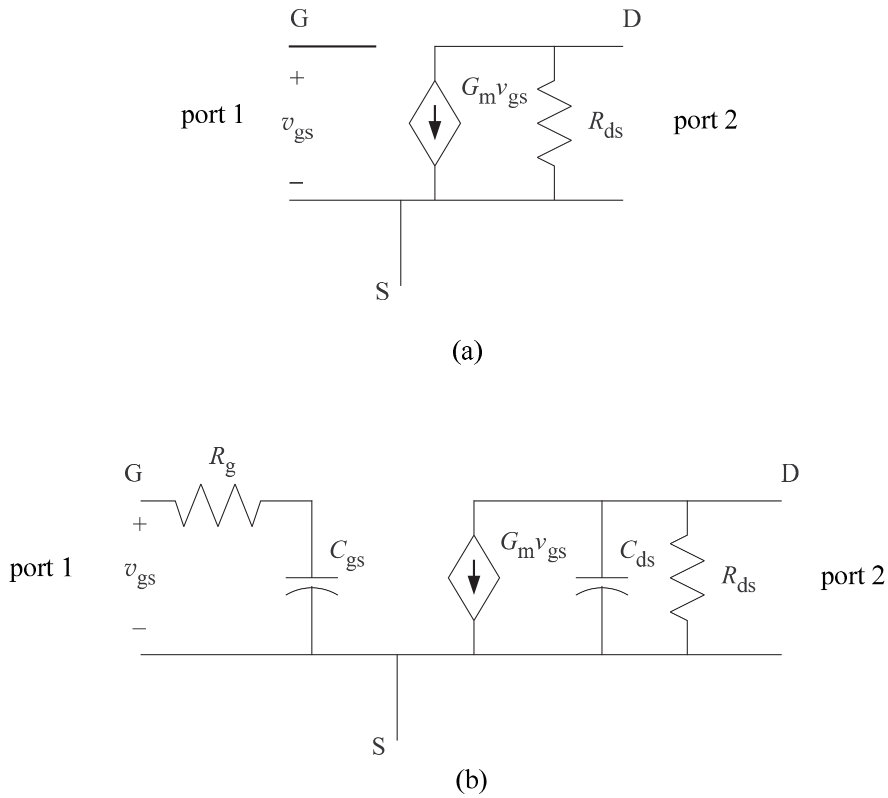



Figure 8.2: Unilateral (a) low and (b) high-frequency equivalent circuit models of a microwave MESFET. No package parasitics are included, so this model is referred to as the *intrinsic* device model.

where $g_m = G_m Z_0$ is the normalized transconductance of the MESFET, and $g_{\rm ds} = Z_0/R_{\rm ds}$ is the normalized drain-source conductance. The $S$-parameters are then obtained by using the conversion formulas mentioned previously:

$$\mathbf{S} = \begin{bmatrix} 1 & 0 \\ \frac{-2g_m}{1+g_{ds}} & \frac{1-g_{ds}}{1+g_{ds}} \end{bmatrix}, \tag{8.4}$$

From the measured values of $S_{21}$ and $S_{22}$ cited above, we find that $g_m = 3.33$ and $g_{ds} = 0.33$ (these conductance values are normalized to $Y_0 = 1/50 = 0.02\,\text{S}$.), so the voltage gain is found to be about $A_V = g_m/g_{ds} = 10$. Note that this simple model has an infinite input impedance and thus cannot be matched at the input.

At higher frequencies, capacitances need to be included in the model. The circuit in Fig. 8.2(b) is called the "intrinsic" equivalent circuit because additional parasitics from the package are not included. Although the $S$-parameters are now given by more complicated formulas than (8.4), it is still true that this model is unilateral ($S_{12} = 0$). 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 $R_g = 10 - 15\,\Omega$. 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 (via the $RC$ time constant). An experimentally obtained formula is

$$f_{max} = \frac{33 \times 10^3}{l}\,\text{Hz},$$

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 are used primarily at lower microwave frequencies, and 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 its complicated layered semiconductor structure. The details of operation of the latter two devices are beyond the scope of this course.

## 8.2 Transistors as Bilateral Two-ports

Before examining the behavior of a transistor as a bilateral two-port network, let us recall some basic properties of power transfer from circuit theory. Consider the circuit of Fig. 8.3. A generator is repre-

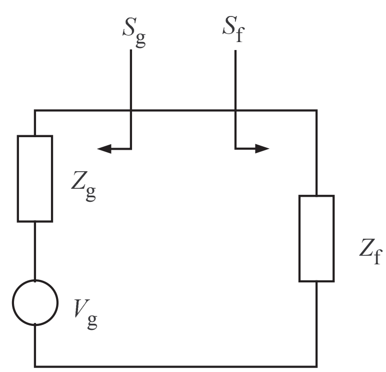



Figure 8.3: Power transfer in a simple circuit.

sented by a Thévenin equivalent generator $V_g$ and a Thévenin impedance $Z_g$. This generator drives an impedance $Z_f$. It is well known that out of all possible values of $Z_f$, the maximum power is supplied to it when $Z_f = Z_g^*$; i. e., the load is conjugate matched. This maximum power is called the available power of the generator, and is given by

$$P_{\text{avl}} = \frac{|V_g|^2}{4R_g} \tag{8.5}$$

where $R_g = \text{Re}\,(Z_g)$. Alternatively, we could define reflection coefficients looking into the generator or the load as

$$S_g = \frac{Z_g - Z_0}{Z_g + Z_0}; \qquad S_f = \frac{Z_f - Z_0}{Z_f + Z_0} \tag{8.6}$$

with respect to a real characteristic impedance $Z_0$. Thus, the maximum power condition can be expressed as $S_f = S_g^*$. Note that in general the conjugate match condition has nothing to do with matching to suppress reflection on a transmission line; its function is to maximize power delivered to $Z_f$, even if that means reflections must be present to accomplish this.

Now insert a transistor between the generator and a load impedance, which we will now call $Z_L$, resulting in the network shown in Fig. 8.4. Two-port transistor $S$-parameters are usually given for the

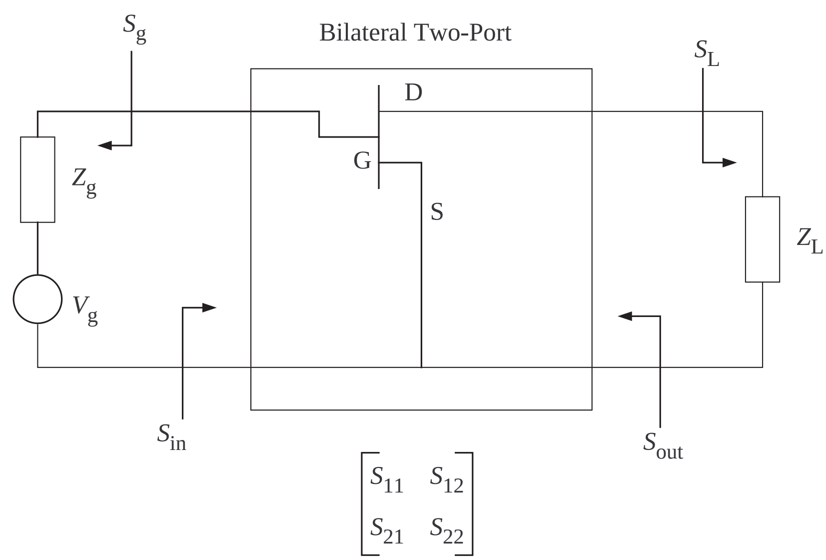



Figure 8.4: Input and output scattering parameters of a bilateral two-port transistor network.

common-source configuration. 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).

The generator ($V_g$, $Z_g$) now drives the input port of the two-port network representing the bilateral MESFET. The output port is connected to $Z_L$. The total input reflection coefficient looking into port 1 of a two port with given $S$-parameters, terminated in an arbitrary load with reflection coefficient

$$S_L = \frac{Z_L - Z_0}{Z_L + Z_0} \tag{8.7}$$

is found from the reflection coefficient definition to be (see homework problem 1):

$$S_{in} = \frac{b_1}{a_1} = S_{11} + \frac{S_{12}S_{21}S_L}{1 - S_{22}S_L}. \tag{8.8}$$

A similar expression can be derived for the output reflection coefficient $S_{out}$ looking into port 2, by replacing $S_L$ with

$$S_g = \frac{Z_g - Z_0}{Z_g + Z_0}$$

and the appropriate $S$-parameters of the two port:

$$S_{out} = \frac{b_2}{a_2} = S_{22} + \frac{S_{12}S_{21}S_g}{1 - S_{11}S_g}. \tag{8.9}$$

Ordinarily, the generator impedance will not be conjugate-matched to the input of the amplifier, nor will the output of the amplifier be conjugate-matched to the load. These mismatches can be corrected by inserting lossless matching networks after the generator and before the load, as shown in Fig. 8.5. The resulting amplifier is now given by Fig. 8.4, but with $S_g$ replaced by $S_g'$ and $S_L$ replaced by $S_L'$

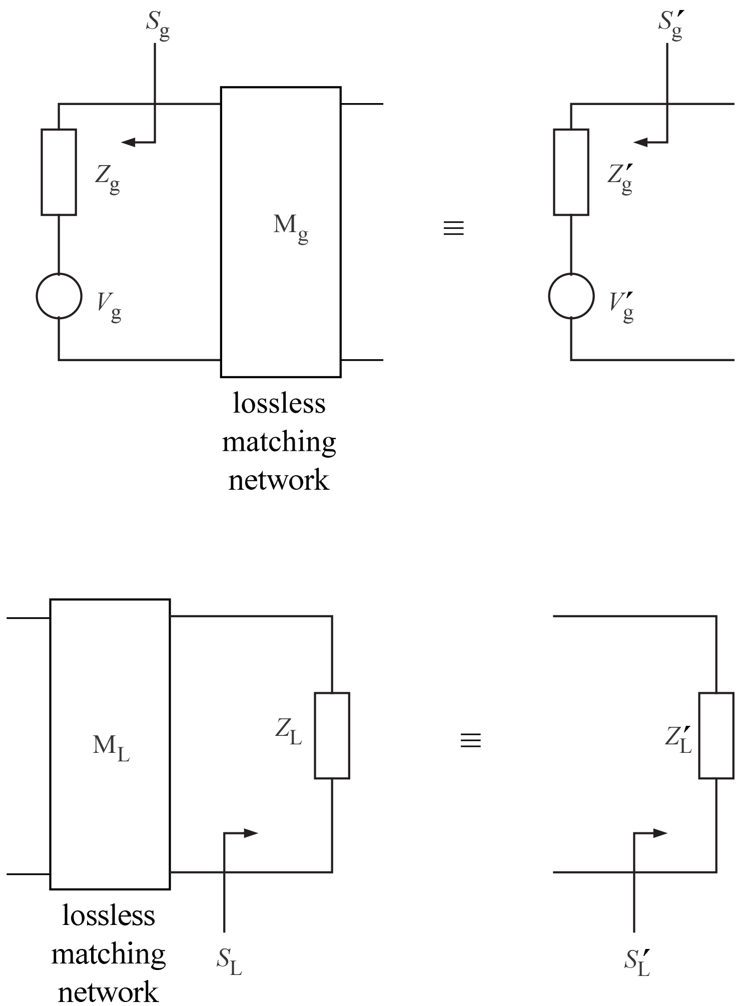



Figure 8.5: Lossless matching networks for the source and load.

respectively. A conjugate match at both input and output is obtained by enforcing

$$S_{in}^* = S_g'; \qquad \text{and} \qquad S_{out}^* = S_L' \tag{8.10}$$

For given transistor parameters $S_{11}$, $S_{21}$, etc., these equations can be solved simultaneously for the required values of $S_g'$ and $S_L'$, from which the desired matching networks $M_g$ and $M_L$ could be designed. Details can be found in the book by G. Gonzalez, *Microwave Transistor Amplifiers: Analysis and Design*, 1984, p. 113.

Since power amplification is an important characteristic of our amplifier, a key question that needs to be answered is how much additional power beyond the available power of the generator can be delivered to $Z_L$? Another important question is: when can we assume that a transistor is unilateral? The answer is: when $S_{12}$ is "small enough", but what does that mean? We will find answers to these questions when we learn about transistor amplifier gain below.

## 8.3 Stability of Microwave Transistor Circuits

If the magnitudes of $S_{in}$ or $S_{out}$ from Fig. 8.4 are greater than unity at some frequency, the circuit is unstable at that frequency and potentially an oscillator (see Section 5.2). Oscillation occurs if we have a nonzero output when the input is zero ($V_g = 0$ in Fig. 8.4). Suppose that $|S_{out}| > 1$ for some generator impedance $Z_g$ corresponding to $S_g$. Then the resonance condition is

$$S_L = \frac{1}{S_{out}} \tag{8.11}$$

as in (5.9). From (8.9), this can be rewritten as

$$1 = S_L S_{22} + S_g S_{11} - S_L S_g \Delta \tag{8.12}$$

where $\Delta$ is the determinant of the device $S$-matrix:

$$\Delta = S_{11}S_{22} - S_{12}S_{21} \tag{8.13}$$

When designing an amplifier, it is important to make sure that oscillation does not occur (that is, (8.12) can never be true), by properly designing the matching networks to modify the device's $S$-parameters. The conditions to be met will be presented next.

An amplifier is *unconditionally stable* when $|S_{in}| < 1$ and $|S_{out}| < 1$ for all passive source and load impedances (i. e., for any $S_L$ and $S_g$ that are $\leq 1$ in magnitude), 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:

$$K = \frac{1 - |S_{11}|^2 - |S_{22}|^2 + |\Delta|^2}{2|S_{12}S_{21}|} > 1 \qquad \text{and} \qquad |\Delta| < 1. \tag{8.14}$$

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 will be done in the lab), RF connections to ground are made using metalized 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 an 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 $S$-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.

## 8.4 Microwave Transistor Amplifiers

A simple graphical explanation of an amplifier is shown in Fig. 8.6(a). 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?

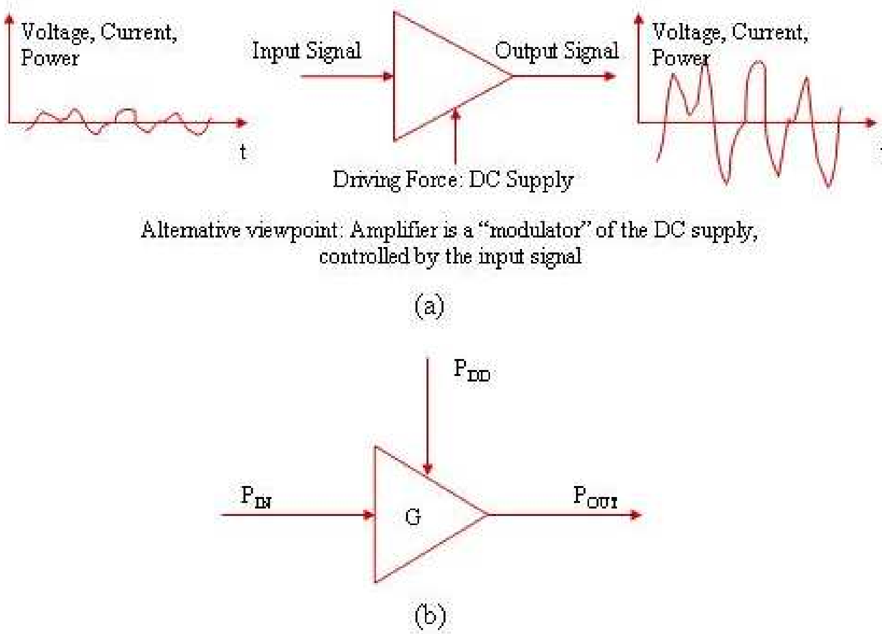



Figure 8.6: (a) General block diagram of an amplifier. (b) Power budget for defining efficiency.

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*, $G_T$:

$$G_T = \frac{P_{\text{load}}}{P_{\text{avl,source}}} \tag{8.15}$$

This definition of gain is the ratio of the power delivered to a generally mismatched load ($S_L \neq 0$) divided by the power available from the source. In the absence of source or load matching networks, $G_T$ can be shown in general to be equal to either of the expressions

$$G_T = \frac{|S_{21}|^2 \left(1 - |S_g|^2\right)\left(1 - |S_L|^2\right)}{|1 - S_{in}S_g|^2 |1 - S_{22}S_L|^2} = \frac{|S_{21}|^2 \left(1 - |S_g|^2\right)\left(1 - |S_L|^2\right)}{|1 - S_{11}S_g|^2 |1 - S_{out}S_L|^2} \tag{8.16}$$

For a given $S_{in}$ with $|S_{in}| \leq 1$, $G_T$ in (8.16) can be maximized by inserting a matching network $M_g$ after the source, such that $S_g \to S_g' = S_{in}^*$, so the maximum possible transducer gain for given $S_{in}$ and $S_L$ is

$$G_{T\max} = \frac{|S_{21}|^2 \left(1 - |S_L|^2\right)}{\left(1 - |S_{in}|^2\right)|1 - S_{22}S_L|^2} \tag{8.17}$$

If the amplifier is not matched to the source, then reflected power is lost and a second definition, the *power gain* $G$, is important, and can be defined as:

$$G = \frac{P_{\text{load}}}{P_{\text{avl,source}} - P_{\text{refl}}} = \frac{P_{\text{load}}}{P_{\text{in}}} \tag{8.18}$$

It can be shown to be equal to

$$G = \frac{|S_{21}|^2 \left(1 - |S_L|^2\right)}{\left(1 - |S_{\text{in}}|^2\right) |1 - S_{22} S_L|^2} \tag{8.19}$$

If $S_{in} = 0$ and $S_L = 0$, then we have simply

$$G = |S_{21}|^2$$

**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. 8.6(b).

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
$$\eta_{\text{total}} = \frac{P_{\text{out}}}{P_{DD} + P_{\text{in}}}.$$

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
$$\eta_{D(C)} = \frac{P_{\text{out}}}{P_{DD(CC)}}.$$

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

$$\text{PAE} = \frac{P_{\text{out}} - P_{\text{in}}}{P_{DD}}.$$

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

$$P_{\text{DISS}} = P_{DD} + P_{\text{in}} - P_{\text{out}}.$$

**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 $\mu$m 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. The inductance can be estimated by using the formula for the inductance of a straight wire conductor of length $l$ and radius $r_0$ given in (2.18). This formula can also be applied to the case of a thin strip conductor of width $w$ by putting $r_0 = w/4$. As an example, consider a wire of length 1 mm and radius 0.2 mm, for which this formula gives an inductance of 0.26 nH. In practical situations, this inductance is usually between 0.5 and 1 nH for standard substrate thicknesses, and the reason most transistor packages have 2 source leads is to halve 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. See sections 3.5 and 8.6 for more information.

This concludes our 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. 8.7. 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 $I$-$V$ curves for a BJT (HBT) are shown in Fig. 8.7 ($I$-$V$ curves for a FET are similar and the discussion here also 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:

$$f = 5\,\mathrm{GHz} : S_{11} = 0.7\angle -100^\circ, S_{12} = 0, S_{21} = 2.5\angle 60^\circ, S_{22} = 0.7\angle -30^\circ.$$

Because we assume $S_{12} = 0$ in the unilateral approximation, we have

$$S_{\mathrm{in}} \simeq S_{11}, \qquad S_{\mathrm{out}} \simeq S_{22} \tag{8.20}$$

If we now assume 50-Ω source and load impedances, the task becomes to transform both $S_{11}$ and $S_{22}$ so as to match 50 Ω (see Figure 8.8). To perform an ordinary conjugate match, we simply want to design matching sections that transform $S_L$ to $S'_L$ and $S_g$ to $S'_g$:

$$S'_g = S^*_{11} \qquad \text{and} \qquad S'_L = S^*_{22} \tag{8.21}$$

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 the

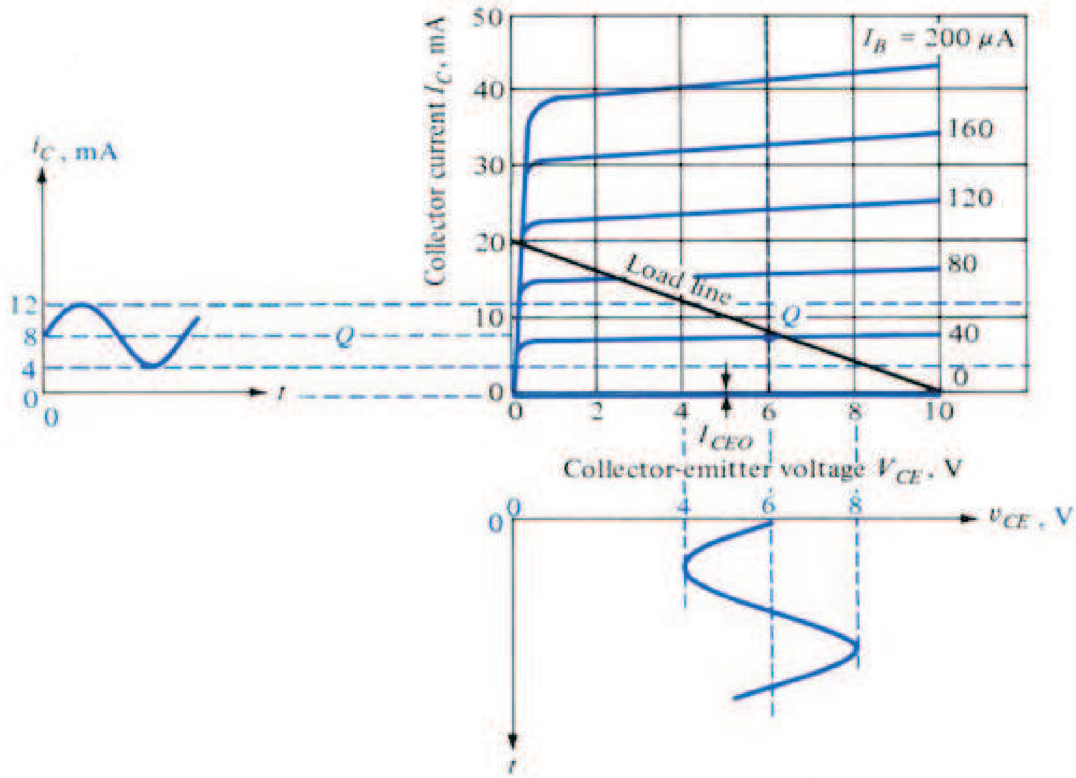



Figure 8.7: *I-V* curves of a transistor showing a load-line, operating point Q and input current and output voltage waveforms.

source and load, we have guaranteed the maximum transducer gain possible for this amplifier, under the unilateral approximation. This matched transducer gain is a function of $S_{21}$ and the matching sections and is found from (8.17) to be

$$G_{T\max} = \frac{|S_{21}|^2 \left(1 - |S_{22}|^2\right)}{\left(1 - |S_{11}|^2\right)\left(1 - |S_{22}|^2\right)^2} = \frac{|S_{21}|^2}{\left(1 - |S_{11}|^2\right)\left(1 - |S_{22}|^2\right)}. \tag{8.22}$$

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 small? 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 used in the previous example. If $S_{12} \neq 0$, it is in most cases mainly 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 $S_g$. Note in this

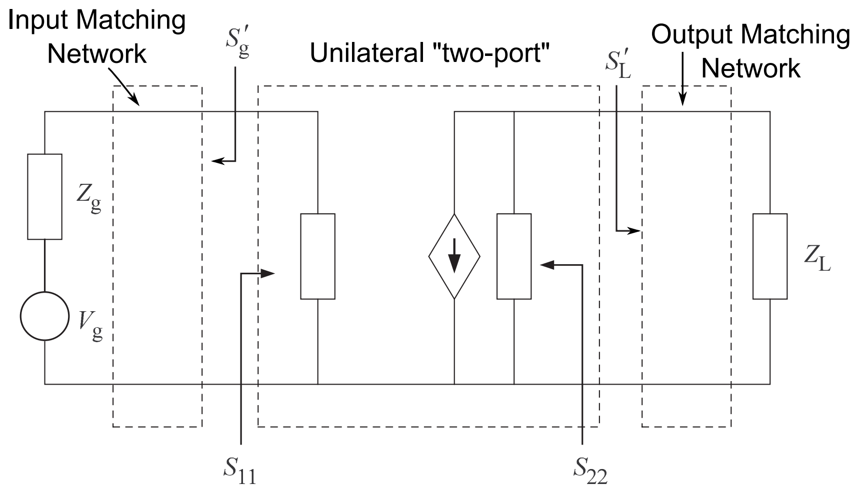



Figure 8.8: Conjugate matching for unilateral amplifier design.

connection that the denominators in the expressions for $S_{in}$ and $S_{out}$ are less than unity. This indicates that for some values of the transistor $S$-parameters, 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).

Therefore, in the bilateral case we have to perform a simultaneous match of the input and output as discussed in section 8.2 (in other words, what we do at the input affects the output match, and vice versa). Without going into the details of the design procedure, we show in Fig. 8.9 how $S_{in}$ and $S_{out}$ differ from $S_{11}$ and $S_{22}$ respectively for a typical case. In this particular example it can be seen how the

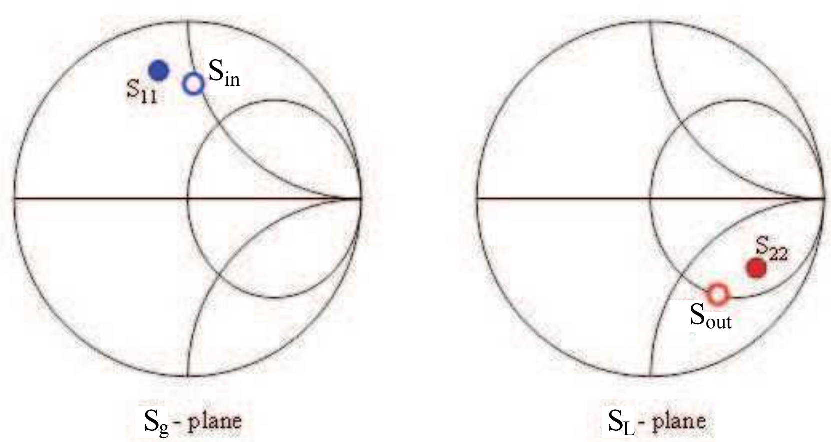



Figure 8.9: Smith chart comparison of the difference in the transistor reflection parameters and the circuit reflections given by $S_{in}$ and $S_{out}$ when $|S_{12}| > 0$.

matching can change 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.

## 8.6 Biasing the Transistor in an Amplifier

In the block diagrams in Fig. 8.6, there is no obvious place where power to the device is added. An amplifier achieves its 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 DC bias supply, in a way that does not change the input and output reflection coefficients. A full amplifier circuit is therefore shown in Fig. 8.10.

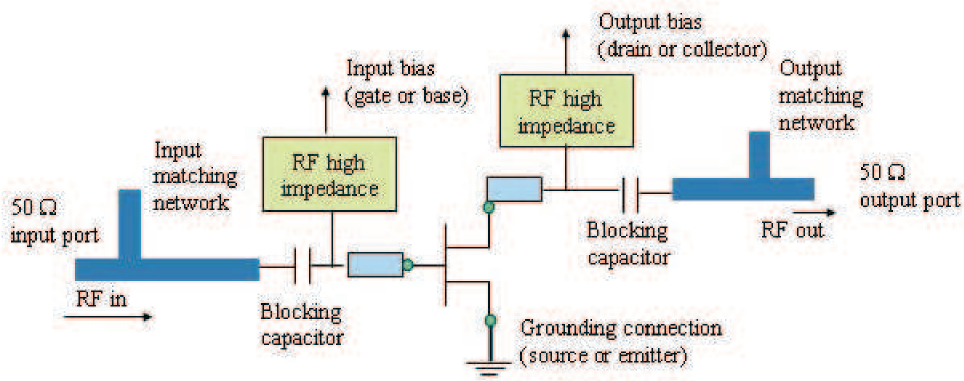



Figure 8.10: Complete circuit schematic of a microwave amplifier including the bias circuits in the gate and drain.

The general principles of supplying DC bias to a microwave circuit were presented in section 3.5. For amplifier circuits, we must additionally be careful not to upset the stability of the circuit. 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 terminal and ground. In the case of coplanar waveguide (CPW) circuits, the connection is more straightforward and the parasitic reactance can be minimized.

Commercial bias Tees can be used as an “add-on” to an amplifier circuit, but are fairly large and have typically SMA connectors at the two RF ports. Often, biasing circuits are realized as an integral part of the amplifier design, and some examples are shown in Fig. 8.11.

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 (right-hand side of Fig. 8.11), 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

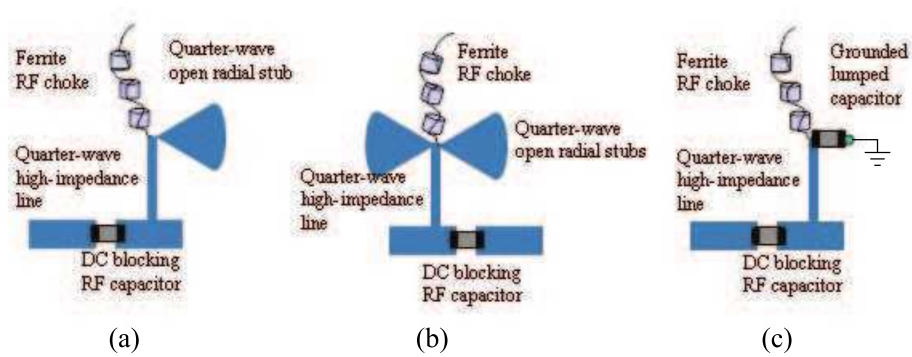



Figure 8.11: Examples of microstrip biasing circuits.

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 to achieve 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 (8.8) and (8.9) starting from Fig. 8.4.

**2.** Derive equation (8.19).

**3.** Derive equation (8.16).

**4.** Derive the $Y$-parameters of the high-frequency MESFET equivalent circuit model shown in Fig. 8.2(b).

**5.** 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.

**6.** Derive (8.14) from (8.8) and (8.9) by allowing $S_g$ and $S_L$ to vary over all possible passive values.[1]

**7.** 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.

**8.** 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?

**9.** 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 $PAE_2 = 70\%$.

**10.** Choose a transistor available in your microwave design software. Simulate and plot the $S$-parameters of the transistor in a 2-port common source configuration for the frequency range 1-6 GHz 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 ($V_{DS}I_{DS}$ is larger)?

**11.** For the transistor specified 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.

**12.** For problem 11 above, perform a bilateral match to improve gain, and input and output match, while maintaining stability over the range where the transistor has gain.

[1] This problem is of greater than average difficulty.

**13.** Design microstrip bias Tee networks of the types shown in Fig. 8.11(a), (b) and/or (c) using microwave design software. Choose a center frequency of 2 GHz and use FR4 substrate ($\epsilon_r = 4.5$, 0.508-mm thick). If necessary, 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 the RF ports, and coupling between RF and DC ports. What is the bandwidth of your bias Tees?

## Lab 8: Microstrip Microwave Transistor Amplifier Measurements

In this lab you will characterize a microwave transistor amplifier using the network analyzer, as well as using a sweeper and a power meter. The steps of designing the amplifier you will be testing are similar to what was done for the pre-lab. It is a conjugate matched MESFET amplifier designed for about 800 MHz, with bias networks and impedance matching networks. You will also examine separate bias tee networks.

## 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 400 to 1200 MHz. 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 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 a bias Tee network. One type of bias Tee network uses a radial stub as part of the low pass filter for the DC input. Bias networks have been fabricated separately so that you can measure them without the transistor connected.

**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 line unconnected for the time being. Capture all the S-parameter data for the bias Tee.

**Q3:** From the measured S-parameters, what is the bandwidth (specify how you have defined it) and optimum operating frequency of the bias Tee? How would you make a more broadband biasing circuit at this optimum frequency?

**Q4:** After disconnecting the bias Tee from the network analyzer, verify that the DC blocking capacitor in line with the RF transmission line is indeed doing its job, and explain in your report.

## Part III. Amplifier Characterization - Small-Signal

The required biasing for the MESFET is:

- VDS = 1 to 4V (positive). Do not exceed 4 V!!!
- Adjust the gate bias so that you are measuring about 60mA of drain current.

The ports corresponding to the gate and drain are labeled on the amplifier board. Before connecting the power supply to the circuit, set the power supply voltage level to -4V for the gate bias and 0V for the drain bias and turn it off. Connect an ammeter to monitor the drain current, which is controlled by adjusting the gate bias voltage. You can use the current limiting feature on the power supply (if it exists) to be sure that too much drain current is not drawn (set this to 80 mA). In practice, you may see oscillations if the drain bias current exceeds 60 mA.

Connect the circuit to the network analyzer as shown in Fig. 8.12(a). 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 |S21| parameter for gain. You may need first to verify that no oscillation is occurring in your amplifier.

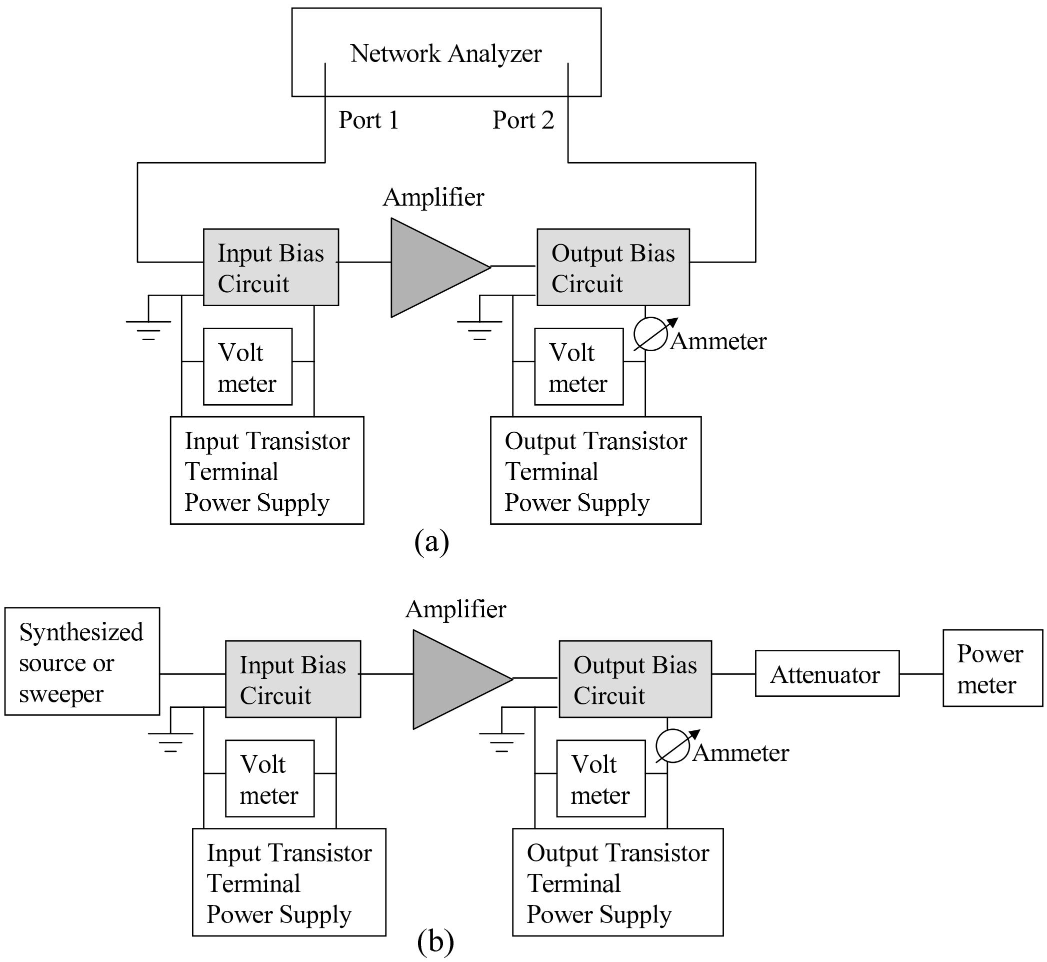



Figure 8.12: Experimental setup for (a) small-signal amplifier characterization and (b) power characterization.

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. You may need first to verify that no oscillation is occurring in your amplifier.

***You will get 20% of the grade for this lab taken off if you burn out the transistor - call the TA over to inspect your setup before turning on your amp.***

Obtain and plot the $S$-parameters for the amplifier. Record values at the design frequency.

**Q5:** Is the amplifier working and how do you know? Explain. If oscillations occur in your amplifier, how can they be suppressed? What is the "best" operating frequency? What is the gain at that frequency? How much current is being drawn from the power supply?

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 1.0V and 4V while keeping the gate voltage at -4V (change the voltage slowly and make sure you don't go over 4V for the bias voltage!).

**Q6:** What trends in output, gain, etc. do you see as you change the bias voltage? Explain what you see.

Fix the drain bias at 3.5V and vary the gate bias over a 1V range.

**Q7:** What trends do you see as you change the gate bias voltage? Explain what you see.

## Part IV. Amplifier Characterization - Power

Because we will be dealing with higher power levels in this part of the experiment, an attenuator should be connected to the output side of the amplifier to protect it and the power sensor. Connect a frequency synthesizer (sweep oscillator) to the input of the amplifier (gate side of the MESFET) and a power meter to the output side, Fig. 8.12(b). Bias the circuit with (3-4 V, -0.5 V) on the drain and gate of the MESFET respectively. Starting at -35 dBm, increase the power delivered to the amplifier until the 1-dB compression point is reached (or until the maximum power of the oscillator is reached, if 1 dB compression is not obtained).

(*Warning*: the power sensor cannot handle more than about 20 dBm, so note the advice above about using an attenuator.)

**Q8:** Record the output and input power levels for the power steps. You will need to measure the oscillator power level directly with the power meter, since the power displayed on the front panel is not necessarily accurate. Plot (1) output power vs. input power and (2) the power gain versus power out, noting the 1-dB compression point.

**Q9:** What are the (1) drain, (2) power-added and (3) total efficiencies of the amplifier? What trends do you see in efficiency when you change the bias voltage?

## Part V. Two-tone Intermodulation Measurement [Optional—requires two oscillators]

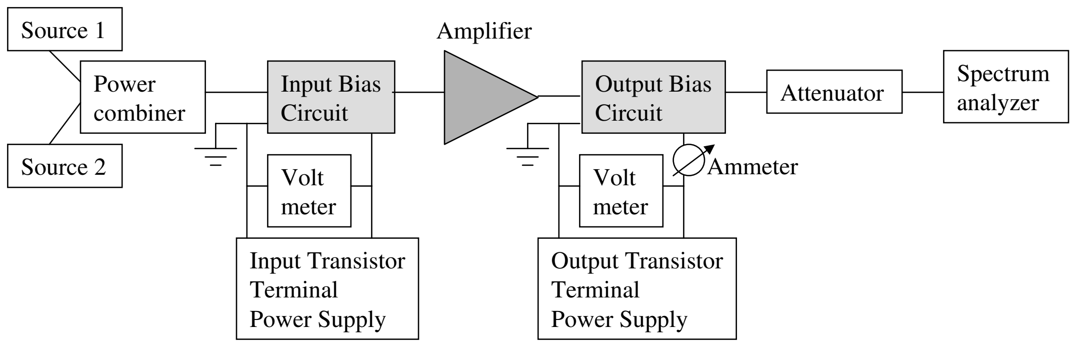



Figure 8.13: Experimental setup for two-tone nonlinearity amplifier characterization.

Connect two sweepers through a combiner circuit to the input of the amplifier, Fig. 8.13. Offset the frequencies on the two sweepers by a few MHz (1 MHz is okay). Connect the output of the amplifier to a spectrum analyzer. Bias the amplifier and set the sweepers to equal small-signal levels (-10 dBm for example).

Q10: Sketch the output of the spectrum analyzer.

Next increase the sweeper power levels, while keeping them the same, and observe the output on the spectrum analyzer.

Q11: 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. At how many dB above the 1-dB compression point does this occur?

# Chapter 9

# Microwave Communication Links: Superheterodyne Systems

## 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

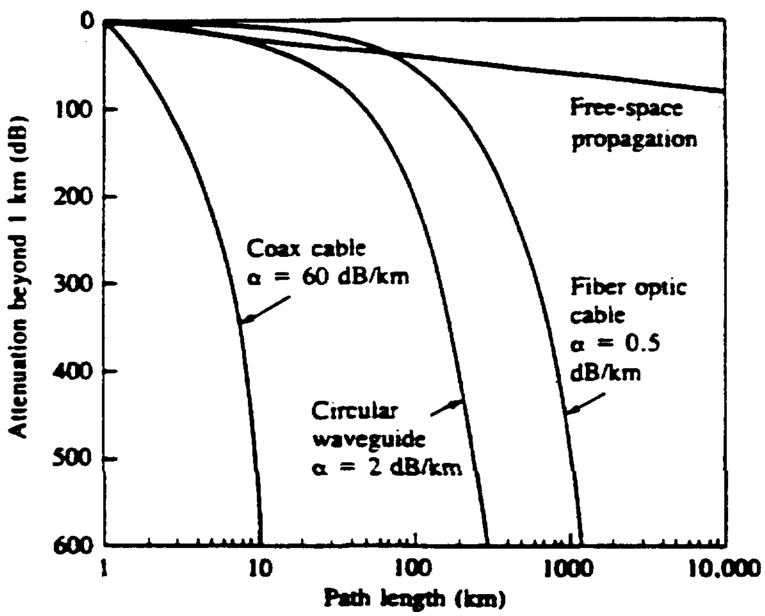



Figure 9.1: Attenuation for different transmission systems.

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 chapter 7. (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

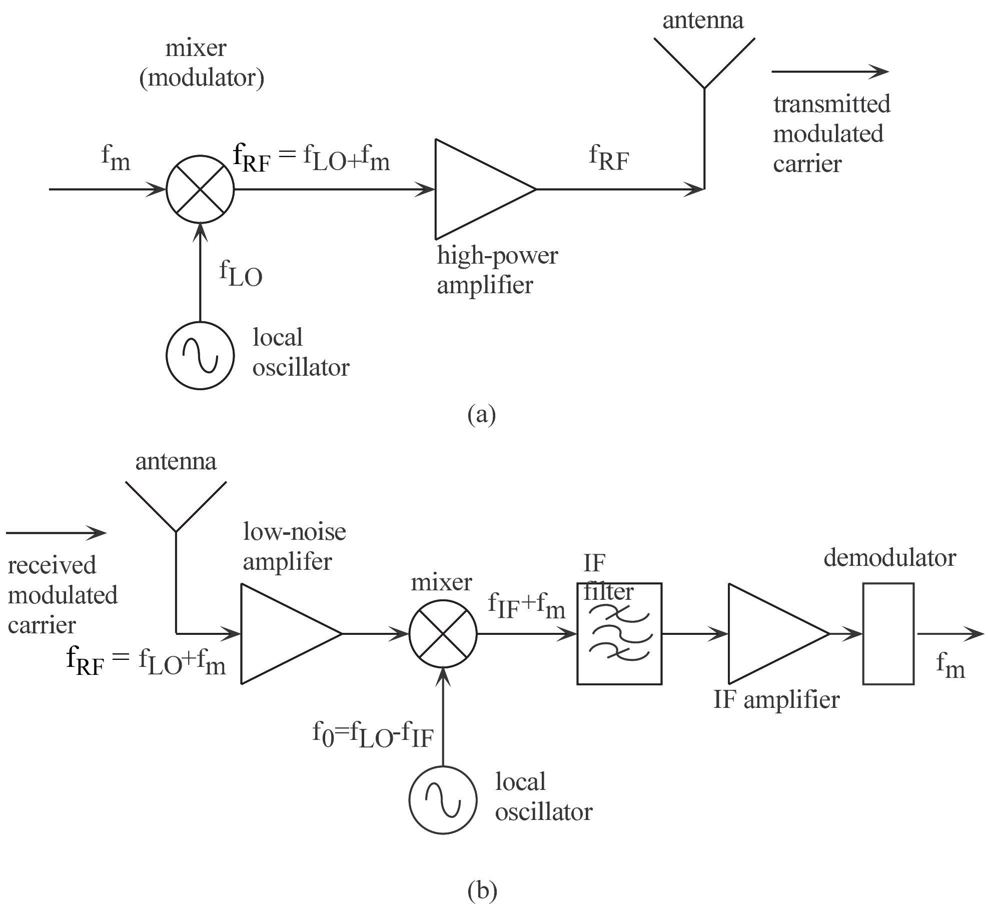



Figure 9.2: Block diagram of an AM microwave transmitter (a) and receiver (b).

(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 $f_{\mathrm{LO}} - f_m$ and one at $f_{\mathrm{LO}} + f_m$ are transmitted. This process is called *up-conversion*. The power amplifier is designed to amplify at the microwave frequency $f_{\mathrm{LO}}$, 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 $f_{\mathrm{IF}}$, which is called the *intermediate frequency (IF)*. We use IF for one or more of the following reasons:

- Amplification is easier to achieve at lower frequencies: stability is more readily maintained and lumped elements can be used.
- If many RF frequencies are to be covered, it is easier to tune them all to the same IF frequency for narrowband amplification rather than to have a wideband RF amplifier employed.
- Filtering is easier to perform at lower frequencies in terms of the percentage of the carrier frequency needed.

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 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 3 and 5, the diode current is a nonlinear function of voltage. To analyze the operation of a mixer, we use the small-signal (quadratic) approximation:

$$I(V) = I_0 + i = I_0 + v_d G_d + \frac{v_d^2}{2} G_d' + \ldots \tag{9.1}$$

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 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:

$$v_d = v_{\mathrm{LO}} + v_{\mathrm{RF}} \tag{9.2}$$

The LO and RF voltages can be written as:

$$v_{\mathrm{LO}} = \sqrt{2} V_0 \cos \omega_0 t \quad \text{and} \quad v_{\mathrm{RF}} = \sqrt{2} V \cos \omega_{\mathrm{RF}} t. \tag{9.3}$$

where usually the received signal voltage is much smaller than the LO voltage: $|V| \ll |V_0|$.

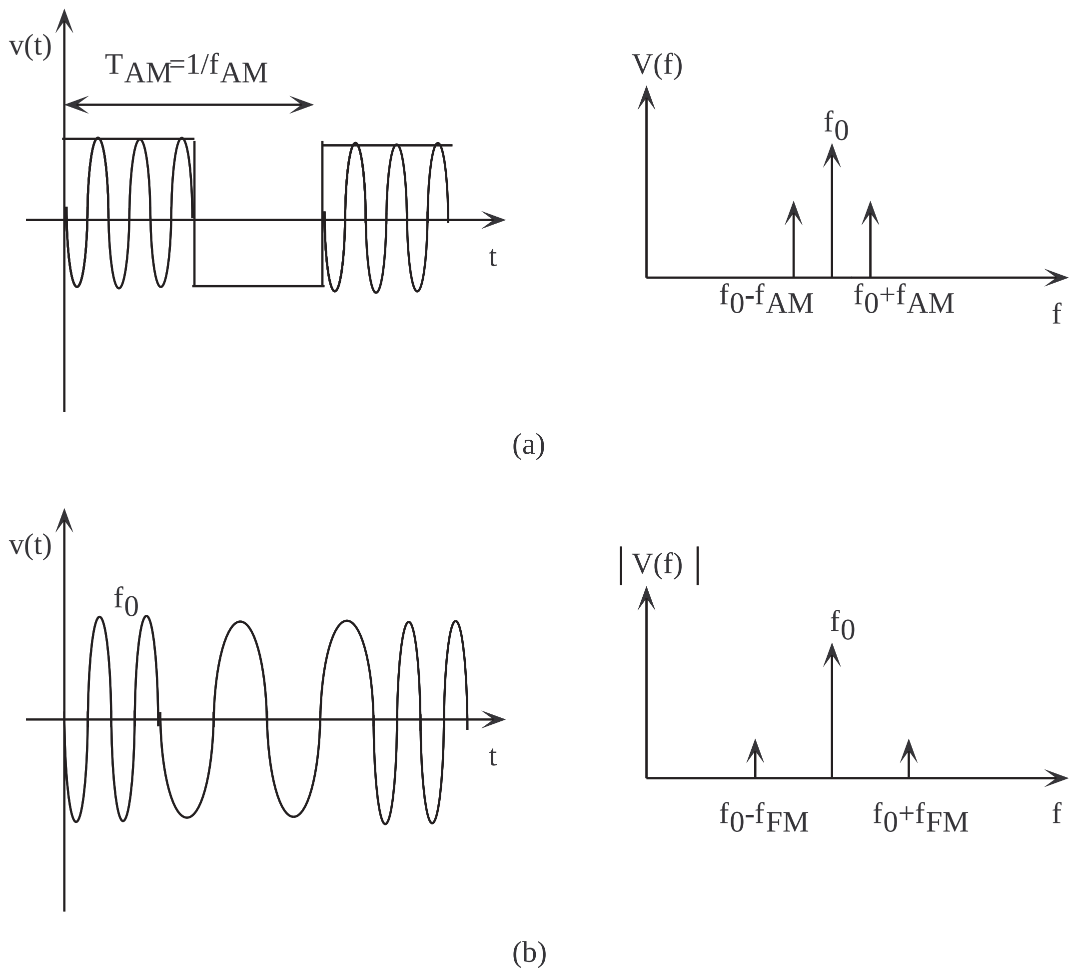



Figure 9.3: Time domain waveforms and spectra of an (a) AM and (b) FM modulated sinusoidal signal of frequency $f_0$.

The diode current expression (9.1) will now have a constant DC bias term, linear terms at both frequencies $\omega_{\rm RF}$ and $\omega_0$, and the $v_d^2$ term of the current which gives the following frequency components:

$$\begin{aligned} i &= \frac{1}{2}G_d' v_d^2 = G_d' \left(V\cos\omega_{\rm RF}t + V_0\cos\omega_0 t\right)^2 \qquad (9.4)\\ &= \frac{G_d'}{2}\left\{V^2 + V_0^2 + V^2\cos 2\omega_{\rm RF}t + V_0^2\cos 2\omega_0 t + 2VV_0\cos\left[(\omega_{\rm RF}-\omega_0)t\right] + 2VV_0\cos\left[(\omega_{\rm RF}+\omega_0)t\right]\right\}\end{aligned}$$

In this expression, the DC terms are not important, and the $\omega$, $\omega_0$, $\omega+\omega_0$, $2\omega$ and $2\omega_0$ terms can be filtered out. The most important term is the one at $\omega_{\rm RF}-\omega_0$. In case of a receiver, the frequency $\omega_{\rm RF}-\omega_0$ is the intermediate (IF) frequency. If the IF, RF and LO signals all act on an impedance $Z_0$, then from (9.4) we get that the IF power is proportional to both the RF and LO powers:

$$P_{\rm IF} = \frac{Z_0^3 (G_d')^2}{2} P_{\rm RF} P_{\rm LO} \qquad (9.5)$$

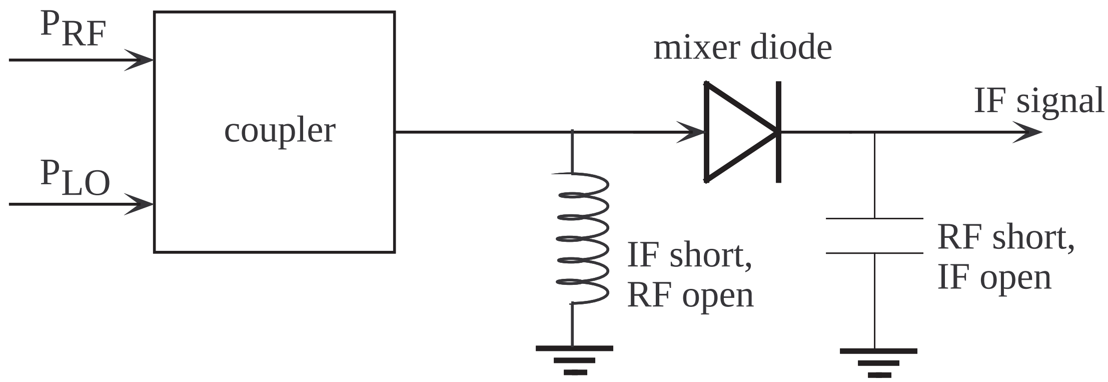



Figure 9.4: A single-ended mixer circuit.

At the transmitter end, the mixer performs the up-conversion. In this case, the IF signal at $\omega_i$ 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 is small compared to $V_0$), and contains frequencies at least up to the IF range. 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

$$v_{\rm LO} = \sqrt{2}\,[V_0 + V_n(t)] \cos\omega_0 t \tag{9.6}$$

Once again, let

$$v_{\rm RF} = \sqrt{2} V \cos\omega_{\rm RF} t \tag{9.7}$$

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

$$\begin{aligned} v_a(t) &= V\cos(\omega_{\rm RF}t - \pi/2) + (V_0 + V_n)\cos(\omega_0 t - \pi) \\ v_b(t) &= V\cos(\omega_{\rm RF}t) - (V_0 + V_n)\cos(\omega_0 t - \pi/2) \end{aligned} \tag{9.8}$$

(keep in mind that $V_n$ is a function of time, but we have replaced $V_n(t)$ by $V_n$ for brevity in the equations). We will now assume that the diodes are identical, and we will look only at the quadratic (nonlinear) terms in the diode currents, $i_a = Kv_a^2$ and $i_b = -Kv_b^2$, 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 ($\omega_i = \omega_{\rm RF} - \omega_0$) terms:

$$\begin{aligned} i_a(t)|_{\rm LPF} &= \frac{K}{2}\Big[V^2 + (V_0+V_n)^2 - 2V(V_0+V_n)\sin\omega_i t\Big], \\ i_b(t)|_{\rm LPF} &= -\frac{K}{2}\Big[V^2 + (V_0+V_n)^2 + 2V(V_0+V_n)\sin\omega_i t\Big]. \end{aligned} \tag{9.9}$$

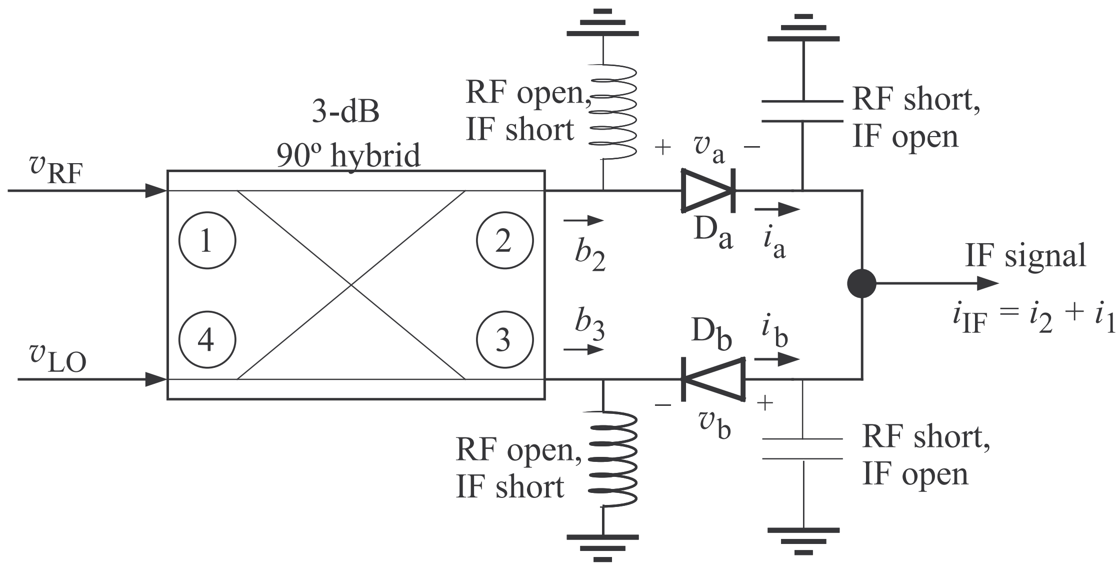



Figure 9.5: A balanced mixer circuit, using a 90° hybrid.

If we remove the DC components from only *one* of these diode currents ($i_a$, say), and keep only the most significant terms in $V$ and $V_n$, we would have

$$i_{a,\text{IF}}(t) \simeq KV_0 \left[V_n(t) - V \sin \omega_i t\right]$$

wherein the noise term is additive to the desired signal. This is essentially what happens in the single-ended mixer. If on the other hand we combine the two diode currents as indicated in the balanced mixer circuit, we find that the $V_0 V_n$ terms cancel, and the total output IF current is now

$$i_{\text{IF}}(t) = i_a + i_b = -2KV(V_0 + V_n) \sin \omega_i t \simeq -2KVV_0 \sin \omega_i t, \tag{9.10}$$

again using the fact that $V, V_n \ll V_0$ so that $VV_n \ll VV_0$. This means that the noise terms have been canceled to first order in the balanced mixer circuit, 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

$$L = \frac{P_{\text{IF}}}{P_{\text{RF}}}, \tag{9.11}$$

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 (see Chapter 8), 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 $\omega_i$ when mixed with an LO at $\omega_0$. Usually, only one of these RF frequencies is wanted, and the other one is called the *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 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. Draw diagrams of a single-ended and balanced mixer and explain the advantage of one over the other.

4. Which quantities are important for describing a mixer?

5. Why is it important for a mixer to have low added noise?

6. Write a summary of the application of baluns for balanced and double-balanced mixers, including an explanation as to why they are required. Find at least two references for your summary. The summary should not exceed one page in 12-point font, single spaced, including figures.

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?

## 9.4 Homework Problems

**1.** Consider a balanced mixer with a 90° hybrid, as shown in Fig. 9.5. The hybrid has the scattering matrix (6.11), where $A = B = 1/\sqrt{2}$. Assume the sources of $P_{\text{RF}}$ and $P_{\text{LO}}$ are matched. If $V_{\text{RF}}$ is the RF phasor input voltage, $V_{\text{LO}} = 0$ and $\rho$ is the reflection coefficient of each of the diodes (assumed to be identical—see Fig. 9.6), express the RF voltages reflected from the two diodes,

$$V_{Ra} = \rho V_a^i \qquad \text{and} \qquad V_{Rb} = \rho V_b^i,$$

in terms of $V_{\text{RF}}$. Write down the expressions for the two reflected waves, $V_R^{\text{LO}}$ and $V_R^{\text{RF}}$,


Figure 9.6: Reflected voltage waves from diodes.

when they combine at the RF and LO input ports of the mixer, assuming $|\rho| \ll 1$ (i. e., find $b_1$ and $b_4$). Is the input RF port well matched? Does the reflected RF signal affect the LO signal?

**2.** A diode can be used not only to detect the power of an unmodulated signal or as a mixer, but also to detect the modulation (or baseband) signal 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 $\omega_m$ is the low modulation frequency, $\omega$ is the high microwave frequency (usually called the carrier frequency), and $m$ is the modulation index (it describes how strongly the carrier is modulated).

**3.** Consider the diode equivalent circuit from Fig. 3.11 and use the element values given in problem 4 of Lecture 3. Input the element values for the biased diode (based on $I_0 = 60\,\mu\text{A}$) into your microwave design software. 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.

**4.** Suppose that a phase-modulated signal is applied to a Schottky diode. Find the diode current, if the PM signal can be written as

$$v(t) = \sqrt{2}V_0 \cos\left(\omega t + m_f \sin\omega_m t\right),$$

where $\omega_m$ is the modulation frequency, $\omega$ is the carrier frequency, and $m_f$ is the phase modulation index. Use the formulas

$$\cos\left(m_f \sin\theta\right) = J_0(m_f) + 2\sum_{n=1}^{\infty} J_{2n}(m_f)\cos 2n\theta$$

$$\sin\left(m_f \sin\theta\right) = 2\sum_{n=1}^{\infty} J_{2n+1}(m_f)\cos(2n+1)\theta$$

where the coefficients $J_k$ are called Bessel functions that are functions of $m_f$. Your expression for the diode current will contain an infinite number of harmonics—show only the first few of them.

# Lab 9: 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. 9.7.

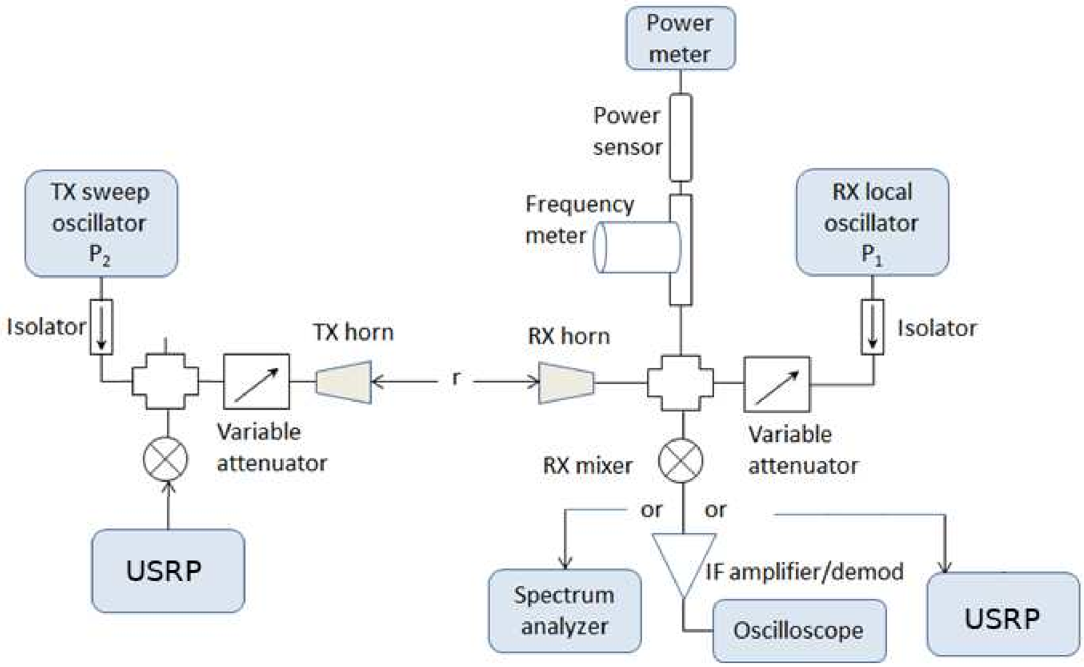



Figure 9.7: The experimental setup for the superheterodyne microwave link.

- The transmitter, on the right hand side, is a sweep oscillator.
- The sweeper is connected through an isolator and a variable attenuator to a standard gain horn antenna, which acts as a transmitting antenna. The wide-range attenuator means that the measurement has a very large dynamic range, i. e., that the ratio of the largest and smallest signal it can generate is very high.
- The receiver, on the right-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 (a diode detector to measure power may be substituted for these if desired).
- 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 narrowband IF amplifier/demodulator and the detected signal can be observed on a scope.

**Q0:** Over-exposure to microwaves can cause cataracts in the eye. The occupational safety limit in Canada is set at 25 mWcm$^{-2}$, rabbits have been reliably affected with 100 mWcm$^{-2}$, however one study with monkeys found no cataracts induced from 10hr/day of exposure to 150mWcm$^{-2}$ at 9.3

GHz over 3 months. Our sweeper can output up to 26 dBm, and these horns have around 10 dBi of gain. Does the maximum power density at the edge of the far-field pose any danger? What do you think right at the aperture?

## Part I: AM link setup

In the first part of the lab, you will look at an amplitude modulated (AM) link, with the two horn antennas set at a fixed distance at the edge of the far field. The transmit sweeper has the possibility of internal amplitude modulation with a 1-kHz square-wave. You turn this on with a button that has a little square-wave picture on it.

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 30MHz before detection.

- Set both attenuators to 0 dB.
- Temporarily turn the RF sweeper off **(just the output)** so you can measure just the LO
- Set the LO frequency to about 9.03 GHz and the frequency meter to something different
- Adjust the power of the LO sweeper P1 so that the power meter is reading -8dBm (or as high as possible) for the receiver attenuator set at 0 dB.
- Tune the frequency meter very slowly around 9.03 GHz, until the power meter dips by about 1 dB. This is your actual LO frequency. Write it down.
- Temporarily turn the LO output off, and turn the RF output back on.
- Set the output power of the transmitting sweeper so that -10dBm is incident at the power meter (and therefore also at the diode).
- Set the frequency meter down 30 MHz from the LO frequency. This is your desired RF frequency.
- Then set the frequency of the transmitting sweeper to about 9 GHz and tune it very slowly until the power meter dips by about 1 dB. Your RF should now be an accurate 30 MHz below your LO.

You could also adjust the frequency of the RF 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. Keep in mind that the output of the IF amp/demodulator is in fact the demodulated version of the amplified IF signal (the 1 kHz square wave) and not the 30MHz IF signal itself. A spectrum analyzer can also be used to help locate the frequencies of the two sweep oscillators.

Now turn on the square-wave modulation and make sure the output of the IF amp is connected to the oscilloscope. Adjust the attenuation on the IF amp so its needle is somewhere in the middle of the range.

**Q1:** What do you expect to observe on the oscilloscope (explain why)? Include a plot of the scope screen, including all the scales.

## Part II: LO and RF power dependence

Now you will measure the dependence of the IF output power on the LO power. Turn off the internal square wave modulation of the transmitter sweeper, disconnect the diode output from the IF amplifier input and read the power level of the 30 MHz IF signal from a spectrum analyzer using the Marker and Peak Search functions (a 1-2 GHz analyzer is sufficient for this purpose). Set the Frequency to 30 MHz, Span to 10 MHz, and Resolution Bandwidth to 10 kHz, for a good compromise between speed, bandwidth, and sensitivity.

**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?

Set the TX attenuator for a reading of -15dBm at the power meter, and measure how the output IF power changes (round to nearest 0.5dB) when you change the LO power by changing the attenuation on the variable attenuator on the receiver side. Fill in Table L9.1., with the last row containing the LO power that makes the IF roughly even with the noise-floor of the spectrum analyzer. Plot your result on a graph of IF power versus LO power. Comment if you think something unusual is happening.

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 9.1. Plot IF power versus LO power on a graph. Comment if you think something unusual is happening.

| RF Power (dBm or mW) | LO Power (dBm or mW) | IF Power (dBm or mW) |
|---|---|---|
| -15 dBm | -8 dBm - 0 dB = -8 dBm | |
| -15 dBm | -8 dBm - 3 dB = -11 dBm | |
| -15 dBm | -8 dBm - 5 dB = -13 dBm | |
| -15 dBm | -8 dBm - 10 dB = -18 dBm | |
| -15 dBm | -8 dBm - 15 dB = -23 dBm | |
| -15 dBm | -8 dBm - 20 dB = -28 dBm | |
| -15 dBm | -8 dBm - 25 dB = -33 dBm | |
| -15 dBm | -8 dBm - 30 dB = -38 dBm | |

Table 9.1: Table for determining the IF versus LO power dependence.

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 -13 dBm (5dB attenuation), and the RF at -10dBm at the power meter. Using the waveguide attenuator, reduce the RF power until the signal and the noise on the spectrum analyzer display have about the same amplitudes (this should be about -50dBm of RF power). Repeat for an LO power of -13dBm. Fill in Table L9.2 and plot the results.

**Q3:** Is there a difference in the behavior of the IF power in the two graphs? Why do you think this is so?

## Part III: LOS link attenuation

Now we will look at an AM microwave link. Re-connect the output of the diode detector to the input of the IF amplifier whose output is connected to the oscilloscope.

| LO Power (dBm or mW) | RF Power (dBm or mW) | IF Power (dBm or mW) |
|---|---|---|
| -13 dBm | -10 dBm - 0 dB = -10 dBm | |
| -13 dBm | -10 dBm - 3 dB = -13 dBm | |
| -13 dBm | -10 dBm - 5 dB = -15 dBm | |
| -13 dBm | -10 dBm - 10 dB = -20 dBm | |
| -13 dBm | -10 dBm - 15 dB = -25 dBm | |
| -13 dBm | -10 dBm - 20 dB = -30 dBm | |
| -13 dBm | -10 dBm - 25 dB = -35 dBm | |
| -13 dBm | -10 dBm - 30 dB = -40 dBm | |

Table 9.2: Table for determining the IF versus LO power dependence.

**Q4:** What is the minimum distance r between the horn apertures needed for far field operation at 9 GHz?

**Q5:** If you got -15 dBm of received RF power for a given TX setting in Part 2 with the horns 30 cm apart, and you need at least -60 dBm received to detect a signal, what is the maximum distance you could have between the horns for the same TX settings?

Make sure the transmit sweeper is still set for -10 dBm at the power meter, and then increase the attenuation of the transmitter attenuator until you observe the signal and the noise amplitudes on the scope as being equal.

**Q6:** How much attenuation did you end up with? What is the maximum value of the distance between the horns rmax for S/N=1, if the attenuator was reset to 0 dB?

**Q7:** Consider displacing the two horns at an angle, as shown in Fig. 9.8 (view looking from the top). The signal would bounce from the transmitting horn to the receiving horn using a metal reflecting sheet. With the previous settings, you will probably not see the signal on the oscilloscope. How much additional TX power would you need to get SNR ¿ 1, accounting only for increased free-space travel? Come back to this part and actually measure it after the horns have been turned for **Q13**.

## Part IV: Voice Link

Now that you have a working AM link measured on the oscilloscope, let us see if you can hear the 1-kHz tone. Take the BNC cable from the scope and plug it into the back of the audio amp which is connected to the speaker. You should hear a relatively high-pitched single tone. Turn the audio amplifier off if it works because nobody wants to hear that.

Next you will listen to a voice microwave radio link. We provide a tape player and a selection of awesome music that will provide you with some culture. Choose wisely.

Connect the audio source to the FM modulation input at the back of the transmitting sweeper through one 10 dB attenuator. Disable the internal 1-kHz modulation of the transmitting sweeper. Check that the sweeper frequencies are still set 30MHz apart (some thermal drifting may have occurred since the beginning of the lab period). You will be using the IF amplifier as a radio to receive the 30MHz mixer product. Turn the audio amplifier back on and enjoy the crystal clear serenade coming from the speakers.

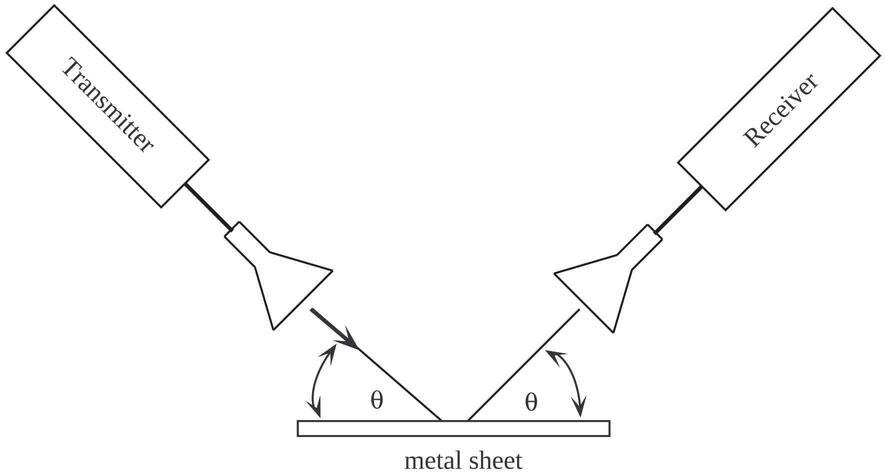



Figure 9.8: Position of horns for **Q7**.

**Q8:** What did you hear through the speakers? Try different levels of RF and LO attenuation. Does the best sound occur when the IF power is maximum?

**Q9:** It is a bit of a mystery why an AM demodulator can demodulate an FM signal. A way to think about this is to consider that the IF amplifier is a filter, and that by operating at the edge of the filter, the response changes quickly with frequency. You can test this by using the Vernier button on the sweeper tuning, which enables very fine frequency control. What happens to the quality of the sound you are receiving as you get to the edge of your IF amplifier frequency response? What happens if the IF frequency is centered directly in the middle of the IF amplifier passband (tune until the IF meter is maximized, you may have to increase the attenuation on it)?

## Part V: Digital modulation link

Now we will use a National Instruments Software Defined Radio (generic: SDR, or brand-name: USRP) to send a digitally-modulated signal over a wireless link and observe multipath fading effects on signal quality.

For this part, we will be using a Transmitted Carrier (TC) scheme, similar to conventional AM radio. This means the pure RF carrier is transmitted along with the signal, so that it can be used for downconversion, eliminating the need for an LO. Connect the TX SDR to the diode on the TX side. Turn the transmit sweeper power all the way up, to 26 dBm. Turn the LO **(output only)** off and connect the RX SDR to the diode on the RX side.

**Q10:** Explain why demodulation still works without an LO. Given this context, and the s-parameters below, why is port 1 of the TX magic tee fitted with a short?

$$[S] = \frac{-j}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{bmatrix} \tag{9.12}$$

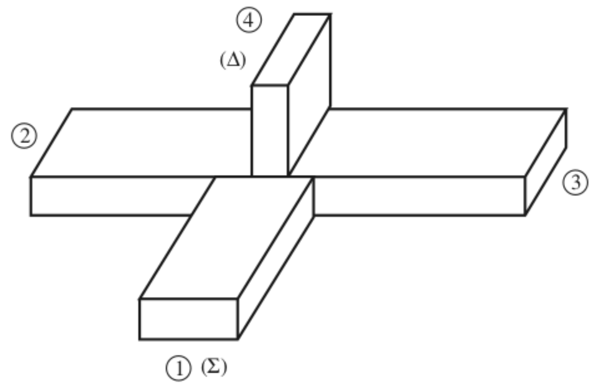



Figure 9.9: S-parameters for waveguide Magic-T, reproduced from Pozar.

In order to drive the SDRs, we will use QAM example projects for LabView Communications. If they are not already open on the computer, look for two shortcuts on the desktop called QAM TX and QAM RX. The icons look like notebooks. If you get error pop-ups, click continue and check if the program is still running. Press the green play button if it has stopped.

**Q11:** Set the frequency to 500MHz with a 16-QAM signal modulation, and the gain to 100 dB (it wont actually go this high, but you will get max power out) for both TX and RX. Press the green play button on each program to enable the SDRs. You will see the symbol constellation in a polar plot at digital baseband at both transmit and receive sides (note the TX plot is static, but the RX plot should be changing). Point the two copolarized antennas towards each other. Record the signals on both ends. Insert the polarizer between the antennas and rotate it. What happens to the receive signal?

**Q12:** With both RX and TX set to 16-QAM, make sure the received constellation is good. Then, increase TX attenuation, or move the antennas apart until the received signal is unusable. Switch to 4-QAM and repeat. How much more attenuation(or distance) can the 4-QAM signal sustain? Why do you think 4G/wifi is slower when the signal is weaker?

Next, misalign the horns like in Fig. 9.8. Be extremely careful with this process. Make sure the stands continue to support the waveguides, and dont let the long low pass filter dangle from the transmit mixer diode. You will likely have to turn/shift at least one of the sweepers to keep from stressing the ends of the coaxial cables. Take your time.

**Q13:** Hold the metallized corner reflector in front of the antennas. There is some signal that is propagating with line-of-sight, and some of the signal is bouncing off the reflection, providing a fading channel. Move the reflector and explain quantitatively what happened, including any plots you feel are relevant.

Now stop both LabView programs and disconnect the SDRs

**Bonus:** Change the frequency and perform measurements to confirm your conclusions from Q13. You can also change the multipath channel. Clearly write down what you did for extra credit.

# Chapter 10

# Radar Fundamentals

Radar can be viewed as a type of communication link, in which the transmitting antenna sends a wave, which eventually reflects off some object, called a "target" or "scatterer". The currents induced on the target re-radiate and this signal is picked up by a receiving antenna. Most commonly, the radar uses the same antenna for transmit and receive; this is called a monostatic radar and is shown in Fig. 10.1(a). For very high-power radars, different antennas located near each other may be used for transmit and receive, and this is called a bi-static radar. In either case, by analyzing the 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-borne (satellite) SAR radars 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 monostatic radar is as follows. We will assume for simplicity that the antenna and target are lossless (100% efficient). The radar transmitter radiates a wave with power $P_T$ toward a target. At the target, the incident power density is

$$S_{\text{inc}}(\theta, \phi) = P_T \frac{D(\theta, \phi)}{4\pi r^2} \tag{10.1}$$

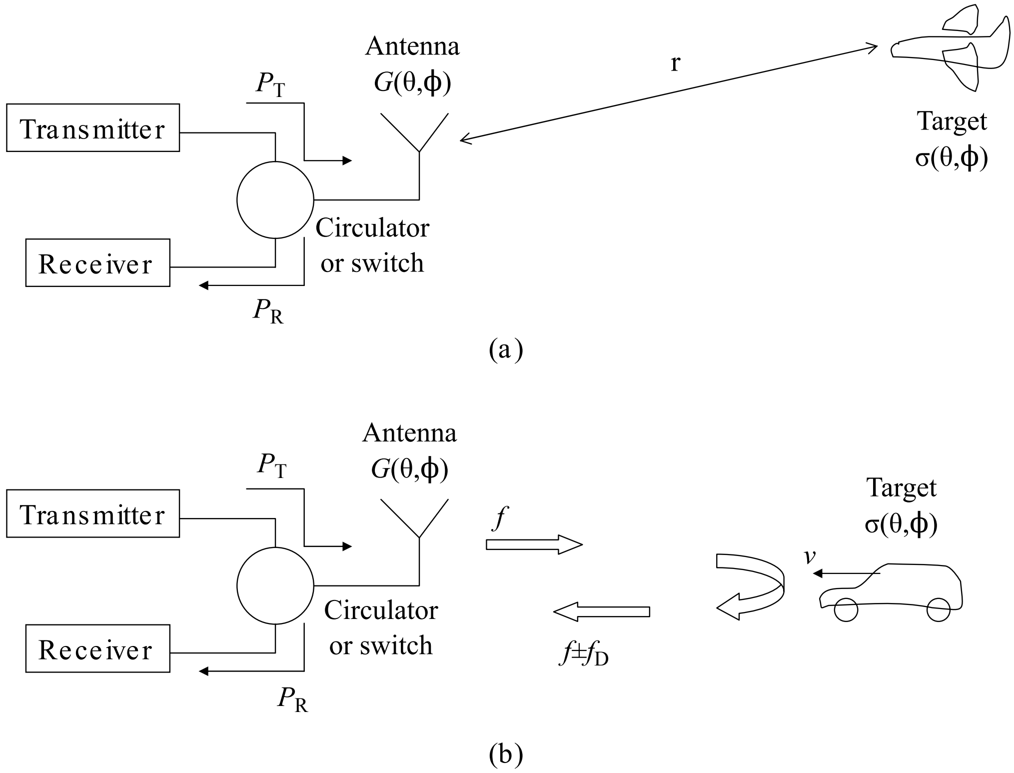
where $r$ is the distance to the target, and $D(\theta, \phi)$ is the radar antenna's directivity in the direction $(\theta, \phi)$ of the target. The target intercepts a portion of the incident power density

$$P_{\text{target}} = \sigma_{\text{tot}}(\theta, \phi) S_{\text{inc}} \tag{10.2}$$

and completely re-radiates it (acting simultaneously as a receiving antenna and a transmitting antenna). The quantity $\sigma_{\text{tot}}$ is called the *total scattering cross section* and has units of area ($\text{m}^2$). The target, as a transmitting antenna, has a directivity $D_{\text{target}}(\theta, \phi)$, so it radiates a reflected power density back towards the radar antenna of

$$S_r = P_{\text{target}} \frac{D_{\text{target}}(\theta, \phi)}{4\pi r^2} = P_T \frac{D(\theta, \phi)}{4\pi r^2} \frac{D_{\text{target}}(\theta, \phi) \sigma_{\text{tot}}(\theta, \phi)}{4\pi r^2} = P_T \frac{D(\theta, \phi)}{4\pi r^2} \frac{\sigma_b(\theta, \phi)}{4\pi r^2} \tag{10.3}$$

where $\sigma_b(\theta, \phi) = D_{\text{target}}(\theta, \phi) \sigma_{\text{tot}}(\theta, \phi)$ is called the *backscattering radar cross section* of the target, a measure of the amount of re-radiation from the target. Finally, the scattered power received by the




[ht]

Figure 10.1: (a) Basic monostatic radar operation. (b) Basic Doppler radar schematic.

radar antenna can be written using the Friis formula and (7.41) as

$$P_R = S_r A(\theta,\phi) = S_r \frac{\lambda^2 D(\theta,\phi)}{4\pi} \tag{10.4}$$

or

$$P_R = P_T \frac{\lambda^2 D^2(\theta,\phi)\sigma_b(\theta,\phi)}{(4\pi)^3 r^4} \tag{10.5}$$

which is known as the *radar equation*. 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 is 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 FM-CW 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. Commercial radar applications include: 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. The reason for the Doppler shift is as follows. Imagine a plane wave of frequency $f$ propagating in free space with velocity $c$. Think of a set of plane wavefronts separated by distances equal to the wavelength $\lambda = c/f$. A fixed observer sees these wavefronts arrive at time intervals $T = \lambda/c = 1/f$ and therefore interprets the wave as having a frequency $f$, as expected. Now, however, suppose the observer is moving towards the source of the plane wave with a velocity $v$. To this observer, the wavefronts appear to arrive more frequently, with a modified period

$$T' = \frac{\lambda}{c+v} \tag{10.6}$$

and thus this observer interprets the wave as having a frequency

$$f' = \frac{1}{T'} = f\left(1 + \frac{v}{c}\right) > f \tag{10.7}$$

instead of $f$. If $v < 0$ (the observer is moving away from the source), the apparent frequency is reduced ($f' < f$). According to the principle of relativity, if we fix the observer and move the source of the plane wave towards the observer with velocity $v$, the same effect will be observed. Note that this explanation is a simplified one, valid in the approximation $v/c \ll 1$. The exact theory requires the use of relativity theory.

Now, if a stationary antenna radiates a transmitted signal that hits a moving target, that incident signal appears to the scatterer to have a modified frequency $f'$, and so the induced currents on the scatterer will also have that frequency. This moving target then re-radiates a signal which is received by the stationary antenna as having the further modified frequency

$$f'' = f'\left(1 + \frac{v}{c}\right) \simeq f\left(1 + 2\frac{v}{c}\right) \tag{10.8}$$

if $v \ll c$ (relativistic effects have been neglected throughout this treatment). Thus, by measuring the Doppler shift, the speed of the target can be detected with a radar system, as shown in Fig. 10.1(b). A wave of frequency $f$ is transmitted, and the wave reflected from a target moving at velocity $v$ towards the transmitter will have a frequency

$$f + f_D \qquad \text{where} \quad f_D = \frac{2vf}{c} \tag{10.9}$$

If the target is moving towards the radar, the frequency is shifted up, and if it is moving away, it is shifted down. Note that it is only the velocity of the object in the direction *normal* to the observer that contributes to the Doppler shift; tangential motion leaves the frequency of the signal unaffected. The change in frequency can be measured by feeding the incident and reflected signals to a mixer to obtain $f_D$ as a mixer product.

## 10.3 FM-CW Ranging Radar

In one type of ranging radar, the frequency of the transmitter is linearly frequency modulated (FM), as shown in Fig. 10.2. The transmitter signal frequency is varied linearly between a minimum frequency $f_1$

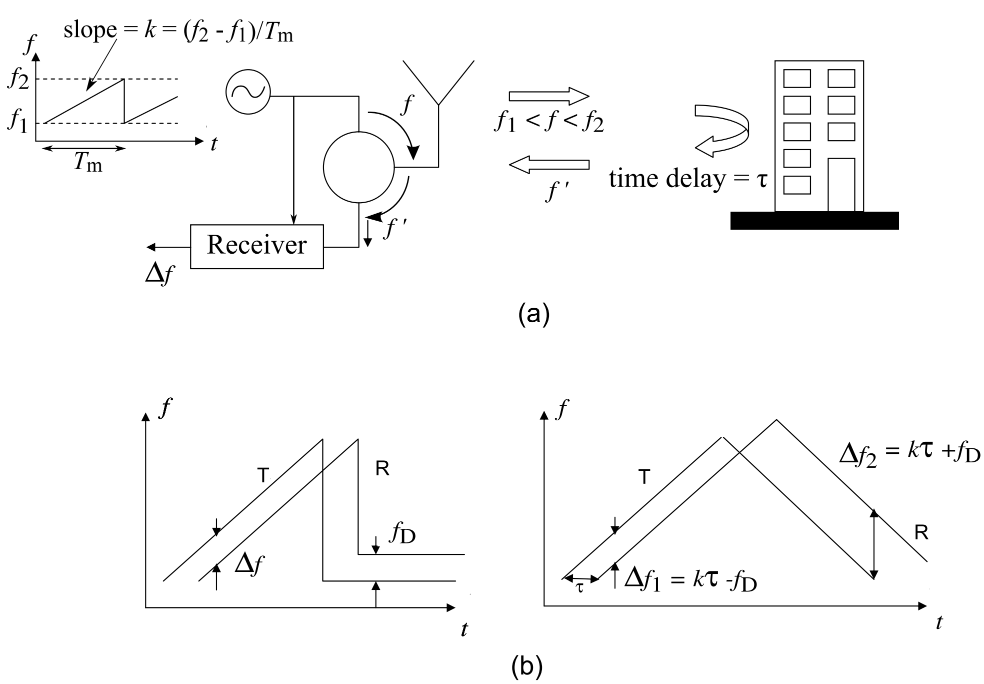



Figure 10.2: (a) Operational principle of FM-CW ranging radar. (b) Correcting for Doppler shift in FM-CW radar: left – transmitting a fixed frequency for some time after the FM part, and right – triangular frequency modulation.

and a maximum frequency $f_2$ as shown in Fig. 10.2(a). The difference $f_2 - f_1$ is called the *modulation bandwidth*. The ramp (or modulation) period $T_\mathrm{m}$ is the time interval in which the frequency variation is repeated. If a frequency $f$ is transmitted at a certain instant of time, then by the time this wave returns to the radar, the transmitted radar frequency has changed to $f'$, 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 $\Delta f$, often referred to as the "beat frequency", is directly related to the distance from the target. An obvious limitation of this system is that there should be at least one complete period of the beat frequency during the ramp period, or in other words $\Delta f T_m > 1$. Otherwise, there would not be sufficient information to determine $\Delta f$ accurately. Since the radar signal is being transmitted continuously, this type of ranging radar is called FM-CW, which stands for Frequency Modulated-Continuous Wave.

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. 10.2(a), enabling measurement of the Doppler frequency 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 right-hand side of Fig. 10.2(b), 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 FM-CW 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 form that the radar equation (10.5) will take 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 backscattering radar cross section of $\sigma_b = 100\lambda^2$, 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 the speed limit 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 10 GHz center-frequency FM-CW ranging-only radar, calculate the required linear frequency modulation bandwidth of the receiver 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 $f_2 - f_1$ is at most 10% (i. e., only frequencies in the range 10 GHz $\pm$5% are detectable). Note that there is more than one solution given the constraints, and you should state whatever choices you make to achieve a single solution.

**5.** For a 10 GHz FM-CW ranging radar with a waveform as shown on the right-hand side in Fig. 10.2(b), with the same slope as in problem 4, determine the difference in the beat frequency that would be recorded if the target were moving towards the transmitter at 100 km/hour.

**6.** An object of mass $m$ slides down the frictionless surface of an inclined wedge of base length $L$ and height $H$ as shown in Fig. **??**. Find an expression for the velocity of this object as a function of its horizontal distance $x$ from the top of the wedge if it has zero velocity at $x = 0$ and moves only under the force of gravity? How does this result change if the object is a cylinder of radius $R$ that rolls down the wedge without slipping (Hint: the kinetic energy now has a part due to the rotational motion)?

**7.** A pendulum consists of a mass $m$ suspended from a string of length $l$ and oscillating back and forth. Review your basic physics and derive the formula for the oscillation period of this pendulum, and the maximum velocity attained by the pendulum during its oscillation. What approximations are used in this derivation?

**8.** A Salisbury screen can be used to reduce the RCS of a target. It consists of a planar surface which is a thin film of conductance $G = Y$, where $Y = 1/Z$ is the admittance of free space (1/377 S). Explain what is meant by the film having a conductance. A reflector is placed parallel to the thin film at a distance of a quarter of a free-space wavelength behind it at a given operating frequency. A plane wave at this frequency is normally incident on the film/reflector. Show that this wave will be completely absorbed by the screen. However, in general the absorbed power will depend on frequency. Find the bandwidth as a percentage of the nominal operating frequency. The bandwidth is given by points where the reflected power is one tenth of the incident power.

**9.** A common feature of practical radar systems is the comparison of the return signal (which is phase delayed and possibly differs in frequency) to the transmitted signal. The basic idea is to effect a mixing of these two signals so that the difference frequency can be observed, and if necessary, the phase shift as well. In Fig. 10.3 we show a very simple configuration. There is no "closed loop" in this arrangement, and relatively few parts are needed. Describe in detail the paths of the transmitted and received signals in the setup shown in Fig. 10.3. In view of problem 14 of Lecture 6, how much of the transmitted and received signals are delivered to the mixer diode? Comment on the possible advantages and disadvantages of this setup.

**10.** In Fig. 10.4, we have replaced the T-junction in the setup of problem 9 above with a pair of magic tee hybrid couplers in order to better separate the transmitted and received signals before combining them at the mixer. This setup features a sort of closed-loop signal path. Describe in detail the paths of the transmitted and received signals in this arrangement. Comment on the possible advantages and disadvantages of this setup as compared to that of Fig. 10.3.

**11.** A further variation of the monostatic radar circuits considered in problems 9 and 10 above is shown in Fig. 10.5. The function of the hybrid couplers in Fig. 10.4 has been replaced by that of a circulator in order to prevent the strong direct signal from overwhelming the return signal at the mixer. The portion of the transmitted signal needed at the mixer is now obtained from the fact that the circulator is not ideal, and that some small fraction of signal incident at port 1 of the circulator will emerge at port 3. Describe in detail the paths of the transmitted and received signals in this arrangement. Comment on the possible advantages and disadvantages of this setup as compared to those of Figs. 10.3 and 10.4.

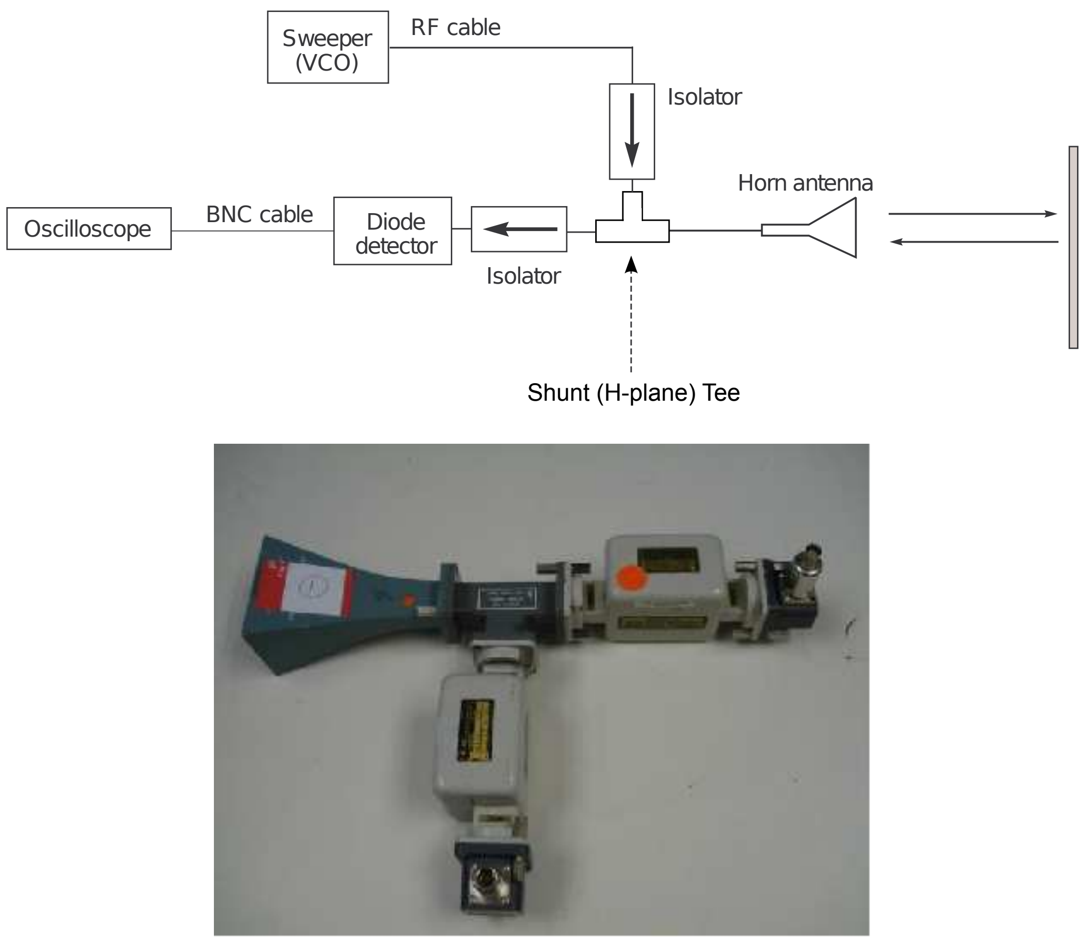



Figure 10.3: Setup for simple monostatic radar. Waveguide components include: horn antenna, ferrite isolators, H-plane (shunt) T-junction, coaxial-to-waveguide adapters and a diode detector (mixer). A CW microwave oscillator is used as the signal source, and an oscilloscope for viewing the output of the mixer.

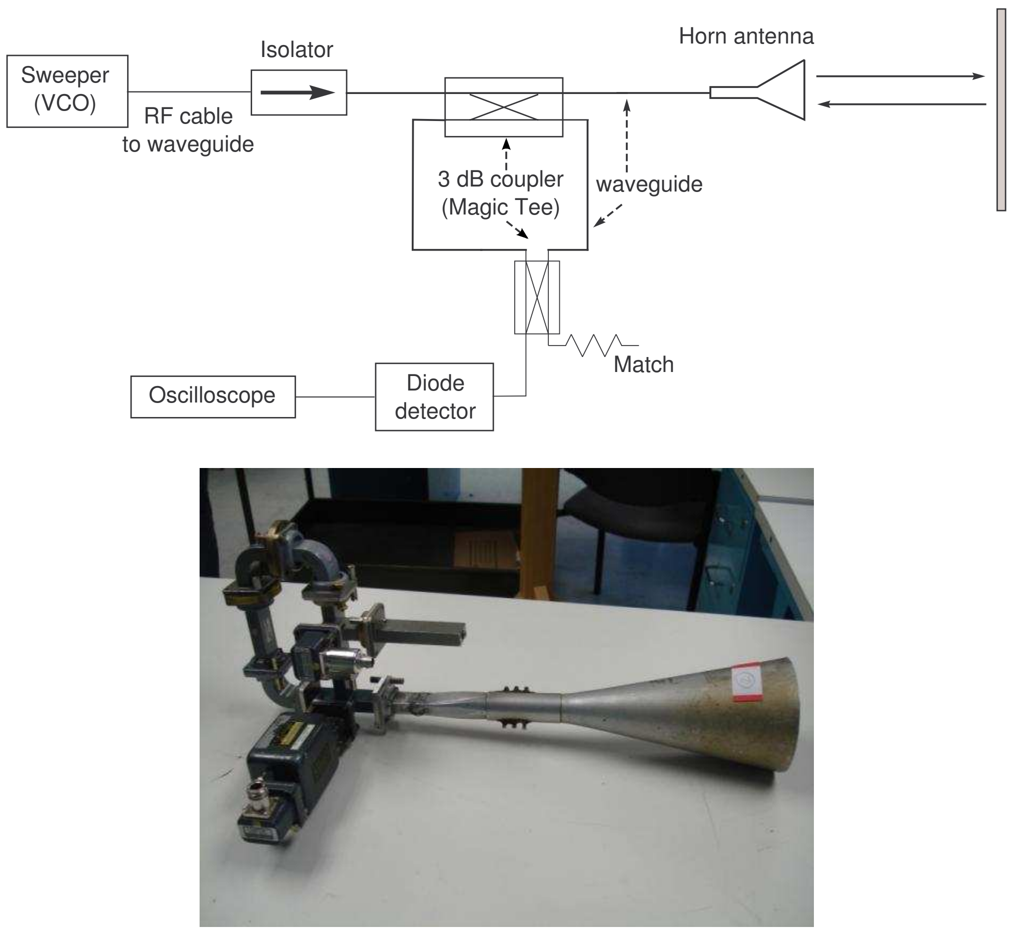



Figure 10.4: Setup for refined version of monostatic radar. Waveguide components include: horn antenna, ferrite isolator, magic tees, coaxial-to-waveguide adapters and a diode detector (mixer).

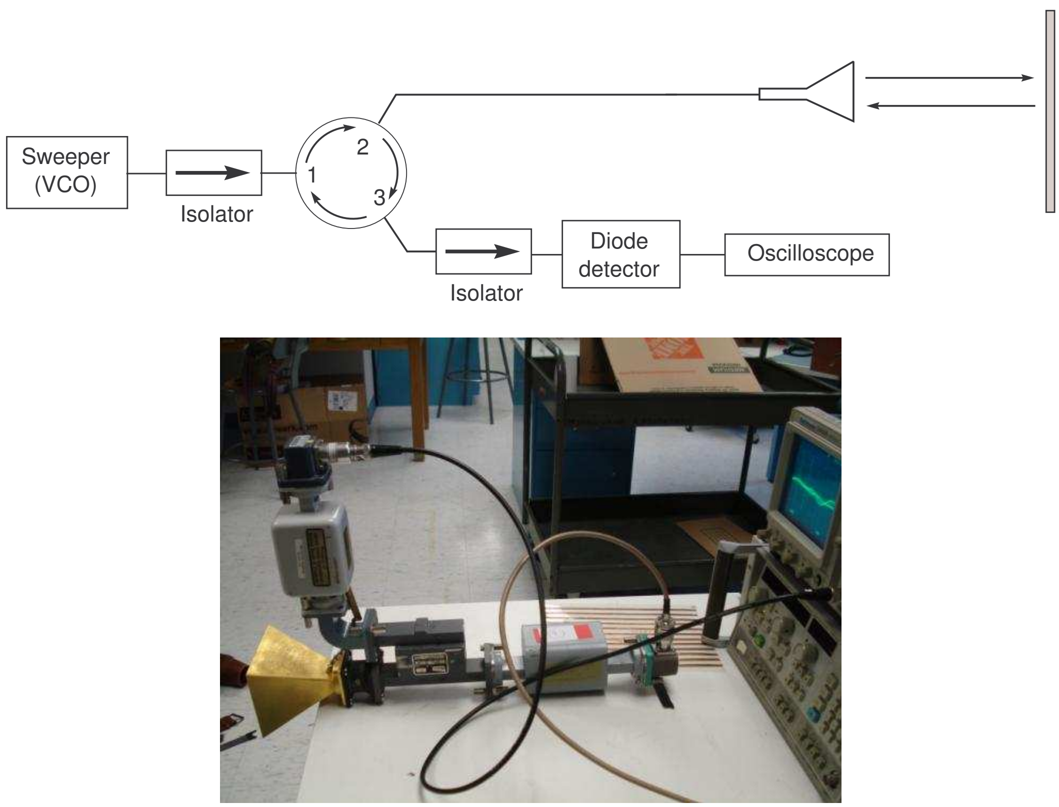



Figure 10.5: Setup for the monostatic radar using a circulator but no feedback loop. Waveguide components include: horn antenna, circulator, ferrite isolator, coaxial-to-waveguide adapters and a diode detector (mixer).

## Lab 10: Doppler and FM-CW Ranging Radar

In the lab, you will put characterize a Continuous-Wave Frequency- Modulation (CW-FM) radar at X-band, and put together a Doppler radar using WR-90 waveguide components. This simple Doppler radar is 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. 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: CW-FM Ranging Radar

**Q1:** Explain the functionality of all waveguide components used in the setup showed in Fig. 10.6. 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?

**Q2:** 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?

**Q3:** 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. 10.6(a). 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.

**Q4:** 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?

**Q5:** 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.

**Q6:** 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.

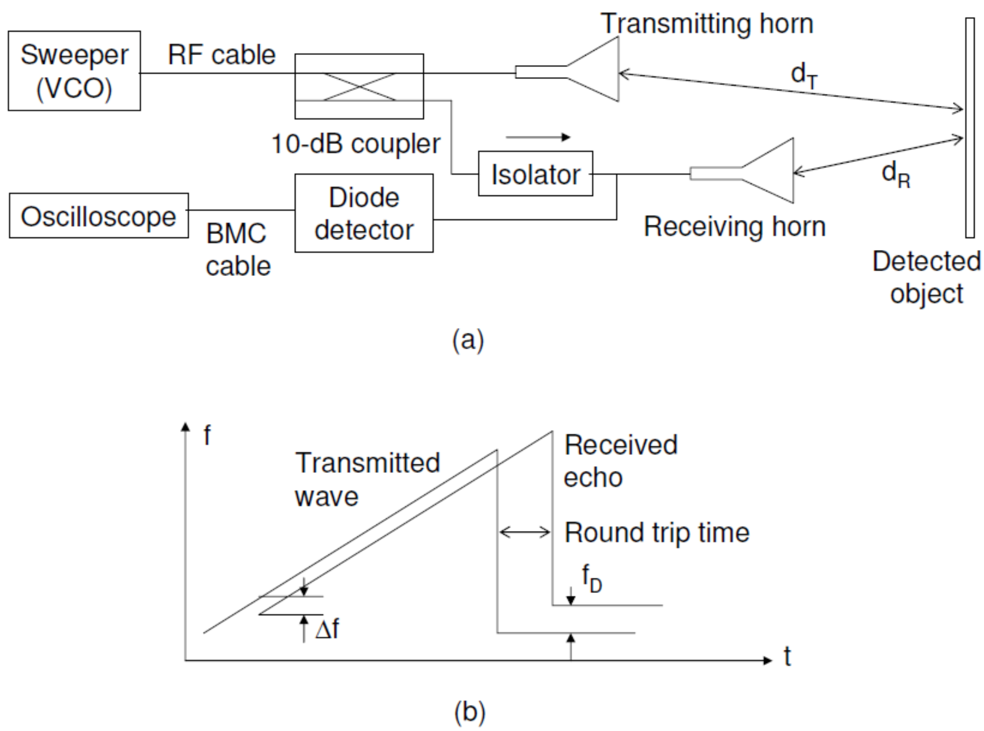



Figure 10.6: (a) Setup for CW-FM radar experiment. The RF signal is generated with a HP8350B (or equivalent) sweeper in internal controlled sweep mode (Sweep Trigger - INT, Sweep - INT). Use a sweep period of 0.020 seconds, a start frequency of 8 GHz and a stop frequency of 9 GHz. Waveguide components 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.

**Q7:** 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.

**Q8:** 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. 10.7. 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.

**Q9:** 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?

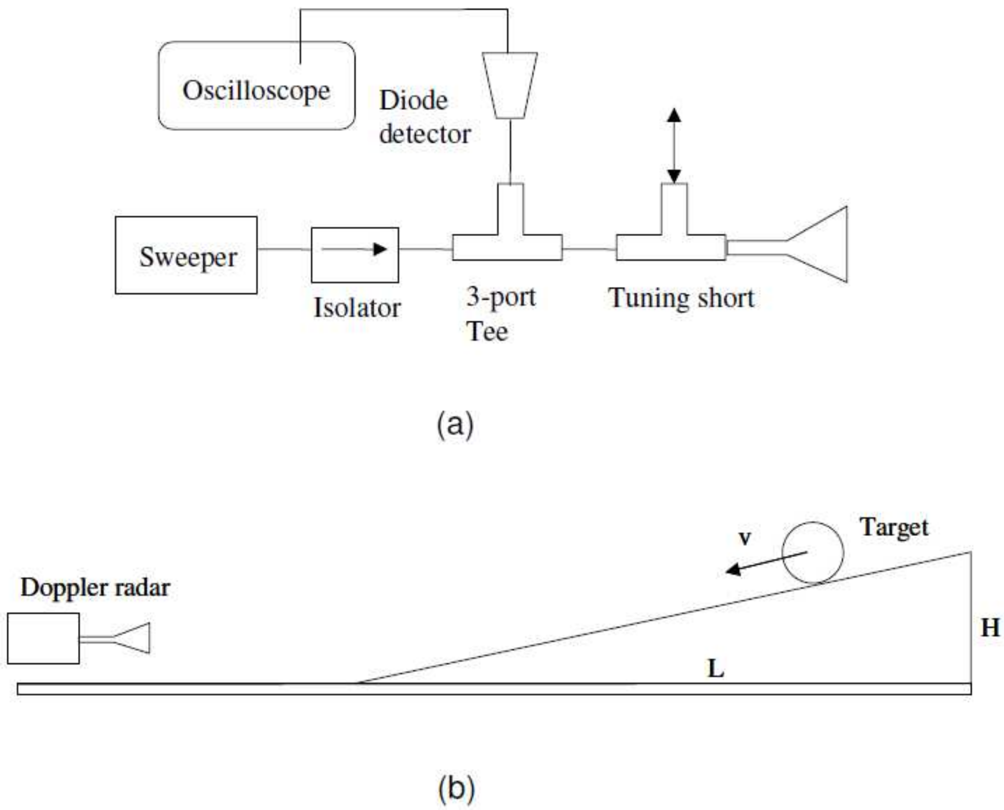



Figure 10.7: (a) Setup for Doppler radar experiment. Use an X-band sweeper, waveguide isolator, waveguide 3-port Tee, waveguide shorted stub tuner, waveguide mounted diode detector (mixer), horn antenna and oscilloscope for detecting the Doppler frequency shift. (b) Inclined plane and cylindrical target. What is the velocity of the target at the foot of the hill?

## Part II. Doppler Radar

**Q10:** The block diagram of the setup is shown in Fig. 10.7(a). 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?

**Q11:** 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. ??(b). Measure the Doppler frequency. 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.

**Q12:** Measure the velocity of a different target that the instructor or TA will give you with a 8-GHz and 12-GHz (at edges of X-band) Doppler radar. What can you conclude?

# Appendix A

# Capturing data from the Agilent Vector Network Analyzers

In this appendix, we describe how to capture data from an Agilent vector network analyzer (VNA) to a file on a computer. The specific instructions here are for the 8753ES VNA, but are similar for other VNAs.

## Method 1: Instrument Data Capture (icon on desktop)

Go to "Instrument" and choose 8753ES. You will get a plot with poor colors, but you can save it immediately as a .png file.

## Method 2: Intuilink VNA Excel Add-In (icon on desktop)

- Open a Microsoft Excel spreadsheet.
- Go to "Add in.
- Click on the "Agilent VNA" icon.
- Choose "model 8753ES" and hit "OK" (this may already have been chosen).
- Click the "Connect to VNA" icon.
- Under the GPIB address, click on "Identify instrument". VNA will show up in the instrument box on the right.
- Click on "Connect". You should see the message "connected" shortly.
- Close that window (or you can keep it in the background to check connection status).
- Click the "Get Data" icon (black with yellow trace) when you want to acquire data.
- Check the boxes for the things you want plotted.
- Click on "Get Data", and a spreadsheet will be opened with all your data and a graph.

Note: You can also export the data as a .sNp file and then plot it in, say, AWR Microwave Office.

# Appendix B

# Capturing data from the HP/Agilent Spectrum Analyzers

In this appendix, we describe how to capture data from an HP/Agilent spectrum analyzer to a file on a computer. The specific instructions here are for the 8590 series analyzers, but are similar for other instruments.

We will use a public domain program by KE5FX (John Miles) called Spectrum Surveillance Monitor (SSM). Before opening the program, you must connect the spectrum analyzer to the GPIB bus that is attached to a computer with SSM installed, and then determine the analyzer's GPIB address. After the analyzer has been powered on, choose

`CONFIG` → `ANALYZER ADDRESS`

to get the GPIB address of the analyzer. If this address conflicts with that of another instrument already connected to the GPIB bus, change the address to remove the conflict.

Now open SSM, and from its menu choose

`Acquire` → `Acquire data from supported device at address` $n$,

where $n$ is the GPIB address of the analyzer that you found above. The program window will display the measured data from the spectrum analyzer. To capture numerical data from the program, from its menu choose

`File` → `Export to CSV file`

to obtain a spreadsheet file of the amplitude in dBm versus frequency in Hz. To obtain a screen capture (bitmap) do

`File` → `Save ... screen shot`

and choose the desired file format. Other settings in the program are fairly straightforward and are found in the menu.

# Appendix C

# Capturing data from circuit modeling software

Circuit, RF and microwave modeling software can not only plot data versus time and frequency in the programs themselves, but also allow the plotted data to be exported so they can be plotted in other programs, manipulated by computations, etc. In this appendix, we indicate how various of these programs allow their data to be exported, since this is not always obvious from the documentation. Once exported, the data can then be pasted into a text file if necessary, or directly into a program like EXCEL, Matlab, etc. for editing and plotting in any way you wish.

- *PSPICE from OrCAD/Cadence (formerly Microsim)*: After running a simulation, run Probe to display the desired voltages, currents, etc. 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.
- *LTSpice IV from Linear Technology Corporation*: After plotting the data, from the main menu choose File → Export to create a text file of the desired trace(s) in the plot.
- *AWR Design Environment (Microwave Office)*: After graphing the data, select a graph and from the main menu, choose Graph → Export Trace Data, to create a tab-delimited text file of the data in the graph.
- *Qucs*: After graphing the data, right click on the desired trace on the graph and select "Export to CSV", to create a CSV file of the data in that trace.

# Appendix D

# Notes on SPICE Usage

Here is some information about using SPICE (a program for circuit modeling and analysis), and especially LTspice, available free of charge at

http://www.linear.com/designtools/software/

Other freely available implementations of SPICE can also be found.

Many of you will have been introduced to SPICE in your circuits or electronics classes, or possibly from a specialized book such as *SPICE: A Guide to Circuit Simulation and Analysis Using PSpice* (Third Edition, 1995) by P. W. Tuinenga. Be aware that such sources often only give the very basics of SPICE syntax, especially regarding transmission lines and controlled sources. For a complete description of all forms of SPICE statements supported by any given implementation of SPICE, see the complete reference, which is generally included with the software. We present here some less familiar element statements that are particularly relevant to our needs. You should examine the complete reference for other questions about SPICE syntax and usage that are not addressed here. Note in particular that some implementations of SPICE offer nonstandard features such as behavioral modeling for the use of which the specific software documentation must be consulted. Features such as these referred to below are those of PSPICE, but are similar for LTspice (check the documentation).

The following provides the command syntax formats. < > indicates a required item in a command line. For example, <model name> in a command line means that the model name parameter is required. < >* indicates that the item shown in italics must occur one or more times in the command line. [ ] indicates an optional item. [ ]* indicates that there is zero or more occurrences of the specified subject. < | > means that you must specify one of the given choices. [ | ] means that you may specify zero or one of the given choices. The symbol + at the beginning of a line indicates a continuation of a statement from the previous line.

## D.1 Transmission Lines

An ideal (lossless) transmission line is characterized by two ports (terminal pairs) which can be thought of as the input and output sides of the transmission line section. Port A's (+) and (−) nodes are 1 and 2, and port B's (+) and (−) nodes are 3 and 4, respectively. The element statement has several variants:

```
T<name> <A port (+) node> <A port (-) node>
+ <B port (+) node> <B port (-) node>
+ [model name]
+ Z0=<value> [TD=<value>] [F=<value> [NL=<value>]]
+ IC= <near voltage> <near current> <far voltage> <far current>
```

*Comments:*

For the ideal line, IC sets the initial guess for the voltage or current across the ports. The value <near voltage> is the voltage across A(+) and A(−) and <far voltage> is the voltage across B(+) and B(−) . The <near current> is the current through A(+) and A(−) and the <far current> is the current through B(+) and B(−). For the ideal case, Z0 is the characteristic impedance. The transmission line's length can be specified either by TD, a delay in seconds, or by F and NL, a frequency and a relative wavelength at F. NL has a default value of 0.25 (F is the quarter-wave frequency). Although TD and F are both shown as optional, one of the two must be specified. Both Z0 (Z-zero) and ZO (Z-O) are accepted by the simulator.

If you are using Tuinenga's SPICE book for information about the transmission line element statement in your SPICE files, beware of the following inaccuracies on p. 269:

1) The meaning of the parameter NL in the ideal transmission line element statement is not "wavelength" as stated, or even "relative wavelength", but the relative length of the transmission line section *in* wavelengths. Thus, a section of line whose length is $0.4\lambda$ would have NL=0.4.

2) Two of the examples he gives use frequencies such as F=2.25MHz. Note that, according to the conventions of SPICE, this would mean 2.25 millihertz, not 2.25 megahertz!

Note that the nodes of the A port and those of the B port are considered to be floating with respect to each other, unless an external circuit connection is provided. In normal practice, this means that you should generally connect each of the (−) port terminals to ground in order to get SPICE to run without errors.

## D.2 Voltage-controlled sources

The voltage-controlled voltage source (E) and the voltage-controlled current source (G) devices have the same syntax. For a voltage-controlled current source just substitute G for E. G generates a current, whereas E generates a voltage. There are seven variants of the element statement for this type of source:

```
E<name> <(+) node> <(-) node>
+ <(+) controlling node> <(-) controlling node> <gain>

E<name> <(+) node> <(-) node> POLY(<value>)
+ < <(+) controlling node> <(-) controlling node> >*
+ < <polynomial coefficient value> >*

E<name> <(+) <node> <(-) node> VALUE = { <expression> }

E<name> <(+) <node> <(-) node> TABLE { <expression> } =
+ < <input value>,<output value> >*

E<name> <(+) node> <(-) node> LAPLACE { <expression> } =
+ { <transform> }

E<name> <(+) node> <(-) node> FREQ { <expression> } = [KEYWORD]
+ < <frequency value>,<magnitude value>,<phase value> >*
+ [DELAY = <delay value>]

E<name> <(+) node> <(-) node> CHEBYSHEV { <expression> } =
+ <[LP] [HP] [BP] [BR]>,<cutoff frequencies>*,<attenuation>*
```

*Keywords:*

`POLY(<value>)`: Specifies the number of dimensions of the polynomial. The number of pairs of controlling nodes must be equal to the number of dimensions.

(+) and (−) nodes: Output nodes. Positive current flows from the (+) node through the source to the (−) node.

The `<(+) controlling node>` and `<(-) controlling node>` are in pairs and define a set of controlling voltages. A particular node can appear more than once, and the output and controlling nodes need not be different. The `TABLE` form has a maximum size of 2048 input/output value pairs.

`FREQ`: If a `DELAY` value is specified, the simulator modifies the phases in the `FREQ` table to incorporate the specified delay value. This is useful for cases of tables which the simulator identifies as being noncausal. When this occurs, the simulator provides a delay value necessary to make the table causal. The new syntax allows this value to be specified in subsequent simulation runs, without requiring the user to modify the table. If a `KEYWORD` is specified for `FREQ` tables, it alters the values in the table. The `KEYWORD` can be one of the following:

- `MAG` causes magnitude of frequency response to be interpreted as a raw value instead of dB.
- `DB` causes magnitude to be interpreted as dB (the default).
- `RAD` causes phase to be interpreted in radians.
- `DEG` causes phase to be interpreted in degrees (the default).
- `R_I` causes magnitude and phase values to be interpreted as real and imaginary magnitudes.

*Examples:*

```
EBUFF 1 2 10 11 1.0
EAMP 13 0 POLY(1) 26 0 0 500
ENONLIN 100 101 POLY(2) 3 0 4 0 0.0 13.6 0.2 0.005
ESQROOT 5 0 VALUE = {5V*SQRT(V(3,2))}
ET2 2 0 TABLE {V(ANODE,CATHODE)} = (0,0) (30,1)
ERC 5 0 LAPLACE {V(10)} = {1/(1+.001*s)}
ELOWPASS 5 0 FREQ {V(10)}=(0,0,0)(5kHz, 0,0)(6kHz -60, 0) DELAY=3.2ms
ELOWPASS 5 0 CHEBYSHEV {V(10)} = LP 800 1.2K .1dB 50dB

GBUFF 1 2 10 11 1.0
GAMP 13 0 POLY(1) 26 0 0 500
GNONLIN 100 101 POLY(2) 3 0 4 0 0.0 13.6 0.2 0.005
GPSK 11 6 VALUE = {5MA*SIN(6.28*10kHz*TIME+V(3))}
GT ANODE CATHODE VALUE = {200E-6*PWR(V(1)*V(2),1.5)}
GLOSSY 5 0 LAPLACE {V(10)} = {exp(-sqrt(C*s*(R+L*s)))}
```

*Comments:*

The first form and the first two examples of `E`-elements apply to the linear case; the second form and the third example are for the nonlinear case. The last five forms and examples of `E`-elements are analog behavioral models (ABM) that can have an expression, a look up table, a Laplace transform expression, a frequency response, or a filter to describe the behavior of the device. Refer to your SPICE users guide for more information on analog behavioral modeling.

Chebyshev filters have two attenuation values, given in dB, which specify the pass band ripple and the stop band attenuation. They can be given in either order, but must appear after all of the cutoff frequencies have been given. Low pass (`LP`) and high pass (`HP`) have two cutoff frequencies, specifying the pass band and stop band edges, while band pass (`BP`) and band reject (`BR`) filters have four. Again, these can be given in any order.

You can get a list of the filter Laplace coefficients for each stage by enabling the `LIST` option in the Simulation Settings dialog box. (Click the Options tab, then select the Output file Category and select Device Summary.) The output is written to the .out file after the simulation is complete.

For the linear case, there are two controlling nodes and these are followed by the gain. For all cases, including the nonlinear case (`POLY`), refer to your SPICE users guide.

Expressions cannot be used for linear and polynomial coefficient values in a voltage-controlled voltage source device statement.

## D.3 Current-controlled sources

The Current-Controlled Current Source (`F`) and the Current-Controlled Voltage Source (`H`) devices have the same syntax as do the voltage-controlled sources. For a Current-Controlled Voltage Source just substitute an `H` for the `F`. The `H` device generates a voltage, whereas the `F` device generates a current. There are two variants of these statements:

```
F<name> <(+) node> <(-) node>
+ <controlling V device name> <gain>

F<name> <(+) node> <(-) node> POLY(<value>)
+ <controlling V device name>*
+ < <polynomial coefficient value> >*
```

*Arguments and options:*

(+) and (−): Output nodes. A positive current flows from the (+) node through the source to the (−) node. The current through the controlling voltage source determines the output current. The controlling source must be an independent voltage source (`V` device), although it need not have a zero DC value.

`POLY(<value>)`: Specifies the number of dimensions of the polynomial. The number of controlling voltage sources must be equal to the number of dimensions.

*Examples:*

```
FSENSE 1 2 VSENSE 10.0
FAMP 13 0 POLY(1) VIN 0 500
FNONLIN 100 101 POLY(2) VCNTRL1 VCINTRL2 0.0 13.6 0.2 0.005
```

*Comments:*

The first general form and the first two examples apply to the linear case. The second form and the last example are for the nonlinear case. For the linear case, there must be one controlling voltage source and its name is followed by the gain. For all cases, including the nonlinear case (`POLY`), refer to your SPICE users guide. In a current-controlled current source device statement, expressions (such as in the `VALUE` keyword version of a voltage-controlled source statement) cannot be used for linear and polynomial coefficient values.

If you are using OrCAD (PSPICE), here are instructions on displaying $S$-parameter data in the Probe (graphical output) window:

`http://www.orcad.com/documents/community.an/pspice/tn10.aspx`

Although this applies specifically to PSPICE, the concepts are applicable more generally.

## Appendix E

# Usage Notes on the DAMS Software