RADIO-FREQUENCY AND MICROWAVE COMMUNICATION CIRCUITS

Analysis and Design

Second Edition

Devendra K. Misra
University of Wisconsin–Milwaukee

WILEY-INTERSCIENCE

A JOHN WILEY & SONS, INC., PUBLICATION

RADIO-FREQUENCY AND MICROWAVE COMMUNICATION CIRCUITS

RADIO-FREQUENCY AND MICROWAVE COMMUNICATION CIRCUITS

Analysis and Design

Second Edition

Devendra K. Misra
University of Wisconsin–Milwaukee

WILEY-INTERSCIENCE

A JOHN WILEY & SONS, INC., PUBLICATION

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

Misra, Devendra, 1949–
Radio-frequency and microwave communication circuits : analysis and design / Devendra K. Misra.—2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-471-47873-3 (Cloth)
1. Radar circuits–Design and construction. 2. Microwave circuits–Design and construction. 3. Electronic circuit design. 4. Radio frequency. I. Title.
TK6560 .M54 2004
621.384′12–dc22

2003026691

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

CONTENTS

PREFACE

Wireless technology continues to grow at a tremendous rate, with new applications still reported almost daily. In addition to the traditional applications in communications, such as radio and television, radio-frequency (RF) and microwaves are being used in cordless phones, cellular communication, local area networks, and personal communication systems. Keyless door entry, radio-frequency identification, monitoring of patients in a hospital or a nursing home, cordless mice or keyboards for computers, and wireless networking of home appliances are some of the other areas where RF technology is being employed. Although some of these applications have traditionally used infrared technology, RF circuits are taking over, because of their superior performance. The present rate of growth in RF technology is expected to continue in the foreseeable future. These advances require the addition of personnel in the areas of radio-frequency and microwave engineering. Therefore, in addition to regular courses as a part of electrical engineering curriculums, short courses and workshops are regularly conducted in these areas for practicing engineers.

This edition of the book maintains the earlier approach of a presentation based on a basic course in electronic circuits. At the same time, a new chapter on electromagnetic fields has been added, following several constructive suggestions from those who used the first edition. It provides the added option of using the book for a traditional microwave engineering course with electromagnetic fields and waves. Or, this chapter can be bypassed so as to follow the approach used in the first edition: that is, instead of using electromagnetic fields as most microwave engineering books do, the subject is introduced via circuit concepts. Further, an overview of communication systems is presented in the beginning to provide the reader with an overall perspective of various building blocks involved.

This edition of the book is organized into thirteen chapters and nine appendixes, using a top-down approach. It begins with an introduction to frequency bands, RF and microwave devices, and applications in communication, radar, industrial, and biomedical areas. The introduction includes a brief description of microwave transmission lines: waveguides, strip lines, and microstrip line. An overview of transmitters and receivers is included, along with digital modulation and demodulation techniques. Modern wireless communication systems, such as terrestrial and satellite communication systems and RF wireless services, are discussed briefly in Chapter 2. After introducing antenna terminology, effective isotropic radiated power, the Friis transmission formula, and the radar range equation are presented. In the final section of the chapter, noise and distortion associated with communication systems are introduced.

Chapter 3 begins with a discussion of distributed circuits and construction of a solution to the transmission line equation. Topics presented in this chapter include RF circuit analysis, phase and group velocities, sending-end impedance, reflection coefficient, return loss, insertion loss, experimental determination of characteristic impedance and the propagation constant, the voltage standing wave ratio, and impedance measurement. The final section in Chapter 3 includes a description of the Smith chart and its application in the analysis of transmission line circuits. Fundamental laws of electromagnetic fields are introduced in Chapter 4 along with wave equations and uniform plane wave solutions. Boundary conditions and potential functions that lead to the construction of solutions are then introduced. The chapter concludes with analyses of various metallic waveguides.

Resonant circuits are discussed in Chapter 5, which begins with series and parallel resistance–inductance–capacitance circuits. This is followed by a section on transformer-coupled circuits. The final two sections of the chapter are devoted to transmission line resonant circuits and microwave resonators. Chapters 6 and 7 deal with impedance-matching techniques. Single reactive element or stub, double-stub, and lumped-element matching techniques are discussed in Chapter 6. Chapter 7 is devoted to multisection transmission line impedance transformers, binomial and Chebyshev sections, and impedance tapers.

Chapter 8 introduces circuit parameters associated with two-port networks. Impedance, admittance, hybrid, transmission, scattering, and chain scattering parameters are presented, along with examples that illustrate their characteristic behaviors. Chapter 9 begins with the image parameter method for the design of passive filter circuits. The insertion-loss technique is introduced next to synthesize Butterworth and Chebyshev low-pass filters. The chapter includes descriptions of impedance and frequency scaling techniques to realize high-pass, bandpass, and bandstop networks. The chapter concludes with a section on microwave transmission line filter design.

Concepts of signal-flow-graph analysis are introduced in Chapter 10, along with a representation of voltage source and passive devices, which facilitates formulation of the power gain relations that are needed in the amplifier design discussed in Chapter 11. The chapter begins with stability considerations using scattering parameters of a two-port network followed by design techniques of various amplifiers.

Chapter 12 presents basic concepts and design of various oscillator circuits. The phase-locked loop and its application in the design of frequency synthesizers are also summarized. The final section of the chapter includes the analysis and design of microwave transistor oscillators using *S*-parameters. Chapter 13 includes the fundamentals of frequency-division multiplexing, amplitude modulation, radio-frequency detection, frequency-modulated signals, and mixer circuits. The book ends with nine appendixes, which include a discussion of logarithmic units (dB, dBm, dBW, dBc, and neper), design equations for selected transmission lines (coaxial line, strip line, and microstrip line), and a list of abbreviations used in the communication area.

Some of the highlights of the book are as follows:

- The presentation begins with an overview of frequency bands, RF and microwave devices, and their applications in various areas. Communication systems are presented in Chapter 2, including terrestrial and satellite systems, wireless services, antenna terminology, the Friis transmission formula, the radar equation, and Doppler radar. Thus, students learn about the systems using blocks of amplifiers, oscillators, mixers, filters, and so on. Students' response has strongly supported this *top-down approach.*
- Since students are assumed to have only one semester of electrical circuits, resonant circuits and two-port networks are included in the book. Concepts of network parameters (impedance, admittance, hybrid, transmission, and scattering) and their characteristics are introduced via examples.
- A separate chapter on oscillator design includes concepts of feedback, the Hartley oscillator, the Colpitts oscillator, the Clapp oscillator, crystal oscillators, phased-locked-loop and frequency synthesizers, transistor oscillator design using *S*-parameters, and three-port *S*-parameter description of transistors and their use in feedback network design.
- A separate chapter on detectors and mixers includes amplitude- and frequency-modulated signal characteristics and their detection schemes, single-diode mixers, RF detectors, double-balanced mixers, conversion loss, intermodulation distortion in diode-ring mixers, and field-effect-transistor mixers.
- Appendixes include logarithmic units, design equations for selected transmission lines, and a list of abbreviations used in the communication area.
- There are 153 solved examples with a step-by-step explanation. Therefore, practicing engineers will find the book useful for self-study as well.
- There are 275 class-tested problems at the ends of chapters. Supplementary material is available to instructors adopting the book.

Acknowledgments

I learned this subject from engineers and authors who are too many to include in this short space, but I gratefully acknowledge their contributions. I would like to thank my anonymous reviewers, instructors here and abroad who used the first edition of this book and provided a number of constructive suggestions, and my former students, who made useful suggestions to improve the presentation. I deeply appreciate the support I received from my wife, Ila, and son, Shashank, during the course of this project. The first edition of this book became a reality only because of enthusiastic support from then-senior editor Philip Meyler and his staff at Wiley. I feel fortunate to continue getting the same kind of support from current editor Val Moliere and from Kirsten Rohstedt.

Devendra K. Misra

1

INTRODUCTION

Scientists and mathematicians of the nineteenth century laid the foundation of telecommunication and wireless technology, which has affected all facets of modern society. In 1864, James C. Maxwell put forth fundamental relations of electromagnetic fields that not only summed up the research findings of Laplace, Poisson, Faraday, Gauss, and others but also predicted the propagation of electrical signals through space. Heinrich Hertz subsequently verified this in 1887, and Guglielmo Marconi transmitted wireless signals across the Atlantic Ocean successfully in 1900. Interested readers may find an excellent discussion of the historical developments of radio frequencies (RFs) and microwaves in the *IEEE Transactions on Microwave Theory and Technique* (Vol. MTT-32, September 1984).

Wireless communication systems require high-frequency signals for the efficient transmission of information. Several factors lead to this requirement. For example, an antenna radiates efficiently if its size is comparable to the signal wavelength. Since the signal frequency is inversely related to its wavelength, antennas operating at RFs and microwaves have higher radiation efficiencies. Further, their size is relatively small and hence convenient for mobile communication. Another factor that favors RFs and microwaves is that the transmission of broadband information signals requires a high-frequency carrier signal. In the case of a single audio channel, the information bandwidth is about 20 kHz. If amplitude modulation (AM) is used to superimpose this information on a carrier, it requires at least this much bandwidth on one side of the spectrum. Further,

Radio-Frequency and Microwave Communication Circuits: Analysis and Design, Second Edition,
By Devendra K. Misra
ISBN 0-471-47873-3

TABLE 1.1 Frequency Bands Used in Commercial Broadcasting

	Channels	Frequency Range	Wavelength Range
AM	107	535–1605 kHz	186.92–560.75 m
TV	2–4	54–72 MHz	4.17–5.56 m
	5–6	76–88 MHz	3.41–3.95 m
FM	100	88–108 MHz	2.78–3.41 m
TV	7–13	174–216 MHz	1.39–1.72 m
	14–83	470–890 MHz	33.7–63.83 cm

commercial AM transmission requires a separation of 10-kHz between the two transmitters. On the other hand, the required bandwidth increases significantly if frequency modulation (FM) is used. Each FM transmitter typically needs a bandwidth of 200 kHz for audio transmission. Similarly, each television channel requires about 6 MHz of bandwidth to carry the video information as well. Table 1.1 shows the frequency bands used for commercial radio and television broadcasts.

In the case of digital transmission, a standard monochrome television picture is sampled over a grid of 512×480 elements called *pixels*. Eight bits are required to represent 256 shades of the gray display. To display motion, 30 frames are sampled per second; thus, it requires about 59 Mb/s ($512 \times 480 \times 8 \times 30 =$ 58,982,400). Color transmission requires even higher bandwidth (on the order of 90 Mb/s).

Wireless technology has been expanding very fast, with new applications reported every day. In addition to the traditional applications in communication, such as radio and television, RF and microwave signals are being used in cordless phones, cellular communication, local, wide, and metropolitan area networks and personal communication service. Keyless door entry, radio-frequency identification (RFID), monitoring of patients in a hospital or a nursing home, and cordless mice or keyboards for computers are some of the other areas where RF technology is being used. Although some of these applications have traditionally used infrared (IR) technology, current trends favor RF, because RF is superior to infrared technology in many ways. Unlike RF, infrared technology requires unobstructed line-of-sight connection. Although RF devices have been more expensive than IR, the current trend is downward because of an increase in their production and use.

The electromagnetic frequency spectrum is divided into bands as shown in Table 1.2. Hence, AM radio transmission operates in the medium-frequency (MF) band, television channels 2 to 12 operate in the very high frequency (VHF) band, and channels 18 to 90 operate in the ultrahigh-frequency (UHF) band. Table 1.3 shows the band designations in the microwave frequency range.

In addition to natural and human-made changes, electrical characteristics of the atmosphere affect the propagation of electrical signals. Figure 1.1 shows

TABLE 1.2 IEEE Frequency Band Designations

Band Designation	Frequency Range	Wavelength Range (in Free Space)
VLF	3–30 kHz	10–100 km
LF	30–300 kHz	1–10 km
MF	300–3000 kHz	100 m–1 km
HF	3–30 MHz	10–100 m
VHF	30–300 MHz	1–10 m
UHF	300–3000 MHz	10 cm–1 m
SHF	3–30 GHz	1–10 cm
EHF	30–300 GHz	0.1–1 cm

TABLE 1.3 Microwave Frequency Band Designations

Frequency Bands	Old (Still Widely Used)	New (Not So Commonly Used)
500–1000 MHz	UHF	C
1–2 GHz	L	D
2–4 GHz	S	E
3–4 GHz	S	F
4–6 GHz	C	G
6–8 GHz	C	H
8–10 GHz	X	I
10–12.4 GHz	X	J
12.4–18 GHz	Ku	J
18–20 GHz	K	J
20–26.5 GHz	K	K
26.5–40 GHz	Ka	K

various layers of the ionosphere and the troposphere that are formed due to the ionization of atmospheric air. As illustrated in Figure 1.2(*a*) and (*b*), an RF signal can reach the receiver by propagating along the ground or after reflection from the ionosphere. These signals may be classified as *ground* and *sky waves*, respectively. The behavior of a sky wave depends on the season, day or night, and solar radiation. The ionosphere does not reflect microwaves, and the signals propagate line of sight, as shown in Figure 1.2(*c*). Hence, curvature of the Earth limits the range of a microwave communication link to less than 50 km. One way to increase the range is to place a human-made reflector up in the sky. This type of arrangement is called a *satellite communication system*. Another way to increase the range of a microwave link is to place repeaters at periodic intervals. This is known as a *terrestrial communication system*.

Figures 1.3 and 1.4 list selected devices used at RF and microwave frequencies. Solid-state devices as well as vacuum tubes are used as active elements in RF and

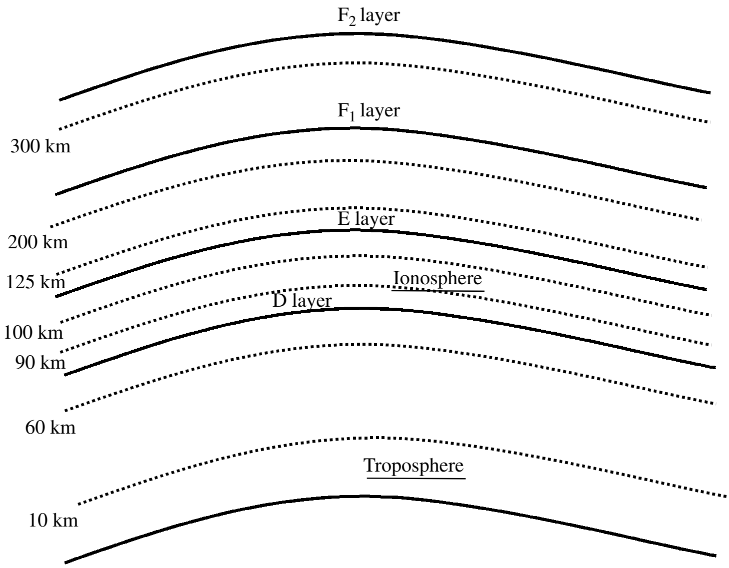


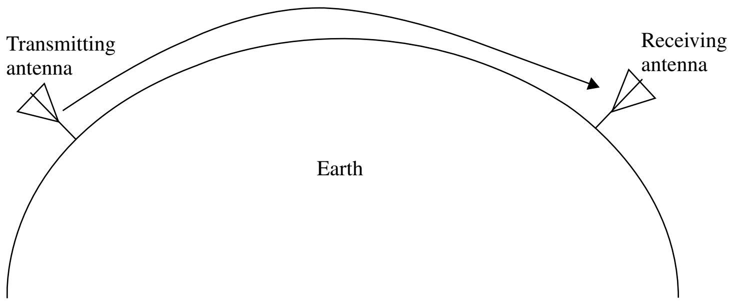
Figure 1.1 Atmosphere surrounding Earth.

microwave circuits. Predominant applications for microwave tubes are in radar, communications, electronic countermeasures (ECMs), and microwave cooking. They are also used in particle accelerators, plasma heating, material processing, and power transmission. Solid-state devices are employed primarily in the RF region and in low-power microwave circuits such as low-power transmitters for local area networks and receiver circuits. Some applications of solid-state devices are listed in Table-1.4.

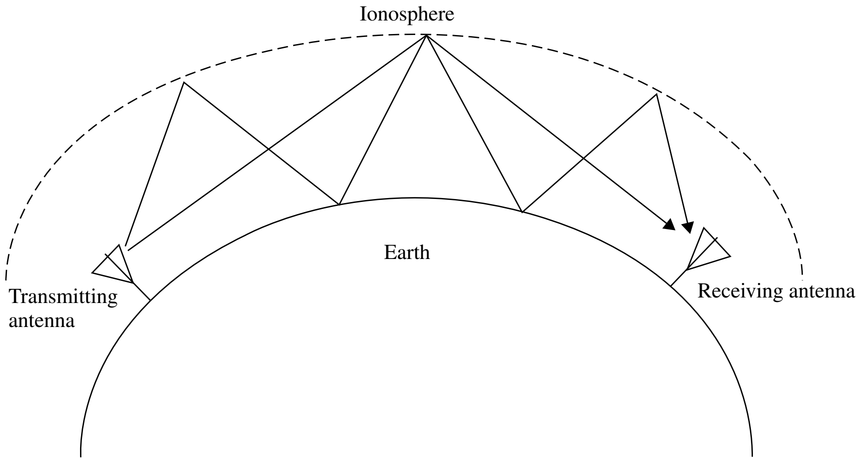
Figure 1.5 lists some applications of microwaves. In addition to terrestrial and satellite communications, microwaves are used in radar systems as well as in various industrial and medical applications. Civilian applications of radar include air traffic control, navigation, remote sensing, and law enforcement. Its military uses include surveillance; guidance of weapons; and command, control, and communication (C^3). Radio-frequency and microwave energy are also used in industrial heating and for household cooking. Since this process does not use a conduction mechanism for the heat transfer, it can improve the quality of certain products significantly. For example, the hot air used in a printing press to dry ink affects the paper adversely and shortens the product's life span. By contrast, in microwave drying only the ink portion is heated, and the paper is barely affected. Microwaves are also used in material processing, telemetry, imaging, and hyperthermia.

1.1 MICROWAVE TRANSMISSION LINES

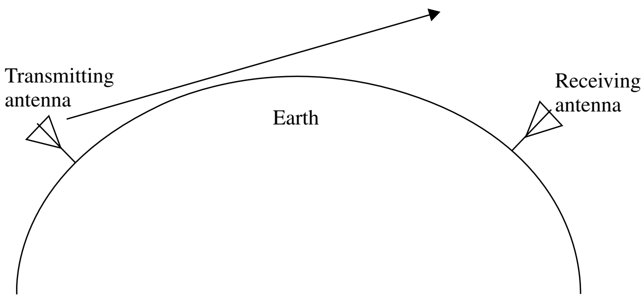
Figure 1.6 shows selected transmission lines used in RF and microwave circuits. The most common transmission line used in the RF and microwave range is the



(*a*) Signal propagation along the ground



(*b*) Propagation of a sky wave



(*c*) Line-of-sight propagation

Figure 1.2 Modes of signal propagation.

coaxial line. A low-loss dielectric material is used in these transmission lines to minimize signal loss. Semirigid coaxial lines with continuous cylindrical conductors outside perform well in the microwave range. To ensure single-mode transmission, the cross section of a coaxial line must be much smaller than the signal wavelength. However, this limits the power capacity of these lines. In

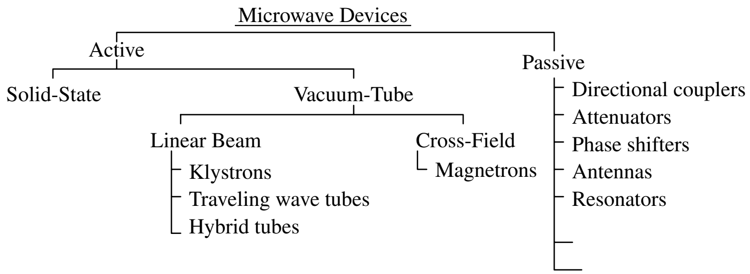


Figure 1.3 Microwave devices.

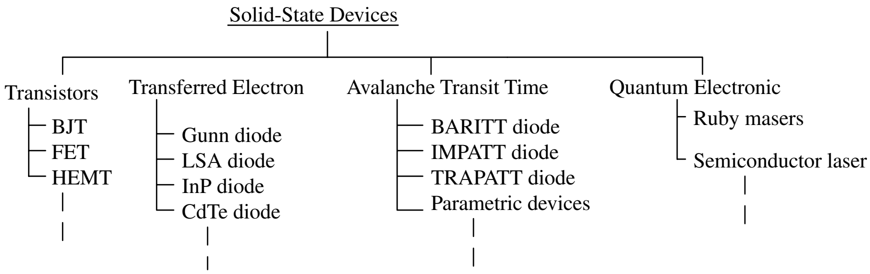


Figure 1.4 Solid-state devices used at RF and microwave frequencies.

TABLE 1.4 Selected Applications of Microwave Solid-State Devices

Devices	Applications	Advantages
Transistors	L-band transmitters for telemetry systems and phased-array radar systems; transmitters for communication systems	Low cost, low power supply, reliable, high-continuous-wave (CW) power output, lightweight
Transferred electron devices (TED)	C-, X-, and Ku-band ECM amplifiers for wideband systems; X- and Ku-band transmitters for radar systems, such as traffic control	Low power supply (12 V), low cost, lightweight, reliable, low noise, high gain
IMPATT diode	Transmitters for millimeter-wave communication	Low power supply, low cost, reliable, high-CW power, lightweight
TRAPATT diode	S-band pulsed transmitter for phased-array radar systems	High peak and average power, reliable, low power supply, low cost
BARITT diode	Local oscillators in communication and radar receivers	Low power supply, low cost, low noise, reliable

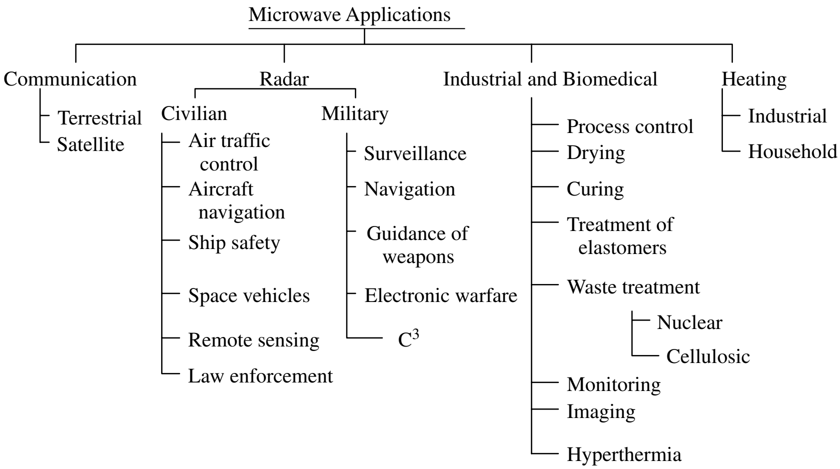


Figure 1.5 Some applications of microwaves.

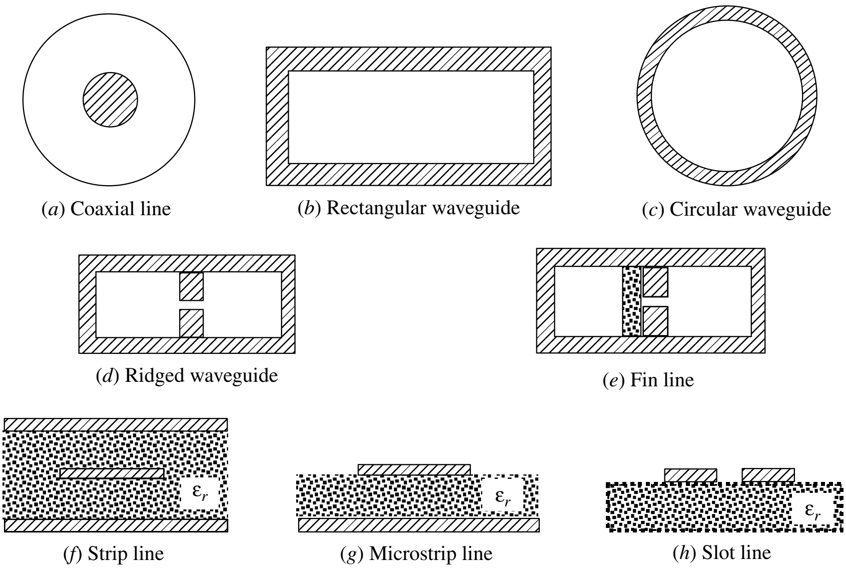


Figure 1.6 Transmission lines used in RF and microwave circuits.

high-power microwave circuits, waveguides are used in place of coaxial lines. Rectangular waveguides are commonly employed for connecting high-power microwave devices because these are easy to manufacture compared with circular waveguides. However, certain devices (e.g., rotary joints) require a circular cross section. In comparison with a rectangular waveguide, a ridged waveguide provides broadband operation. The fin line shown in Figure 1.6(*e*) is commonly used in the millimeter-wave band. Physically, it resembles a slot line enclosed in a rectangular waveguide.

The transmission lines illustrated in Figure 1.6(*f*) to (*h*) are most convenient in connecting the circuit components on a printed circuit board (PCB). The physical dimensions of these transmission lines depend on the dielectric constant ε_r of insulating material and on the operating frequency band. The characteristics and design formulas of selected transmission lines are given in the appendixes.

1.2 TRANSMITTER AND RECEIVER ARCHITECTURES

Wireless communication systems require a transmitter at one end to send the information signal and a receiver at the other to retrieve it. In one-way communication (such as a commercial broadcast), a transmitting antenna radiates the signal according to its radiation pattern. The receiver, located at the other end, receives this signal via its antenna and extracts the information, as illustrated in Figure 1.7. Thus, the transmitting station does not require a receiver, and vice versa. On the other hand, a transceiver (a transmitter and a receiver) is needed at both ends to establish a two-way communication link.

Figure 1.7 is a simplified block diagram of a one-way communication link. At the transmitting end, an information signal is modulated and mixed with a local oscillator to up-convert the carrier frequency. Bandpass filters are used before

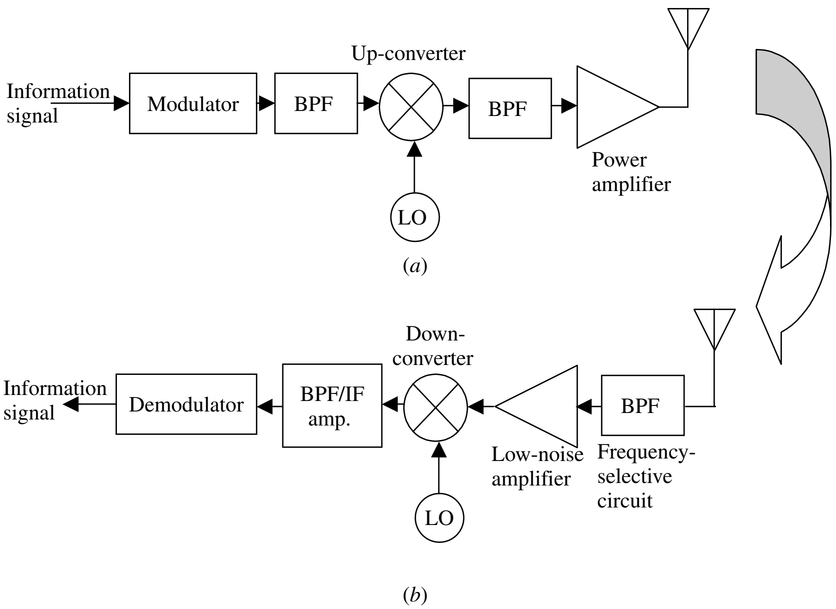


Figure 1.7 Simplified block diagram of the transmitter (*a*) and receiver (*b*) of a wireless communication system.

and after the mixer to stop undesired harmonics. The signal power is amplified before feeding it to the antenna. At the receiving end the entire process is reversed to recover the information. Signal received by the antenna is filtered and amplified to improve the signal-to-noise ratio before feeding it to the mixer for down-converting the frequency. A frequency-selective amplifier (tuned amplifier) amplifies it further, before feeding it to a suitable demodulator, which extracts the information signal.

In an analog communication system, the amplitude or angle (frequency or phase) of the carrier signal is varied according to the information signal. These modulations are known as amplitude modulation (AM), frequency modulation (FM), and phase modulation (PM), respectively. In a digital communication system, the input passes through channel coding, interleaving, and other processing before it is fed to the modulator. Various modulation schemes are available, including on–off keying (OOK), frequency-shift keying (FSK), and phase-shift keying (PSK). As these names suggest, a high-frequency signal is turned on and off in OOK to represent the logic states 1 and 0 of a digital signal. Similarly, two different signal frequencies are employed in FSK to represent the two signal states. In PSK, the phase of the high-frequency carrier is changed according to the 1 or 0 state of the signal. If only two phase states (0° and 180°) of the carrier are used, it is called binary phase-shift keying (BPSK). On the other hand, a phase shift of 90° gives four possible states, each representing 2 bits of information (known as a *dibit*). This type of digital modulation, known as quadrature phase-shift keying (QPSK), is explained further below.

Consider a sinusoidal signal S_{mod}, as given by equation (1.2.1). Its angular frequency and phase are ω radians per second and ϕ, respectively:

$$S_{\text{mod}}(t) = \sqrt{2}\cos(\omega t - \phi) = \sqrt{2}(\cos\omega t\cos\phi + \sin\omega t\sin\phi) \qquad (1.2.1)$$

This can be simplified further as follows:

$$S_{\text{mod}}(t) = S_I\cos\omega t + S_Q\sin\omega t \qquad (1.2.2)$$

where

$$S_I = \sqrt{2}\cos\phi \qquad (1.2.3)$$

and

$$S_Q = \sqrt{2}\sin\phi \qquad (1.2.4)$$

The subscripts I and Q represent in-phase and quadrature-phase components of S_{mod}. Table 1.5 shows the values of S_I and S_Q for four different phase states together with the corresponding dibit representations. This scheme can be implemented easily for a polar digital signal (positive peak value representing logic 1 and the negative peak logic 0). It is illustrated in Figure 1.8. The demultiplexer simultaneously feeds one bit of the digital input S_{BB} to the top and the other to the bottom branch of this circuit. The top branch multiplies the signal by cos ωt

TABLE 1.5 QPSK Scheme

ϕ	S_I	S_Q	Dibit Representation
$\pi/4$	1	1	11
$3\pi/4$	-1	1	01
$5\pi/4$	-1	-1	00
$7\pi/4$	1	-1	10

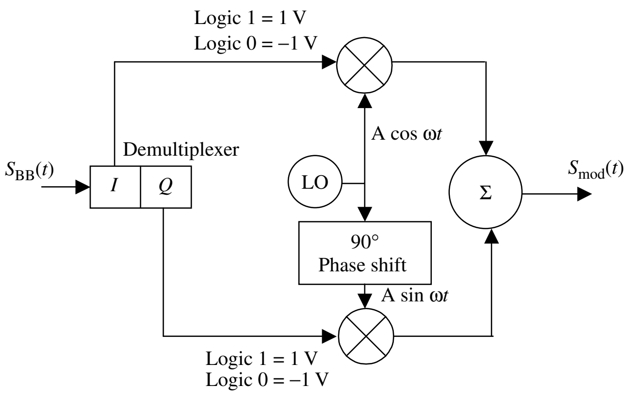


Figure 1.8 Block diagram of a QPSK modulation scheme.

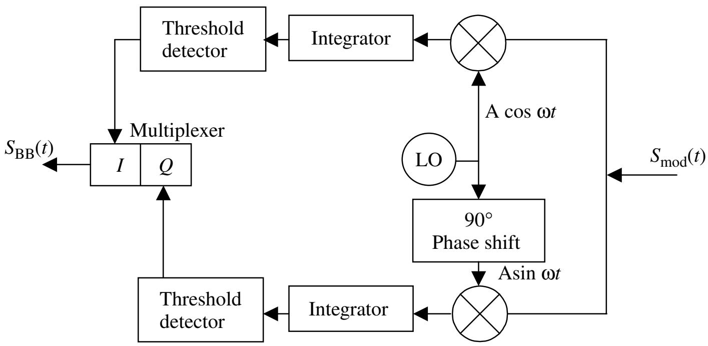


Figure 1.9 Block diagram of a QPSK demodulation scheme.

and the bottom signal by sin ωt. The outputs of the two mixers are then added to generate S_{mod}.

Demodulation inverts the modulation to retrieve S_{BB}. As illustrated in Figure 1.9, S_{mod} is multiplied by cos ωt in the top branch and by sin ωt in the bottom branch of the demodulator. The top integrator stops S_Q, and the bottom integrator, S_I. Two threshold detectors generate the corresponding logic states that are multiplexed by the multiplexer unit to recover S_{BB}. Chapter 2 provides an overview of wireless communication systems and their characteristics.

2

COMMUNICATION SYSTEMS

Modern communication systems require RF and microwave signals for the wireless transmission of information. These systems employ oscillators, mixers, filters, and amplifiers to generate and process various types of signals. The transmitter communicates with the receiver via antennas placed on each side. Electrical noise associated with the systems and the channel affects the performance. A system designer needs to know about the channel characteristics and system noise in order to estimate the required power levels. The chapter begins with an overview of microwave communication systems and RF wireless services to illustrate the applications of circuits and devices that are described in the following chapters. It also covers the placement of various building blocks in a given system.

A short discussion on antennas is included to help in understanding signal behavior when it propagates from a transmitter to a receiver. The Friis transmission formula and the radar range equation are important to understanding the effects of frequency, range, and operating power levels on the performance of a communication system. Note that radar concepts now find many other applications, such as proximity or level sensing in an industrial environment. Therefore, a brief discussion of Doppler radar is also included. Noise and distortion characteristics play a significant role in analysis and design of these systems. Minimum detectable signal (MDS), gain compression, intercept point, and the dynamic range of an amplifier (or receiver) are introduced later. Other concepts associated with noise and distortion characteristics are also introduced in this chapter.

Radio-Frequency and Microwave Communication Circuits: Analysis and Design, Second Edition,
By Devendra K. Misra
ISBN 0-471-47873-3

2.1 TERRESTRIAL COMMUNICATION

As mentioned in Chapter 1, microwave signals propagate along the line of sight. Therefore, the Earth's curvature limits the range over which a microwave communication link can be established. A transmitting antenna sitting on a 25-ft-high tower can typically communicate only up to a distance of about 50 km. Repeaters can be placed at regular intervals to extend the range. Figure 2.1 is a block diagram of a typical repeater.

A repeater system operates as follows. A microwave signal arriving at antenna A works as input to port 1 of a circulator. It is directed to port 2 without loss, assuming that the circulator is ideal. Then it passes through a receiver protection circuit that limits the magnitude of large signals but passes those of low intensity with negligible attenuation. The purpose of this circuit is to block excessively large signals from reaching the receiver input. The mixer following it works as a down-converter that transforms a high-frequency signal to a low-frequency signal, typically in the range of 70 MHz. A Schottky diode is generally employed in the mixer because of its superior noise characteristics. This frequency conversion facilitates amplification of the signal economically. A bandpass filter is used at the output of the mixer to stop undesired harmonics. An intermediate-frequency (IF) amplifier is then used to amplify the signal. It is generally a low-noise solid-state

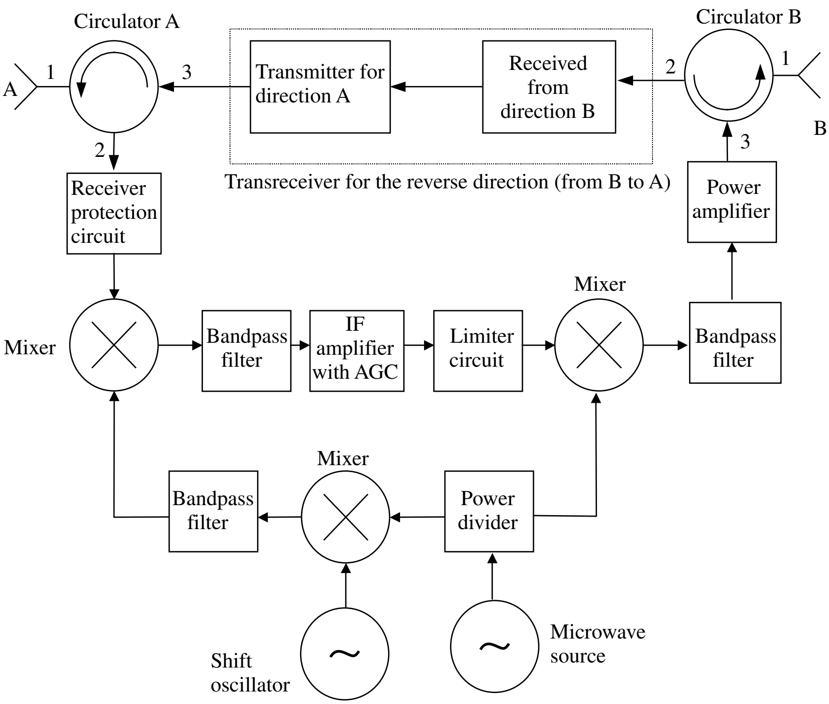


Figure 2.1 Block arrangement of a repeater system.

amplifier with ultralinear characteristics over a broadband. The amplified signal is mixed with another signal for up-conversion of frequency. After filtering out undesired harmonics introduced by the mixer, it is fed to a power amplifier stage that feeds circulator B for onward transmission through antenna B. This up-converting mixer circuit generally employs a varactor diode. Circulator B directs the signal entering at port 3 to the antenna connected at its port 1. Similarly, the signal propagating upstream is received by antenna B and the circulator directs it toward port 2. It then goes through the processing as described for the downstream signal and is radiated by antenna A for onward transmission. Hence, the downstream signal is received by antenna A and transmitted in the forward direction by antenna B. Similarly, the upstream signal is received by antenna B and forwarded to the next station by antenna A. The two circulators help channel the signal in the correct direction.

A parabolic antenna with tapered horn as primary feeder is generally used in microwave links. This type of composite antenna system, known as a *hog horn*, is fairly common in high-density links because of its broadband characteristics. These microwave links operate in the frequency range 4 to 6 GHz, and signals propagating in two directions are separated by a few hundred megahertz. Since this frequency range overlaps with C-band satellite communication, the interference of these signals needs to be taken into design consideration. A single frequency can be used twice for transmission of information using vertical and horizontal polarization.

2.2 SATELLITE COMMUNICATION

The ionosphere does not reflect microwaves as it does RF signals. However, one can place a conducting object (satellite) up in the sky that reflects them back to Earth. A satellite can even improve the signal quality using on-board electronics before transmitting it back. The gravitational force needs to be balanced somehow if this object is to stay in position. An orbital motion provides this balancing force. If a satellite is placed at low altitude, greater orbital force will be needed to keep it in position. These low- and medium-altitude satellites are visible from a ground station only for short periods. On the other hand, satellites placed at an altitude of about 36,000 km over the equator, called *geosynchronous* or *geostationary satellites* are visible from their shadows at all times.

C-band geosynchronous satellites use between 5725 and 7075 MHz for their uplinks. The corresponding downlinks are between 3400 and 5250 MHz. Table 2.1 lists the downlink center frequencies of a 24-channel transponder. Each channel has a total bandwidth of 40 MHz; 36 MHz of that carries the information, and the remaining 4 MHz is used as a guard band. It is accomplished with a 500-MHz bandwidth using different polarization for the overlapping frequencies. The uplink frequency plan may be found easily after adding 2225 MHz to these downlink frequencies. Figure 2.2 illustrates the simplified block diagram of a C-band satellite transponder. A 6-GHz signal received from the Earth station

TABLE 2.1 C-Band Downlink Transponder Frequencies

Horizontal Polarization		Vertical Polarization	
Channel	Center Frequency (MHz)	Channel	Center Frequency (MHz)
1	3720	2	3740
3	3760	4	3780
5	3800	6	3820
7	3840	8	3860
9	3880	10	3900
11	3920	12	3940
13	3960	14	3980
15	4000	16	4020
17	4040	18	4060
19	4080	20	4100
21	4120	22	4140
23	4160	24	4180

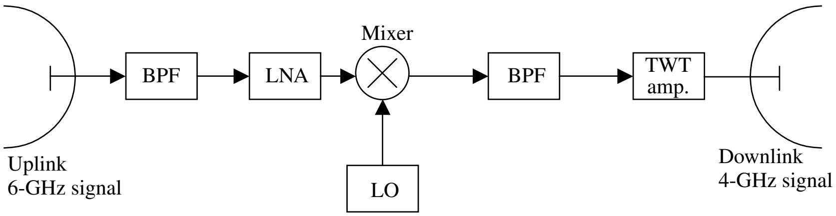


Figure 2.2 Simplified block diagram of a transponder.

is passed through a bandpass filter before amplifying it through a low-noise amplifier (LNA). It is then mixed with a local oscillator (LO) signal to bring down its frequency. A bandpass filter that is connected right after the mixer filters out the unwanted frequency components. This signal is then amplified by a traveling wave tube (TWT) amplifier and transmitted back to Earth.

Another frequency band in which satellite communication has been growing continuously is the Ku-band. The geosynchronous Fixed Satellite Service (FSS) generally operates between 10.7 and 12.75 GHz (space to Earth) and 13.75 to 14.5 GHz (Earth to space). It offers the following advantages over the C-band:

- The size of the antenna can be smaller (3 ft or even smaller, with higher-power satellites against 8 to 10 ft for C-band).
- Because of higher frequencies used in the up- and downlinks, there is no interference with C-band terrestrial systems.

Since higher-frequency signals attenuate faster while propagating through adverse weather (rain, fog, etc.), Ku-band satellites suffer from this major drawback. Signals with higher powers may be used to compensate for this loss. Generally, this power is on the order of 40 to 60 W. The high-power direct broadcast satellite (DBS) system uses power amplifiers in the range 100 to 120 W.

The National Broadcasting Company (NBC) has been using the Ku-band to distribute the programming to its affiliates. Also, various news-gathering agencies have used this frequency band for some time. Convenience stores, auto parts distributors, banks, and other businesses have used the very small aperture terminal (VSAT) because of its small antenna size (typically, on the order of 3 ft in diameter). It offers two-way satellite communication; usually back to hub or headquarters. The Public Broadcasting Service (PBS) uses VSATs for exchanging information among public schools.

Direct broadcast satellites (DBSs) have been around since 1980, but early DBS ventures failed for various reasons. In 1991, Hughes Communications entered into the direct-to-home (DTH) television business. DirecTV was formed as a unit of GM Hughes, with *DBS-1* launched in December 1993. Its longitudinal orbit is at 101.2°W, and it employs a left-handed circular polarization. *DBS-2*, launched in August 1994 uses a right-handed circular polarization, and its orbital longitude is at 100.8°W. DirecTV employs a digital architecture that can utilize video and audio compression techniques. It complies with Motion Picture Experts Group (MPEG)-2. By using compression ratios of 5 to 7, over 150 channels of programs are available from the two satellites. These satellites include 120-W TWT amplifiers that can be combined to form eight pairs at 240 W of power. This higher power can also be utilized for high-definition television (HDTV) transmission. Earth-to-satellite link frequency is 17.3 to 17.8 GHz; satellite-to-Earth link frequency uses the 12.2- to 12.7-GHz band. Circular polarization is used because it is less affected by rain than is linear orthogonal polarization.

Several communication services are now available that use low-Earth-orbit satellites (LEOSs) and medium-Earth-orbit satellites (MEOSs). LEOS altitudes range from 750 to 1500 km; MEOS systems have an altitude around 10,350 km. These services compete with or supplement cellular systems and geosynchronous Earth-orbit satellites (GEOSs). The GEOS systems have some drawbacks, due to the large distances involved. They require relatively large powers, and the propagation time delay creates problems in voice and data transmissions. The LEOS and MEOS systems orbit Earth faster because of being at lower altitudes, and these are therefore visible only for short periods. As Table 2.2 indicates, several satellites are used in a personal communication system to solve this problem.

Three classes of service can be identified for mobile satellite services:

1. Data transmission and messaging from very small, inexpensive satellites
2. Voice and data communications from big LEOSs
3. Wideband data transmission

Another application of L-band microwave frequencies (1227.60 and 1575.42 MHz) is in global positioning systems (GPSs). A constellation of 24 satellites is used to determine a user's geographical location. Two services

TABLE 2.2 Specifications of Certain Personal Communication Satellites

	Iridium (LEO)	Globalstar (LEO)	Odyssey (MEO)
Number of satellites	66	48	12
Altitude (km)	755	1,390	10,370
Uplink (GHz)	1.616–1.6265	1.610–1.6265	1.610–1.6265
Downlink (GHz)	1.616–1.6265	2.4835–2.500	2.4835–2.500
Gateway terminal uplink (GHz)	27.5–30.0	C-band	29.5–30.0
Gateway terminal downlink (GHz)	18.8–20.2	C-band	19.7–20.2
Average satellite connect time (min)	9	10–12	120
Features of handset			
Modulation	QPSK	FQPSK	QPSK
BER	1E-2 (voice) 1E-5 (data)	1E-3 (voice) 1E-5 (data)	1E-3 (voice) 1E-5 (data)
Supportable data rate (kb/s)	4.8 (voice) 2.4 (data)	1.2–9.6 (voice and data)	4.8 (voice) 1.2–9.6 (data)

are available: the standard positioning service (SPS) for civilian use, utilizing a single-frequency course/acquisition (C/A) code, and the precise positioning service (PPS) for the military, utilizing a dual-frequency P-code (protected). These satellites are at an altitude of 10,900 miles above the Earth, with an orbital period of 12 hours.

2.3 RADIO-FREQUENCY WIRELESS SERVICES

A lot of exciting wireless applications have been reported that use voice and data communication technologies. Wireless communication networks consist of microcells that connect people with truly global, pocket-size communication devices, telephones, pagers, personal digital assistants, and modems. Typically, a cellular system employs a 100-W transmitter to cover a cell 0.5 to 10 miles in radius. The handheld transmitter has a power of less than 3 W. Personal communication networks (PCN/PCS) operate with a 0.01- to 1-W transmitter to cover a cell radius of less than 450 yards. The handheld transmitter power is typically less than 10 mW. Table 2.3 shows the cellular telephone standards of selected systems.

There have been no universal standards set for wireless personal communication. In North America, cordless has been CT-0 (an analog 46/49-MHz standard) and cellular AMPS (Advanced Mobile Phone Service) operating at 800 MHz. The situation in Europe has been far more complex; every country has had its own standard. Although cordless was nominally CT-0, different countries used their own frequency plans. This led to a plethora of new standards. These include,

TABLE 2.3 Selected Cellular Telephones

	Analog Cellular Phone Standard		Digital Cellular Phone Standard			
	AMP	ETACS	NADC (IS-54)	NADC (IS-95)	GSM	PDC
Freq. range (MHz)						
Tx	824–849	871–904	824–849	824–849	880–915	940–956 1477–1501
Rx	869–894	916–949	869–894	869–894	925–960	810–826 1429–1453
Transmitter power (max.)			600 mW	200 mW	1 W	
Multiple access	FDMA	FDMA	TDMA/FDM	CDMA/FDM	TDMA/FDM	TDMA/FDM
Number of channels	832	1000	832	20	124	1600
Channel spacing (kHz)	30	25	30	1250	200	25
Modulation	FM	FM	$\pi/4$ DQPSK	QPSK/BPSK	GMSK	$\pi/4$ DQPSK
Bit rate (kb/s)	—	—	48.6	1228.8	270.833	42

TABLE 2.4 Selected Cordless Telephones

	Analog Cordless Phone Standard		Digital Cordless Phone Standard		
	CT-0	CT-1 and CT-1+	CT-2 and CT-2+	DECT	PHS (Formerly PHP)
Frequency range (MHz)	46/49	CT-1: 914/960; CT-1+: 885–932	CT-1: 864–868; CT-2+: 930/931; 940/941	1880–1900	1895–1918
Transmitter power (max.) (mW)	—	—	10 and 80	250	80
Multiple access	FDMA	FDMA	TDMA/FDM	TDMA/FDM	TDMA/FDM
Number of channels	10–20	CT-1: 40; CT-1+: 80	40	10 (12 users per channel)	300 (four users per channels)
Channel spacing (kHz)	40	25	100	1728	300
Modulation	FM	FM	GFSK	GFSK	$\pi/4$ DQPSK
Bit rate (kb/s)	—	—	72	1152	384

but are not limited to, CT-1, CT-1+, DECT (Digital European Cordless Telephone), PHP (Personal Handy Phone in Japan), E-TACS (Extended Total Access Communication System in the UK), NADC (North American Digital Cellular), GSM (Global System for Mobile Communication), and PDC (Personal Digital Cellular). Specifications for selected cordless telephones are given in Table 2.4.

2.4 ANTENNA SYSTEMS

Figure 2.3 illustrates some of the antennas that are used in communication systems. These can be categorized into two groups: wire antennas and aperture-type antennas. Electric dipole, monopole, and the loop antennas belong to the former group; horn, reflector, and lens belong to the latter category. The aperture antennas can be further subdivided into primary and secondary (or passive) antennas. Primary antennas are directly excited by the source and can be used independently for transmission or reception of signals. On the other hand, a secondary antenna requires another antenna as its feeder. Horn antennas fall in the first category, whereas the reflector and lens belong to the second. Various types of horn antennas are commonly used as feeders in reflector and lens antennas.

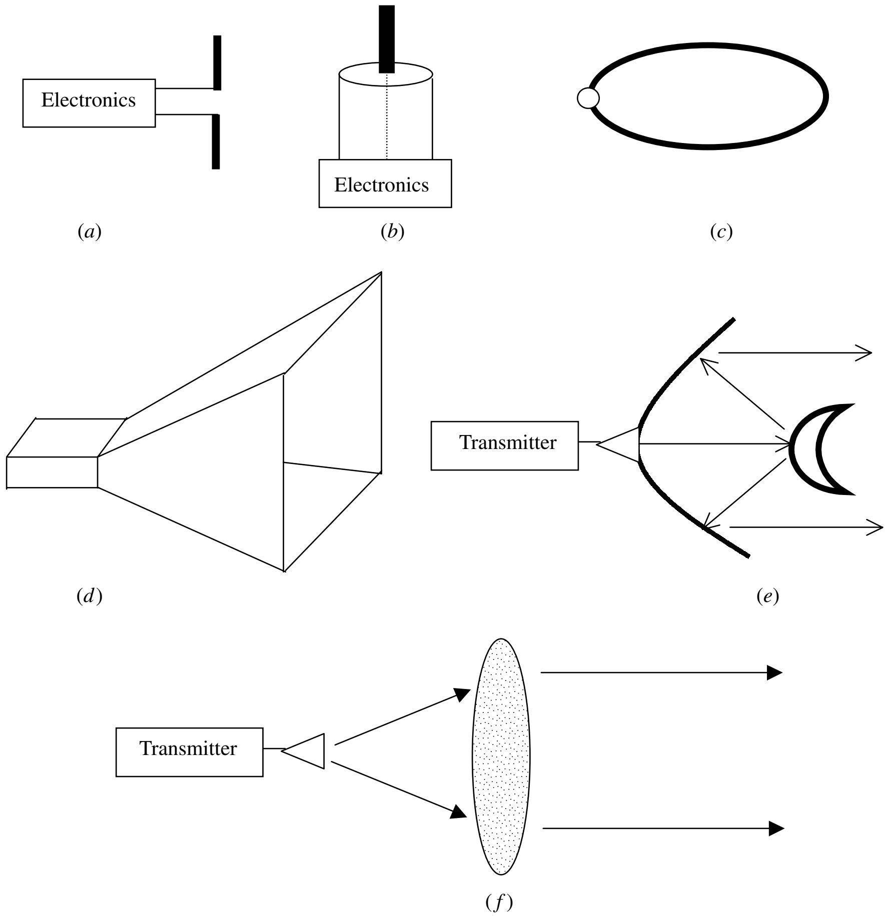


Figure 2.3 Some commonly used antennas: (*a*) electric dipole; (*b*) monopole; (*c*) loop; (*d*) pyramidal horn; (*e*) Cassegrain reflector; (*f*) lens.

When an antenna is energized, it generates two types of electromagnetic fields. Part of the energy stays nearby and part propagates outward. The propagating signal represents the radiation fields, while the nonpropagating is reactive (capacitive or inductive) in nature. Space surrounding the antenna can be divided into three regions. The reactive fields dominate in the nearby region but are reduced in strength at a faster rate than those associated with the propagating signal. If the largest dimension of an antenna is D and the signal wavelength is λ, reactive fields dominate up to about $0.62\sqrt{D^3/\lambda}$ and diminish after $2D^2/\lambda$. The region beyond $2D^2/\lambda$ is called the *far-field* (or *radiation field*) *region*.

Power radiated by an antenna per unit solid angle is known as the *radiation intensity* U. It is a far-field parameter that is related to power density (power per unit area) W_{rad} and distance r as follows:

$$U = r^2 W_{\mathrm{rad}} \tag{2.4.1}$$

Directive Gain and Directivity

If an antenna radiates uniformly in all directions, it is called an *isotropic antenna*. This is a hypothetical antenna that helps in defining the characteristics of a real one. The directive gain D_G is defined as the ratio of radiation intensity due to the test antenna to that of an isotropic antenna. It is assumed that total radiated power remains the same in the two cases. Hence,

$$D_G = \frac{U}{U_0} = \frac{4\pi U}{P_{\mathrm{rad}}} \tag{2.4.2}$$

where U is the radiation intensity due to the test antenna in watts per unit solid angle, U_0 the radiation intensity due to the isotropic antenna in watts per unit solid angle, and P_{rad} the total power radiated in watts. Since U is a directional-dependent quantity, the directive gain of an antenna depends on the angles θ and ϕ. If the radiation intensity assumes its maximum value, the directive gain is called the *directivity* D_0. That is,

$$D_0 = \frac{U_{\max}}{U_0} = \frac{4\pi U_{\max}}{P_{\mathrm{rad}}} \tag{2.4.3}$$

Gain of an Antenna

The *power gain* of an antenna is defined as the ratio of its radiation intensity at a point to the radiation intensity that results from a uniform radiation of the same input power. Hence,

$$\text{gain} = 4\pi \frac{\text{radiation intensity}}{\text{total input power}} = 4\pi \frac{U(\theta, \phi)}{P_{\mathrm{in}}} \tag{2.4.4}$$

Most of the time we deal with relative gain. It is defined as a ratio of the power gain of the test antenna in a given direction to the power gain of a reference antenna. Both antennas must have the same input power. The reference antenna is usually a dipole, horn, or any other antenna whose gain can be calculated or is known. However, the reference antenna is a lossless isotropic radiator in most cases. Hence,

$$\text{gain} = 4\pi \frac{U(\theta, \phi)}{P_{\text{in}}(\text{lossless isotropic antenna})} \tag{2.4.5}$$

When the direction is not stated, the power gain is usually taken in the direction of maximum radiation.

Radiation Patterns and Half-Power Beam Width

Far-field power distribution at a distance r from the antenna depends on the spatial coordinates θ and ϕ. Graphical representations of these distributions on the orthogonal plane (θ-plane or ϕ-plane) at a constant distance r from the antenna are called its *radiation patterns*. Figure 2.4 illustrates the radiation pattern of the vertical dipole antenna with θ. Its ϕ-plane pattern can be found after rotating it about the vertical axis. Thus, a three-dimensional picture of the radiation pattern of a dipole is doughnut shaped. Similarly, the power distributions of other antennas generally show peaks and valleys in the radiation zone. The highest peak between the two valleys is known as the *main lobe*; the others are called *side lobes*. The total angle about the main peak over which power is reduced by 50% of its maximum value is called the *half-power beam width* on that plane.

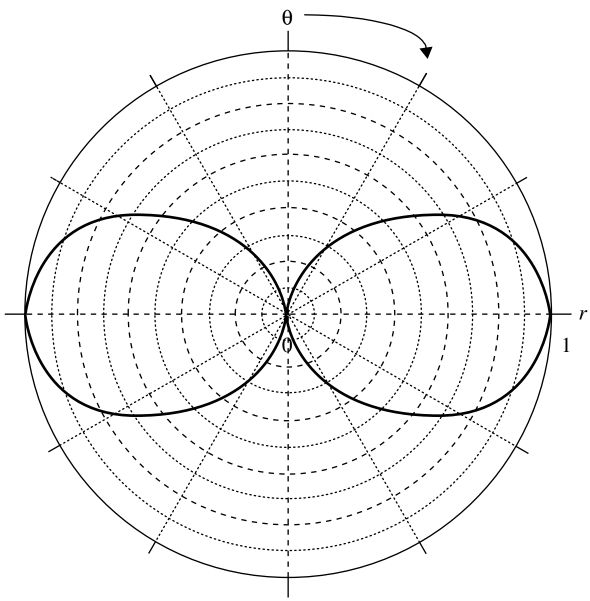


Figure 2.4 Radiation pattern of a dipole in a vertical (θ) plane.

The following relations are used to estimate the power gain G and the half-power beam width (HPBW, or simply BW) of an aperture antenna:

$$G = \frac{4\pi}{\lambda^2} A_e = \frac{4\pi}{\lambda^2} A\kappa \tag{2.4.6}$$

and

$$\mathrm{BW(degrees)} = \frac{65\lambda}{d} \tag{2.4.7}$$

where A_e is the effective area of radiating aperture in square meters, A its physical area ($\pi d^2/4$ for a reflector antenna dish with diameter d), κ the antenna efficiency (ranges from 0.6 to 0.65), and λ the signal wavelength in meters.

Example 2.1 Calculate the power gain (in decibels) and the half-power beam width of a parabolic dish antenna 30 m in diameter that is radiating at 4 GHz.

SOLUTION The signal wavelength and area of the aperture are

$$\lambda = \frac{3 \times 10^8}{4 \times 10^9} = 0.075\,\mathrm{m}$$

and

$$A = \frac{\pi d^2}{4} = \pi \frac{30^2}{4} = 706.8584\,\mathrm{m}^2$$

Assuming that the aperture efficiency is 0.6, the antenna gain and half-power beam width are found as follows:

$$\begin{aligned} G &= \frac{4\pi}{(0.075)^2} \times 706.8584 \times 0.6 = 947{,}482.09 \\ &= 10 \log_{10}(947{,}482.09) = 59.76 \approx 60\,\mathrm{dB} \\ \mathrm{BW} &= \frac{65 \times 0.075}{30} = 0.1625^\circ \end{aligned}$$

Antenna Efficiency

If an antenna is not matched with its feeder, a part of the signal available from the source is reflected back. It is considered to be the reflection (or mismatch) loss. The *reflection* (or *mismatch*) *efficiency* is defined as a ratio of power input to the antenna to that of power available from the source. Since the ratio of reflected power to that of power available from the source is equal to the square of the magnitude of the voltage reflection coefficient, the reflection efficiency e_r is given by

$$e_r = 1 - |\Gamma|^2$$

where

$$\Gamma = \text{voltage reflection coefficient} = \frac{Z_A - Z_0}{Z_A + Z_0}$$

Here Z_A is the antenna impedance and Z_0 is the characteristic impedance of the feeding line.

In addition to mismatch, the signal energy may dissipate in an antenna due to imperfect conductor or dielectric material. These efficiencies are hard to compute. However, the combined conductor and dielectric efficiency e_{cd} can be determined experimentally after measuring the input power P_{in} and the radiated power P_{rad}. It is given as

$$e_{cd} = \frac{P_{\text{rad}}}{P_{\text{in}}}$$

The overall efficiency e_o is a product of the efficiencies above. That is,

$$e_o = e_r e_{cd} \tag{2.4.8}$$

Example 2.2 A 50-Ω transmission line feeds a lossless one-half-wavelength-long dipole antenna. The antenna impedance is 73 Ω. If its radiation intensity, $U(\theta, \phi)$, is given as

$$U = B_0 \sin^3 \theta$$

find the maximum overall gain.

SOLUTION The maximum radiation intensity, $U_{\max}$, is the B_0 value that occurs at $\theta = \pi/2$. Its total radiated power is found as follows:

$$P_{\text{rad}} = \int_0^{2\pi} \int_0^{\pi} B_0 \sin^3 \theta \ \sin\theta \ d\theta \ d\phi = \tfrac{3}{4}\pi^2 B_0$$

Hence,

$$D_0 = 4\pi \frac{U_{\max}}{P_{\text{rad}}} = \frac{4\pi B_0}{\frac{3}{4}\pi^2 B_0} = \frac{16}{3\pi} = 1.6977$$

or

$$D_0(\text{dB}) = 10 \log_{10}(1.6977)\text{dB} = 2.2985\,\text{dB}$$

Since the antenna is lossless, the radiation efficiency e_{cd} is unity (0 dB). Its mismatch efficiency is computed as follows.

The voltage reflection coefficient at its input (it is formulated in Chapter 2) is

$$\Gamma = \frac{Z_A - Z_0}{Z_A + Z_0} = \frac{73 - 50}{73 + 50} = \frac{23}{123}$$

Therefore, the mismatch efficiency of the antenna is

$$e_r = 1 - (23/123)^2 = 0.9650 = 10\log_{10}(0.9650)\text{dB} = -0.1546\,\text{dB}$$

The overall gain G_0 (in decibels) is found as follows:

$$G_0(\text{dB}) = 2.2985 - 0 - 0.1546 = 2.1439\,\text{dB}$$

Bandwidth

Antenna characteristics such as gain, radiation pattern, impedance, and so on are frequency dependent. The bandwidth of an antenna is defined as the frequency band over which its performance with respect to some characteristic (HPBW, directivity, etc.) conforms to a specified standard.

Polarization

Polarization of an antenna is the same as polarization of its radiating wave. It is a property of the electromagnetic wave describing the time-varying direction and relative magnitude of an electric field vector. The curve traced by the instantaneous electric field vector with time is the polarization of that wave. The polarization is classified as follows:

- *Linear polarization.* If the tip of the electric field intensity traces a straight line in some direction with time, the wave is linearly polarized.
- *Circular polarization.* If the end of the electric field traces a circle in space as time passes, that electromagnetic wave is circularly polarized. Further, it may be right-handed circularly polarized (RHCP) or left-handed circularly polarized (LHCP), depending on whether the electric field vector rotates clockwise or counterclockwise.
- *Elliptical polarization.* If the tip of the electric field intensity traces an ellipse in space as time lapses, the wave is elliptically polarized. As in the preceding case, it may be right- or left-handed elliptical polarization (RHEP and LHEP).

In a receiving system, the polarization of the antenna and the incoming wave need to be matched for maximum response. If this is not the case, there will be some signal loss, known as *polarization loss*. For example, if there is a vertically polarized wave incident on a horizontally polarized antenna, the induced voltage available across its terminals will be zero. In this case, the antenna is *cross-polarized with an incident wave*. The square of the cosine of the angle between wave polarization and antenna polarization is a measure of the polarization loss. It can be determined by squaring the scalar product of unit vectors representing the two polarizations.

Example 2.3 The electric field intensity of an electromagnetic wave propagating in a lossless medium in the z-direction is given by

$$\mathbf{E}(\mathbf{r}, t) = \hat{x} E_0(x, y) \cos(\omega t - kz) \qquad \text{V/m}$$

It is incident upon an antenna that is linearly polarized as follows:

$$\mathbf{E}_a(\mathbf{r}) = (\hat{x} + \hat{y}) E(x, y, z) \qquad \text{V/m}$$

Find the polarization loss factor.

SOLUTION In this case, the incident wave is linearly polarized along the x-axis while the receiving antenna is linearly polarized at 45° from it. Therefore, one-half of the incident signal is cross-polarized with the antenna. It is determined mathematically as follows. The unit vector along the polarization of incident wave is

$$\hat{u}_i = \hat{x}$$

The unit vector along the antenna polarization may be found as

$$\hat{u}_a = \frac{1}{\sqrt{2}}(\hat{x} + \hat{y})$$

Hence, the polarization loss factor is

$$|\hat{u}_i \bullet \hat{u}_a|^2 = 0.5 = -3.01\,\text{dB}$$

Effective Isotropic Radiated Power

The effective isotropic radiated power (EIRP) is a measure of the power gain of the antenna. It is equal to the power needed by an isotropic antenna that provides the same radiation intensity at a given point as that of the directional antenna. If power input to the feeding line is P_t and the antenna gain is G_t, the EIRP is defined as

$$\text{EIRP} = \frac{P_t G_t}{L} \tag{2.4.9}$$

where L is the input/output power ratio of the transmission line that is connected between the output of the final power amplifier stage of the transmitter and the antenna. It is given by

$$L = \frac{P_t}{P_{\text{ant}}} \tag{2.4.10}$$

Alternatively, the EIRP can be expressed in dBw as

$$\text{EIRP(dBw)} = P_t(\text{dBw}) - L(\text{dB}) + G(\text{dB}) \tag{2.4.11}$$

Example 2.4 In a transmitting system, the output of the final high-power amplifier is 500 W, and the line feeding its antenna has an attenuation of 20%. If the gain of the transmitting antenna is 60 dB, find the EIRP in dBw.

SOLUTION

$$P_t = 500\,\text{W} = 26.9897\,\text{dBw}$$
$$P_{\text{ant}} = 0.8 \times 500 = 400\,\text{W}$$
$$G = 60\,\text{dB} = 10^6$$

and

$$L = \frac{500}{400} = 1.25 = 10\log_{10}(1.25) = 0.9691\,\text{dB}$$

Hence,

$$\text{EIRP(dBw)} = 26.9897 - 0.9691 + 60 = 86.0206\,\text{dBw}$$

or

$$\text{EIRP} = \frac{500 \times 10^6}{1.25} = 400 \times 10^6\,\text{W}$$

Space Loss

The transmitting antenna radiates in all directions, depending on its radiation characteristics. However, the receiving antenna receives only the power that is incident on it. Hence, the rest of the power is not used and is lost in space. It is represented by the space loss. It can be determined as follows.

The power density w_t of a signal transmitted by an isotropic antenna is given by

$$w_t = \frac{P_t}{4\pi R^2} \qquad \text{W/m}^2 \tag{2.4.12}$$

where P_t is the transmitted power in watts and R is the distance from the antenna in meters. The power received by a unity-gain antenna located at R is found to be

$$P_r = w_t A_{\text{eu}} \tag{2.4.13}$$

where A_{eu} is the effective area of an isotropic antenna.

From (2.4.6), for an isotropic antenna,

$$G = \frac{4\pi}{\lambda^2} A_{\text{eu}} = 1$$

or

$$A_{\text{eu}} = \frac{\lambda^2}{4\pi}$$

Hence, (2.4.12) can be written as

$$P_r = \frac{P_\mathrm{t}}{4\pi R^2}\frac{\lambda^2}{4\pi} \tag{2.4.14}$$

and the space loss ratio is found to be

$$\frac{P_r}{P_t} = \left(\frac{\lambda}{4\pi R}\right)^2 \tag{2.4.15}$$

It is usually expressed in decibels as follows:

$$\text{space loss ratio} = 20\log_{10}\frac{\lambda}{4\pi R} \qquad \text{dB} \tag{2.4.16}$$

Example 2.5 A geostationary satellite is 35,860 km away from Earth's surface. Find the space loss ratio if it is operating at 4 GHz.

SOLUTION

$$R = 35{,}860{,}000\,\text{m} \quad \text{and} \quad \lambda = \frac{3\times 10^8}{4\times 10^9} = 0.075\,\text{m}$$

Hence,

$$\text{space loss ratio} = \left(\frac{4\pi\times 35{,}860{,}000}{0.075}\right)^2 = 2.77\times 10^{-20} = -195.5752\,\text{dB}$$

Friis Transmission Formula and Radar Range Equation

Analysis and design of communication and monitoring systems often require an estimation of transmitted and received powers. The Friis transmission formula and radar range equation provide the means for such calculations. The former is applicable to a one-way communication system where the signal is transmitted at one end and is received at the other end of the link. The radar range equation is applicable when the signal transmitted hits a target and the signal reflected is generally received at the location of the transmitter. We consider these two formulations next.

Friis Transmission Equation

Consider a simplified communication link as illustrated in Figure 2.5. A distance R separates the transmitter and the receiver. Effective apertures of transmitting and receiving antennas are A_{et} and A_{er}, respectively. Further, the two antennas

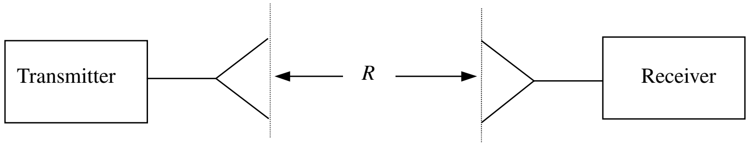


Figure 2.5 Simplified block diagram of a communication link.

are assumed to be polarization matched. If power input to the transmitting antenna is P_t, isotropic power density w_0 at a distance R from the antenna is given as

$$w_0 = \frac{P_t e_t}{4\pi R^2} \tag{2.4.17}$$

where e_t is the radiation efficiency of the transmitting antenna. For a directional transmitting antenna, the power density w_t can be written as

$$w_t = \frac{P_t G_t}{4\pi R^2} = \frac{P_t e_t D_t}{4\pi R^2} \tag{2.4.18}$$

where G_t is the gain and D_t is the directivity of transmitting antenna. The power collected by the receiving antenna is

$$P_r = A_{er} w_t \tag{2.4.19}$$

From (2.4.6),

$$A_{er} = \frac{\lambda^2}{4\pi} G_r \tag{2.4.20}$$

where the receiving antenna gain is G_r. Therefore, we find that

$$P_r = \frac{\lambda^2}{4\pi} G_r w_t = \frac{\lambda^2}{4\pi} G_r \frac{P_t G_t}{4\pi R^2}$$

or

$$\frac{P_r}{P_t} = \left(\frac{\lambda}{4\pi R}\right)^2 G_r G_t = e_r e_t \left(\frac{\lambda}{4\pi R}\right)^2 D_r D_t \tag{2.4.21}$$

If signal frequency is f, for a free-space link,

$$\frac{\lambda}{4\pi R} = \frac{3 \times 10^8}{4\pi f R} \tag{2.4.22}$$

where f is in hertz and R is in meters.

Generally, the link distance is long and the signal frequency is high, such that kilometer and megahertz will be more convenient units than the usual meter and hertz, respectively. For R in kilometers and f in megahertz, we find that

$$\frac{\lambda}{4\pi R} = \frac{3\times 10^8}{4\pi \times 10^6 f_{\mathrm{MHz}} \times 10^3 R_{\mathrm{km}}} = \frac{0.3}{4\pi}\frac{1}{f_{\mathrm{MHz}} R_{\mathrm{km}}} \tag{2.4.23}$$

Hence, from (2.4.21),

$$P_r(\mathrm{dBm}) = P_t(\mathrm{dBm}) + 20\log_{10}\frac{0.3}{4\pi} - 20\log_{10}(f_{\mathrm{MHz}} R_{\mathrm{km}}) + G_t(\mathrm{dB}) + G_r(\mathrm{dB})$$

or

$$P_r(\mathrm{dBm}) = P_t(\mathrm{dBm}) + G_t(\mathrm{dB}) + G_r(\mathrm{dB}) - 20\log_{10}(f_{\mathrm{MHz}} R_{\mathrm{km}}) - 32.4418 \tag{2.4.24}$$

where the power transmitted and the power received are in dBm while the two antenna gains are in decibels.

Example 2.6 A 20-GHz transmitter on board the satellite uses a parabolic antenna that is 45.7 cm in diameter. The antenna gain is 37 dB and its radiated power is 2 W. The ground station that is 36,941.031 km away from it has an antenna gain of 45.8 dB. Find the power collected by the ground station. How much power would be collected at the ground station if there were isotropic antennas on both sides?

SOLUTION The power transmitted, $P_t(\mathrm{dBm}) = 10\log_{10}(2000) = 33.0103$ dBm and

$$20\log_{10}(f_{\mathrm{MHz}} R_{\mathrm{km}}) = 20\log_{10}(20\times 10^3 \times 36{,}941.031) = 177.3708\ \mathrm{dB}$$

Hence, the power received at the Earth station is found as follows:

$$P_r(\mathrm{dBm}) = 33.0103 + 37 + 45.8 - 177.3708 - 32.4418 = -94.0023\ \mathrm{dBm}$$

or

$$P_r = 3.979 \times 10^{-10}\ \mathrm{mW}$$

If the two antennas are isotropic, $G_t = G_r = 1$ (or 0 dB) and therefore

$$\mathrm{P}_r(\mathrm{dBm}) = 33.0103 + 0 + 0 - 177.3708 - 32.4418 = -176.8023\ \mathrm{dBm}$$

or

$$P_r = 2.0882 \times 10^{-18}\ \mathrm{mW}$$

Radar Equation

In the case of a radar system, the signal transmitted is scattered by the target in all possible directions. The receiving antenna collects part of the energy that is scattered back toward it. Generally, a single antenna is employed for both the transmitter and the receiver, as shown in Fig. 2.6. If power input to the transmitting antenna is P_t and its gain is G_t, the power density w_{inc} incident on the target is

$$w_{\text{inc}} = \frac{P_t G_t}{4\pi R^2} = \frac{P_t A_{et}}{\lambda^2 R^2} \tag{2.4.25}$$

where A_{et} is the effective aperture of the transmitting antenna.

The radar cross section σ of an object is defined as the area intercepting that amount of power that when scattered isotropically produces at the receiver a power density that is equal to that scattered by the actual target. Hence,

$$\text{radar cross section} = \frac{\text{scattered power}}{\text{incident power density}} \qquad \text{square meters}$$

or

$$\sigma = \frac{4\pi r^2 w_r}{w_{\text{inc}}} \tag{2.4.26}$$

where w_r is isotropically backscattered power density at a distance r and w_{inc} is power density incident on the object. Hence, the radar cross section of an object is its effective area that intercepts an incident power density w_{inc} and gives an isotropically scattered power of $4\pi r^2 w_r$ for a backscattered power density. Radar cross sections of selected objects are listed in Table 2.5.

Using the radar cross section of a target, the power intercepted by it can be found as follows:

$$P_{\text{inc}} = \sigma w_{\text{inc}} = \frac{\sigma P_t G_t}{4\pi R^2} \tag{2.4.27}$$

Power density arriving back at the receiver is

$$w_{\text{scatter}} = \frac{P_{\text{inc}}}{4\pi R^2} \tag{2.4.28}$$

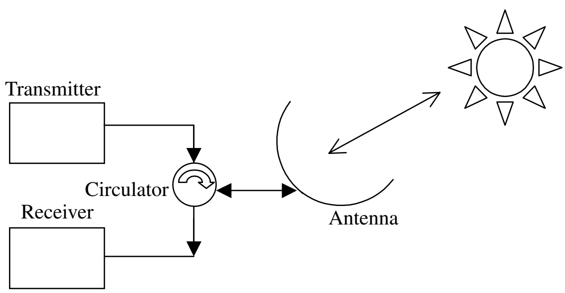


Figure 2.6 Radar system.

TABLE 2.5 Radar Cross Sections of Selected Objects

Object	Radar Cross Section (m^2)
Pickup truck	200
Automobile	100
Jumbo-jet airliner	100
Large bomber	40
Large fighter aircraft	6
Small fighter aircraft	2
Adult male	1
Conventional winged missile	0.5
Bird	0.01
Insect	0.00001
Advanced tactical fighter	0.000001

and power available at the receiver input is

$$P_r = A_{er} w_{\text{scatter}} = \frac{G_r \lambda^2 \sigma P_t G_t}{4\pi (4\pi R^2)^2} = \frac{\sigma A_{er} A_{et} P_t}{4\pi \lambda^2 R^4} \tag{2.4.29}$$

Example 2.7 A distance of 100λ separates two lossless X-band horn antennas (Figure 2.7). Reflection coefficients at the terminals of transmitting and receiving antennas are 0.1 and 0.2, respectively. Maximum directivities of the transmitting and receiving antennas are 16 and 20 dB, respectively. Assuming that the input power in a lossless transmission line connected to the transmitting antenna is 2 W and that the two antennas are aligned for maximum radiation between them and are polarization matched, find the power input to the receiver.

SOLUTION As discussed in Chapter 3, impedance discontinuity generates an echo signal very similar to that of an acoustical echo. Hence, signal power available beyond the discontinuity is reduced. The ratio of the reflected signal voltage to that of the incident is called the *reflection coefficient*. Since the power is proportional to the square of the voltage, the power reflected from the discontinuity is equal to the square of the reflection coefficient times the incident power.

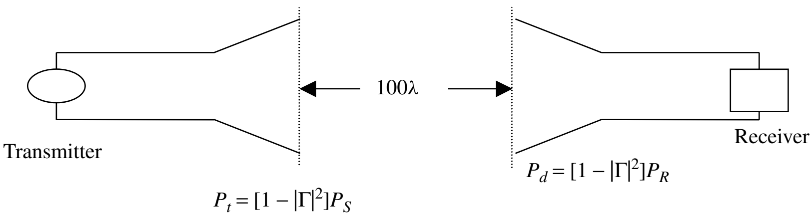


Figure 2.7 Setup for Example 2.7.

Therefore, power transmitted in the forward direction will be given by

$$P_t = (1 - |\Gamma|^2)P_{\text{in}}$$

and the power radiated by the transmitting antenna is found to be

$$P_t = (1 - 0.1^2)2 = 1.98\,\text{W}$$

Since the Friis transmission equation requires the antenna gain as a ratio instead of in decibels, G_t and G_r are calculated as follows:

$$G_t = 16\,\text{dB} = 10^{1.6} = 39.8107 \quad \text{and} \quad G_r = 20\,\text{dB} = 10^{2.0} = 100$$

Hence, from (2.4.21),

$$P_r = \left(\frac{\lambda}{4\pi \times 100\lambda}\right)^2 \times 100 \times 39.8107 \times 1.98$$

or

$$P_r = 5\,\text{mW}$$

and power delivered to the receiver, P_d, is

$$P_d = (1 - 0.2^2)5 = 4.8\,\text{mW}$$

Example 2.8 Radar operating at 12 GHz transmits 25 kW through an antenna of 25 dB gain. A target with its radar cross section at 8 m^2 is located at 10 km from the radar. If the same antenna is used for the receiver, determine the power received.

SOLUTION

$$P_t = 25\,\text{kW}$$

$$f = 12\,\text{GHz} \to \lambda = \frac{3 \times 10^8}{12 \times 10^9} = 0.025\,\text{m}$$

$$G_r = G_t = 25\,\text{dB} \to 10^{2.5} = 316.2278$$

$$R = 10\,\text{km} \qquad \sigma = 8\,\text{m}^2$$

Hence,

$$P_r = \frac{G_r G_t P_t \sigma \lambda^2}{4\pi(4\pi R^2)^2} = \frac{316.2278^2 \times 25,000 \times 8 \times 0.025^2}{(4\pi)^3(10^4)^4} = 6.3 \times 10^{-13}\,\text{W}$$

or

$$P_r = 0.63\,\text{pW}$$

Doppler Radar

An electrical signal propagating in free space can be represented by a simple expression as follows:

$$v(z,t) = A\cos(\omega t - kz) \tag{2.4.30}$$

The signal frequency is ω radians per second and k is its wave number (equal to ω/c, where c is the speed of light in free space) in radians per meter. Assume that there is a receiver located at $z = R$, as shown in Figure 2.5 and R is changing with time (the receiver may be moving toward or away from the transmitter). In this situation, the receiver response $v_o(t)$ is given as follows:

$$v_o(t) = V\cos(\omega t - kR) \tag{2.4.31}$$

The angular frequency, ω_0, of v_o(t) can be determined easily after differentiating the argument of the cosine function with respect to time. Hence,

$$\omega_0 = \frac{d}{dt}(\omega t - kR) = \omega - k\frac{dR}{dt} \tag{2.4.32}$$

Note that k is time independent and that the time derivative of R represents the velocity, v_r, of the receiver with respect to the transmitter. Hence, (2.4.32) can be written

$$\omega_0 = \omega - \frac{\omega v_\mathrm{r}}{c} = \omega\left(1 - \frac{v_r}{c}\right) \tag{2.4.33}$$

If the receiver is closing in, v_r will be negative (negative slope of R), and therefore the signal received will indicate a signal frequency higher than ω. On the other hand, it will show a lower frequency if R is increasing with time. It is the Doppler frequency shift that is employed to design the Doppler radar.

Consider the simplified block diagram shown in Figure 2.8. A microwave signal generated by the oscillator is split into two parts via a power divider. The circulator feeds one part of this power to an antenna that illuminates a target while a mixer uses the remaining fraction as its reference signal. Further, the antenna intercepts a part of the signal that is scattered by an object. It is then directed to the mixer through a circulator. The mixer output includes a difference frequency signal that can be filtered out for further processing. Two inputs to the mixer will have the same frequency if the target is stationary, and therefore the Doppler shift $\delta\omega$ will be zero. On the other hand, the mixer output will have Doppler frequency if the target is moving. Note that the signal travels twice over the same distance, and therefore the Doppler frequency shift in this case will be twice that found via (2.4.33). Mathematically,

$$\omega_0 = \omega\left(1 - \frac{2v_r}{c}\right) \tag{2.4.34}$$

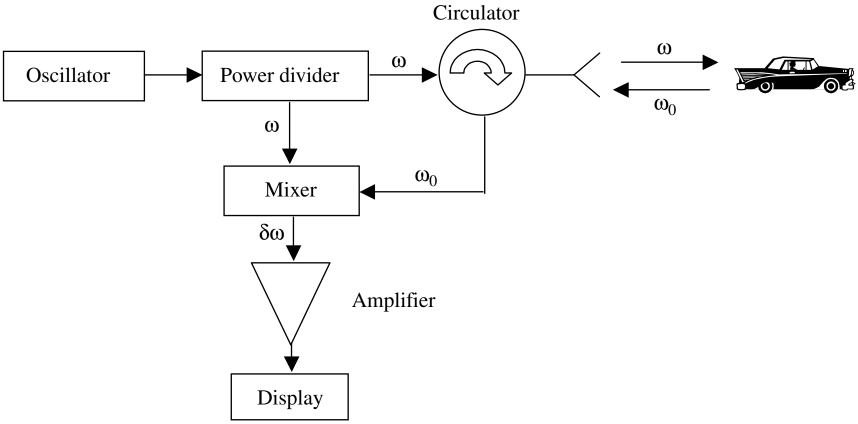


Figure 2.8 Simplified block diagram of a Doppler radar.

and

$$\delta\omega = \frac{2\omega v_r}{c} \tag{2.4.35}$$

2.5 NOISE AND DISTORTION

Random movement of charges or charge carriers in an electronic device generates currents and voltages that vary randomly with time. In other words, the amplitude of these electrical signals cannot be predicted at any time. However, it can be expressed in terms of probability density functions. These signals are termed *noise*. For most applications it suffices to know the mean-square or root-mean-square value. Since the square of the current or the voltage is proportional to the power, mean-square noise voltage and current values are generally called *noise power*. Further, noise power is normally a function of frequency and the power per unit frequency (watts per hertz) is defined as the power spectral density of noise. If the noise power is the same over the entire frequency band of interest, it is called *white noise*. There are several mechanisms that can cause noise in an electronic device. Some of these are as follows:

- *Thermal noise.* This is the most basic type of noise, which is caused by thermal vibration of bound charges. Johnson studied this phenomenon in 1928, and Nyquist formulated an expression for spectral density around the same time. Therefore, it is also known as *Johnson noise* or *Nyquist noise*. In most electronic circuits, thermal noise dominates; therefore, it will be described further because of its importance.
- *Shot noise.* This is due to random fluctuations of charge carriers that pass through the potential barrier in an electronic device. For example, electrons

emitted from the cathode of thermionic devices or charge carriers in Schottky diodes produce a current that fluctuates about the average value I. The mean-square current due to shot noise is generally given by

$$\langle i_{\mathrm{Sh}}^2 \rangle = 2eIB \tag{2.5.1}$$

where e is electronic charge (1.602×10^{-19} C) and B is the bandwidth in hertz.

- *Flicker noise.* This occurs in solid-state devices and vacuum tubes operating at low frequencies. Its magnitude decreases with an increase in frequency. It is generally attributed to chaos in the dynamics of a system. Since the flicker noise power varies inversely with frequency, it is often called *1/f noise*. Sometimes it is referred to as *pink noise*.

Thermal Noise

Consider a resistor R that is at a temperature of T Kelvin. Electrons in this resistor are in random motion with a kinetic energy that is proportional to the temperature T. These random motions produce small, random voltage fluctuations across its terminals. This voltage has a zero average value but a nonzero mean-square value $\langle v_n^2 \rangle$. It is given by Planck's blackbody radiation law as

$$\langle v_n^2 \rangle = \frac{4hfRB}{\exp(hf/kT) - 1} \tag{2.5.2}$$

where h is Planck's constant (6.546×10^{-34} J · s), k the Boltzmann constant (1.38×10^{-23} J/K), T the temperature in kelvin, B the system bandwidth in hertz, and f the center frequency of the bandwidth in hertz.

For frequencies below 100 GHz, the product hf will be smaller than 6.546×10^{-23} J and kT will be greater than 1.38×10^{-22} J if T stays above 10 K. Therefore, kT will be larger than hf for such cases. Hence, the exponential term in equation (2.5.2) can be approximated as follows:

$$\exp\left(\frac{hf}{kT}\right) \approx 1 + \frac{hf}{kT}$$

Therefore,

$$\langle v_n^2 \rangle \approx \frac{4hfRB}{hf/kT} = 4BRkT \tag{2.5.3}$$

This is known as the *Rayleigh–Jeans approximation.*

A Thévenin-equivalent circuit can replace the noisy resistor as shown in Figure 2.9. As illustrated, it consists of a noise-equivalent voltage source in series with a noise-free resistor. This source will supply a maximum power to a load

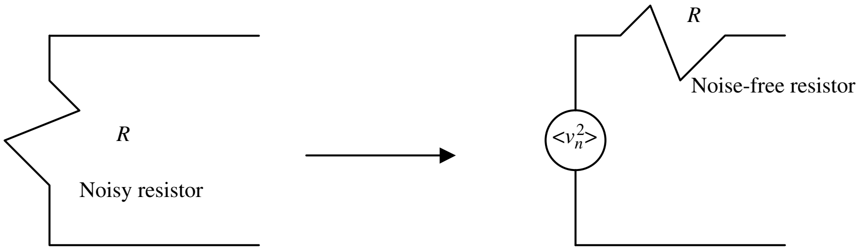


Figure 2.9 Noise-equivalent circuit of a resistor.

of resistance R. The power delivered to that load in a bandwidth B is found as follows:

$$P_n = \frac{\langle v_n^2 \rangle}{4R} = kTB \tag{2.5.4}$$

Conversely, if an arbitrary white noise source with its driving point impedance R delivers a noise power P_s to a load R, it can be represented by a noisy resistor of value R that is at temperature T_e. Hence,

$$T_e = \frac{P_s}{kB} \tag{2.5.5}$$

where T_e is an equivalent temperature selected so that the same noise power is delivered to the load.

Consider the noisy amplifier as shown in Figure 2.10. Its gain is G over the bandwidth B. Let the amplifier be matched to the noiseless source and the load resistors. If the source resistor is at a hypothetical temperature of $T_s = 0\,\mathrm{K}$, the power input to the amplifier P_i will be zero and the output noise power P_o will only be due to noise generated by the amplifier. We can obtain the same noise power at the output of an ideal noiseless amplifier by raising the temperature T_s of the source resistor to T_e:

$$T_e = \frac{P_o}{GkB} \tag{2.5.6}$$

Hence, the output power in both cases is $P_o = GkT_eB$. The temperature T_e is known as the equivalent noise temperature of the amplifier.

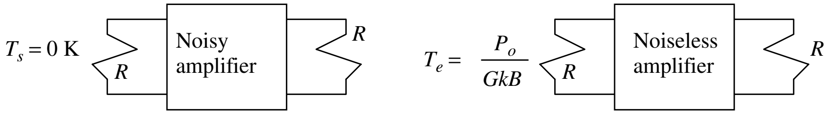


Figure 2.10 Noise-equivalent representation of an amplifier.

Measurement of Noise Temperature by the Y-Factor Method

According to the definition, the noise temperature of an amplifier (or any other two-port network) can be determined by setting the source resistance R at 0 K and then measuring the output noise power. However, a temperature of 0 K cannot be achieved in practice. We can circumvent this problem by repeating the experiment at two different temperatures. This procedure is known as the *Y-factor method.*

Consider an amplifier with a power gain of G over a frequency band of B hertz. Further, its equivalent noise temperature is T_e(K). The input port of the amplifier is terminated by a matched resistor R while a matched power meter is connected at its output, as illustrated in Figure 2.11. With R at temperature T_h, the power meter measures the noise output as P_1. Similarly, the noise power is found to be P_2 when the temperature of R is set at T_c. Hence,

$$P_1 = GkT_hB + GkT_eB$$

and

$$P_2 = GkT_cB + GkT_eB$$

For T_h higher than T_c, the noise power P_1 will be larger than P_2. Therefore,

$$\frac{P_1}{P_2} = Y = \frac{T_h + T_e}{T_c + T_e}$$

or

$$T_e = \frac{T_h - YT_c}{Y - 1} \tag{2.5.7}$$

For T_h larger than T_c, Y will be greater than unity. Further, measurement accuracy is improved by selecting two temperature settings that are far apart. Therefore, T_h and T_c represent hot and cold temperatures, respectively.

Example 2.9 An amplifier has a power gain of 10 dB in the 500-MHz to 1.5-GHz frequency band. The following data are obtained for this amplifier using the

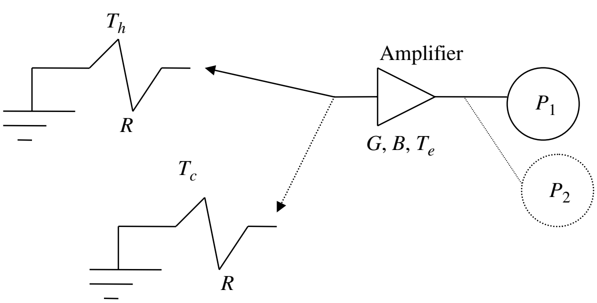


Figure 2.11 Experimental setup for measurement of noise temperature.

Y-factor method: At $T_h = 290\,\text{K}$, $P_1 = -70\,\text{dBm}$; at $T_c = 77\,\text{K}$, $P_2 = -75\,\text{dBm}$. Determine its equivalent noise temperature. If this amplifier is used with a source that has an equivalent noise temperature of 450 K, find the output noise power in dBm.

SOLUTION Since P_1 and P_2 are given in dBm, the difference of these two values will give Y in decibels. Hence,

$$Y = (P_1 - P_2)\ \text{dBm} = (-70) - (-75) = 5\,\text{dB}$$

or

$$Y = 10^{0.5} = 3.1623$$

Therefore,

$$T_e = \frac{290 - (3.1623)(77)}{3.1623 - 1} = 21.51\,\text{K}$$

If a source with an equivalent noise temperature of $T_s = 450\,\text{K}$ drives the amplifier, the noise power input to this will be kT_sB. The total noise power at the output of the amplifier will be

$$\begin{aligned} P_o &= GkT_sB + GkT_eB = 10 \times 1.38 \times 10^{-23} \times 10^9 \times (450 + 21.51) \\ &= 6.5068 \times 10^{-11}\,\text{W} \end{aligned}$$

Therefore,

$$P_o = 10\log(6.5068 \times 10^{-8}) = -71.8663\,\text{dBm}$$

Noise Factor and Noise Figure

The noise factor of a two-port network is obtained by dividing the signal-to-noise ratio at its input port by the signal-to-noise ratio at its output. Hence,

$$\text{noise factor, } F = \frac{S_i/N_i}{S_o/N_o}$$

where S_i, N_i, S_o, and N_o represent the power in input signal, input noise, output signal, and output noise, respectively. If the two-port network is noise-free, the signal-to-noise ratio at its output will be the same as its input, resulting in a noise factor of unity. In reality, the network will add its own noise, while the input signal and noise will be altered by the same factor (gain or loss). It will lower the output signal/noise ratio, resulting in a higher noise factor. Therefore, the noise factor of a two-port network is generally greater than unity. By definition, the input noise power is assumed to be the noise power resulting from a matched resistor at $T_0 = 290\,\text{K}$ (i.e., $N_i = kT_0B$). Using the circuit arrangement illustrated

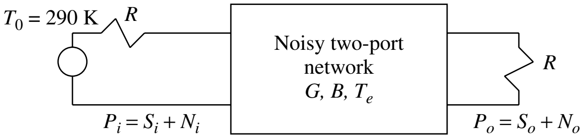


Figure 2.12 Circuit arrangement for determination of the noise factor of a noisy two-port network.

in Figure 2.12, the noise factor of a noisy two-port network can be defined as follows:

$$\text{noise factor} = \frac{\text{total output noise in } B \text{ when input source temperature is } 290\,\text{K}}{\text{output noise of source (only) at } 290\,\text{K}}$$

or

$$F = \frac{kT_0BG + P_{\text{int}}}{kT_0BG} = 1 + \frac{P_{\text{int}}}{kT_0BG} \qquad (2.5.8)$$

where P_{int} represents the output noise power that is generated by the two-port network internally. It can be expressed in terms of noise factor as follows:

$$P_{\text{int}} = k(F - 1)T_0BG = kT_eBG \qquad (2.5.9)$$

where T_e is known as the equivalent noise temperature of a two-port network. It is related to the noise factor as follows:

$$T_e = (F - 1)T_0 \qquad (2.5.10)$$

When the noise factor is expressed in decibels, it is commonly called the *noise figure* (NF). Hence,

$$\text{NF} = 10\log_{10} F \qquad \text{dB} \qquad (2.5.11)$$

Example 2.10 The power gain of an amplifier is 20 dB in the frequency band 10 to 12 GHz. If its noise figure is 3.5 dB, find the output noise power in dBm.

SOLUTION

$$\text{Output noise power} = kT_0BG + P_{\text{int}} = FkT_0BG = N_o$$
$$F = 3.5\,\text{dB} = 10^{0.35} = 2.2387$$
$$G = 20\,\text{dB} = 10^2 = 100$$

Therefore,

$$N_o = 2.2387 \times 1.38 \times 10^{-23} \times 290 \times 2 \times 10^9 \times 100 = 1.7919 \times 10^{-9}\,\text{W}$$

or

$$N_o = 10\log_{10}(1.7919 \times 10^{-6}\,\text{mW})\ \text{dBm} = -57.4669\,\text{dBm} \approx -57.5\,\text{dBm}$$

Noise in Two-Port Networks

Consider the noisy two-port network shown in Figure 2.13(a). $Y_s = G_s + jB_s$ is the source admittance connected at port 1 of the network. The noise it generates is represented by the current source i_s with a root-mean-square value of I_s. This noisy two-port can be replaced by a noise-free network with a current source i_n and a voltage source v_n connected at its input, as shown in Figure 2.13(b). I_n and V_n represent the corresponding root-mean-square current and voltage of the noise. It is assumed that the noise represented by i_s is uncorrelated with that represented by i_n and v_n. However, a part of i_n, i_{nc}, is assumed to be correlated with v_n via the correlation admittance $Y_c = G_c + jX_c$ while the remaining part i_{nu} is uncorrelated. Hence,

$$I_s^2 = \langle i_s^2 \rangle = 4kTBG_S \tag{2.5.12}$$

$$V_n^2 = \langle v_n^2 \rangle = 4kTBR_n \tag{2.5.13}$$

and

$$I_{nu}^2 = \langle i_{nu}^2 \rangle = 4kTBG_{nu} \tag{2.5.14}$$

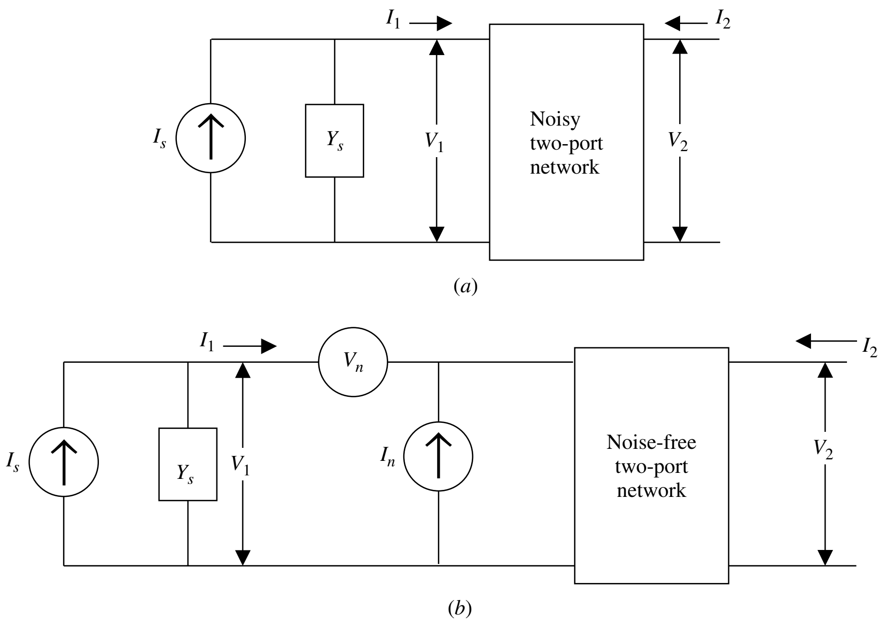


Figure 2.13 Noisy two-port network (a) and its equivalent circuit (b).

Now we find a Norton equivalent for the circuit that is connected at the input of the noise-free two-port network shown in Figure 2.13(*b*). Since these are random variables, the mean-square noise current $< i_{\text{eq}}^2 >$ is found as follows:

$$\langle i_{\text{eq}}^2 \rangle = \langle i_s^2 \rangle + \langle |i_n + Y_s v_n|^2 \rangle = \langle i_s^2 \rangle + \langle |i_{nu} + i_{nc} + Y_s v_n|^2 \rangle$$
$$= \langle i_s^2 \rangle + \langle |i_{nu} + (Y_c + Y_s) v_n|^2 \rangle$$

or

$$\langle i_{\text{eq}}^2 \rangle = \langle i_s^2 \rangle + \langle i_{nu}^2 \rangle + |Y_c + Y_n|^2 \langle v_n^2 \rangle \tag{2.5.15}$$

Hence, the noise factor F is

$$F = \frac{\langle i_{\text{eq}}^2 \rangle}{\langle i_s^2 \rangle} = 1 + \frac{\langle i_{nu}^2 \rangle}{\langle i_s^2 \rangle} + |Y_c + Y_s|^2 \frac{\langle v_n^2 \rangle}{\langle i_s^2 \rangle} = 1 + \frac{G_{nu}}{G_s} + |Y_c + Y_s|^2 \frac{R_n}{G_s}$$

or

$$F = 1 + \frac{G_{nu}}{G_s} + \frac{R_n}{G_s}[(G_s + G_c)^2 + (X_s + X_c)^2] \tag{2.5.16}$$

For a minimum noise factor, F_{min},

$$\frac{\partial F}{\partial G_s} = 0 \Rightarrow G_s^2 = G_c^2 + \frac{G_{nu}}{R_n} = G_{\text{opt}}^2 \tag{2.5.17}$$

and

$$\frac{\partial F}{\partial X_s} = 0 \Rightarrow X_s = -X_c = X_{\text{opt}} \tag{2.5.18}$$

From (2.5.17),

$$G_{nu} = R_n (G_{\text{opt}}^2 - G_c^2) \tag{2.5.19}$$

Substituting (2.5.18) and (2.5.19) into (2.5.16), the minimum noise factor is found as

$$F_{\text{min}} = 1 + 2R_n (G_{\text{opt}} + G_c) \tag{2.5.20}$$

Using (2.5.18) to (2.5.20), (2.5.16) can be expressed as

$$F = F_{\text{min}} + \frac{R_n}{G_s} |Y_s - Y_{\text{opt}}|^2 \tag{2.5.21}$$

Noise Figure of a Cascaded System

Consider a two-port network with gain G_1, noise factor F_1, and equivalent noise temperature T_{e1}. It is connected in cascade with another two-port network, as shown in Figure 2.14(*a*). The second two-port network has gain G_2, noise factor

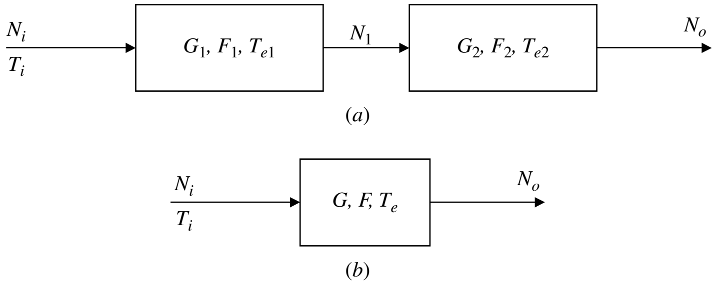


Figure 2.14 Two networks connected in cascade (*a*) and its equivalent system (*b*).

F_2, and noise temperature T_{e2}. Our goal is to find the noise factor F and equivalent noise temperature T_e of the overall system illustrated in Figure 2.14(*b*).

Assume that the noise power input to the first two-port network is N_i. Its equivalent noise temperature is T_i. The output noise power of the first system is N_1, whereas it is N_o after the second system. Hence,

$$N_1 = G_1 k T_i B + G_1 k T_{e1} B \tag{2.5.22}$$

and

$$\begin{aligned} N_o &= G_2 N_1 + G_2 k T_{e2} B = G_1 G_2 k B (T_i + T_{e1} + T_{e2}/G_1) \\ &= G_2[G_1 k B (T_i + T_{e1})] + G_2 k T_{e2} B \end{aligned}$$

or

$$N_o = G_1 G_2 k B (T_e + T_i) = G k B (T_e + T_i) \tag{2.5.23}$$

Therefore, the noise temperature of a cascaded system is

$$T_e = T_{e1} + \frac{T_{e2}}{G_1}$$

and from $T_{e1} = (F_1 - 1)T_i$, $T_{e2} = (F_2 - 1)T_i$, and $T_e = (F - 1)T_i$, we get

$$F = F_1 + \frac{F_2 - 1}{G_1}$$

The equations for T_e and F above can be generalized as follows:

$$T_e = T_{e1} + \frac{T_{e2}}{G_1} + \frac{T_{e3}}{G_1 G_2} + \frac{T_{e4}}{G_1 G_2 G_3} + \cdots \tag{2.5.24}$$

and

$$F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \frac{F_4 - 1}{G_1 G_2 G_3} + \cdots \tag{2.5.25}$$

Example 2.11 A receiving antenna is connected to an amplifier through a transmission line that has an attenuation of 2 dB. The gain of the amplifier is 15 dB, and its noise temperature is 150 K over a bandwidth of 100 MHz. All the components are at an ambient temperature of 300 K.

(**a**) Find the noise figure of this cascaded system.

(**b**) What would be the noise figure if the amplifier were placed before the transmission line?

SOLUTION First we need to determine the noise factor of the transmission line alone. The formulas derived in the preceding section can then provide the noise figures desired for the two cases. Consider a transmission line that is match-terminated at both its ends by resistors R_0, as illustrated in Figure 2.15. Since the entire system is in thermal equilibrium at T(K), noise powers delivered to the transmission line and available at its output are kTB. Mathematically,

$$P_o = kTB = GkTB + GN_{\text{added}}$$

where N_{added} is noise generated by the line as it appears at its input terminals. G is an output-to-input power ratio and B is the bandwidth of the transmission line. Note that input noise is attenuated in a lossy transmission line, but there is noise generated by it as well. Hence,

$$N_{\text{added}} = \frac{1}{G}(1-G)kTB = \left(\frac{1}{G}-1\right)kTB = kT_eB$$

An expression for the equivalent noise temperature T_e of the transmission line is found from it as follows:

$$T_e = \left(\frac{1}{G}-1\right)T$$

and

$$F = 1 + \frac{T_e}{T_0} = 1 + \left(\frac{1}{G}-1\right)\frac{T}{T_0}$$

The gain of the amplifier, $G_{\text{amp}} = 15\,\text{dB} = 10^{1.5} = 31.6228$. For the transmission line, $1/G = 10^{0.2} = 1.5849$. Hence, the noise factor of the line is $1 + (1.5849 - 1)300/290 = 1.6051$. The corresponding noise figure is 2.05 dB. Similarly, the

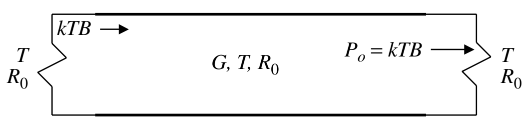


Figure 2.15 Lossy transmission line that is matched terminated at its ends.

noise factor of the amplifier is found to be $1 + 150/290 = 1.5172$. Its noise figure is 1.81 dB.

(**a**) In this case the noise figure of the cascaded system, F_{cascaded}, is found as

$$F_{\text{cascaded}} = F_{\text{line}} + \frac{F_{\text{amp}} - 1}{G_{\text{line}}} = 1.6051 + \frac{1.5172 - 1}{1/1.5849} = 2.4248 = 3.8468\,\text{dB}$$

(**b**) If the amplifier is connected before the line (i.e., the amplifier is placed right at the antenna),

$$F_{\text{cascaded}} = F_{\text{amp}} + \frac{F_{\text{line}} - 1}{G_{\text{amp}}} = 1.5172 + \frac{1.6051 - 1}{31.6228} = 1.5363 = 1.8649\,\text{dB}$$

Note that the amplifier alone has a noise figure of 1.81 dB. Hence, the noisy transmission line connected after it does not alter the noise figure significantly.

Example 2.12 Two amplifiers, each with a 20-dB gain, are connected in cascade as shown in Figure 2.16. The noise figure of amplifier A_1 is 3 dB, while that of A_2 is 5 dB. Calculate the overall gain and noise figure for this arrangement. If the order of two amplifiers is changed in the system, find its resulting noise figure.

SOLUTION The noise factors and gains of two amplifiers are

$$F_1 = 3\,\text{dB} = 10^{0.3} = 2$$
$$F_2 = 5\,\text{dB} = 10^{0.5} = 3.1623$$
$$G_1 = G_2 = 20\,\text{dB} = 10^2 = 100$$

Therefore, the overall gain and noise figure of the cascaded system is found as follows:

$$G = \frac{P_3}{P_1} = \frac{P_3}{P_2} \times \frac{P_2}{P_1} = 100 \times 100 = 10,000 = 40\,\text{dB}$$

and

$$F = F_1 + \frac{F_2 - 1}{G_1} = 2 + \frac{3.1623 - 1}{100} = 2.021623 = 3.057\,\text{dB}$$

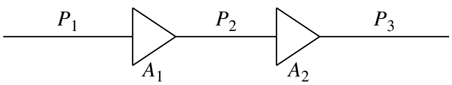


Figure 2.16 Setup for Example 2.12.

If the order of amplifiers is changed, the overall gain will stay the same. However, the noise figure of the new arrangement will change as follows:

$$F = F_2 + \frac{F_1 - 1}{G_2} = 3.1623 + \frac{2 - 1}{100} = 3.1723 = 5.013743\,\text{dB}$$

Minimum Detectable Signal

Consider a receiver circuit with gain G over a bandwidth B. Assume that its noise factor is F. P_I and P_o represent the signal power at its input and output ports, respectively. N_I is the input noise power and N_o is the total noise power at its output, as illustrated in Figure 2.17. Hence,

$$N_o = kT_0FBG \tag{2.5.26}$$

This constitutes the noise floor of the receiver. A signal weaker than this will be lost in noise. N_o can be expressed in dBW as follows:

$$N_o(\text{dBW}) = 10\log_{10} kT_\text{o} + F(\text{dB}) + 10\log_{10} B + G(\text{dB}) \tag{2.5.27}$$

The minimum detectable signal must have power higher than this. Generally, it is taken as 3 dB above this noise floor. Further,

$$10\log_{10} kT_0 = 10\log_{10}(1.38 \times 10^{-23} \times 290) \approx -204\,\text{dBW/Hz}$$

Hence, the minimum detectable signal $P_{o\text{MDS}}$ at the output is

$$P_{o\text{MDS}} = -201 + F(\text{dB}) + 10\log_{10} B_\text{Hz} + G(\text{dB}) \tag{2.5.28}$$

The corresponding signal power $P_{I\text{MDS}}$ at its input is

$$P_{I\text{MDS}}(\text{dBW}) = -201 + F(\text{dB}) + 10\log_{10} B_\text{Hz} \tag{2.5.29}$$

Alternatively, $P_{I\text{MDS}}$ can be expressed in dBm as follows:

$$P_{I\text{MDS}}(\text{dBm}) = -111 + F(\text{dB}) + 10\log_{10} B_\text{MHz} \tag{2.5.30}$$

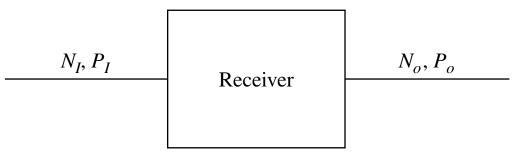


Figure 2.17 Signals at the two ports of a receiver with a noise figure of F decibels.

Example 2.13 The noise figure of a communication receiver is found as 10 dB at room temperature (290 K). Determine the minimum detectable signal power if (**a**) $B = 1\,\text{MHz}$, (**b**) $B = 1\,\text{GHz}$, (**c**) $B = 10\,\text{GHz}$, and (**d**) $B = 1\,\text{kHz}$.

SOLUTION From (2.5.30),

$$P_{I\text{MDS}} = -111 + F(\text{dB}) + 10\log_{10} B_{\text{MHz}}\ \text{dBm}$$

Hence,

(**a**) $P_{I\text{MDS}} = -111 + 10 + 10\log_{10}(1) = -101\,\text{dBm} = 7.94 \times 10^{-11}\,\text{mW}$

(**b**) $P_{I\text{MDS}} = -111 + 10 + 10\log_{10}(10^3) = -71\,\text{dBm} = 7.94 \times 10^{-8}\,\text{mW}$

(**c**) $P_{I\text{MDS}} = -111 + 10 + 10\log_{10}(10^4) = -61\,\text{dBm} = 7.94 \times 10^{-7}\,\text{mW}$

(**d**) $P_{I\text{MDS}} = -111 + 10 + 10\log_{10}(10^{-3}) = -131\,\text{dBm} = 7.94 \times 10^{-14}\,\text{mW}$

These results show that the receiver can detect a relatively weak signal when its bandwidth is narrow.

Intermodulation Distortion

The electrical noise of a system determines the minimum signal level that it can detect. On the other hand, the signal will be distorted if its level is too high. This occurs because of the nonlinear characteristics of electrical devices, such as diodes, transistors, and so on. In this section we analyze the distortion characteristics and introduce the associated terminology.

Consider the nonlinear system illustrated in Figure 2.18. Assume that its nonlinearity is frequency independent and can be represented by the following power series:

$$v_o = k_1 v_i + k_2 v_i^2 + k_3 v_i^3 + \cdots \tag{2.5.31}$$

For simplicity, we assume that the k_i are real and the first three terms of this series are sufficient to represent its output signal. Further, it is assumed that the input signal has two different frequency components that can be expressed as follows:

$$v_i = a\cos\omega_1 t + b\cos\omega_2 t \tag{2.5.32}$$

Therefore, the corresponding output signal can be written as

$$\begin{aligned} v_o \approx\ & k_1(a\cos\omega_1 t + b\cos\omega_2 t) + k_2(a\cos\omega_1 t + b\cos\omega_2 t)^2 \\ & + k_3(a\cos\omega_1 t + b\cos\omega_2 t)^3 \end{aligned}$$

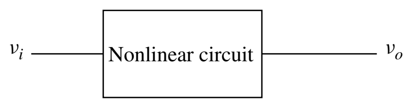


Figure 2.18 Nonlinear circuits with input signal v_i that produces v_o at its output.

After simplifying and rearranging it, we get

$$v_o = k_1(a\cos\omega_1 t + b\cos\omega_2 t) + k_2\left[\frac{a^2}{2}(1+\cos 2\omega_1 t) + \frac{b^2}{2}(1+\cos 2\omega_2 t)\right.$$
$$\left. + \frac{ab}{2}\{\cos(\omega_1+\omega_2)t + \cos(\omega_1-\omega_2 t)\}\right]$$
$$+ k_3\begin{bmatrix} \frac{3}{4}a^3\cos\omega_1 t + \frac{3}{2}ab^2\cos\omega_1 t + \frac{a^3}{4}\cos 3\omega_1 t \\ + \frac{3}{4}ab^2\cos(\omega_1 - 2\omega_2)t + \frac{3}{4}a^2 b\cos(2\omega_1-\omega_2)t \\ + \frac{3}{2}a^2 b\cos\omega_2 t + \frac{3}{4}b^3\cos\omega_2 t + \frac{b^3}{4}\cos 3\omega_2 t \\ + \frac{3}{4}a^2 b\cos(2\omega_1+\omega_2)t + \frac{3}{4}ab^2\cos(\omega_1+2\omega_2)t \end{bmatrix} \quad (2.5.33)$$

Therefore, the output signal has several frequency components in its spectrum. Amplitudes of various components are listed in Table 2.6.

Figure 2.19 illustrates the input–output characteristic of an amplifier. If the input signal is too low, it may be submerged under the noise. The output power rises linearly above the noise as the input is increased. However, it deviates from the linear characteristic after a certain level of input power. In the linear region, output power can be expressed in dBm as follows:

$$P_{\text{out}}(\text{dBm}) = P_{\text{in}}(\text{dBm}) + G(\text{dB})$$

The input power for which output deviates by 1 dB below its linear characteristic is known as 1-dB compression point. In this figure, it occurs at an input power

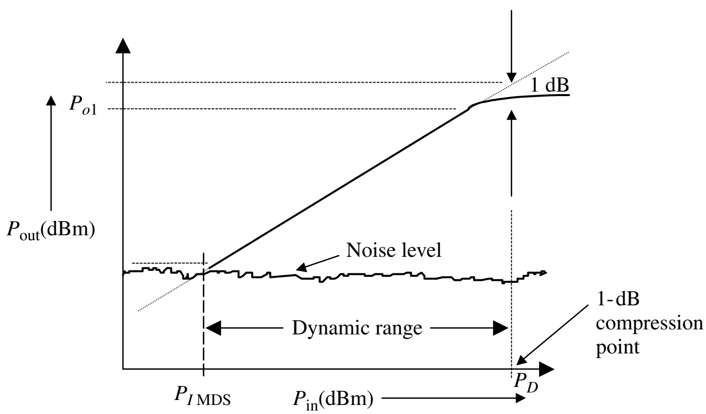


Figure 2.19 Gain characteristics of an amplifier.

TABLE 2.6 Amplitudes of Various Harmonics in the Output

Harmonic Components	Amplitude
ω_1	$k_1 a + k_3\left(\frac{3}{4}a^3 + \frac{3}{2}ab^2\right)$
ω_2	$k_1 b + k_3\left(\frac{3}{4}b^3 + \frac{3}{2}a^2 b\right)$
$\omega_1 - \omega_2$	$k_2\frac{ab}{2}$
$\omega_1 + \omega_2$	$k_2\frac{ab}{2}$
$2\omega_1$	$k_2\frac{a^2}{2}$
$2\omega_2$	$k_2\frac{b^2}{2}$
$3\omega_1$	$k_3\frac{a^3}{4}$
$3\omega_2$	$k_3\frac{b^3}{4}$
$2\omega_1 - \omega_2$	$\frac{3}{4}k_3 a^2 b$
$\omega_1 - 2\omega_2$	$\frac{3}{4}k_3 ab^2$
$2\omega_1 + \omega_2$	$\frac{3}{4}k_3 a^2 b$
$\omega_1 + 2\omega_2$	$\frac{3}{4}k_3 ab^2$

of P_D(dBm) that produces an output of P_{o1}(dBm). From the relation above, we find that

$$P_{o1}(\text{dBm}) + 1 = P_D(\text{dBm}) + G(\text{dB})$$

or

$$P_D(\text{dBm}) = P_{o1}(\text{dBm}) + 1 - G(\text{dB}) \tag{2.5.34}$$

The difference between the input power at 1-dB compression point and the minimum detectable signal defines the *dynamic range* (DR). Hence,

$$\text{DR} = P_D(\text{dBm}) - P_{I\text{MDS}}$$

From (2.5.30) and (2.5.34), we find that

$$\text{DR} = P_{o1}(\text{dBm}) + 112 - G(\text{dB}) - F(\text{dB}) - 10\log_{10} B_{\text{MHz}} \tag{2.5.35}$$

Gain Compression

Nonlinear characteristics of the circuit (amplifier, mixer, etc.) compress its gain. If there is only one input signal [i.e., b is zero in (2.5.32)], the amplitude a_1 of $\cos\omega_1 t$ in its output is found from Table 2.6 to be

$$a_1 = k_1 a + \tfrac{3}{4}k_3 a^3 \tag{2.5.36}$$

The first term of a_1 represents the linear (ideal) case, while its second term results from the nonlinearity. At the 1-dB compression point,

$$\begin{aligned}\left(k_1 a + \tfrac{3}{4}k_3 a^3\right)\Big|_{\mathrm{dB}} = k_1 a|_{\mathrm{dB}} - 1 \rightarrow 20\log\frac{k_1 a + \frac{3}{4}k_3 a^3}{k_1 a}\\ = -1 = -20\log\,(1.122)\end{aligned}$$

or

$$20\log\left(\frac{k_1 + \frac{3}{4}k_3 a^2}{k_1}\times 1.122\right) = 0 \rightarrow \frac{k_1 + \frac{3}{4}k_3 a^2}{k_1}\times 1.122 = 1$$

Therefore,

$$k_3 = -0.145\frac{k_1}{a^2} \tag{2.5.37}$$

This indicates that k_3 is a negative constant for a positive k_1, or vice versa. Therefore, it tends to reduce a_1, resulting in a lower gain. The single-tone gain compression factor may be defined as follows:

$$A_{1C} = \frac{a_1}{k_1 a} = 1 + \frac{3k_3}{4k_1}a^2 \tag{2.5.38}$$

Let us now consider the case when both of the input signals are present in (2.5.32). The amplitude of $\cos\omega_1 t$ in the output now becomes $k_1 a + k_3\left(\frac{3}{4}a^3 + \frac{3}{2}ab^2\right)$. If b is large compared with a, the term with k_3 may dominate (undesired) over the first (desired) one. Since k_3 and k_1 have opposite signs, the desired output $k_1 a$ may be *blocked* completely for a certain value of b.

Second Harmonic Distortion

Second harmonic distortion occurs due to k_2. If b is zero, the amplitude of the second harmonic will be $k_2\left(\dfrac{a^2}{2}\right)$. Since power is proportional to the square of the voltage, the desired term in the output can be expressed as

$$P_1 = 10\log_{10}\left(\zeta k_1\frac{a}{2}\right)^2 = 20\log_{10} a + C_1 \tag{2.5.39}$$

where ζ is the proportionality constant and

$$C_1 = 20 \log \frac{\zeta k_1}{2}$$

Similarly, power in the second harmonic component can be expressed as

$$P_2 = 10 \log_{10} \left(\varsigma k_2 \frac{a^2}{2} \right)^2 = 40 \log_{10} a + C_2 \tag{2.5.40}$$

and the input power is

$$P_{in} = 10 \log_{10} \left(\varsigma \frac{a}{2} \right)^2 = 20 \log_{10} a + C_3 \tag{2.5.41}$$

Proportionality constants ς and k_2 are embedded in C_2 and C_3. From (2.5.39)–(2.5.41) we find that

$$P_1 = P_{in} + D_1 \tag{2.5.42}$$

and

$$P_2 = 2P_{in} + D_2 \tag{2.5.43}$$

where D_1 and D_2 replace $C_1 - C_3$ and $C_2 - 2C_3$, respectively. Equations (2.5.42) and (2.5.43) indicate that both the fundamental and second harmonic signals in the output are linearly related with input power. However, the second harmonic power increases at twice the rate of the fundamental (the desired) component.

Intermodulation Distortion Ratio

From Table 2.6 we find that the cubic term produces intermodulation frequencies $2\omega_1 \pm \omega_2$ and $2\omega_2 \pm \omega_1$. If ω_1 and ω_2 are very close, $2\omega_1 + \omega_2$ and $2\omega_2 + \omega_1$ will be far away from the desired signals, and therefore, these can be filtered out easily. However, the other two terms, $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$, will be so close to ω_1 and ω_2 that these components may be within the passband of the system. It will distort the output. This characteristic of a nonlinear circuit is specified via the intermodulation distortion. It is obtained after dividing the amplitude of one of the intermodulation terms by the desired output signal. For an input signal with both ω_1 and ω_2 (i.e., a two-tone input), the intermodulation distortion ratio (IMR) may be found as

$$\text{IMR} = \frac{\frac{3}{4} k_3 a^2 b}{k_1 a} = \frac{3k_3}{4k_1} ab \tag{2.5.44}$$