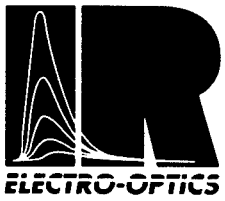
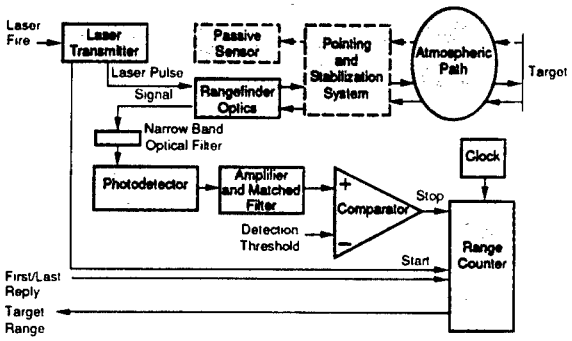


The Infrared & Electro-Optical Systems Handbook

VOLUME 6

Active Electro-Optical Systems

Clifton S. Fox, *Editor*



Active Electro-Optical Systems

VOLUME 6

The Infrared and Electro-Optical Systems Handbook

The Infrared and Electro-Optical Systems Handbook

Joseph S. Accetta, David L. Shumaker, *Executive Editors*

Chapter 3. Optomechanical System Design, Daniel Vukobratovich

Chapter 4. Infrared Imaging System Testing, Gerald C. Holst

Chapter 5. Tracking and Control Systems, Robert E. Nasburg

Chapter 6. Signature Prediction and Modeling, John A. Conant, Malcolm A. LeCompte

■ **VOLUME 5.** **Passive Electro-Optical Systems,**
Stephen B. Campana, *Editor*

Chapter 1. Infrared Line Scanning Systems, William L. McCracken

Chapter 2. Forward-Looking Infrared Systems, George S. Hopper

Chapter 3. Staring-Sensor Systems, Michael J. Cantella

Chapter 4. Infrared Search and Track Systems, Joseph S. Accetta

■ **VOLUME 6.** **Active Electro-Optical Systems,** Clifton S. Fox, *Editor*

Chapter 1. Laser Radar, Gary W. Kamerman

Chapter 2. Laser Rangefinders, Robert W. Byren

Chapter 3. Millimeter-Wave Radar, Elmer L. Johansen

Chapter 4. Fiber Optic Systems, Norris E. Lewis, Michael B. Miller

■ **VOLUME 7.** **Countermeasure Systems,** David Pollock, *Editor*

Chapter 1. Warning Systems, Donald W. Wilmot, William R. Owens, Robert J. Shelton

Chapter 2. Camouflage, Suppression, and Screening Systems, David E. Schmieder, Grayson W. Walker

Chapter 3. Active Infrared Countermeasures, Charles J. Tranchita, Kazimieras Jakstas, Robert G. Palazzo, Joseph C. O'Connell

Chapter 4. Expendable Decoys, Neal Brune

Chapter 5. Optical and Sensor Protection, Michael C. Dudzik

Chapter 6. Obscuration Countermeasures, Donald W. Hoock, Jr., Robert A. Sutherland

■ **VOLUME 8.** **Emerging Systems and Technologies,**
Stanley R. Robinson, Editor

Chapter 1. Unconventional Imaging Systems, Carl C. Aleksoff, J. Christopher Dainty, James R. Fienup, Robert Q. Fugate, Jean-Marie Mariotti, Peter Nisenson, Francois Roddier

Chapter 2. Adaptive Optics, Robert K. Tyson, Peter B. Ulrich

Chapter 3. Sensor and Data Fusion, Alan N. Steinberg

Chapter 4. Automatic Target Recognition Systems, James W. Sherman, David N. Spector, C. W. "Ron" Swonger, Lloyd G. Clark, Edmund G. Zelnio, Terry L. Jones, Martin J. Lahart

Chapter 5. Directed Energy Systems, Gary Golnik

Chapter 6. Holography, Emmett N. Leith

Chapter 7. System Design Considerations for a Visually-Coupled System, Brian H. Tsou

Copublished by

Infrared Information Analysis Center
Environmental Research Institute of Michigan
Ann Arbor, Michigan USA

and

SPIE OPTICAL ENGINEERING PRESS
Bellingham, Washington USA

Sponsored by

Defense Technical Information Center, DTIC-DF
Cameron Station, Alexandria, Virginia 22304-6145

Active Electro-Optical Systems

Clifton S. Fox, *Editor*
Night Vision and Electronic Sensors Directorate

VOLUME 6

The Infrared and Electro-Optical Systems Handbook

Joseph S. Accetta, David L. Shumaker, *Executive Editors*
Environmental Research Institute of Michigan

Library of Congress Cataloging-in-Publication Data

The Infrared and electro-optical systems handbook / Joseph S. Accetta, David L. Shumaker, executive editors.
p. cm.
Spine title: IR/EO systems handbook.
Cover title: The Infrared & electro-optical systems handbook.
Completely rev. ed. of: Infrared handbook. 1978
Includes bibliographical references and indexes.
Contents: v. 1. Sources of radiation / George J. Zissis, editor — v. 2. Atmospheric propagation of radiation / Fred G. Smith, editor — v. 3. Electro-optical components / William D. Rogatto, editor — v. 4. Electro-optical systems design, analysis, and testing / Michael C. Dudzik, editor — v. 5. Passive electro-optical systems / Stephen B. Campana, editor — v. 6. Active electro-optical systems / Clifton S. Fox, editor — v. 7. Countermeasure systems / David Pollock, editor — v. 8. Emerging systems and technologies / Stanley R. Robinson, editor.
ISBN 0-8194-1072-1
1. Infrared technology—Handbooks, manuals, etc. 2. Electrooptical devices—Handbooks, manuals, etc. I. Accetta, J. S. II. Shumaker, David L. III. Infrared handbook. IV. Title: IR/EO systems handbook. V. Title: Infrared & electro-optical systems handbook.
TA1570.I5 1993
621.36'2—dc20 92-38055
CIP

Copublished by

Infrared Information Analysis Center
Environmental Research Institute of Michigan
P.O. Box 134001
Ann Arbor, Michigan 48113-4001

and

SPIE Optical Engineering Press
P.O. Box 10
Bellingham, Washington 98227-0010

PRINTED IN THE UNITED STATES OF AMERICA

Preface

The Infrared and Electro-Optical Systems Handbook is a joint product of the Infrared Information Analysis Center (IRIA) and the International Society for Optical Engineering (SPIE). Sponsored by the Defense Technical Information Center (DTIC), this work is an outgrowth of its predecessor, *The Infrared Handbook,* published in 1978. The circulation of nearly 20,000 copies is adequate testimony to its wide acceptance in the electro-optics and infrared communities. *The Infrared Handbook* was itself preceded by *The Handbook of Military Infrared Technology.* Since its original inception, new topics and technologies have emerged for which little or no reference material exists. This work is intended to update and complement the current *Infrared Handbook* by revision, addition of new materials, and reformatting to increase its utility. Of necessity, some material from the current book was reproduced as is, having been adjudged as being current and adequate. The 45 chapters represent most subject areas of current activity in the military, aerospace, and civilian communities and contain material that has rarely appeared so extensively in the open literature.

Because the contents are in part derivatives of advanced military technology, it seemed reasonable to categorize those chapters dealing with systems in analogy to the specialty groups comprising the annual Infrared Information Symposia (IRIS), a Department of Defense (DoD) sponsored forum administered by the Infrared Information Analysis Center of the Environmental Research Institute of Michigan (ERIM); thus, the presence of chapters on active, passive, and countermeasure systems.

There appears to be no general agreement on what format constitutes a "handbook." The term has been applied to a number of reference works with markedly different presentation styles ranging from data compendiums to tutorials. In the process of organizing this book, we were obliged to embrace a style of our choosing that best seemed to satisfy the objectives of the book: to provide derivational material data, descriptions, equations, procedures, and examples that will enable an investigator with a basic engineering and science education, but not necessarily an extensive background in the specific technology, to solve the types of problems he or she will encounter in design and analysis of electro-optical systems. Usability was the prime consideration. In addition, we wanted each chapter to be largely self-contained to avoid time-consuming and tedious referrals to other chapters. Although best addressed by example, the essence of our handbook style embodies four essential ingredients: a brief but well-referenced tutorial, a practical formulary, pertinent data, and, finally, example problems illustrating the use of the formulary and data.

The final product represents varying degrees of success in achieving this structure, with some chapters being quite successful in meeting our objectives and others following a somewhat different organization. Suffice it to say that the practical exigencies of organizing and producing a compendium of this magnitude necessitated some compromises and latitude. Its ultimate success will be judged by the community that it serves. Although largely oriented toward system applications, a good measure of this book concentrates on topics endemic and fundamental to systems performance. It is organized into eight volumes:

Volume 1, edited by George Zissis of ERIM, treats sources of radiation, including both artificial and natural sources, the latter of which in most military applications is generally regarded as background radiation.

Volume 2, edited by Fred Smith of OptiMetrics, Inc., treats the propagation of radiation. It features significant amounts of new material and data on absorption, scattering, and turbulence, including nonlinear propagation relevant to high-energy laser systems and propagation through aerodynamically induced flow relevant to systems mounted on high-performance aircraft.

Volume 3, edited by William Rogatto of Santa Barbara Research Center, treats traditional system components and devices and includes recent material on focal plane array read-out electronics.

Volume 4, edited by Michael Dudzik of ERIM, treats system design, analysis, and testing, including adjunct technology and methods such as trackers, mechanical design considerations, and signature modeling.

Volume 5, edited by Stephen Campana of the Naval Air Warfare Center, treats contemporary infrared passive systems such as FLIRs, IRSTs, IR line scanners, and staring array configurations.

Volume 6, edited by Clifton Fox of the Night Vision and Electronic Sensors Directorate, treats active systems and includes mostly new material on laser radar, laser rangefinders, millimeter-wave systems, and fiber optic systems.

Volume 7, edited by David Pollock, consultant, treats a number of countermeasure topics rarely appearing in the open literature.

Volume 8, edited by Stanley Robinson of ERIM, treats emerging technologies such as unconventional imaging, synthetic arrays, sensor and data fusion, adaptive optics, and automatic target recognition.

Acknowledgments

It is extremely difficult to give credit to all the people and organizations that contributed to this project in diverse ways. A significant amount of material in this book was generated by the sheer dedication and professionalism of many esteemed members of the IR and EO community who unselfishly contributed extensive amounts of precious personal time to this effort and to whom the modest honorarium extended was scarcely an inducement. Their contributions speak elegantly of their skills.

Directly involved were some 85 authors and editors from numerous organizations, as well as scores of technical reviewers, copyeditors, graphic artists, and photographers whose skill contributed immeasurably to the final product.

We acknowledge the extensive material and moral support given to this project by various members of the managements of all the sponsoring and supporting organizations. In many cases, organizations donated staff time and internal resources to the preparation of this book. Specifically, we would like to acknowledge J. MacCallum of DoD, W. Brown and J. Walker of ERIM, and J. Yaver of SPIE, who had the foresight and confidence to invest significant resources in the preparation of this book. We also extend our appreciation to P. Klinefelter, B. McCabe, and F. Frank of DTIC for their administrative support during the course of this program.

Supporting ERIM staff included Ivan Clemons, Jenni Cook, Tim Kellman, Lisa Lyons, Judy Steeh, Barbara Wood, and the members of their respective organizations that contributed to this project.

We acknowledge Lorretta Palagi and the publications staff at SPIE for a professional approach to the truly monumental task of transforming the manuscripts into presentable copy and the patience required to interact effectively with the authors.

We would like to pay special tribute to Nancy Hall of the IRIA Center at ERIM who administrated this at times chaotic project with considerable interpersonal skill, marshaling the numerous manuscripts and coordinating the myriad details characteristic of a work of this magnitude.

We properly dedicate this book to the people who created it and trust it will stand as a monument to their skills, experience, and dedication. It is, in the final analysis, a product of the community it is intended to serve.

Joseph S. Accetta
David L. Shumaker
Ann Arbor, Michigan

January 1993

Notices and Disclaimer

This handbook was prepared by the Infrared Information Analysis Center (IRIA) in cooperation with the International Society for Optical Engineering (SPIE). The IRIA Center, Environmental Research Institute of Michigan, is a Defense Technical Information Center-sponsored activity under contract DLA-800-C-393 and administrated by the Defense Electronics Supply Center, Defense Logistics Agency.

This work relates to the aforementioned ERIM contract and is in part sponsored by the Department of Defense; however, the contents do not necessarily reflect the position or the policy of the Department of Defense or the United States government and no official endorsement should be inferred.

The use of product names does not in any way constitute an endorsement of the product by the authors, editors, Department of Defense or any of its agencies, the Environmental Research Institute of Michigan, or the International Society for Optical Engineering.

The information in this handbook is judged to be from the best available sources; however, the authors, editors, Department of Defense or any of its agencies, the Environmental Research Institute of Michigan, or the International Society for Optical Engineering do not assume any liability for the validity of the information contained herein or for any consequence of its use.

Contents

Introduction

This volume is devoted to the technology of "active" electro-optical systems, which employ an artificial source of optical radiation (typically a laser) and a receiver that collects and detects some fraction of that radiation either directly or after reflection from a target scene. This is as opposed to "passive" systems, which sense naturally emitted radiation or reflected sunlight or moonlight. In most cases, it is preferable to use a passive sensor for a particular application if the need can thereby be met. Frequently, however, active sensors, with their unique capabilities, are needed to satisfy the requirement at hand. Accurate determination of target range, range rate, three-dimensional geometry, vibrational characteristics, spectral reflectance, and many other attributes is possible using the appropriate active sensor. Active sensors are also frequently used in conjunction with passive sensors to take advantage of the best features of both. A simple example of this is the use of a laser rangefinder with a thermal imaging system.

Although millimeter-wave (MMW) radiation is not in the optical portion of the electromagnetic spectrum, MMW radar is also addressed in this volume due to the current strong interest in the use of that technology in conjunction with electro-optical sensors and in order to highlight the similarities and differences between laser and MMW radars.

In some cases (for example, fiber optic communication systems), a signal is encoded on the transmitted beam and routed as directly as practical to a receiver at a remote location where it is detected. In other cases (for example, laser and MMW radars and rangefinders), the beam is transmitted to a target and a small portion of the reflected radiation is collected and detected by the receiver, which in most cases is collocated with the transmitter. In these cases, the desired information to be sensed by the receiver is one or more characteristics of the target. This information is impressed on the reflected radiation by the target itself. In most of these cases, a specific waveform (for example, pulsed, cw, AM, or FM) is used on the outgoing beam in order to most effectively extract the desired information from the target.

Chapter 1 addresses laser radar. Coherent detection is emphasized since so many laser radars now employ coherent (heterodyne) detection receivers. Laser radars are currently being developed and used for a wide variety of applications, from short-range sensors employing low-power solid-state or semiconductor diode-type laser transmitters and direct detection receivers (for example, airborne wire and obstacle detectors with typical ranges to hundreds of meters or less) to very long-range sensors employing relatively high-power gas lasers and coherent detection receivers (for example, laser radars for strategic target discrimination with ranges to hundreds of kilometers). In many cases, coherent

detection provides substantially higher sensitivity and versatility than a direct detection sensor of equal transmitter power. Coherent detection provides excellent capability to sense radial motion, including both gross target motion (Doppler shift) and vibration (micro-Doppler).

Chapter 2 addresses laser rangefinders and emphasizes direct detection since most laser rangefinders employ direct detection receivers. Laser rangefinders are currently being developed and used for a very wide range of military and civilian applications. Military use is primarily in fire control systems for combat vehicles, aviation platforms, and air defense systems as well as a wide variety of manportable weapons systems. Pulsed solid-state lasers are the most common type of radiation source used in currently fielded laser rangefinders, although some semiconductor diode and gas laser types also exist. Most of the currently fielded devices operate at a wavelength of 1.06 μm or less and are not eyesafe. However, increasing numbers of eyesafe varieties operating at wavelengths longer than 1.4 μm are expected to be fielded in larger numbers in the near future.

Chapter 3 addresses millimeter-wave radar. Recent advances in component and systems level technology, especially at 35- and 95-GHz frequencies, have created a strong interest and substantial system development activities for military target acquisition and fire control applications. The generally superior foul weather penetration capability of radiation at millimeter-wave frequencies compared to optical radiation, combined with the much better angular resolution achievable with millimeter-wave compared to microwave radars of a given aperture size, make this type of radar an attractive candidate for a variety of applications.

Chapter 4 addresses fiber optic systems. Fiber optic communication systems provide tremendous improvement in bandwidth and, therefore, information handling capacity, compared to conventional conductive wire communication systems. Substantial improvements in the optical and mechanical quality of fiber optic cable, improved performance of diode laser transmitters and receivers, and the reduced cost of componentry have made fiber optic communication systems the preferred choice for closed-channel applications.

Clifton S. Fox
Fort Belvoir, Virginia

January 1993

CHAPTER 1

Laser Radar

Gary W. Kamerman
Teledyne Brown Engineering
Huntsville, Alabama

CONTENTS

1.1 INTRODUCTION

The term *radar* originated during World War II as an acronym for *ra*dio *d*etection *a*nd *r*anging. At that time, it referred to the technique of monitoring reflected, radio frequency, electromagnetic radiation to locate remote objects. Since that time, the basic radar technique has been applied to progressively shorter (and in some cases, longer) wavelengths so that the term radar no longer applies only to systems that operate at radio frequencies. Laser radar is simply radar that operates at optical frequencies and uses a laser as its source of electromagnetic radiation.[a]

1.1.1 Types of Laser Radars

Laser radars can be grouped according to the type of measurement made by the laser radar, the detection technique, the type of interferometer employed in a coherent laser radar (if appropriate), the modulation technique, the demodulation technique, the type of laser or the wavelength of operation, the function performed, the type of data collected, or the data format. In addition, laser radar can be classed as monostatic or bistatic, depending on whether it uses a single aperture to transmit and to receive or separate apertures. Some of these groupings are summarized in Table 1.1. The name given to a particular system is seldom sufficient to completely identify what it does and is certainly not sufficient to identify how well it performs.

As can be seen in Table 1.1, there are many types of laser radars. The variety found among laser radar systems is one of the primary reasons for their versatility. Unfortunately, it can also create some confusion. For example, wavelength-dependent technological limitations frequently prevent simple parametric extrapolation of performance from one type of system to another. These limitations can make routine performance at a one laser wavelength well beyond the state of the art (and possibly beyond fundamental physical limitations) at another wavelength. Extreme care must be exercised when extrapolating the performance of one type of laser radar to another.

The use of very wavelength specific technology and components represents a significant difference from passive optical systems or conventional radar systems. The availability of laser sources makes only a finite (and small) number of wavelengths practical alternatives for laser radars. Passive optics and conventional [radio frequency (rf) through millimeter-wave (MMW)] radars can select the wave band to optimize performance without major changes in technology. Laser radars often must change technologies completely (e.g., electrically pumped gas lasers versus optically pumped solid-state lasers) to effect even small changes in operating wavelength.

[a]The term *laser radar* is used exclusively in this text. Optical radar, LADAR (*la*ser *d*etection *a*nd *r*anging), and LIDAR (*li*ght *d*etection *a*nd *r*anging) are sometimes used in the literature to describe laser radar. In the strictest sense, optical radar or lidar do not require that the source of electromagnetic radiation be a laser. However, the physical behavior of the device is the same regardless of the type of source used and the treatment in this text is appropriate to all radars operating at optical frequencies. As a result of the state of technology as of this writing, all known optical radars employ one or more lasers as the light source. The term laser radar is used, therefore, without any loss of generality.

Table 1.1 Types of Laser Radars

Types of Lasers (Typical)	Carrier Wavelength
CO_2	9.2 μm -- 11.2 μm
Er:YAG	2 μm
Raman Shifted Nd:YAG	1.54 μm
Nd:YAG	1.06 μm
GaAlAs	0.8 μm -- 0.904 μm
HeNe	0.63 μm
Frequency Doubled Nd:YAG	0.53 μm

Detection Technique	Interferometer Type	Modulation Technique
Direct Detection	Not Applicable	Pulsed Amplitude Modulation (AM)
Coherent Detection	Heterodyne Homodyne Offset Homodyne	Pulsed Amplitude Modulation (AM) Frequency Modulation (AM) Hybrid (AM/FM, Pulse Burst) None (CW)

Functions	Measurements
Tracking Moving Target Indication (MTI) Machine Vision Velocimetry Wind Shear Detection Target Identification Imaging Vibration Sensing	Amplitude (Reflectance) Range (time delay) Velocity (Doppler shift or differential range) Angular Position Vibration Spectra

1.1.2 Principles of Operation

The operation of laser radar is functionally identical to conventional radar. Laser radar transmits a signal that is reflected by a target and then collected by the laser radar receiver. Range to the target is determined by measuring the round-trip time of the reflected light. Radial velocity of the target is measured by either determining the Doppler shift of the reflected light or by making two (or more) range measurements and calculating the rate of change of range.

A block diagram of a typical heterodyne or coherent detection laser radar is shown in Fig. 1.1. An optical signal is generated by the transmitter laser.

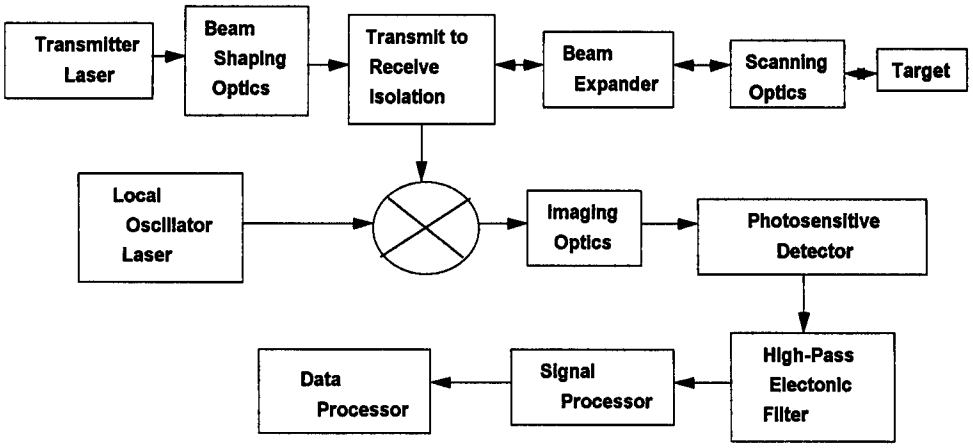


Fig. 1.1 Block diagram of a typical laser radar.

The divergence and beam diameter of this optical signal are then matched to the rest of the system by beam-shaping optics. This matching is optional because some systems are designed to operate with the unmodified transmitter laser beam. In a monostatic system, the transmitted laser signal enters a transmit-to-receive (T/R) switch. The T/R switch permits the laser radar transmitter and receiver to operate through a common optical aperture. The laser radar signal then enters the beam expander or output telescope and the scanning optics that direct the optical signal to the target.

In a monostatic system, radiation reflected from the target is collected by the scanning optics and the beam expander, which now acts as an optical receiver. The T/R switch directs the received radiation to an optical mixer, where it is combined with an optical reference signal, which is the local oscillator. The combined signal is then focused onto a photosensitive detector by the imaging optics. The photosensitive detector generates an electrical signal in response to the received optical signal. The electrical signal is then high-pass filtered to remove any low-frequency components, such as those from background sources and from the local oscillator-induced dc signal. The high-frequency components of this electrical signal contain the target information obtained by the laser radar. Metric information is then extracted from the electrical signal by signal and data processors.

In a bistatic system, the T/R switch is omitted. A separate beam expander and scanning optics are then dedicated to the receiver. The remainder is identical to a monostatic system, as previously described.

1.1.3 Applications and Advantages

Laser radars are the result of the hybridization of technology. They are formed by the merger of conventional radar technology with optics. This combination results in several advantages over either "parent" technology. Laser radars have inherently higher resolution than microwave radars, and most applications for laser radars result from this advantage.[1] Laser radar has a narrower beamwidth than conventional radar because of the higher carrier frequency

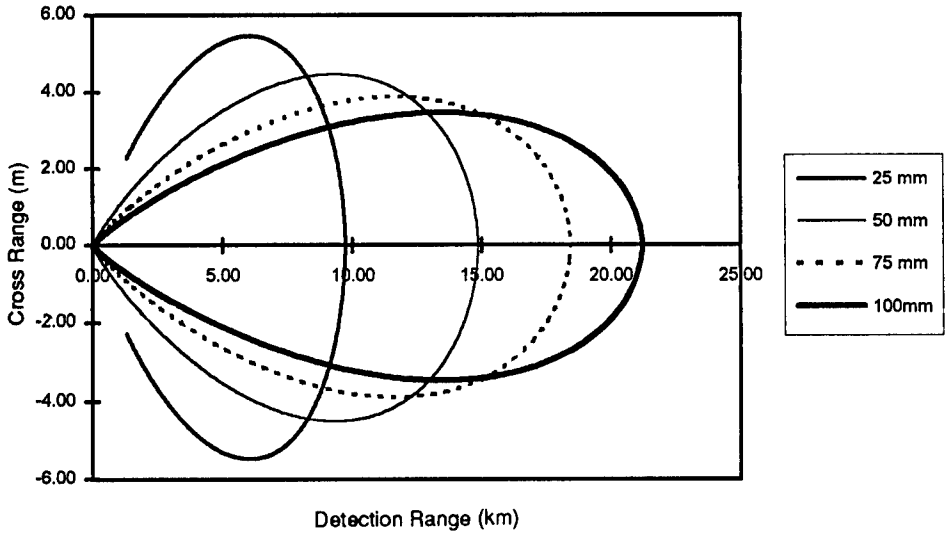


Fig. 1.2 Coherent laser radar detection range for several matched transmitter and receiver apertures (Airy local oscillator, *f*/4 receiver).

(e.g., 27,000 GHz for a 10.6-μm laser as compared to 95 GHz for a typical millimeter-wave radar).[b] The narrower beamwidth, in turn, provides for higher angular resolution, lower sidelobes, higher countermeasure (jamming and spoofing) resistance, and covertness. The higher carrier frequency also supports higher resolution and accuracy measurements of range, velocity, and angular position.

The narrow beamwidth of laser radars is ideally suited to applications such as target tracking. Figure 1.2 shows the detection range for a typical CO_2 heterodyne laser radar as a function of the cross-range position of a target. The performance of the monostatic CO_2 laser radar in Fig. 1.2 is based on the receiver and the transmitter having equal aperture diameters. Four different aperture diameters, ranging from 25 to 100 mm, are shown in the figure. The effective beamwidth is quite small in all cases, which supports extremely precise tracking.

In addition to the narrow beamwidth, the sensitivity of a coherent laser radar receiver also falls off rapidly with increasing off-axis position. The receiver in the cases shown in Fig. 1.2 employed a matched Airy pattern local oscillator to illuminate the receiver detector and effect heterodyne detection. These two effects combine to create the very narrow "antenna" patterns demonstrated in Fig. 1.2.

These extremely narrow antenna patterns are not limited to the specific design features of the system shown in Fig. 1.2. Figure 1.3 shows the antenna

[b]All units used in this chapter and in the accompanying formulas and equations are SI unless explicitly noted. Manufacturer's data are often presented in unit systems other than SI, which may be more convenient for a particular application or technology. Conversion of these data to SI units is required before the formulas and equations presented here can be used with such data.

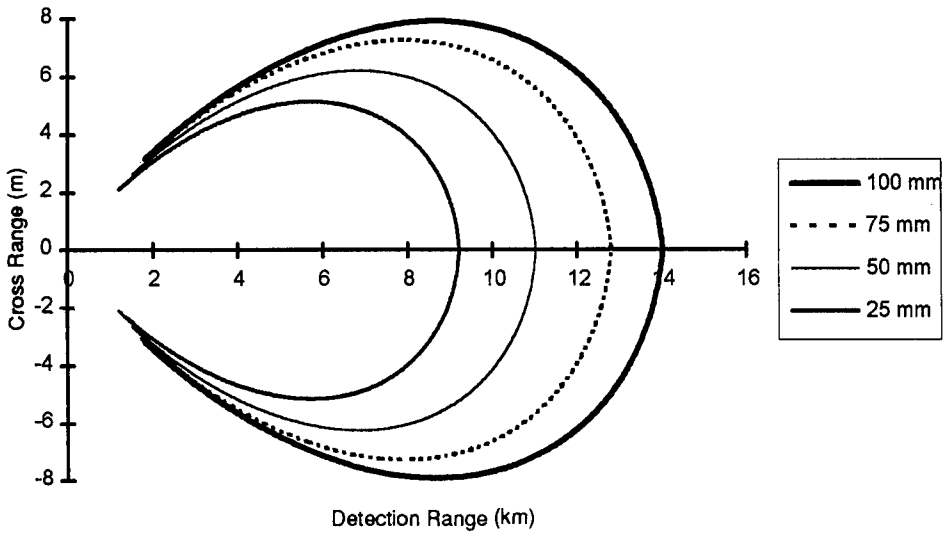


Fig. 1.3 Coherent laser radar detection range for several transmitter aperture diameters (100-mm receiver aperture diameter, Airy local oscillator, *f*/4 receiver).

patterns that result from mixed transmitter and receiver sizes and uniform local oscillator illumination of the detector. In these cases, the maximum range is significantly reduced but the basic "pencil beam" pattern is retained. Note that in both Figs. 1.2 and 1.3 the cross-range axes are scaled in meters, whereas the detection-range axes are scaled in kilometers. The shape of the antenna pattern of laser radar is significantly narrower than is immediately apparent from casual inspection of these figures.

The shape of the antenna pattern is also largely independent of transmitted laser power and laser cross section of the target. Increases in power or cross section increase both the down-range and cross-range coverage of laser radar, but not the basic shape. The pattern retains its "pencil beam" contour for almost all configurations.

The narrowness of the antenna pattern makes surveillance with laser radar very difficult. The large search volumes associated with surveillance operations generally require that laser radar operate with a very high pulse repetition rate so that the entire volume can be interrogated within a prescribed time. Alternatively, laser radar may employ multiple beams, which operate in parallel through the same output aperture, to enlarge the search volume to be covered. Systems that use multiple beams operating in parallel are generally more optically complicated. However, this may be a simpler overall solution than the complexity of a high-pulse-repetition-rate transmitter. The use of multiple beams may permit slower scan rates, which reduces lag-angle effects (see Sec. 1.6.4).

The higher resolution of laser radars permits the recognition and identification of targets. Subareas of the target's surface can be resolved in angle, range, and/or velocity. These resolved measurements permit the detection of features that are unique to a particular target or target type. These unique features include shape, size, velocity, spin or rotation rate, and vibration.

The ability to measure unique features of the target makes high-resolution systems ideal for automatic target detection (ATD), which is based on computer interpretation of the data collected. High-resolution laser radars are often called *imaging laser radars* because the data collected are often compatible with conventional video display techniques and, consequently, human visual interpretation of the data collected. The data display technique is not normally critical to the laser radar performance, but it is frequently critical to human interpretation of the data collected. Imaging actually requires only that the data collected be a set of ordered triples (e.g., azimuth angle, elevation angle and range or range, and velocity and intensity). Laser radars frequently make more than three measurements simultaneously. If the data are to be displayed as an image, then multiple-image planes are required to display all the data. Generally, any requirement for human interpretation is often more of a constraint to performance than an asset. As a result, imaging laser radars are a subset of high-resolution systems.

The high carrier (optical) frequency of laser radars permits the detection and interrogation of small objects. Wires and projectiles[2] have small cross sections at microwave and millimeter-wave frequencies because they have dimensions that are smaller than the radar wavelength. As a result, these objects are difficult to detect with conventional radars. However, because all dimensions of these objects are larger than laser radar wavelengths, the laser cross section of these objects is much larger. Laser radars are well suited for the detection of these small objects, such as is required for wire and obstacle detection for low-altitude aircraft and for tracking small-caliber projectiles.

Airborne aerosols and rain also have significantly larger laser cross sections than radar cross sections. Range and velocity measurements of airborne aerosols are an important technique for wind field mapping, which is used for a number of meteorological applications. These include wind shear detection, wind flow field mapping, clear air turbulence detection, pollution monitoring, and ceiliometry. Except for those applications that are based on the spectroscopic analysis of chemical species, the design principles for laser radar for atmospheric sensing applications are the same as those used for laser radar for "hard target" applications. Spectroscopic chemical analysis is normally accomplished by differential absorption lidar (DIAL) techniques, which are not discussed in this chapter.

The high carrier frequency also provides enhanced vibration detection sensitivity. The detection of vibratory motion is based on the detection of the time-varying Doppler shift of the reflected radiation. A surface that is experiencing vibration will modulate this shift. However, the modulation index is proportional to the carrier frequency. For this reason, laser radars are generally two to three orders of magnitude more sensitive than a typical MMW radar for vibration detection.

The high carrier frequency also permits laser radars to be packaged into a relatively small size. This smaller size permits laser radars to be employed in aircraft applications and other applications where space and weight are at a premium.

The primary disadvantage of laser radars, compared to rf or MMW radars, is their inability to penetrate weather. The atmosphere absorbs and scatters optical radiation more strongly than it does microwaves or millimeter waves.

As a result, the performance of laser radar is much more weather sensitive. However, laser radar has superior weather capability over a passive optical system because laser radar can employ range and/or velocity gating to reject signals that have been determined to be only clutter.

1.2 LASER RADAR RANGE EQUATION

A fundamental problem in laser radar analysis is the determination of the total optical power density that is present at the receiver aperture and, consequently, the total optical power incident on the photosensitive element of the receiver. The laser radar range equation is commonly used to determine the power received by laser radar under specific conditions and against a particular target.

1.2.1 General Forms of the Range Equation

The general form of the laser radar range equation is given by

$$P_r = \frac{4KP_sT_{A1}\eta_t}{\pi\varphi^2 r_1^2}\Gamma\frac{T_{A2}}{4\pi r_2^2}\frac{\pi D^2\eta_r}{4}, \tag{1.1}$$

where

P_r = received signal power in watts
P_s = source laser power in watts
K = beam profile function
T_{A1} = atmospheric transmission from the source to the target
η_t = transmitter optical efficiency
φ = beamwidth in radians
r_1 = range from transmitter to target in meters
Γ = target laser cross section in square meters
T_{A2} = atmospheric transmission from the target to the receiver
r_2 = range from target to receiver in meters
D = receiver aperture diameter in meters
η_r = receiver optical efficiency.

Equation (1.1) computes the power received by laser radar as the product of four terms. These terms are associated with each step in the physical process of (1) propagation of the light to the target, (2) reflection of light by the target, (3) propagation of the scattered light to the receiver, and (4) collection of the scattered light by the receiver.

The first term of Eq. (1.1) represents the optical power density (W m^{-2}) that is incident on the target. The beam profile function K is a normalization factor that accounts for the intensity pattern of the beam and the target's position in the beam. The beam profile function is discussed in more detail in Sec. 1.3. The second term of Eq. (1.1) is the laser radar cross section (LCS) of the target. The LCS has the units of area and is used to describe the apparent size of the target. The LCS is discussed in more detail in Sec. 1.5. The product of the first and second terms of Eq. (1.1) describes the total optical power scattered by the target in the direction of the receiver.

The third term is the inverse of the area of a sphere (further reduced by any atmospheric attenuation), which has a radius equal to the distance from the target to the laser radar receiver. The product of the first three terms is the optical power density of the reflected radiation at the receiver aperture. The fourth term is the area of the receiver collection aperture, reduced by an optical efficiency factor. The product of all four terms is the total power received and collected by the laser radar receiver.

Equation (1.1) is appropriate for a target in the far field of the transmitter aperture and for a receiver in the far field of the target. It is appropriate for both monostatic and bistatic systems. The target is modeled as a point source; i.e., its angular extent is much smaller than both the transmitter beamwidth and the instantaneous field of view of the receiver. However, with the substitution of the appropriate laser radar cross section Γ, it can be applied to extended targets.

Because laser radar obeys the same physical laws as microwave radar, the range equation used for microwave systems can also be applied. The microwave radar range equation is given by[39]

$$P_r = P_s \eta_t \frac{G_t T_{A1}}{4\pi r_1^2} \Gamma \frac{T_{A2}}{4\pi r_2^2} \frac{\pi D^2 \eta_r}{4} , \tag{1.2}$$

where G_t is the transmitter antenna gain $= 16K/\varphi^2$ and all other terms are as previously defined.

Equation (1.2) uses the concept of antenna gain to describe the transmitted beam. Antenna gain is simply the ratio of the intensity on the beam axis to the intensity that would have been present if the same amount of power were radiated isotropically. This concept is common in electrical engineering but is not generally used in physics. With the exception that Eq. (1.2) does not explicitly account for pointing error effects [as Eq. (1.1) does in the beam shape function], Eqs. (1.1) and (1.2) are exactly equivalent. However, small pointing errors, when combined with the very narrow beamwidths associated with laser radars, often have a dominant effect on laser radar system performance. Similar pointing errors in microwave systems have no measurable effect because the beamwidth of microwave systems is generally much broader. If Eq. (1.2) is used to compute the received laser radar power, then provision for pointing error losses should also be included.

1.2.2 Special Forms of the Range Equation

For a stationary (or slowly moving) monostatic system, $r_1 = r_2 = r$. For small relative velocities and, therefore, negligible Doppler shifts, for exoatmospheric operation, or for particulate-dominated atmospheric attenuation, $T_{A1} = T_{A2} = T$. Under these conditions, Eq. (1.1) can be simplified. Combining terms and rearranging yields

$$P_r = \frac{K P_s T_A^2 \Gamma D^2 \eta_t \eta_r}{16\pi \varphi^2 r^4} , \tag{1.3}$$

where all terms are as previously defined. This is an important result. It applies to a large number of systems and it illustrates the inverse fourth-power de-

pendence of the received power with range for monostatic laser radars when operating against an unresolved target. An unresolved target is an object whose angular extent is smaller than the laser radar beamwidth or the instantaneous field of view (IFOV) of the photosensitive detector element in the receiver. An unresolved target is sometimes referred to as a *point* target.

If the angular extent of the target is larger than the transmitter beamwidth (in both azimuth and elevation), then the laser cross section of the illuminated region of the target is proportional to the area of the illuminated region, i.e., Γ is proportional to both φ^2 and r^2 (see Sec. 1.5). In this case, the beam profile function can be approximated by a uniform intensity distribution (i.e., constant over the illuminated region) and Eq. (1.1) can be further simplified for Lambertian targets to yield

$$P_r = \frac{P_s T_A^2 \rho D^2 \eta_t \eta_r}{4r^2} , \tag{1.4}$$

where ρ is the total hemispherical target reflectance and all other terms are as previously defined. Equation (1.4) illustrates the dependence of the received power with the inverse square of the range for extended, or large, targets. This behavior is not commonly encountered in conventional radar systems.

Instead of determining the optical power received at a particular range, it is often necessary to establish the maximum range at which a specified optical power (usually the minimum detectable power) is received. The nonlinear nature of the range equation makes algebraic inversion with respect to range impossible, except in some very special cases. In general, the determination of maximum range can be easily accomplished by iteration using Newton's method or another suitable numerical technique. However, in the case of negligible atmospheric losses and for a monostatic system, Eq. (1.1) can be inverted to yield

$$r = \left(\frac{K P_s \Gamma D^2 \eta_t \eta_r}{4\pi \varphi^2 P_r} \right)^{1/4} , \tag{1.5}$$

where all terms are as previously defined.

1.2.3 Solved Problems

Problem 1.1. Calculate the received power for a monostatic, 5-W laser radar operating against an extended target having a total hemispheric reflectivity of 4% at a range of 1 km. The transmitter optics are 80% efficient. The receiver has a clear aperture of 50 mm and has an efficiency of 65%. The atmospheric transmission at 1 km is 50%.

Solution 1.1. The received power for a laser radar against an extended target is given by Eq. (1.4):

$$P_r = \frac{P_s T_A^2 \rho D^2 \eta_t \eta_r}{4r^2} .$$

For this problem,

$P_s = 5$ W
$T_a = 0.5$
$\rho = 0.04$
$D = 0.05$ m
$\eta_t = 0.8$
$\eta_r = 0.65$
$r = 1000$ m.

Substitution into Eq. (1.4) yields

$$P_r = \frac{5(0.5)^2(0.04)(0.05)^2(0.8)(0.65)}{4(1000)^2}$$

$$\therefore P_r = 1.625 \times 10^{-11} \text{ W} = 16.25 \text{ pW} .$$

Problem 1.2. Find the range in space at which a 100-W CO_2 laser radar will receive 10 pW of reflected light from a 1-m^2 target. The laser radar has a beam divergence of 1 mrad and a uniform beam profile. Both the transmitter and receiver have an optical efficiency of 80% and a clear aperture of 1 in.

Solution 1.2. Because there is no range-dependent atmospheric loss in space, the range at which a prescribed power is received can be calculated by Eq. (1.5):

$$r = \left(\frac{KP_s\Gamma D^2\eta_t\eta_r}{16\pi\varphi^2 P_r}\right)^{1/4} .$$

In this problem,

$P_s = 100$ W
$D = 0.0254$ m
$\Gamma = 1$ m^2
$\eta_t = 0.8$
$\eta_r = 0.8$
$\varphi = 1 \times 10^{-3}$ rad
$P_r = 10 \times 10^{-12}$ W.

A uniform beam profile implies that $K = 4$ (from Sec. 1.3). Substitution yields

$$r = \left[\frac{4(100)(0.0254)^2(1)(0.8)(0.8)}{16\pi(10^{-3})^2(10 \times 10^{-12})}\right]^{1/4} ,$$

$$\therefore r = 4258 \text{ m} .$$

1.3 TRANSMITTER CHARACTERISTICS

The transmitted laser beam is characterized by total optical power, divergence or beamwidth, intensity profile across its width, and direction of propagation with respect to the target. For an unresolved or point target, its position in the beam relative to the intended line of sight can significantly affect the

intensity at the target. The target's position with respect to the beam center is less important for an extended target. The effects these factors have on the received signal power, as predicted by Eq. (1.1), are contained within the beam profile function K.

The transmitter is also characterized by modulation format, modulation characteristics, and coherence time. These features do not normally affect the received signal power as predicted by the range equation [Eq. (1.1)]. Transmitter modulation effects are discussed in more detail in Sec. 1.11.

1.3.1 Beam Shape and the Beam Profile Function

The intensity of an optical beam is not constant across its diameter at all ranges. Some typical beam profiles encountered in laser radar system design and analysis are shown in Fig. 1.4. The most common beam profile encountered in laser radar systems is the Gaussian. This is a result of the fact that the lowest order mode (TEM_{00}) of a conventional laser resonator is Gaussian. Waveguide lasers (both circular and rectangular bores) do not produce true Gaussian beams. For analytical purposes, it is a convenient, common practice

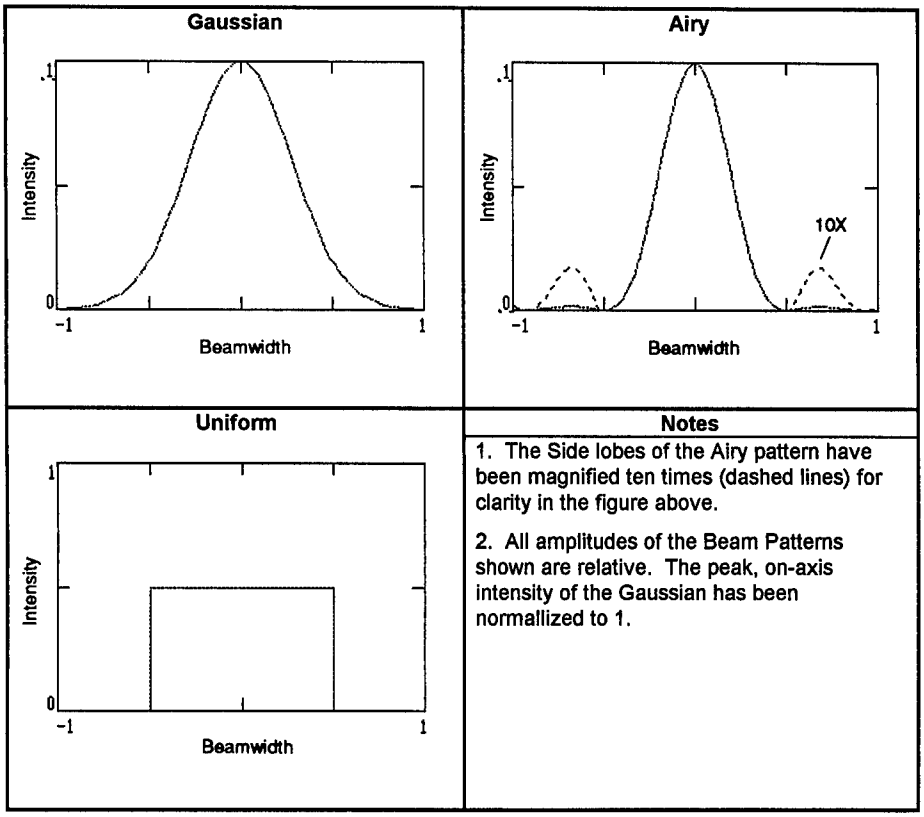


Fig. 1.4 Common laser radar transmitted beam shape functions.

to approximate these beam profiles with a Gaussian. In some special applications, other profiles may be employed.

To simplify initial performance estimation, a uniform beam profile is often assumed, i.e., the beam is represented as having constant intensity everywhere within the beam and zero intensity elsewhere. Although this is physically unrealistic, it does provide a simplified approach that is useful for the initial estimation of laser radar performance, laser power, or receiver aperture requirements. This approach neglects a detailed assessment of pointing error effects but provides a more pessimistic estimate than assuming perfect transmitter pointing and a real beam profile. It also permits a more rapid estimate of performance than a detailed analysis. The beam profile function for a uniform beam is given by

$$K(\psi,\varphi) = \begin{cases} 1, & \psi \leq \varphi \\ 0, & \psi > \varphi \end{cases}, \tag{1.6}$$

where ψ is the line-of-sight pointing error and φ is the beamwidth.

In some applications, the transmitter aperture may be uniformly illuminated. In this case, the far-field beam takes on an intensity profile described by an Airy function. The beam profile function for an Airy beam is given by[3]

$$K(\psi,\varphi) = 4.181\left[\frac{\varphi J_1(2.44\pi\psi/\varphi)}{2.44\pi\psi}\right], \tag{1.7}$$

where J_1 is the first-order Bessel function of the first kind and all other terms are as previously defined.

Laser radars frequently use lasers operating in the TEM_{00} mode as the transmitter source. This mode has a Gaussian intensity profile. For small transmitter line-of-sight errors, the beam profile function of a Gaussian beam is given by

$$K(\psi,\varphi) = 2\exp\left(-\frac{2r^2\psi^2}{\omega^2}\right), \tag{1.8}$$

where $\omega^2 = \omega_0^2[1 + (\lambda r/\pi\omega_0^2)^2]$, ω_0 is the Gaussian waist radius, and all other terms are as previously defined. Depending on the specific conditions, the Gaussian beamwidth is related to the beam radius, ω, by Eq. (1.10) or (1.11). The on-axis intensity of a Gaussian beam is twice the intensity of a uniform beam whose beamwidth is equal to the Gaussian e^{-2} beamwidth. The on-axis intensity of a Gaussian is approximately equal to the on-axis intensity of an Airy beam.

A laser source operating in the TEM_{00} mode maintains a perfectly Gaussian intensity profile in the far field only for the theoretical case wherein the clear aperture diameter of the system is infinite. Real laser radars have finite clear transmitter apertures that truncate the Gaussian beam. In most applications, the beam profile function of a truncated Gaussian beam can be approximated by a perfect Gaussian, provided that the clear aperture diameter is at least three times larger than the beam waist radius. If the truncation is strong, then a far-field beam profile function, which is a combination of Gaussian and Airy patterns, results.

1.3.2 Beamwidth

Beamwidth is a measure of the angular extent of the transmitted laser beam profile. Unfortunately, there is no uniform or standard definition for the specification or measurement of beamwidth. Generally, beamwidth is based on those positions within the beam where the intensity falls to a specified fraction of the peak value or on the angular subtense of a circle that inscribes a specified fraction of the total transmitted power. In addition, some conventions are based on half width instead of the full beamwidth.

For example, the beamwidth of a conventional radar system is specified as the full width at half maximum (FWHM). This specification is independent of the beam profile. However, for an optical system, the manner in which the beamwidth is specified is dependent on the beam profile. For example, the width of an Airy pattern is taken as the angular subtense of the first dark ring (first minimum), whereas for a Gaussian the width is defined as the angular separation of the e^{-2} intensity points.

Regardless of how beamwidth is specified, the minimum beamwidth or divergence, in the far field, is limited by the transmitter output aperture diameter as a result of diffraction. The actual, transmitted beamwidth is normally larger than the diffraction limit. Imperfections in the transmitter and the transmitter optical train distort the phase front, which, in turn, results in a broader far-field divergence. The deviation of the performance of a real system from a theoretical optimum is typically characterized by a beam quality, a Strehl ratio, or both.

The definition of the beamwidth for a laser radar system is based on accepted optical practices. This definition differs from conventional radar systems where the beamwidth is usually defined as the full width at half of the maximum.

1.3.2.1 Diffraction-Limited Beamwidth. The beamwidth of a diffraction-limited, transmitted beam that uniformly illuminates a circular output aperture is given by

$$\varphi = \frac{2.44\lambda}{d}, \tag{1.9}$$

where

φ = transmitter beamwidth in radians
λ = carrier (optical) wavelength in meters
d = transmitter clear aperture diameter in meters.

This definition of beamwidth is based on the angular width of the first null in an Airy pattern, which is also known as the Airy disk. Eighty-four percent of the transmitted optical power is inside this beamwidth in a diffraction-limited system. This definition of beamwidth is also commonly used in conjunction with a uniform beam profile function for first-order system performance analysis.

The effective beamwidth of a laser radar beam with a Gaussian profile is defined as the full width at the e^{-2} irradiance points. The effective beamwidth of a diffraction-limited, transmitted beam that has a Gaussian profile is given by

$$\varphi = 2 \arctan \frac{\varphi_0}{r} \left[1 + \left(\frac{\lambda r}{\pi \omega_0} \right)^2 \right]^{1/2} , \tag{1.10}$$

where ω_0 is the Gaussian beam waist radius in meters, and all other terms are as previously defined.

In the far field, i.e., $r \geq \pi\omega_0/\lambda$, the beamwidth of a Gaussian beam can be approximated by

$$\varphi = \frac{2\lambda}{\pi \omega_0} , \tag{1.11}$$

where all terms are as previously defined.

As mentioned earlier, the beamwidth of a Gaussian is defined by optical conventions as the full width across the beam measured to the e^{-2} irradiance points, whereas the conventional radar definition of beamwidth is the FWHM. If the beamwidth of a particular system that has a Gaussian beam profile is specified by its FWHM, it can be converted to an equivalent e^{-2} width by

$$\varphi = 1.699 \varphi_{\text{FWHM}} , \tag{1.12}$$

where φ_{FWHM} is the full beamwidth at half maximum intensity and all other terms are as previously defined.

1.3.2.2 Beam Quality. Various definitions have been used for beam quality. The performance of real systems is limited by the state of the art of components and by imperfect assembly of these components. A rigorous analysis of the performance of these systems must include an assessment of the actual beam profile as a function of range. A functional representation of the actual beam profile can then be included in laser radar range equation calculations. Rigorous analysis is time consuming and is frequently replaced by a variety of approximation techniques. Beam quality is one example.

A rigorous definition of beam quality is not known to exist. However, in this chapter, beam quality is defined by Eq. (1.13) as the ratio of the actual, far-field beamwidth to the theoretical, diffraction-limited beamwidth. As mentioned, the beamwidth may be measured in two different ways. It is either the width of the beam measured between the points at which the intensity is a specified fraction of the peak value (e.g., FWHM) or it is the angular subtense of a circle that inscribes a specified fraction of the total transmitted power. The exact interpretation of beam quality, therefore, depends on the technique used to define beamwidth.

$$Q = \frac{\varphi_M}{\varphi_T} , \tag{1.13}$$

where

Q = beam quality
φ_M = actual transmitter beamwidth in radians
φ_T = theoretical, diffraction-limited transmitter beamwidth.

Diffraction-limited performance naturally implies $Q \equiv 1$ in all cases. However, because beam quality does not have a universal definition, other values of Q do not have specific meaning without further explanation. Uniform illumination across the transmitter aperture is commonly assumed when computing the theoretical beamwidth for beam quality determination. A superior approach is to use the aperture illumination profile employed in the design of the system under analysis (e.g., Gaussian). Because the beam quality is also an indicator of the quality of the transmitter optical path, the use of the design aperture profile tends to give a better indication of the overall quality of the system.

1.3.2.3 Divergence Aperture Product. It is common practice to specify the product of divergence (in milliradians) and aperture diameter (in millimeters) for solid-state lasers (e.g., Nd:YAG, Nd:GSGG, etc.) instead of the beam quality. For these systems, Eq. (1.13) can be expressed as

$$Q = \frac{x}{2.44\lambda} , \tag{1.14}$$

where $x = \varphi_M d_1$ = divergence aperture product in meter-radians, d_1 = laser beam diameter in meters, and all other terms are as previously defined. Note that most manufacturers do not use SI units when specifying the divergence aperture product.

1.3.2.4 Strehl Ratio. The Strehl ratio provides another way to approximate the performance of laser radar systems. The Strehl ratio [Eq. (1.15)] is defined as the ratio of the intensity of the actual beam along the optical axis to the on-axis intensity of a diffraction-limited beam[4]:

$$S = \frac{I_M}{I_T} , \tag{1.15}$$

where

S = Strehl ratio
I_M = measured peak transmitter beam intensity in W m^{-2}
I_T = theoretical peak transmitter beam intensity in W m^{-2}.

As with beam quality, uniform illumination of the transmitter aperture is often assumed and the theoretical, on-axis intensity used is an Airy pattern. Also as with beam quality, diffraction-limited performance implies that $S \equiv 1$.

The Strehl ratio can be approximated by[5]

$$S = 1 - 4\pi^2\left(\frac{\text{OPD}}{\lambda}\right)^2, \quad \text{OPD} < \frac{\lambda}{10} , \tag{1.16}$$

where S is the Strehl ratio, OPD is the rms optical path difference, and all

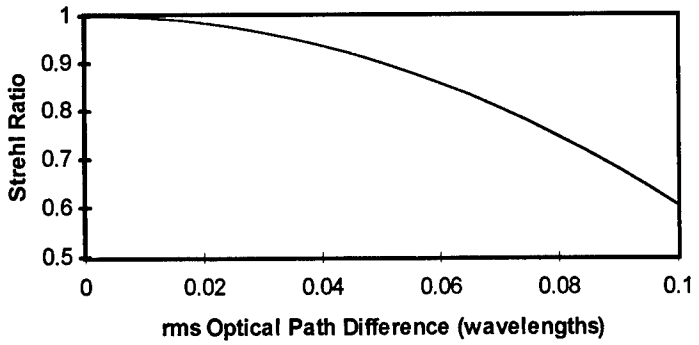


Fig. 1.5 Strehl ratio as a function of rms optical path difference.

other terms are as previously defined. This approximation is valid only for Strehl values greater than 0.64.

Both the beam quality and the Strehl ratio are related to the wave-front error of the transmitted beam. The wave-front error is also known as the optical path difference (OPD) and is measured in fractions of a wavelength. Figure 1.5 shows the Strehl ratio of a uniform beam as a function of rms OPD values. For rms OPD values less than $\lambda/14$, the beam is essentially diffraction limited. For rms OPD values greater than $\lambda/3$, beam intensity profiles deviate strongly from the theoretical, and Strehl ratio and beam quality are less meaningful parameters.

1.3.2.5 Non-Diffraction-Limited Beamwidth. Large departures from diffraction-limited performance are seldom the case for coherent laser radar systems. These are typically high-performance systems that mandate high-performance optics, so optical losses are minimized. However, small deviations from nonideal conditions are common and must be addressed to assess real-world performance and limitations.

The beamwidth of a non-diffraction-limited transmitted beam that uniformly illuminates the output aperture is given by

$$\varphi = \frac{2.44\lambda}{d} Q \ , \tag{1.17}$$

where all terms are as previously defined. Equation (1.17) is also frequently used to describe the beamwidth of a non-diffraction-limited system that is modeled as having a uniform transmitted beam profile.

The beamwidth of a non-diffraction-limited transmitted beam that has a Gaussian profile is given by

$$\varphi = 2Q \arctan\frac{\omega_0}{r}\left[1 + \left(\frac{\lambda r}{\pi\omega_0}\right)^2\right]^{1/2} . \tag{1.18}$$

Both Eqs. (1.17) and (1.18) are valid only for values of Q close to 1.

1.3.3 Transmitter Pointing Error

The target is seldom located at the center of the illuminating beam. Imperfect knowledge of target location and the inability to direct the beam to the desired position contribute to this effect. The target is illuminated by a region of the beam that has a lower intensity than the beam center. Less energy is reflected and received. As a result, performance estimates, which presume that the target is located at the center of the beam, are too optimistic.

Pointing error has less effect on the performance of a system that operates against extended targets. Because an extended target is larger than the beam, the target still intercepts all the transmitted energy, even in the presence of the pointing error. However, the pointing error tends to corrupt the information obtained. For example, a system that is intended to obtain highly accurate range information should also maintain small pointing errors so that the target position in three-dimensional space can be determined. Pointing errors result in cross-range errors, even when the radial range measurement is perfect.

1.3.4 Solved Problems

Problem 1.3. Calculate the effective beamwidth in the far field of a 10.6-μm, Gaussian beam having a 33-mm beam waist diameter and a beam quality of 1.2.

Solution 1.3. The diffraction-limited beamwidth of a Gaussian beam in the far field is given by Eq. (1.11):

$$\varphi = \frac{2\lambda}{\pi\omega_0} .$$

Here, the wavelength is 10.6×10^{-6} m and the Gaussian waist radius is 16.5×10^{-3} m. (Note that the problem identified the beam waist diameter.) The diffraction-limited beam width is then

$$\varphi = \frac{2(10.6 \times 10^{-6}\ \text{m})}{\pi(16.5 \times 10^{-3}\ \text{m})} = 4.09 \times 10^{-4}\ \text{rad} .$$

From Eq. (1.13), the measured beamwidth is

$$\varphi_M = Q\varphi_T .$$

For a system with a beam quality of 1.2

$$\varphi_M = 1.2 \times 4.09 \times 10^{-4}\ \text{rad} .$$

The actual beamwidth is then 4.91×10^{-4} rad.

Problem 1.4. Calculate the beam quality of a 0.830-μm beam that uniformly illuminates a 25.4-mm aperture and has a measured far-field divergence of 100 μrad.

Solution 1.4. The diffraction-limited beamwidth of a uniformly illuminated aperture is given by Eq. (1.9):

$$\varphi = \frac{2.44\lambda}{d} .$$

For a wavelength of 0.83×10^{-6} m and an aperture diameter of 0.0254 m, the theoretical, diffraction-limited beamwidth is 79.7 μrad.

The beam quality is defined by Eq. (1.13) as the ratio of the measured beamwidth to the theoretical:

$$Q = \frac{\varphi_M}{\varphi_T} .$$

For a measured beamwidth of 100 μrad, the beam quality is

$$Q = \frac{100}{79.7} ,$$

$$\therefore Q = 1.254 .$$

Problem 1.5. Calculate the Strehl ratio of a uniform, 10.6-μm beam having an rms OPD of 0.7 μm.

Solution 1.5. The Strehl ratio for small values of OPD can be calculated by using Eq. (1.16):

$$S = 1 - 4\pi^2\left(\frac{\text{OPD}}{\lambda}\right)^2 .$$

Inserting into Eq. (1.16) for the wavelength and the rms OPD, yields

$$S = 1 - 4\pi^2\left(\frac{0.7 \times 10^{-6}\ \text{m}}{10.6 \times 10^{-6}\ \text{m}}\right)^2 ,$$

$$\therefore S = 0.83 .$$

1.4 ATMOSPHERIC PROPAGATION

The performance of a laser radar system is often limited by the propagation characteristics of the earth's atmosphere. The phenomenology of atmospheric propagation for laser radars is similar to that of other optical systems. This section discusses those features of atmospheric propagation that are particularly important to laser radar systems.

In principle, the atmospheric transmission for either active or passive optical systems can be calculated by Beer's law:

$$T_A = \exp\left[-\int_0^r \int_{\lambda_1}^{\lambda_2} \alpha_d(w,l)\, dw\, dl \right] , \tag{1.19}$$

where

α_d = atmospheric extinction coefficient density function in inverse square meters
r = range to the target in meters
l = range integration variable
w = wavelength integration variable
λ_1 = shortest wavelength of interest in meters
λ_2 = longest wavelength of interest in meters.

The atmospheric extinction coefficient density function has units of reciprocal path length and reciprocal wavelength (not reciprocal area). The difficulty with this approach is the determination of the coefficient appropriate for the laser line and the atmospheric conditions over the specific optical path. Because the transmitted radiation of a laser radar has a very narrow linewidth, the extinction coefficient used in the calculation must be appropriate for exactly that laser wavelength.

If the atmospheric extinction coefficient density function is constant over the beam path and spectral region, then Eq. (1.19) reduces to the more familiar form of Beer's law as

$$T_A = e^{-\alpha r} , \tag{1.20}$$

where α is the atmospheric extinction coefficient in inverse meters, and all other terms are as previously defined.

The atmospheric extinction coefficient is also often expressed in units of decibels per kilometer. In these cases,

$$\alpha = \frac{\gamma \ln 10}{10^4} , \tag{1.21}$$

where γ is the atmospheric extinction coefficient in decibels per kilometer and all other terms are as previously defined.

1.4.1 Molecular Absorption and Scattering

Some features of the interaction of laser radars with the atmosphere are different than those encountered in routine practice with conventional optical systems. Most of these differences are the result of the interaction of the highly monochromatic laser radiation with the fine absorption structure of the atmosphere. The fine structure of the atmosphere is the result of molecular absorption and scattering, both of which are strongly wavelength dependent. Conventional (passive) optical systems typically operate over bandwidths that are large compared to the width of most molecular absorption lines. As a result, the response of passive systems is integrated over the entire band and the effects of the fine structure are "averaged out."

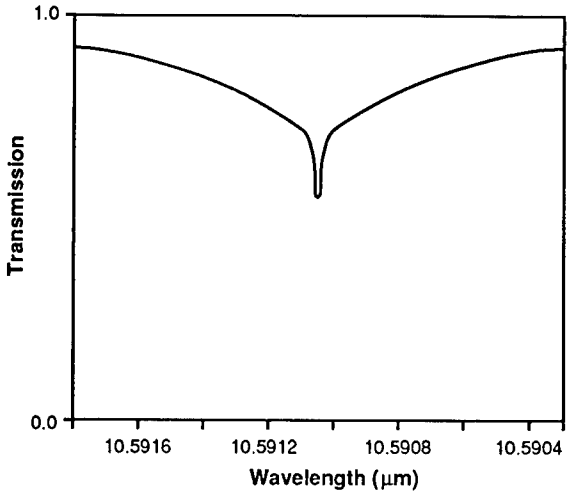


Fig. 1.6 Atmospheric transmission around 10.591 μm (vertical path from 3.65 km above sea level to space).

The effect of the fine structure of the atmosphere on laser radar performance is most severe for systems that operate over long ranges and use a naturally occurring atmospheric gas as the laser gain medium. In these cases, there is an unavoidable coincidence of the laser line with an atmospheric absorption line. For example, Fig. 1.6 shows the transmission through the atmosphere, from the ground vertically to space, in the spectral region around the P20 line of a normal isotope CO_2 laser (10.591 μm). The P20 line coincides with the sharp drop in transmission near the center of the figure. The average transmission over the wave band shown is over 90%. The actual transmission at the P20 line is only 60%.

The fine structure of the atmospheric transmission can also have significantly different effects on the atmospheric transmission of the transmit and receive paths. If the transmitter or the target is in motion, the Doppler effect will shift the carrier wavelength on transmission or reflection respectively. The absorption line shown in Fig. 1.6, which coincides with the P20 CO_2 laser line, has a FWHM of 175 MHz. The Doppler shift of 10.591-μm radiation reflected off an object traveling at 6 km/s (a common reentry velocity) is 1.13 GHz. This places the received radiation for such a laser radar completely away from the absorption of atmospheric CO_2 molecules.

Alternatively, a transmitted laser line, which normally does not coincide with an atmospheric absorption line, can be Doppler shifted so that it experiences significantly higher attenuation. Although this is a rare occurrence, the system design of the laser radar must consider the manner in which it will be used to ensure that such conditions are avoided.

1.4.2 Weather

Atmospheric transmission, particularly in the lower atmosphere, is a function of prevailing weather. As a result, "all-weather" operation requires that the

design of laser radar accommodate the worst conditions that are expected to be encountered. The nature of these extreme conditions is dependent on the geographic location of operation, season of the year, and the time of day.

Prevailing atmospheric conditions are frequently recorded as visibility and weather type (rain, fog, snow, etc.). Visibility is a subjective evaluation of the greatest distance at which a prominent dark object can be seen above the horizon during the daytime or the greatest distance at which a moderately bright light can be seen at night.[6] This may have meaning to a pilot or an air traffic controller, but, by itself, does little to quantify the atmospheric transmission losses of laser radar. In some cases, however, it is the only information available to the laser radar designer concerning the frequency of occurrence of weather at a particular location.

In addition to being vaguely defined, visibility is also highly subjective. It is based on the opinion of a human observer. Different observers at the same site perceive visibility differently. These observers typically use a few discrete landmarks to determine the visibility. Therefore, the range of values reported for visibility tends to be discrete and not continuous. Because the range to observable landmarks varies from site to site, the set of discrete values reported also varies from site to site. However, 1, 5, 7, 15, and 30 miles tend to be popular reported values.

This measurement approach tends to underestimate visibility. For example, an observer who can see a tower at 5 miles but who cannot see the top of a mountain at 10 miles will report the visibility as being 5 miles, even though the visibility could have been anything between 5 and 10 miles. Only the range to the farthest identifiable object is reported.

Meteorological range is an attempt to provide a more rigorous quantification of atmospheric propagation than is provided by visibility. It is defined as the range at which the atmospheric transmission is 0.02 or 2%. Empirically, it has been found that the meteorological range at midvisible wavelengths (around 0.5 μm) averages 8 to 10% longer than the reported visibility. Although meteorological range is well defined, it is rarely measured. Nearly all site climatologies are compiled from visibility and not from meteorological range.

If the absorption in the midvisible can be related to the absorption at the laser wavelength of operation, then reasonable estimates of the atmospheric transmission may be possible. Figure 1.7 shows the clear air attenuation coef-

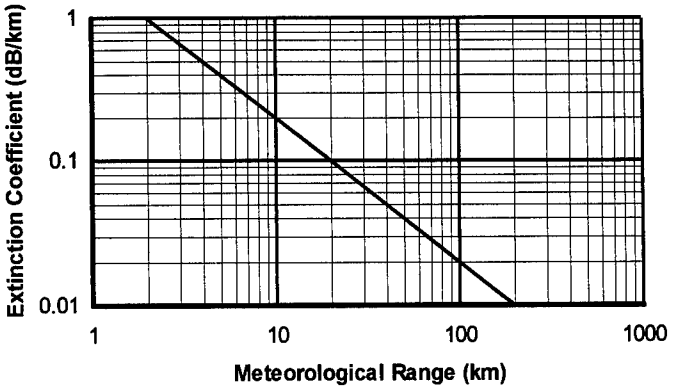


Fig. 1.7 Clear air extinction coefficients[7] at 10.6 μm.

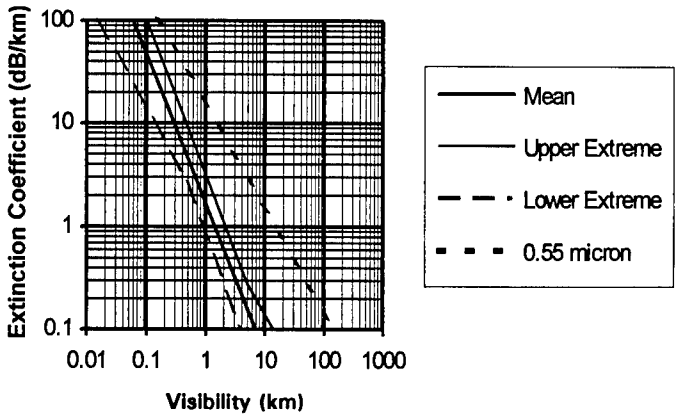


Fig. 1.8 Fog extinction coefficients[8–11] at 10.6 μm.

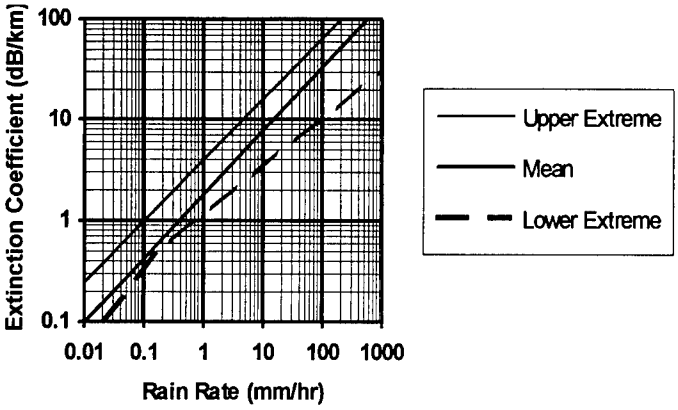


Fig. 1.9 Rain extinction coefficients[9,14,15,41] at 10.6 μm.

ficient at 10.6 μm as a function of meteorological range. Attenuation versus meteorological range in fog at 10.6 μm is shown in Fig. 1.8. The attenuation at 0.55 μm (midvisible) is also plotted for comparison.

The determination of weather type is less subjective than the determination of visibility. However, it is less straightforward than is normally expected from routine experience. Rain must be characterized by rain rate. Also, two distinct types of fogs and over 30 different types of snow have been identified. Each of these can have significantly different propagation characteristics.

Precipitation also has a strong effect on atmospheric attenuation. However, because rain drops tend to be much larger than the laser wavelengths, the attenuation from rain is not strongly wavelength dependent. Rensch and Long[12] have shown that the actual attenuation caused by rain does not vary significantly from 0.63 to 10.6 μm. Attenuation resulting from rain at 10.6 μm is shown in Fig. 1.9. Snow attenuation is shown in Fig. 1.10.

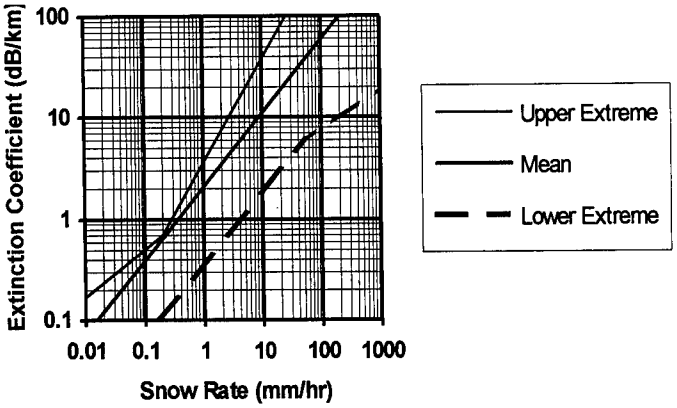


Fig. 1.10 Snow extinction coefficients[8,16,17] at 10.6 μm.

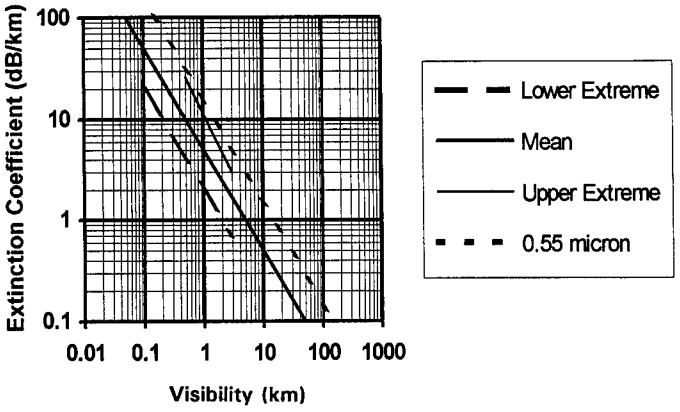


Fig. 1.11 Dust extinction coefficients[8,18] at 10.6 μm.

Along with natural atmospheric conditions, laser radar may be required to operate through or in the presence of natural or man-made obscurants. These include dirt (such as is ejected into the air by the explosion of a mortar round), dust (thrown up by moving vehicles), and smoke (from fires, factory smoke stacks, smoke grenades, or battlefield smoke generators). Figure 1.11 shows the attenuation caused by dust as a function of meteorological range.

Extensive weather data bases have been developed and are maintained[13] to assist in the determination of the climatic extremes and nominal conditions that a system may encounter. Table 1.2 shows the statistical occurrence of various weather types in central Europe. Such statistical data are required to establish most severe operating requirements to ensure a particular probability of performance. A variety of atmospheric models and empirical formulas have been developed to assist in the estimation of weather effects on system performance. Some of these empirical relationships are summarized in Table 1.3.

Table 1.2 Statistical Occurrence of Weather in Central Europe (from Ref. 19)

		Monthly Percentage												
Weather Condition	Visibility	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Mean Value
Very Fair Weather	v>18 km	3	4	5	10	12	15	18	15	5	4	3	2	8.0
Fair Weather	9km<v< 18km	16	19	22	33	38	40	40	39	39	28	20	19	29.4
Light Haze	4km<v< 9km	38	38	40	41	37	34	28	30	37	37	36	34	35.8
Haze	1.8km<v < 4km	26	26	23	12	10	9	10	12	13	18	24	25	17.3
Light Fog	0.9km<v < 1.8km	8	7	6	2	2	2	2	2	3	4	8	8	4.5
Fog	0.5km<v < 0.9km	3	2	2	1	1	0.5	1	1	1	2	3	3	1.7
Heavy Fog	v<0.5km	6	4	2	1	0.5	0.5	0.5	1	2	7	6	9	3.3

Table 1.3 Attenuation Coefficients at 10.59 μm (from Refs. 8 and 20)

Condition	Attenuation (dB/km)
Clear	$1.084 \times 10^{-5} p(P + 193\,p)\left(\frac{296}{T}\right)^{5.25} + 625\left(\frac{296}{T}\right)^{1.5} \times 10^{\frac{-970}{T}} + \frac{1.4}{V}$
Rain	$1.9R^{0.63}$
Snow	$2S^{0.75}$
Fog	$\frac{1.7}{V^{1.5}}$
Dust	$\frac{5}{V}$
where	V = visibility (km), S = snowfall rate (mm/hr), R = rainfall (mm/hr), T = temperature (°K), P = atmospheric pressure (millibars), and p = partial pressure of water vapor (millibars).

1.4.3 Computer Models

Computer models have been developed to aid in the calculation of atmospheric transmission.[21–23] The most common atmospheric models are LOWTRAN, MODTRAN, and FASCODE, which were developed and are maintained by the Air Force Phillips Laboratory. However, many computer models of atmospheric propagation, such as LOWTRAN, calculate the mean transmission over a small, but finite, wave band using Beer's law and an absorption coefficient that is averaged over the wave band. The actual absorption at a specific laser radar wavelength may be much stronger than that predicted by the average absorption over a wider bandwidth. As a result, these low-resolution computer models that neglect the fine structure are useful for laser radar design and analysis only under a restricted set of conditions. Those conditions occur when the total atmospheric extinction, at the specific laser line, is dominated by the broadband aerosol absorption and scattering and when the molecular components, which tend to be very narrow band phenomena, can be neglected.

Specifically, LOWTRAN computes atmospheric transmission over relatively wide spectral bands (typically 20 cm^{-1}). This low spectral resolution may give false indication of the transmission of laser radar because it integrates all spectral transmission fine structure over a bandwidth of 20 cm^{-1}. However, LOWTRAN can be used effectively for laser radar propagation modeling when the atmospheric extinction is dominated by airborne aerosol absorption and scattering. In this case, the predominant optical loss is broadband and the integrated absorption over the band is approximately equal to the wavelength specific absorption.

FASCODE provides much finer spectral resolution than LOWTRAN because it computes transmission on a molecular, line-by-line basis. FASCODE calculations require an extensive catalog of absorption lines for all atmospheric gases. This modeling approach makes FASCODE very comprehensive. The primary limitation on the utility of FASCODE is the need for large amounts of computer memory to support the absorption line data base and long processing times required to complete propagation calculations.

MODTRAN provides a modeling approach with performance intermediate to LOWTRAN and FASCODE. MODTRAN predicts atmospheric transmission with a resolution of 2 cm^{-1}. However, this transmission is still much wider than most laser lines.

Another useful atmospheric transmission model is the Electro-Optical Systems Atmospheric Effects Library (EOSAEL), which was developed by the U.S. Army Atmospheric Sciences Laboratory. EOSAEL is not just a single model but a collection of models and data bases. In addition to "classical" transmission models, EOSAEL includes transmission through dust and dirt, various types of smoke, and other battlefield obscurants.

1.4.4 Solved Problems

Problem 1.6. Calculate the performance degradation for a 10.6-μm system resulting from 5 mm/hr rain rate at a range of 1.5 km; at 10 mm/hr; at 5 km.

Solution 1.6. The attenuation coefficient as the result of 5 mm/hr rain rate can be found from Table 1.3 to be $1.9 \times 5^{0.63}$ or 5.24 dB/km. For a target at 1.5-km range, the round-trip path length is 3 km. The total extinction is 3 ×

5.24 or 15.72 dB. The signal transmission factor is then $10^{-(15.72/10)}$ or 0.027. At 10 mm/hr, the attenuation increases to $1.9 \times 10^{0.63}$ or 8.11 dB/km. Over a 3-km path length, the total attenuation is 3×8.11 or 24.33 dB. The signal transmission factor in this case is $10^{-2.433}$ or 0.0037. Increasing the range for the case of 5 mm/hr rain rate increases the total signal loss but not the loss per kilometer. The attenuation at 5 km for 5 mm/hr rain rate is 10×5.24 or 52.4 dB. The signal transmission factor is, therefore, $10^{-5.24}$ or 0.00000575.

Problem 1.7. Determine the fraction of the time the attenuation of 0.55-μm radiation exceeds 10 dB/km.

Solution 1.7. The meteorological range is defined as the range at which the transmission is 0.02. This is an optical power loss of 17 dB. A total path length of 1.7 km is required to produce this optical loss under the conditions of 10 dB/km attenuation. Because meteorological range is approximately 10% longer than visibility, the visibility is then 1.7/1.1, or approximately 1.5 km (1545 m).

From Table 1.2, 1.5-km visibility is typical of light fog, which occurs 4.5% of the time. Two conditions with lower visibility (higher attenuation) are fog and heavy fog with 1.7 and 3.2% probabilities of occurrence, respectively. Because 1.5-km visibility is an intermediate value for light fog (light fog is defined as 0.9- to 1.8-km visibility), 4.5% should be considered as an upper bound on the percentage of the time that the visibility is between 0.9 and 1.5 km. The actual percentage will probably be less than 4.5%. The exact amount cannot be determined from Table 1.2.

The probability that 10 dB/km or higher attenuation will occur for this location is the cumulative probability for these three conditions. It will not exceed 9.4% (4.5% + 1.7% + 3.2%) nor be less than 4.9% (1.7% + 3.2%).

1.5 TARGET LASER CROSS SECTION

The laser radar cross section (LCS) of an object is defined as the cross-sectional area of a perfectly reflecting sphere that produces an intensity at the receiver equal to that of the object.[c] The laser radar cross section of a sphere is independent of orientation or aspect angle. It is given by[24]

$$\Gamma = \pi \rho z^2 , \tag{1.22}$$

where

Γ = laser radar target cross section in square meters
ρ = reflectivity of the sphere's surface
z = radius of the sphere in meters.

Although the LCS of an object has the units of area, it may have little relationship to any physical dimensions of the object. In general, the more nearly the object resembles a sphere, the stronger the relationship of the LCS with the physical dimensions of the object.

[c]The laser cross section or laser radar cross section has the dimensions of area. Other definitions, based on "standard" targets other than a sphere, are occasionally encountered. However, the LCS is only loosely associated with the physical dimensions of the targets and the numerical values ascribed to the LCS are somewhat arbitrary.

Most laser radars are sensitive to only one component of polarization. This is especially true of coherent systems because the optical heterodyne mixing process requires that the local oscillator and the received signal have matching phase fronts and polarization (see Sec. 1.8). However, many target surfaces tend to depolarize the return signal. For a polarization-sensitive receiver, the LCS of an object is reduced proportionally by the depolarization of the object's surface.

1.5.1 Specularly Reflecting Targets

The perfectly reflecting sphere, which is used to define LCS, is only one type of specular target. Any spherical object, whose rms surface roughness is significantly less than the laser wavelength, will generate a specular return. The unique feature of a spherical target, which is uniformly illuminated by the laser beam, is that its reflection is equal and specular in all directions. A specular return is frequently referred to as a glint.

1.5.1.1 Cube Corner Reflector. Another important target of this class is the cube corner reflector. Cube corner reflectors are not just specular, they are also retroreflective, i.e., all the incident energy is returned along the illuminating line of sight, within the limits of optical diffraction. Hollow corner reflectors can be formed by joining three mirrors at right angles to each other for broad optical band operation. Another common corner reflector is a right triangular pyramid made from solid glass. This type of corner reflector is also known as a reflex prism and exploits total internal reflection to return the incident energy. Its use is limited to the optical passband of the material used. The LCS of a corner reflector is approximately proportional to the cosine of the angle of incidence (for small angles) because the projected area of the reflector decreases as the incident angle increases.

As a result of diffraction, the solid angle of the returned radiation depends on the diameter of the corner reflector and the wavelength of the illumination. Arrays of corner reflectors increase the target area proportionally to the number of reflectors, but do not decrease the return-angle divergence. As the angle of incidence increases, the reflecting area changes from a circle (or triangle or hexagon, etc.) to a narrow slit at extreme incidence angles.

The LCS of a cube corner reflector is given by[1,24]

$$\Gamma = 4\pi \frac{l^4}{3\lambda^2} , \tag{1.23}$$

where l is the length of the edge of the cube corner reflector in meters and all other terms are as previously defined. Equation (1.23) is valid if the curvature of the transmitted signal is less than $\lambda/4$ across the corner reflector.

1.5.1.2 Reflective Sheeting. Reflective sheeting consists of an array of small glass beads, reflectively coated on the rear halves of the spheres. The return divergence is determined by the inherent properties of the glass beads, so increasing the size of the sheeting does not reduce the return divergence. Reflective sheeting comes in paint and sheet form, and has a cross section from 100 to 1500 times larger than ordinary diffuse white paint.

Unlike a hollow retroreflector, which may use first surface mirrors and be useful into the far infrared, reflective sheeting is useful only in the visible and near infrared. At longer wavelengths, the glass beads become opaque.

1.5.2 Diffusely Reflecting Targets

If the rms roughness of the target surface is of the order of the laser wavelength or larger, then the reflected signal will be scattered over a large volume. The distribution and amplitude of the reflected light are characterized by the bi-directional reflectance distribution function (BRDF). The BRDF of a diffuse target depends on a large number of material-dependent and sample-specific factors. Experimental measurement is generally required to determine the BRDF of a specific material. However, the BRDF of many materials are similar enough that general classes can be assigned.

1.5.2.1 Lambertian Surfaces. A Lambertian surface is one in which the intensity of the scattered radiation obeys Lambert's cosine law. Specifically, the intensity (i.e., flux per unit solid angle) reflected in any given direction from the surface of the material is proportional to the cosine of the angle between that direction and the normal to the surface. Lambertian surfaces are sometimes mistakenly called *isotropic*. This error is the result of the fact that when the angular subtense of the illuminated area is equal to or greater than the instantaneous field of view of the receiver, then the power of the radiation observed by the receiver is independent of the angle between the receiver and the normal to the surface of the target. This effect is a result of the fact that the area of the illuminated target surface is inversely proportional to the cosine of the angle between the target normal and the direction to the receiver, whereas the reflected intensity is proportional to the cosine. The actual area of the target that contributes to the power at the receiver increases while the contribution from each infinitesimal element of the target to the received power decreases in exactly the same proportion. Thus, the two effects cancel, and no angular dependence is observed.

The laser radar cross section of a flat Lambertian disk that is smaller than the transmitted beam is given by

$$\Gamma = 4\pi\rho z^2 \cos\theta \ , \tag{1.24}$$

where

ρ = total hemispherical reflectance of the Lambertian surface
z = radius of Lambertian disk in meters
θ = angle of incidence on Lambertian surface in radians,

and all other terms are as previously defined.

Perfect Lambertian surfaces are rare in nature but are often assumed in analysis. The best natural example is snow. However, many surfaces do approximate Lambertian behavior. BRDF plots of real surfaces typically display a combination of specular and Lambertian (diffuse) reflection components. Perfect Lambertian surfaces depolarize incident light perfectly. The polarized LCS of a true Lambertian target is, therefore, half of the unpolarized LCS.

Equations (1.23) and (1.24) are for targets that are much smaller than the beam diameter. They are appropriate for use with any of the beam profile functions.

When the target is larger than the beamwidth, the beam profile function and pointing error are less important factors in determining maximum range performance. Airborne systems pointed at the ground routinely operate under this type of condition. However, as the result of the narrow beamwidths of laser radars, it is very common for an isolated target (e.g., a land vehicle or an airplane) to be larger than the beam. Under these conditions the laser cross section of a Lambertian surface is given by

$$\Gamma = \pi\rho\varphi^2 r^2 \ , \tag{1.25}$$

where all terms are as previously defined.

Wires represent a special case of an extended target. They are generally larger than the beam in only one direction. The laser cross section of a diffusely reflecting wire or thin target is given by

$$\Gamma = 4\rho D_w \varphi r \ , \tag{1.26}$$

where D_w is the diameter of the wire in meters and all other terms are as previously defined. Both Eqs. (1.25) and (1.26) are based on uniform illumination over the target surface; therefore, uniform beam profiles should be used in conjunction with these formulas.

Wires and cables frequently also exhibit a strong specular reflection component, depending on the particular wavelength of the illuminating laser. The angular position and subtense of this specular component has been shown to be related to the angle of twist of the wire strands in uninsulated wire and cable. Experimental measurements on the specific target of interest are recommended before a final system design is completed.

1.5.2.2 Aerosols and Volumetric Scatterers. In most laser radar systems, airborne aerosols, rain, snow and other particulates are clutter to be overcome. However, in a few specialized cases, these particulates are the laser target. As a target of laser radar, they exhibit the properties of a diffusely reflecting target.

The LCS of aerosols or volumetric scatterers is closely coupled to the propagation losses of the laser beam through that aerosol. As a result, the LCS cannot be selected independently of propagation losses. The strength of the return from an aerosol is characterized by the backscatter coefficient β. The backscatter coefficient is uniquely related to the LCS for a monostatic system by

$$\Gamma = \pi^2 \varphi^2 r^2 \beta \Delta r \ , \tag{1.27}$$

where β is the atmospheric backscatter coefficient in $m^{-1}\,sr^{-1}$, Δr is the range resolution of the laser radar in meters, and all other terms are as previously defined.

The backscatter coefficient β is also related to α, the Beer's law atmospheric extinction coefficient (see Sec.1.4). This relationship is dependent on the phys-

ical characteristics of the aerosol (number density, size distribution, shape, composition, etc.). However, although the aerosol scattering component of α has been shown to be almost wavelength independent in the visible and infrared,[12] β is not. The backscatter can be as much as 7 dB higher in the visible than in the infrared. The simplest relationship between α and β occurs for rain. In this case, the material and its optical characteristics are known and the number density, size distribution, and shape are a strong function of rain rate.

1.5.3 High-Resolution Target Cross Sections

The appropriate laser cross section is one that falls within the resolution limits of the receiver. In a low-resolution system (e.g., most rf radar applications), the target seldom extends beyond a single range, velocity, or angular resolution cell. This is not the case for many laser radars. If the target extends beyond the resolution cell limit of the system, then only a portion of the target contributes to the signal in that resolution cell.

For example, a short-pulse laser illuminating a physically deep target does not illuminate all portions of the target at the same time. Instead, successive "slices" of the target are illuminated as the laser pulse propagates down the length of the target. As a result, the signal generated at the receiver at any particular time later can only result from the reflection of a piece of the target. Under these conditions, the laser cross section is not characterized by a single value. It is a function of range (target depth) and has different values for each piece of the target.

1.5.4 Computer Models

Laser radar cross sections can be calculated for simple geometric shapes.[25–28] Computer models have been developed to calculate the laser cross section for complex shapes and/or for high-resolution laser radar systems. The best examples of these types of models are CALIBER 3 and DELTAS™, both of which have been developed by the U.S. Army Space and Strategic Defense Command (USASSDC), Huntsville, Alabama. All such codes make some form of simplifying assumption and, as a result, have some limits on their applicability.

CALIBER assumes that the target is illuminated uniformly over its surface. This assumption is reasonable if the target is small compared to the beam. The target is modeled as a collection of flat, triangular facets that are joined at their edges to form a closed geometric surface. This modeling approach represents a sphere by a polyhedron resembling a geodesic dome. Although this approach may not be aesthetically pleasing, it has shown very good agreement for objects without small radii of curvature and whose surface material does not have a strong specular reflection component.

DELTAS™ is a much more comprehensive code that permits incorporation of arbitrary illumination intensity profiles. The target surfaces are modeled as smooth geometric shapes instead of polyhedrons. However, the computational complexity of this code limits its use to large, mainframe-type machines. In addition, even though the illumination profile and target surface contours are modeled as continuous functions, numerical integration over the target surface is performed to compute the LCS. This results in a discrete sampling

of the beam and the surface, which negates much of the advantage of the modeling approach.

Both CALIBER and DELTAS™ rely on measured BRDF data on the target surface materials. Extensive BRDF data bases exist for a wide variety of materials. The Environmental Research Institute of Michigan (ERIM) maintains one of the largest BRDF data bases for materials of military significance. However, many of the data sets in these and other files are inadequately documented. It is common to find BRDF data recorded with only an extremely brief description of the test sample (e.g., olive drab green paint on plywood, flame sprayed aluminum, etc.), with no mention of method of the sample's preparation, list of materials used to prepare the sample, age of the sample, current condition (e.g., wet, dry, dirty, scratched from improper storage, etc.), calibration standards, or calibration procedure. Care should be taken when using these computer models for target LCS because the quality of the signature prediction will be no better than the quality of the data that the models employ.

1.5.5 Speckle

The laser radar cross section of a diffuse target for a coherent laser radar is dominated by the phenomenon known as *speckle*. Speckle arises from the reflection of coherent illumination from a diffuse surface. The diffuse surface creates an array of scatterers that is independent and randomly phased. This results in constructive (bright spots) and destructive (dark spots) interference at the observation plane.

In many laser radar systems, the receiver aperture is smaller than the mean transverse dimension of the speckle lobes (bright spots). In this case, the target cross section fluctuates rapidly as a Rayleigh distributed random variable. This condition is referred to as target fading in conventional rf radar systems and has been previously analyzed by Swerling.[29] Swerling evaluated four different target cases. These cases differed by the target cross-section probability density function assumed and by the rate of target cross-section fluctuation. A speckle target has a negative exponential density function and fluctuates rapidly (i.e., the target cross section is uncorrelated from one laser pulse to the next). A speckle target is functionally identical to the Swerling case 2.

If the receiver aperture is larger than the mean transverse dimension of the speckle lobe, then aperture averaging of the signal occurs. Averaging of the speckle lobes results in milder amplitude fluctuations, but it also results in a lower collection efficiency (see Sec. 1.6.2). Because the individual lobes must add coherently, their vector sum is less than their scalar sum.

1.5.6 Laser Radar Cross-Section Standards

Target standards are required to conduct laser radar system calibration. Unfortunately, target standards that are traceable to the National Institute of Standards and Technology standards are not generally available. The experimentalist is forced to improvise with available materials. Spherical calibration targets frequently used include steel ball bearings, steel ball bearings with a vapor-deposited gold coating, gold-coated convex mirrors, and flame-sprayed aluminum spheres.

When properly prepared, flat-surface calibration targets have been shown to closely approximate a Lambertian reflectance distribution. These include fresh flowers of sulfur, salt, silicon carbide sand paper, White Velvet Nextel™ paint,[d] Black Velvet Nextel™ paint, and flame-sprayed aluminum. Gold-coated sand paper can also be used as a high-reflectivity Lambertian target, but the BRDF should be experimentally verified because many samples have exhibited strong specular reflection components.

1.5.7 Solved Problems

Problem 1.8. Calculate the laser cross section of the ground at a range of 2 km for a 10.6-μm laser radar that has a beam waist diameter of 33 mm.

Solution 1.8. The ground is an extended target that is much larger than the beam at such a relatively short range. The divergence of a Gaussian beam in the far field is normally given by Eq. (1.11), but to compute the laser cross section of an extended target, Eq. (1.9) is often used instead. There is no effect on the computation of received power for extended targets as long as the same definition of beamwidth is used consistently. The beamwidth, from Eq. (1.11), is

$$\varphi = \frac{2\lambda}{\pi\omega_0} .$$

The cross section of an extended target is given by Eq. (1.25) as

$$\Gamma = \pi\rho\varphi^2 r^2 .$$

Substitution of Eq. (1.11) into Eq. (1.25) yields

$$\Gamma = \pi\rho\left[\frac{2\lambda}{\pi\omega_0}\right]^2 r^2 .$$

Dirt has a total reflectivity of approximately 4%. At a range of 2000 m, this results in a cross section of

$$\Gamma = \pi(0.04)\left[\frac{2(10.6 \times 10^{-6})}{\pi(33 \times 10^{-3}/2)}\right]^2 (2000)^2 ,$$

$$\therefore \Gamma = 0.084 \text{ m}^2 .$$

Problem 1.9. Calculate the laser cross section of a perfectly reflecting corner reflector that is 5 cm on a side and is illuminated by a CO_2 laser operating at 11.12 μm.

Solution 1.9. The laser cross section of a corner reflector is given by Eq. (1.23):

$$\Gamma = 4\pi\frac{l^4}{3\lambda^2} .$$

[d]Nextel is a registered trademark of the 3-M Corporation.

In this problem, $l = 0.05$ m and $\lambda = 11.12 \times 10^{-6}$ m. Substitution yields

$$\Gamma = 4\pi \frac{(0.05)^4}{3(11.12 \times 10^{-6})^2} ,$$

$$\therefore \Gamma = 2.12 \times 10^5 \text{ m}^2 .$$

Problem 1.10. Calculate the laser cross section of a 50-cm-diam polished aluminum sphere at 1.06 μm.

Solution 1.10. The laser cross section of a sphere is given by Eq. (1.22):

$$\Gamma = \pi \rho z^2 .$$

In this problem, $z = 0.25$ m. Bare aluminum has a reflectivity of approximately 0.99 at 1.06 μm. Substitution yields

$$\Gamma = \pi \times 0.99 \times 0.25^2 ,$$

$$\therefore \Gamma = 0.194 \text{ m}^2 .$$

1.6 RECEIVER CHARACTERISTICS

The optical receiver in laser radar is characterized by the detection technique used, collection area, optical efficiency, and detection efficiency. Optical efficiency and detection efficiency are derived quantities that depend on a large number of specific design parameters.

1.6.1 Receiver Architectures

Laser radars employ two different and distinct types of optical detection techniques. These techniques are direct detection and heterodyne or coherent detection.

1.6.1.1 Direct Detection. In direct detection laser radar (Fig. 1.12), the received optical energy is focused onto a photosensitive element that generates a voltage (or current) that is directly proportional to the optical power that strikes it. This process is identical to a conventional, passive optical receiver or to a typical laser rangefinder.

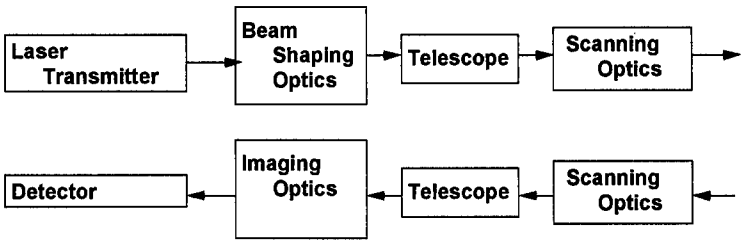


Fig. 1.12 The block diagram of a direct detection laser radar is similar in operation to a laser range finder.

1.6.1.2 Heterodyne Detection. In heterodyne (or coherent) detection laser radar (Fig. 1.1), the received optical energy is first mixed with a reference optical signal or local oscillator. The local oscillator is commonly generated by a continuous wave (cw) laser. This mixed signal is then focused onto a photosensitive element. If the signal and the local oscillator are spatially and temporally coherent, are well aligned with respect to each other, and have identical polarizations, then the two optical signals generate temporal interference at a frequency equal to the frequency difference between them. This is the heterodyne signal.

The optical heterodyne mixing process is illustrated in Fig. 1.13. If the local oscillator and the received signal have matched phase fronts over the detector surface, it can be shown that the intensity of the radiation falling on the detector is given by

$$I(t) = c\varepsilon_0[E_{lo}^2 + E_s^2 + 2E_{lo}E_s \cos(\delta\omega t)] \ , \tag{1.28}$$

where

$I(t)$ = intensity incident on the detector in watts per square meter
E_{lo} = local oscillator electric field amplitude in volts per meter
E_s = received signal electric field amplitude in volts per meter
$\delta\omega$ = $\omega_s - \omega_{lo}$
ω_s = received signal electric field frequency in Hertz
ω_{lo} = local oscillator electric field frequency in Hertz
ε_0 = electric permittivity of free space
c = speed of light in a vacuum.

As with direct detection, the photosensitive element in the receiver generates a voltage directly proportional to the optical power that strikes it. However, the optical power on the detector is now composed of three components: the original received signal (E_s^2), the local oscillator signal (E_{lo}^2), and the heterodyne signal. The heterodyne signal is at the difference frequency between the

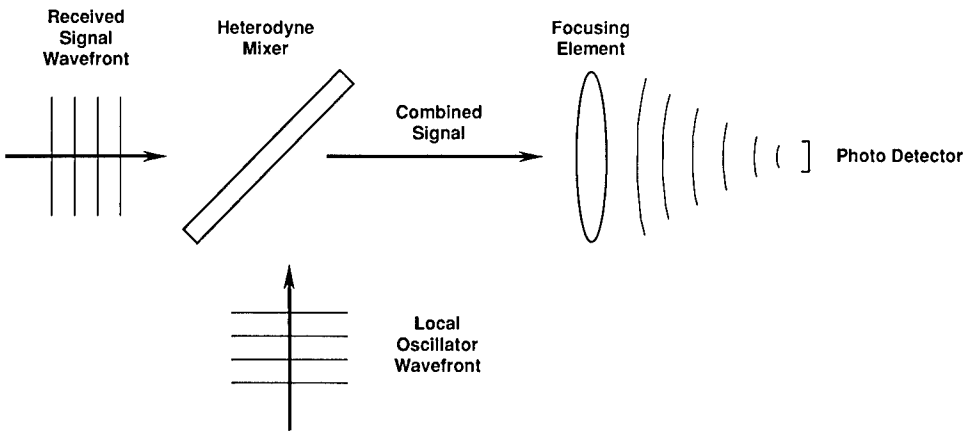


Fig. 1.13 Optical heterodyne mixing process.

local oscillator and the received signal. If the target is in motion relative to the laser radar, then the frequency of the received optical signal is altered by the Doppler shift. If the transmitted frequency and the local oscillator frequency are known, then the Doppler shift of the received signal can be determined and, as a result, the radial velocity of the target can be determined. The heterodyne signal preserves the phase information of the received signal carrier, whereas the direct detection signal preserves only the amplitude information of the received signal carrier envelope.

The detector also generates a dc electrical signal in response to the cw local oscillator. The heterodyne signal consists of an ac electrical signal that weakly modulates this large dc signal. Filtering of this combined signal is required to isolate the heterodyne signal, which carries all target information, from the local oscillator signal. This is easily accomplished with a high-pass electronic filter.

The power of the electrical (heterodyne) signal generated by the photodetector is proportional to the electric field strength of the received signal and to the electric field strength of the local oscillator. Likewise, the power of the shot noise generated by the dc electrical signal created in response to the cw local oscillator radiation absorbed by the detector is proportional to the electric field strength of the local oscillator. Because the local oscillator can be made arbitrarily intense (within the damage threshold and dynamic range of the detector), the heterodyne signal can be made to dominate over the direct detection signal. The local-oscillator-generated shot noise can then be made to dominate all other noise sources. Under these conditions, the heterodyne signal is limited only by shot noise. The SNR of the heterodyne detector then no longer depends on the intensity of the local oscillator.

There are, however, limits to this process. First, the local oscillator cannot in practice be made arbitrarily intense. All detector elements have some optical power damage limit. For a mercury cadmium telluride (HgCdTe) detector operating at 77 K, the damage threshold is approximately 10 W/cm^2. For a 100-μm detector element, the total optical power striking the detector must be less than about 1 mW.

1.6.1.3 Homodyne Detection. A conventional heterodyne receiver, as shown in Fig. 1.1, requires a separate laser source to serve as the local oscillator. In some cases, part of the laser radiation from the transmitter source can also be used as the local oscillator for the receiver. This technique, known as homodyne detection, eliminates the need for a separate laser source. A block diagram of a conventional homodyne laser radar transmitter and receiver is shown in Fig. 1.14. A portion of the laser beam is split off to form the local oscillator, which is mixed with the received optical signal. The optical detection process in a homodyne receiver is identical to that in a heterodyne receiver.

As with a heterodyne receiver, detection requires that the local oscillator illuminate the detector simultaneously with the received signal. Because the range to the target is generally not known, the time of arrival is also not known. As a result, the transmitter in a homodyne laser radar is generally a cw source. An unmodulated cw source is well suited for velocity measurements but is incapable of measuring range. If range information is required, the transmitted beam can be modulated after the local oscillator beam is split off but before the transmit beam enters the transmit to receive isolator.

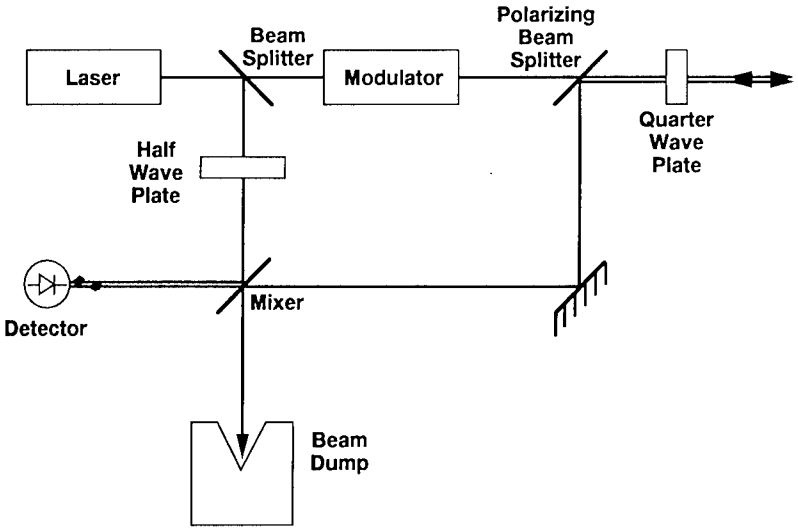


Fig. 1.14 Block diagram of a conventional homodyne laser radar.

Another approach to range measurement in a homodyne system is to frequency modulate (FM) the transmitter laser source (see Secs. 1.11.3 and 1.12.3.1). As a consequence, the frequency of the local oscillator is also modulated. Because the transmitter frequency is changing as a function of time, the round trip causes the received frequency to be different from the currently transmitted frequency and the current local oscillator frequency. The difference in frequency (neglecting any Doppler shift) between the received signal and the local oscillator signal is a measure of the range to the target. In principle, any frequency modulation format can be used (sinusoidal, sawtooth, triangular, etc.). In practice, saw-toothed and triangular waveforms are the most common.

If the local oscillator and the transmitted signal are not frequency modulated in a homodyne laser radar, then the electronic signal frequency is entirely the result of the Doppler shift. In this case, measurement of the homodyne signal frequency is a direct measure of the radial target velocity. As a result, homodyne systems that are not frequency modulated are limited in their ability to interrogate stationary objects, because the frequency of the homodyne signal from a stationary object is zero. In the case of a stationary target, the homodyne signal is directly competing with the dc local-oscillator-induced signal.

Homodyne laser radars have two primary advantages. First, they are simpler (and therefore more reliable) than heterodyne laser radars because they eliminate the need for a second laser source. The second advantage is that frequency stability requirements of the signal laser can be relaxed. The rate at which the laser frequency may drift can be determined by dividing the acceptable frequency error by the round-trip transit time of the laser radiation. Even at ranges in excess of 10 km, this round-trip time is less than 100 μs. The required stability is usually well within the natural behavior of many common laser sources.

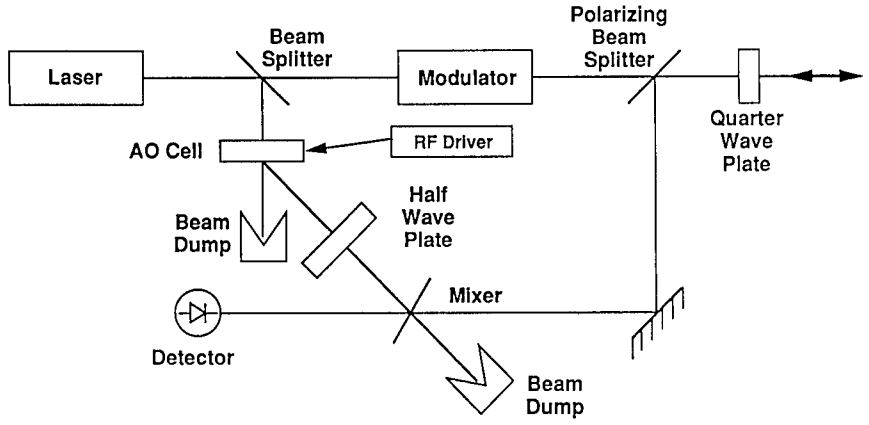


Fig. 1.15 Block diagram of a frequency offset homodyne laser radar using a single acousto-optic cell.

1.6.1.4 Offset Homodyne Detection. The difficulty in measuring stationary objects can be overcome with an offset homodyne receiver. In an offset homodyne receiver, the local oscillator beam is frequency shifted from the transmitter beam. Stationary objects generate a signal at the shift frequency, which can be made to be well separated from the dc local oscillator induced signal.

In principle, the frequency shift can be accomplished by inserting a frequency shifting device between the beamsplitter and the mixing element in the local oscillator path shown in Fig. 1.15. The shifting device is commonly an acousto-optic modulator (AOM), which requires an rf input signal at a frequency equal to the desired frequency shift of the optical beam. In practice, it is often found that the electromagnetic coupling (EMC) between the radio frequency AOM driver and the laser radar photodetector and/or the receiver electronics is so strong that two or more AOMs, each operating at a different rf frequency, have to be used. This technique is employed so that the total frequency shift of the optical radiation is not at the frequency of any of the rf drivers, nor at any of their harmonics, nor at any of the subharmonics that result from the mixing of the various rf drive frequencies.

1.6.1.5 Three-Frequency Heterodyne. The Doppler shift of a moving target provides very precise information on the velocity of the target. The target-induced Doppler shift can also require that the receiver have a very wide bandwidth so that the frequency of the returning, heterodyne-detected signal falls within the bandwidth of the receiver electronics. Increasing the bandwidth to accommodate the range of Doppler shifts that may be present increases the noise level within the receiver. This increased noise reduces the probability of detection or increases the probability of a false alarm, or both. A three-frequency heterodyne receiver overcomes this limitation.

A block diagram of a three-frequency heterodyne system is shown in Fig. 1.16. As can be seen, the three-frequency technique involves more than just the receiver. Two separate laser transmitter sources that are frequency offset

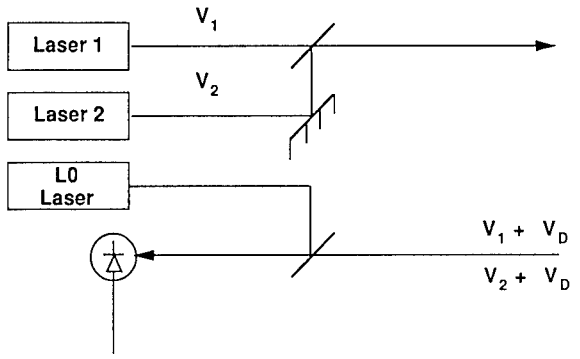


Fig. 1.16 Block diagram of a three-frequency heterodyne laser radar.

from one another by a small, but finite and fixed, amount are used. These two laser beams are then combined along a single optical axis and both beams are transmitted to the target. Both beams are Doppler shifted by the target's motion and reflected back to the receiver. At the receiver, the two reflected signals are mixed with the local oscillator and imaged onto a photosensitive detector as in a conventional heterodyne receiver.

Each received signal generates its own, separate heterodyne signal in the photosensitive detector. These two electronic signals create additional signals at their sum and difference frequencies. Because the frequency difference between the two beams is small, the Doppler shift of each reflected signal is almost the same. Any difference in Doppler shift between the two beams is generally so small that it is not measurable. As a result, the difference in frequency between the two electronic signals is the same as the frequency difference between the two transmitter lasers, regardless of the target velocity. The electronics that follow the photodetector are required to be sensitive to only a narrow frequency band near the frequency difference between the two lasers and not to the full spectral range of the target Doppler shift.

1.6.2 Receiver Aperture Diameter

In a coherent laser radar, speckle limits the effective receiver aperture diameter. This is a departure from ordinary experience with conventional optics and from what might be expected by simple inspection of the laser radar range equation [Eq. (1.1)]. The apparent received power increases linearly with the receiver aperture area (square of the aperture diameter) until the physical (not effective) diameter of the receiver aperture is approximately equal to the average diameter of a speckle lobe. The apparent received power increases only linearly with the diameter thereafter. A discussion of target speckle effects is found in Sec. 1.5.

This reduction in collection efficiency is the result of a loss in coherence of the reflected signal over the aperture area. In the strictest sense, this is a target signature effect and not a loss in receiver efficiency. However, it is simpler to treat the phenomenon here. Because individual speckle lobes are coherent with respect to themselves, but not with respect to each other, only that portion of each lobe that adds in phase contributes to the heterodyne

signal. The lobes average together over a large aperture, reducing speckle-induced fluctuations but also generating a weaker electronic signal, on the average, than a perfectly coherent optical signal with the same average power.

When the receiver aperture diameter is larger than the mean diameter of a speckle lobe, the effective diameter of the receiver can be expressed by

$$D = \sqrt{D_r d_s} \ , \tag{1.29}$$

where

D = effective receiver aperture diameter in meters
D_r = receiver clear aperture diameter in meters
d_s = mean speckle lobe diameter in meters.

It is common in optical analyses to use an effective receiver aperture diameter smaller than the actual physical aperture diameter. This reduction is used to account for losses in the receiver optical train. The laser radar range equation [Eq. (1.1)] is based on the physical aperture diameter or the speckle lobe reduced diameter, whichever is smaller. Separating the receiver aperture size from the collection efficiency of the receiver reduces the possibility of confusion when evaluating system performance. Receiver optical losses are considered separately as a part of the receiver optical efficiency.

1.6.3 Performance Limitations

The ability of a laser radar receiver to detect incoming radiation is limited by the performance of the individual components and the quality of the design. Of primary importance in a heterodyne system is the quality of the optical components, the sensitivity and frequency response of the detector(s), and the alignment of the target with the optical axis of the receiver (boresight error).

1.6.3.1 Optical Efficiency and Polarization Sensitivity. The laser radar receiver does not transfer all the energy it intercepts to the detector element. The received energy experiences optical losses as the result of absorption and scattering by its own components just as do conventional passive optical receivers. These optical losses are included in the optical efficiency term of the laser radar range equation.

In addition to the conventional optical losses, some laser radar receivers experience an additional, polarization-specific loss. Many direct detection receivers and all heterodyne receivers are sensitive to only one component of the polarization of light. Many real targets tend to depolarize the reflected optical radiation. Analytically, this depolarization effect can be accounted for either by a reduction in the effective target laser cross section or by a reduction in the effective optical efficiency of the receiver. Failure to include this effect can result in a significant overestimation of system performance.

1.6.3.2 Heterodyne Efficiency. Heterodyne efficiency does not affect the magnitude of the optical radiation that reaches the detector and, as a result, does not appear in the laser radar range equation [Eq. (1.1)]. Heterodyne efficiency is a measure of the receiver's ability to detect coherent radiation. It

is included here because it is directly dependent on the receiver design. Heterodyne efficiency is an integral element of the determination of the received SNR discussed in Sec. 1.7. Heterodyne efficiency is discussed in more detail in Sec. 1.8.

In addition to losing the light the receiver aperture intercepted, the optical receiver may also corrupt the phase front of the received radiation. Phase-front errors reduce the heterodyne efficiency of the receiver. In a direct detection receiver, all the light that strikes the photosensitive element contributes to the electronic signal. However, in a heterodyne or homodyne system, only the light that matches the phase front of the local oscillator generates a heterodyne signal. All other light is either not detected or may, in some cases, actually intefere with the detection of the heterodyne signal.

1.6.3.3 Detector Quantum Efficiency. Detector quantum efficiency is a measure of the ability of a specific detector to convert incident photons into electrons. As with heterodyne efficiency, detector quantum efficiency does not affect the magnitude of the power incident on the receiver. It is, however, a major factor that determines the system's overall sensitivity and the SNR achieved for any particular signal. The detector quantum efficiency is generally expressed as a percentage or as a fraction between 0 and 1.

1.6.3.4 Detector Bandwidth. The quantum efficiency of the detector is a strong function of the modulation bandwidth of the received optical signal. Typically, the measured quantum efficiency is highest under cw illumination and decays monotonically as the frequency of amplitude modulation of the illumination is increased. The quantum efficiency measured under cw conditions is sometimes referred to as the dc quantum efficiency. When conducting system analyses, it is important to consider the quantum efficiency of the detector at the highest operating frequency to be encountered in the specific application.

1.6.4 Line-of-Sight Errors and the Lag-Angle Effect

The sensitivity of a laser radar receiver depends on the alignment of the receiver with the target. Any difference between the optical axis of the receiver and the line of sight between the receiver and the target is referred to as the receiver line-of-sight error. In a direct-detection receiver, the loss of sensitivity is primarily the result of the received target image falling off the sensitive area of the photodetector. In a heterodyne receiver, the sensitivity loss is the result of this effect and of a reduction in the heterodyne efficiency of the receiver for off-axis signals.

Scanned laser radars introduce a range-dependent, line-of-sight error between a monostatic transmitter and the receiver. The round-trip time and the scan rate cause the received signal to arrive from a different position than the transmitter is currently pointing. The receiver must point behind or lag the transmitter if the reflected radiation is to arrive on the receiver optical axis. This lag angle (Fig. 1.17) is a result of the laser radar scanner motion only. Target motion, without scanner motion, does not generate an angular displacement of the received signal. The lag angle is independent of target type and occurs for extended targets (e.g., ground returns from an airborne system)

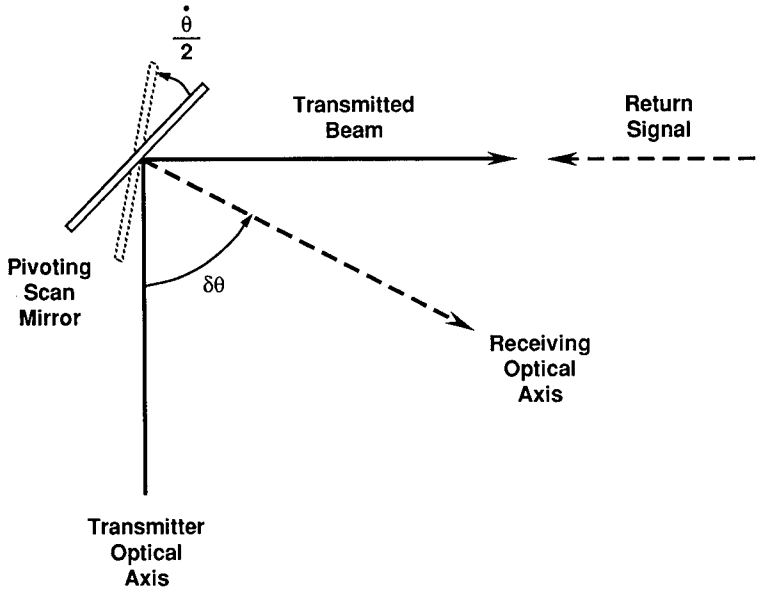


Fig. 1.17 The lag-angle effect.

and for point targets (e.g., returns from an orbiting satellite). The lag angle is given by

$$\theta = \dot{\theta}\tau = \frac{2\dot{\theta}nr}{c}, \tag{1.30}$$

where

θ = lag angle in radians
$\dot{\theta}$ = angular rate or scan rate in radians per second
τ = round-trip time in seconds
r = range to target in meters
c = speed of light in meters per second
n = mean index of refraction of the propagation media.

1.6.5 Field of View

The receiver is also commonly characterized by its angular coverage. Several terms have often been used interchangeably to describe this coverage. The following definitions are recommended.

The instantaneous field of view (IFOV) is the far-field angular subtense of a single detector element or the width of the receiver point-spread function, whichever is larger. The field of view (FOV) is the total far-field angular subtense of all detectors in the focal plane. The scanned field of view (SFOV) is the angular coverage of the detector(s) in a single scan cycle (e.g., the angular size of an image frame in a raster-scanned system). The field of regard (FOR) is the total angular extent over which the FOV can be positioned.

1.6.6 Solved Problem

Problem 1.11. Find the receiver optical efficiency of a catadioptric Cassegrainian receiver having a central obscuration of 50% (linear) of the clear aperture diameter. Each of the mirrors of the Cassegrain has a reflectivity of 98.5%. The focal lens has a transmission of 99% per surface.

Solution 1.11. A catadioptric Cassegrain has a primary and a secondary mirror and a refracting eyepiece. The eyepiece is assumed to be a singlet. The secondary, which is half of the diameter of the primary, creates a central obscuration and blocks 25% of the incoming light. The receiver optical efficiency is the product of the one minus the obscuration ratio and the reflectivity (or transmissivity) of each element.

$$\therefore \eta_r = (1 - 0.25)(0.99)(0.99)(0.985)(0.985) = 0.71 \ .$$

1.7 SIGNAL DETECTION IN NOISE

Laser radar receivers are limited by the same noise sources (ambient backgrounds, detector dark current, thermal noise, quantum shot noise) as conventional optical systems. However, because the laser radar system controls the illumination (transmitter) process, additional techniques, which are not normally available to conventional optical systems, are available for noise reduction in laser radars. The most important laser radar technique for noise reduction is optical heterodyne detection, as discussed in Sec. 1.6.

1.7.1 Signal-to-Noise Ratio

SNR is defined as the ratio of the peak signal power to rms noise power. The theoretical SNR of laser radar employing coherent (heterodyne) detection is given by

$$\text{SNR} = \frac{P_r \eta_{\text{HQ}}}{h \vartheta B f \tau} \ , \tag{1.31}$$

where

SNR = signal to noise ratio
P_r = received power in watts
η_{HQ} = heterodyne quantum efficiency
h = Planck's constant = 6.626196×10^{-34} J s
ϑ = optical (carrier) frequency in hertz
B = effective receiver bandwidth in hertz
f = transmitter pulse repetition frequency in hertz
τ = effective pulse length in seconds.

Measurements of the SNR achieved by coherent laser radars have consistently fallen short of the theoretical value predicted by Eq. (1.31). After accounting for all known losses, actual performance has empirically been shown to be typically between 3 and 8 dB lower than that theoretically predicted. At the present time, the discrepancy between theory and experiment is unre-

solved. The laser radar designer or analyst is cautioned to include such systematic losses in performance estimates to prevent insufficient performance in the field.

1.7.2 Signal Detection

All radar systems establish some criterion to determine the presence of a return signal. This criterion commonly takes the form of a voltage threshold. If the signal (and noise) exceed this threshold, then a return signal has been detected. However, a finite probability exists that the noise will exceed this threshold, even without a return signal. Likewise, the probability is less than unity that the sum of the noise and the signal will exceed the threshold, even if the signal voltage is equal to or greater than the threshold. The likelihood that the signal and noise will exceed the threshold is the probability of detection. The probability of the noise alone exceeding the threshold during any one observation opportunity (i.e., a resolution cell) is the probability of a false alarm.

For a given type of signal and type of noise, the relationship between the SNR, the probability of detection, and the probability of a false alarm is unique, i.e., specifying any two of the three uniquely defines the third. The probability of detection for a sinusoidal, nonfluctuating signal in Gaussian white noise is given by

$$P_d = \frac{1}{\pi}\int_{V_T}^{\infty} x \exp\{-[(x^2 + A^2)/2]\}\left[\int_0^{\pi} \exp(xA\cos y)\, dy\right] dx \ , \tag{1.32}$$

where

P_d = probability of detection
A = $(2\text{SNR})^{1/2}$
V_T = $[-2\ln(P_{fa})]^{1/2}$
P_{fa} = probability of a false alarm,

and x and y are integration variables.[30] Figure 1.18 is a plot of the probability of detection as a function of the SNR for several values of the probability of a false alarm for a constant amplitude signal.

Equation (1.32) is difficult to evaluate analytically; numerical techniques or approximations are normally employed. The probability of detection can be approximated by[31]

$$P_d = \frac{1}{2}\left(1 + \text{erf}\left\{\left(\frac{1}{2} + \text{SNR}\right)^{1/2} - \left[\ln\left(\frac{1}{P_{fa}}\right)\right]^{1/2}\right\}\right) \ , \tag{1.33}$$

where all terms are as previously defined. This approximation is valid for most cases encountered in laser radar design and analysis. Although the relative difference between Eqs. (1.32) and (1.33) is large for small SNRs, numerical studies have shown that the absolute error of Eq. (1.32) is less than 0.02 if $10^{-12} < P_{fa} < 10^{-3}$ and SNR > 2. Laser radar systems seldom operate outside these domains.

If the received signal strength fluctuates, as it does when the target return is dominated by speckle induced fluctuations, then Eq. (1.31) is no longer valid.

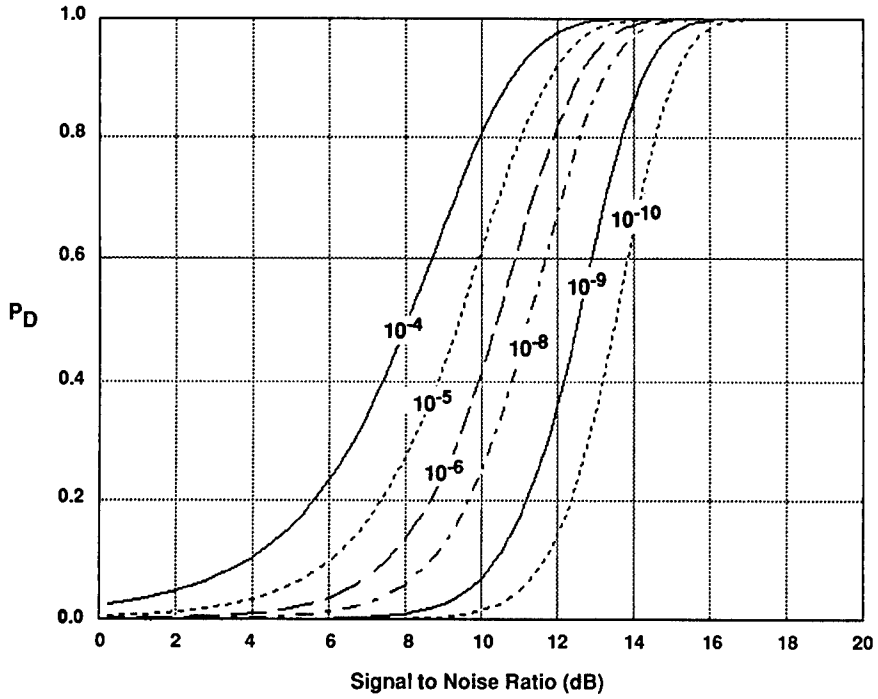


Fig. 1.18 Probability of detection of a sine wave in Gaussian white noise for several values of the probability of a false alarm. The probability of a false alarm presented here is per resolution cell.

Swerling has calculated the detection probabilities for four different fluctuation models.[30] A higher mean SNR is normally required to achieve the same probability of detection and probability of a false alarm than with a constant signal strength.

The exception to the requirement for higher SNR occurs for systems that operate with a low (less than approximately 0.4), single shot probability of detection. In this case, the signal fluctuations enhance the probability of detection. A lower mean SNR is required for such systems under these conditions of signal fluctuations. Figure 1.19 illustrates the additional SNR required to achieve a particular probability of detection.

1.7.3 Detection Laws

In some cases, external constraints or the state of the art may limit the SNR to values that do not permit acceptable performance from a single measurement attempt. In these cases, it may be possible to combine the received signal from several measurement attempts to increase the net system performance.

1.7.3.1 Signal Averaging. One approach to increasing the probability of detection is signal averaging. If n measurements are attempted and the received signals are added together, then the net SNR increases if the received signals are correlated and the noise is uncorrelated. The probability of detection

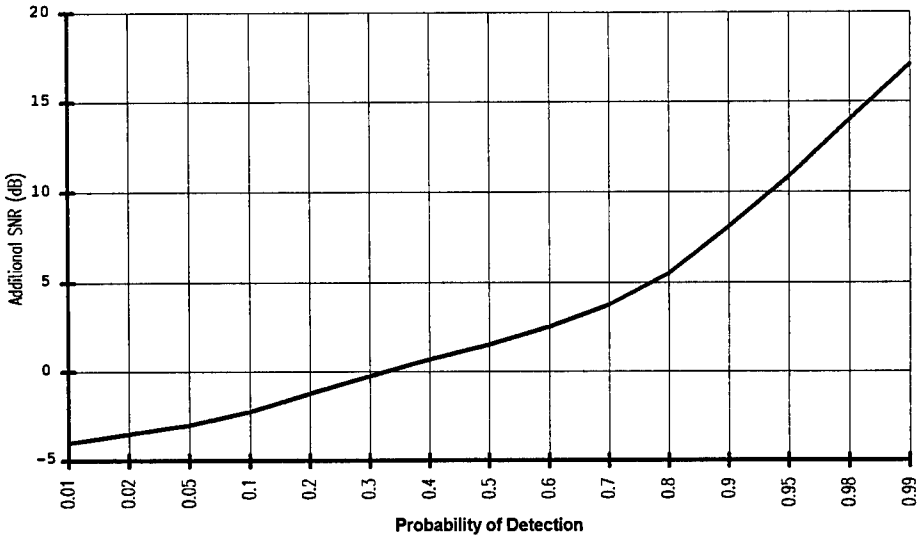


Fig. 1.19 Additional SNR required to achieve a specified probability of detection, when the received signal fluctuates (as compared with a nonfluctuating signal). A target return dominated by speckle is equivalent to Swerling[30] case 2.

is increased and the SNR-dependent measurement errors are reduced. The net SNR resulting from combining n independent measurement attempts when the received signals are perfectly correlated and the noise is uncorrelated (e.g., white noise) is given by

$$\mathrm{SNR}_{\mathrm{net}} = \sqrt{n}\mathrm{SNR} \ , \tag{1.34}$$

where

$\mathrm{SNR}_{\mathrm{net}}$ = net SNR
n = number of independent measurements
SNR = SNR of a single measurement.

Signal averaging can be implemented by transmitting a preselected number of pulses or with a variable number. If the number of pulses is varied, then the received signals are added until the averaged signal results in a detection decision or until a maximum number of measurements have been attempted.

1.7.3.2 *m*-Out-of-*n* Detection. Another approach to increasing the overall probability of detection without increasing the SNR is n-of-m detection. In this technique, a fixed number of measurement attempts are made and "detection" (e.g., a threshold crossing) is required in only a few of these attempts. A system-level detection is made if a specified number of these measurement attempts are successful. This is also referred to as a double-threshold detector and as a binary moving-window detector.[32] Here m is the number of successful measurements, and n is the number of attempts. Thus, a system that attempts to make a system-level detection decision on each measurement can be referred to as using a 1-of-1 detection law.

The advantage of an n-of-m system is partially offset by the lower probability of a false alarm, which is required on a per measurement basis, when compared to a 1-of-1 system, to achieve the same system-level false alarm rate. For a given SNR, this results in a lower probability of detection for each measurement attempt.

As with signal averaging, a fixed number of measurement attempts may be made or m can be treated as an upper limit on the number of attempts. If n successful measurements are made before m attempts, the system can disengage the current target and initiate operation against the next target. This strategy, which is often used to conserve laser radar resources, is commonly referred to as shoot-look-shoot. Although the m-out-of-n detection technique enhances the probability of detection, it does not necessarily reduce the range-dependent measurement errors.

1.7.3.3 *m*-Out-of-*n* Detection with Averaging. Signal averaging and m-out-of-n detection are not mutually exclusive. Individual measurement attempts can be evaluated individually to determine detection of an m-of-n basis and can be simultaneously averaged to determine if the composite signal has crossed the detection threshold. This preserves the best of both techniques and has been demonstrated to be more sensitive (i.e., produces a higher probability of detection) than either n of m or averaging alone. The principal disadvantage is that the receiver electronics required to implement this technique are significantly more complex.

1.7.4 Receiver Bandwidth

The bandwidth of the electronics that follow the optical detector has a strong effect on the SNR of the entire system. If the bandwidth is too large, excess noise is admitted into the system. If it is too narrow, then a portion of the signal is excluded. Either excluding signal or admitting excess noise reduces the SNR. This condition gives rise to the need to "match" the bandwidth of the receiver to the frequency content of the received signal.

The frequency content of the received signal is the result of several sources. In most systems, the spectral content of the transmitted signal defines the bandwidth required for the receiver. The bandwidth of the transmitted modulation is discussed in Sec. 1.11.6. In systems where the transmitter is very narrow band, most notably velocimetry systems, the receiver bandwidth is defined by target characteristics. The preamplifier, which immediately follows the detector, is typically the most important component in defining the operational bandwidth of the receiver.

1.7.5 Solved Problems

Problem 1.12. Find the received power required to achieve a probability of detection of 0.99 with a probability of false alarm of 10^{-6} for a cw laser radar against a nonfluctuating target at a laser wavelength of 11.15 μm over a 1-MHz bandwidth when the receiver has a heterodyne quantum efficiency of 20%.

Solution 1.12. From Fig. 1.18, the SNR required to achieve a 99% probability of detection with a 10^{-6} probability of a false alarm is approximately 14 dB or 25. The exact value can be found more precisely by iteration of Eq. (1.33). The SNR of a coherent laser radar is given by Eq. (1.31):

$$\text{SNR} = \frac{P_r \eta_{\text{HQ}}}{h\vartheta B f_\tau} .$$

Solving Eq. (1.31) for received power yields

$$P_r = \frac{h\vartheta B f_\tau \text{SNR}}{\eta_{\text{HQ}}} .$$

In this problem

$$\begin{aligned} \text{SNR} &= 25 \\ \lambda &= 11.15 \times 10^{-6} \text{ m} \\ \vartheta &= c/\lambda = 2.69 \times 10^{13} \text{ Hz} \\ \eta_{\text{HQ}} &= 0.2 \\ B &= 1 \times 10^{6} \text{ Hz}. \end{aligned}$$

For a cw system, the pulse repetition frequency and the pulse duration can be regarded as reciprocals of each other. Substitution yields

$$P_r = 2.23 \times 10^{-12} \text{ W} .$$

Problem 1.13. Find the average SNR required to achieve an average probability of detection of 0.95 with a probability of false alarm of 10^{-4} per resolution cell against a target whose return signal is dominated by speckle.

Solution 1.13. From Fig. 1.18, the SNR required to achieve a probability of detection of 95% and a probability of false alarm of 0.0001 is approximately 12 dB. A target whose return is dominated by speckle produces a returning signal that fluctuates rapidly. Figure 1.19 indicates that an average of 11 dB more signal is required to preserve the probability of detection at 0.95 under these conditions. The average SNR must be 12 dB + 11 dB or 23 dB to maintain the same level of system performance against a speckle target as was achieved at 12 dB for a nonspeckle target.

Problem 1.14. Find the SNR theoretically achieved by a cw 10.6-μm received signal having a mean power of 3.5×10^{-8} W into a receiver having an optical efficiency of 0.8, a heterodyne efficiency of 0.5, a detector quantum efficiency of 0.6, and an effective receiver bandwidth of 100 kHz.

Solution 1.14. The SNR of a coherent laser radar is given by Eq. (1.31):

$$\text{SNR} = \frac{P_r \eta_{\text{HQ}}}{h\vartheta B f_\tau} .$$

In this problem, 3.5×10^{-8} W of power are incident on the receiver but only 80% of it reaches the detector as the result of the optical efficiency of the receiver:

$$\therefore P_r = 2.8 \times 10^{-8} \text{ W} .$$

In this problem,

$\eta_{HQ} = \eta_H \times \eta_Q = 0.6 \times 0.5 = 0.3$
$\lambda = 10.6 \times 10^{-6}$ m
$\vartheta = c/\lambda = 2.83 \times 10^{13}$ Hz
$B = 100{,}000$ Hz
$f = 1/\tau$.

Substitution yields

$$\text{SNR} = 4.48 \times 10^6 \quad \text{or } 66.5 \text{ dB} .$$

1.8 HETERODYNE EFFICIENCY

Heterodyne quantum efficiency is a measure of the degradation of the actual heterodyne response of a laser radar receiver from the optimum response. In general, it depends on the point-by-point responsivity of the detector, the signal and local oscillator field distributions, and any phase front misalignment between the signal and local oscillator fields.[33–35]

The heterodyne quantum efficiency for a circularly symmetrical detector, local oscillator, and received signal is given by[33]

$$\eta_{HQ} = \frac{\left[\int_0^{z_0} \eta_Q(z) U_s(z) U_{lo}(z) J_0(kz\theta) z \, dz\right]^2}{\left\{\int_0^{z_0} \eta_Q(z)[U_{lo}(z)]^2 z \, dz\right\}\left\{\int_0^{\infty} [U_s(z)]^2 z \, dz\right\}} , \tag{1.35}$$

where

η_{HQ} = heterodyne quantum efficiency
η_Q = point-by-point quantum efficiency of the detector
z_0 = detector radius in meters
U_s = signal field amplitude distribution
U_{lo} = local oscillator field amplitude distribution
J_0 = zero-order Bessel function of the first kind
k = magnitude of the received signal wave vector
θ = angle between the local oscillator and received signal phase fronts (see Fig. 1.20).

In contrast, the heterodyne efficiency (or mixing efficiency) is dependent only on the signal and local oscillator field distributions and their relative alignment. These two parameters, heterodyne efficiency and heterodyne quan-

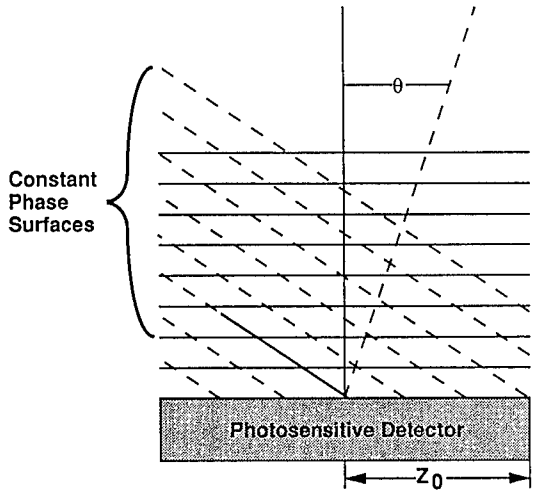


Fig. 1.20 Local oscillator, received signal, and detector surface geometry—side view. The surfaces of constant phase for the local oscillator (solid lines) and received signal (dashed lines) are shown. The angle between the direction of propagation of the local oscillator and received signal is shown as θ.

tum efficiency, are often confused. The heterodyne efficiency is, in effect, one of the components that contributes to the heterodyne quantum efficiency. Because the point-by-point quantum efficiency over the detector surface is seldom available, the heterodyne efficiency is generally computed for modeling and performance prediction purposes. In these cases, the quantum efficiency is normally approximated by its average value over the detector area.

The heterodyne (mixing) efficiency can be calculated from Eq. (1.35) by setting the quantum efficiency equal to 1. The heterodyne efficiency is therefore given by

$$\eta_H = \frac{\left[\int_0^{z_0} U_s(z)U_{lo}(z)J_0(kz\theta)z\,dz\right]^2}{\left\{\int_0^{z_0}[U_{lo}(z)]^2 z\,dz\right\}\left\{\int_0^{\infty}[U_s(z)]^2 z\,dz\right\}}, \tag{1.36}$$

where all terms are as previously defined. In those cases where the detector quantum efficiency is constant over the detector surface, the heterodyne quantum efficiency is then

$$\eta_{HQ} = \eta_H \eta_Q ,$$

where all terms are as previously defined.

However, the performance of actual laser radar systems is dependent on both the mixing efficiency and the quantum efficiency. If the detector quantum efficiency is not constant over the detector area, then the response of a heterodyne receiver is a nonlinear function of the mixing efficiency and the quan-

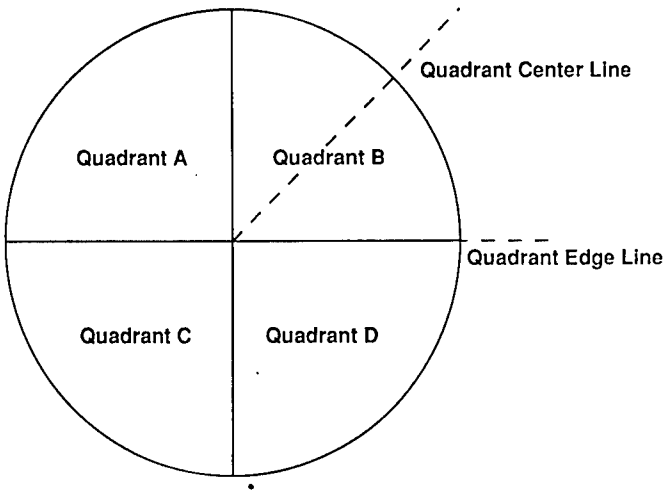


Fig. 1.21 Quadrant detector geometry for heterodyne efficiency calculations. The local oscillator was perfectly aligned with the center of the detector array in all cases. The target image was displaced either along a line that bisected a single quadrant (quadrant center line) or along a line that bounded a quadrant (quadrant edge line).

tum efficiency.[33] It is the heterodyne quantum efficiency that is normally measured experimentally. Some disagreements between measured and predicted performance may be attributable to the confusion between heterodyne efficiency and heterodyne quantum efficiency and the techniques required to measure each.[33]

Heterodyne efficiency is also dependent on the alignment of the received signal, the detector, and the local oscillator. This is particularly important for tracking laser radars because the target image is seldom centered on the detector at the initial acquisition. Figures 1.20 and 1.21 show the detector and local oscillator geometry for a quadrant detector array. Figures 1.22 through 1.27 illustrate the decrease in the calculated heterodyne efficiency with target image displacement at the focal plane for such a detector and for various arrangements of Airy-patterned signals and Airy and uniform pattern local oscillators.

1.9 LASER RADAR MEASUREMENTS

Laser radars are capable of making a variety of measurements. This capability presents the immediate possibility of multidimensional information. However, not all possible measurements can be made simultaneously to an arbitrary level of accuracy. The quality of the individual measurements is dependent on many factors, including the SNR, laser wavelength, and modulation and demodulation techniques.

1.9.1 Range

Conceptually, the simplest type of laser radar determines range by transmitting a light pulse and measuring its round-trip time of flight. Range is then

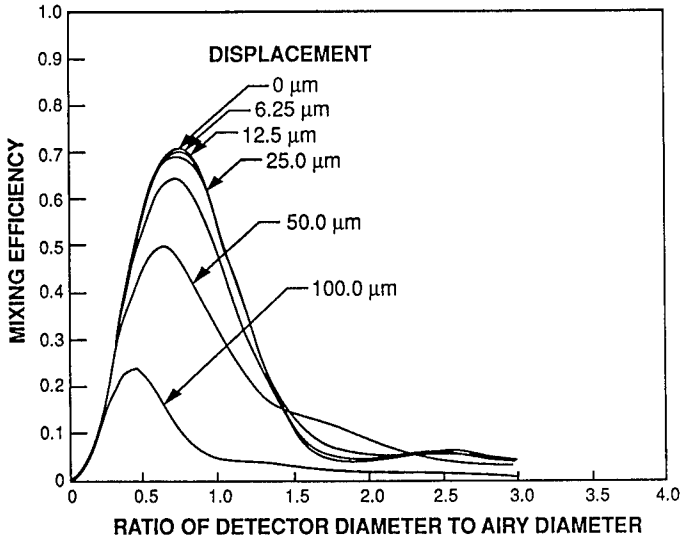


Fig. 1.22 Heterodyne efficiency for a quadrant array illuminated by an Airy signal and a uniform local oscillator. The sum of all four quadrant elements of a 200-μm-diam detector is shown for several displacements of the Airy signal from the center of the array as a function of Airy disk diameter.

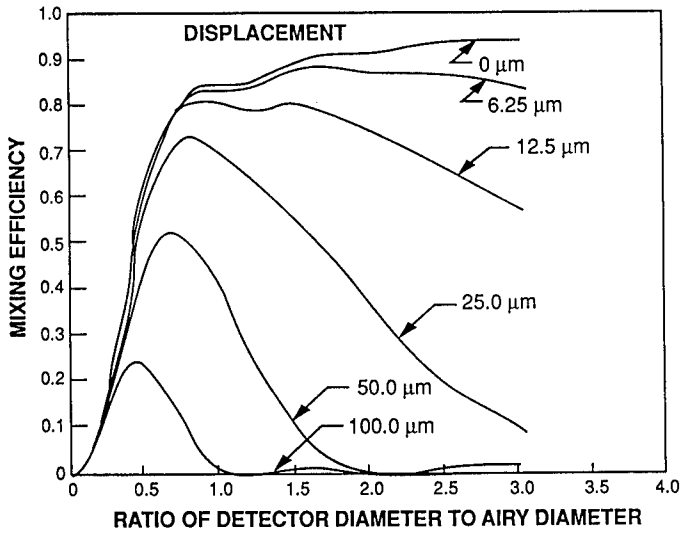


Fig. 1.23 Heterodyne efficiency for a quadrant array illuminated by a matched Airy signal and local oscillator. The sum of all four quadrant elements of a 200-μm-diam detector is shown for several displacements of the Airy signal from the center of the array as a function of Airy disk diameter.

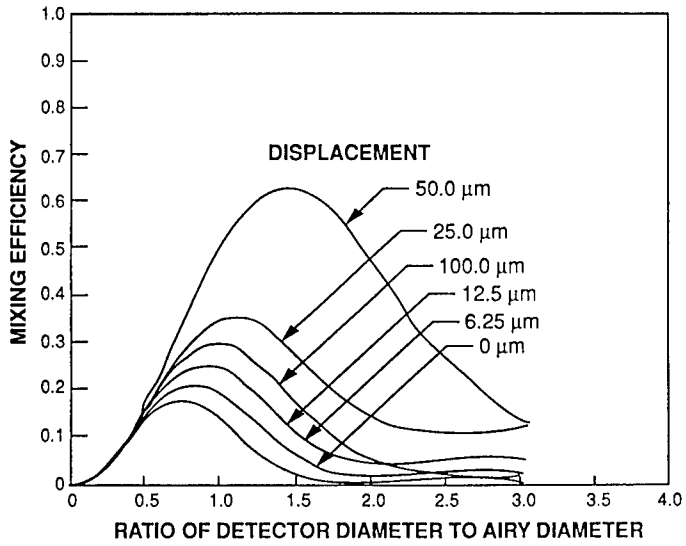


Fig. 1.24 Heterodyne efficiency for a single element of a quadrant array illuminated by an Airy signal and a uniform local oscillator as a function of Airy disk diameter. The heterodyne efficiency for a single quadrant element is shown for several displacements of the Airy signal along the quadrant center line for a 200-μm detector.

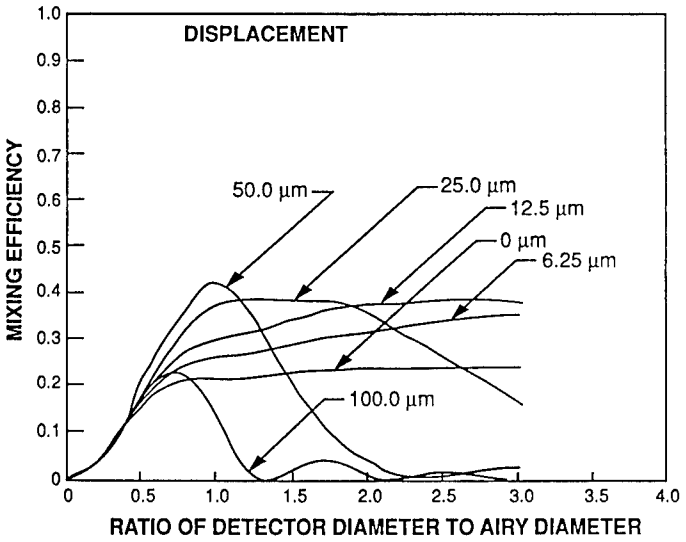


Fig. 1.25 Heterodyne efficiency for a single element of a quadrant array illuminated by a matched Airy signal and local oscillator as a function of Airy disk diameter. The heterodyne efficiency for a single quadrant element is shown for several displacements of the Airy signal along the quadrant center line for a 200-μm detector.

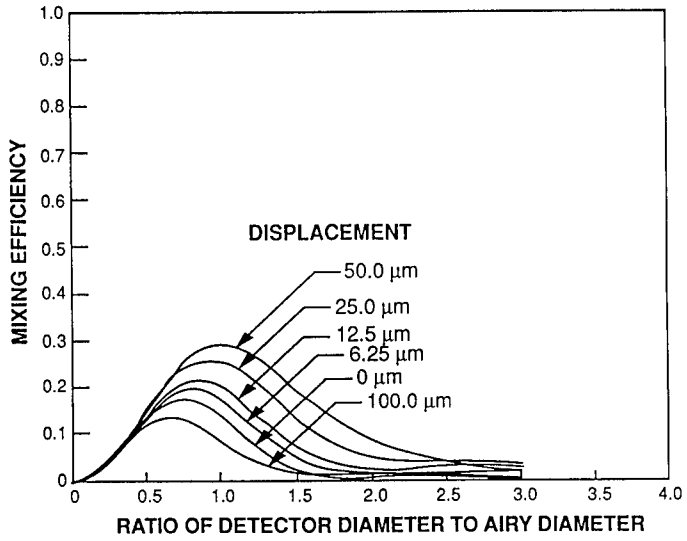


Fig. 1.26 Heterodyne efficiency for a single element of a quadrant array illuminated by an Airy signal and a uniform local oscillator as a function of Airy disk diameter. The heterodyne efficiency for a single quadrant element is shown for several displacements of the Airy signal along the quadrant edge line for a 200-μm detector.

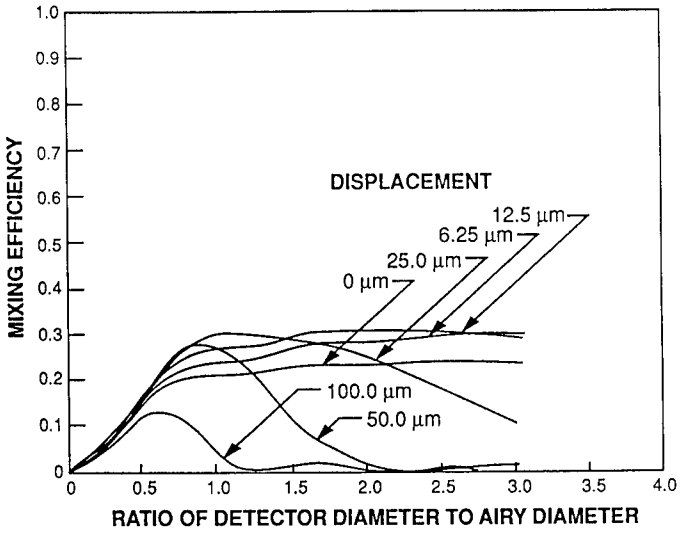


Fig. 1.27 Heterodyne efficiency for a single element of a quadrant array illuminated by a matched Airy signal and local oscillator as a function of Airy disk diameter. The heterodyne efficiency for a single quadrant element is shown for several displacements of the Airy signal along the quadrant edge line for a 200-μm detector.

determined by dividing the round-trip time by twice the speed of light in the medium. In practice, many forms of modulation can be imposed on the light to determine the round-trip time. The range to a target is given by

$$r = \frac{c\Delta t}{2n} , \tag{1.37}$$

where

r = range in meters
c = speed of light in a vacuum in meters per second = 2.997925×10^8 m/s
Δt = round-trip time of the laser radar signal in seconds
n = mean index of refraction of the propagation media.

1.9.2 Velocity

Two different approaches are commonly used to determine the velocity of a target. The simplest method is to take multiple range measurements at different times and compute the rate of change of the range. Although conceptually simple, time differentiation of range tends to amplify errors in the measurement of range and in the measurement of the time between the range measurements. A more sensitive technique is to measure the Doppler shift of the reflected light using a heterodyne receiver.

A heterodyne receiver has been used[36] to measure the differential Doppler shifts from multiple corner reflectors mounted on a rigid body. This differential velocity measurement permits the measurement of pitch rate as well as the center-of-mass translation. The Doppler shift frequency of the laser radar radiation reflected from a moving target is given by

$$\vartheta_D = \frac{\vartheta}{c}(v_1 \cos\theta_1 + v_2 \cos\theta_2) , \tag{1.38}$$

where

ϑ_D = Doppler shift frequency in hertz
ϑ = optical (carrier) frequency in hertz
θ_1 = angle between the transmitter line of sight and the target velocity vector in radians (see Fig. 1.26)
θ_2 = angle between the receiver line of sight and the target velocity vector in radians
v_1 = magnitude of the target velocity vector in an inertial frame where the transmitter is at rest
v_2 = magnitude of the target velocity vector in an inertial frame where the receiver is at rest,

and all other terms are as previously defined. Equation (1.38) is valid as long as the velocities of the target, the illuminator, and the receiver are all much less than the speed of light (nonrelativistic case). For a monostatic laser radar, $v_1 = v_2 = v$ and $\theta_1 = \theta_2 = \theta$. In this case, Eq. (1.38) reduces to

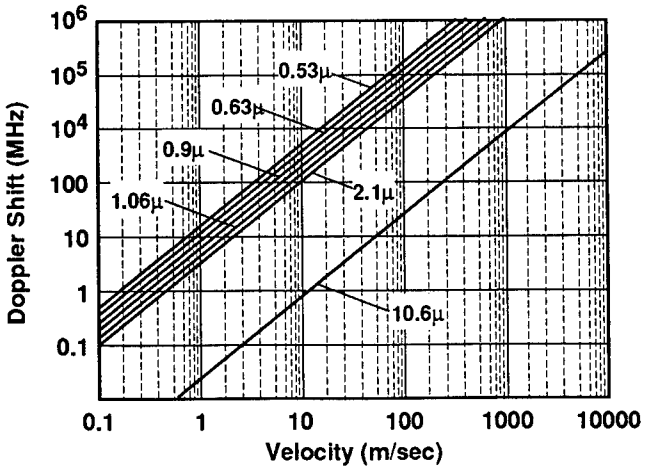


Fig. 1.28 Doppler shift of common laser wavelengths.

$$\vartheta_D = \frac{2v\vartheta \cos\theta}{c} = \frac{2v_{\parallel}\vartheta}{c} , \tag{1.39}$$

where $v_{\parallel} = v \cos\theta$ = radial velocity or velocity along the line of sight and all other terms are as previously defined.

Because the Doppler shift is proportional to the carrier frequency, the Doppler shift for laser radar is much larger than that observed by microwave or millimeter-wave radar. The Doppler shift observed by several common laser radar wavelengths is shown in Fig. 1.28. As shown in this figure, a radial target velocity of only 1.0 m/s produces a Doppler shift of approximately 189 kHz for a CO_2 laser radar operating at 10.6 μm. In comparison, the Doppler shift for a millimeter-wave system operating at 35 GHz at this same velocity is only 233 Hz. This larger Doppler shift permits much more accurate and precise measurement of velocity (see Sec. 1.10). This superior measurement capability is one of the principal advantages of laser radar.

The radial velocity of a target, measured by monostatic laser radar is given by

$$v_{\parallel} = \frac{c\vartheta_D}{2\vartheta} = \frac{\lambda\vartheta_D}{2} , \tag{1.40}$$

where λ is the transmitted wavelength and all other terms are as previously defined.

1.9.3 Intensity

The strength of the return signal is often used to test the validity of a received signal by postdetection signal-processing algorithms. Low-amplitude signals, although above the detection threshold, may not be used in some image-processing functions or may be subjected to additional processing (e.g., median filters) to

verify their validity. The strength of the return is also necessary to determine bearing to the target as discussed in Sec. 1.9.4.

1.9.4 Bearing

The angular position of a target can be determined by measuring the direction of arrival of the received laser radar signal. Because laser radar is inherently highly directional, this measurement is often unnecessary. However, in those cases where even finer angular position data are required, they can be obtained by imaging the received signal on a four-quadrant photodetector (or multielement array). The received light generates an electronic signal in more than one detector element. The ratio of the signal strength between adjoining elements is a measure of the centroid of the target image and, as a result, of the target's angular position.

1.9.5 Target Signature

It is often necessary to determine the type of target in addition to its location. This is commonly accomplished by making multiple measurements of range, velocity, and/or intensity, which are simultaneously resolved in one of the other measurement parameters (angle, range, etc.). If these resolved features or target signature is unique, then the type of target can be determined unambiguously.

This type of signature data is frequently referred to as an image, even though the data taken may bear little resemblance to a photograph. Mathematically, an image is a set of ordered triples. Laser radar signature data frequently satisfy this criterion but may not be compatible with photographic display or human interpretation based on a photographic display.

One type of target signature data that can, at times, be amenable to human interpretation is angle-resolved range imagery (ARRI). ARRI is commonly obtained by making a series of range measurements, with successive measurements taken along a scanned line of sight. In many cases, the scan pattern follows a raster format that permits video-type display by gray shading or color coding the measured range. The target signature data are a measure of target shape with the data collected in spherical coordinates (angle, angle, range). Shape is a distinguishing feature that supports target identification. If the ARRI system employs sinusoidal amplitude modulation, then the range data collected will be ambiguous in multiples of half of the modulation wavelength (see Secs. 1.10.1.5 and 1.11.2). An example of the data collected by this type of system is shown in Fig. 1.29. The "fringe pattern" seen in this figure is the result of the ambiguous range measurements.

Angle-resolved measurements of velocity can be made instead of, or in addition to, range measurements. Angle-resolved velocity measurements can easily isolate a moving object from terrain and other background sources. This technique forms a basis for moving target indicator (MTI) systems. It is also used for wind flow mapping systems, wind shear warning systems, and atmospheric research.[37]

Range-Doppler imaging (RDI) is a conceptually more complex form of target signature measurement. If a rigid target undergoes rotation, then the Doppler shift changes across the target. This results from the Doppler shift's dependence on the line-of-sight velocity and not on the total velocity. The velocity along

Fig. 1.29 Ambiguous range angle-resolved laser radar image of a military tank displays characteristic "fringe" pattern. (Courtesy Robert Y. Li, University of Nebraska)

the line of sight from the laser radar to a point on the surface of the target varies for different sections of the target as the object rotates. If the target is illuminated with a modulated waveform with sufficient bandwidth to simultaneously resolve range, then a range-Doppler image may be generated (range, velocity, signal strength). Figure 1.30 shows the type of data collected by this type of system. When such information is displayed in a video format it may or may not bear any resemblance to the object's photographic image.

Vibration spectra are another form of target signature measurements available to laser radar systems. Vibrations of the target's surface modulate the Doppler shift of the optical carrier. This modulation appears as sidebands that surround the bulk Doppler shift. The spectrum of this modulation can, in some cases, be a feature unique to the particular target and can, therefore, be used for target identification.

1.9.6 Solved Problems

Problem 1.15. Find the round-trip time of a 10-ns laser radar pulse that is reflected off a target (a) located at a distance of 3 km from a monostatic laser radar, (b) located at a distance of 1000 km, and (c) located on the moon.

Solution 1.15. The round-trip time is independent of pulse length. The range to a target is given by Eq. (1.37):

$$r = \frac{c\Delta t}{2n} .$$

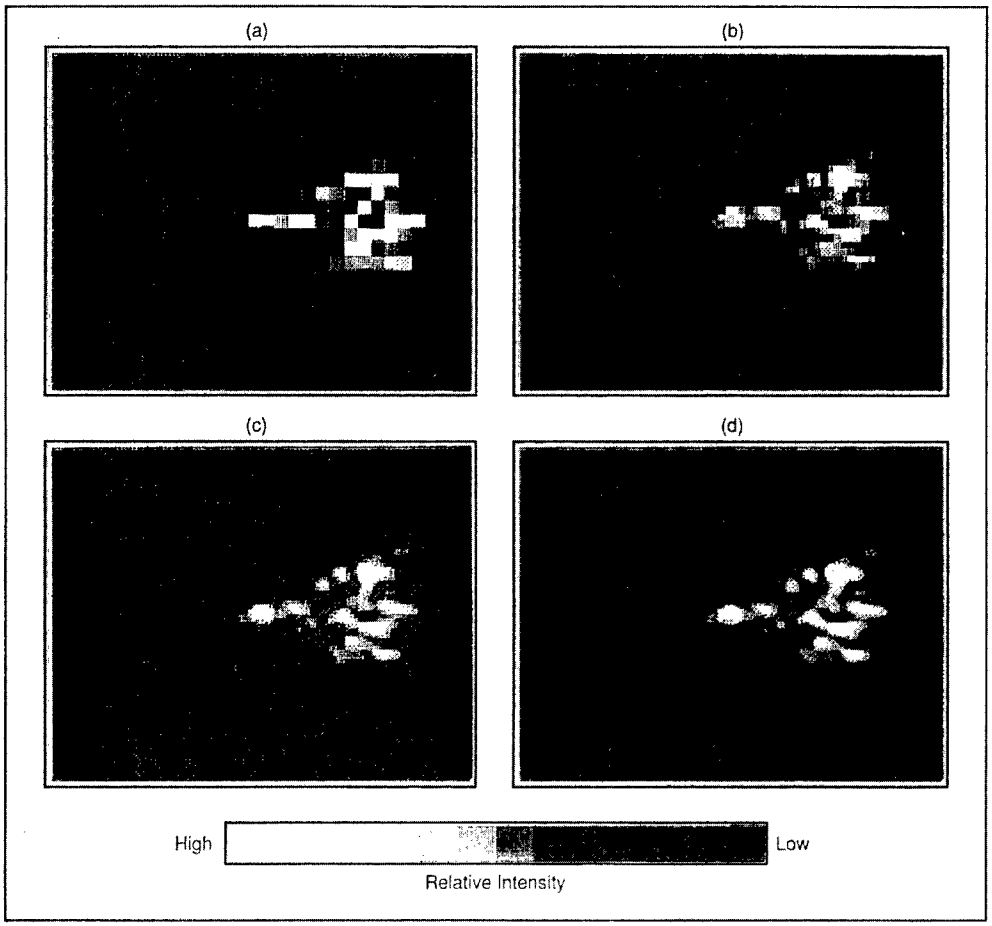


Fig. 1.30 Range-Doppler laser radar images[40] of a spinning cone at 14.5-deg aspect. The four views represent four different display formats of the same data. (Used by permission of Lincoln Laboratory)

Solving for the round-trip time yields

$$\Delta t = \frac{2rn}{c} .$$

The index of refraction of air is 1.00028 or approximately 1. At 3000 m, the round-trip time is 20 μs. At 1000 km it is 6.67 ms. At a range of 250,000 miles it is 2.7 s.

Problem 1.16. Find the Doppler shift of the received ground return for an airborne, monostatic laser radar operating at 10.6 μm. The laser radar is at an altitude of 1000 m and is traveling at 200 m/s. The laser radar is pointed along the aircraft ground path but depressed 30 deg from the horizontal.

Solution 1.16. The velocity along the line of sight is just the cosine of the angle between the line of sight and the velocity vector and the magnitude of the velocity vector. The Doppler shift is given by Eq. (1.39):

$$\vartheta_D = \frac{2v\vartheta \cos\theta}{c} = \frac{2v_{\|}\vartheta}{c} .$$

In this problem,

$\theta = 30$ deg
$v = 200$ m/s
$v_{\|} = 173$ m/s
$\vartheta = c/\lambda$
$\lambda = 10.6 \times 10^{-6}$ m.

Substitution yields

$$\vartheta = \frac{2(200) \cos 30 \text{ deg}}{10.6 \times 10^{-6}} .$$

The Doppler shift is then 32.68 MHz.

Problem 1.17. Find the Doppler shift of the laser radar described in Problem 1.16 when the system is pointed directly to port but depressed by 45 deg to the horizontal.

Solution 1.17. With the laser radar pointed directly to one side or another, the angle between the line of sight and the velocity vector is always 90 deg, independent of the depression angle. The observed Doppler shift under these conditions is zero.

1.10 MEASUREMENT CHARACTERISTICS

The quality of a laser radar measurement, i.e., angular position, intensity of the return, range or velocity, can be described by several parameters. Each has a very strict technical meaning and, unfortunately, they are sometimes misused and interchanged even in the technical literature. This leads to significant confusion. Care should be exercised when consulting the open literature to verify the meaning of those terms as used in the text.

1.10.1 Measurement Parameters

Laser radar measurements of range, velocity, direction of arrival, and/or the strength of a signal return are characterized by resolution, accuracy, precision, error, and ambiguity. Each of these parameters is strongly dependent on the design characteristics of the laser radar and on the SNR of the laser radar return. These parameters are discussed here as they relate to laser radar systems.

1.10.1.1 Resolution. Resolution is the limit of the ability to measure the separation between two different sources (or targets) in a single signal (or laser radar return).[38] The best known example of this is angular resolution from

classical optics. In this case, the angular resolution is the minimum angular separation of two, equal intensity point sources that an observer can identify as two sources. Different criteria exist for angular resolution (e.g., Sparrow criterion versus Rayleigh criterion). Likewise, different criteria exist for range and velocity resolution. However, the half width at half maximum of the ambiguity function, measured along the range or velocity axis, is commonly taken as a measure of the range and velocity resolution, respectively.

Resolution in range or velocity is dependent on the modulation waveform used in the laser radar system. Range resolution is the limit of the ability to measure the distance between two separate objects located along a common line of sight. It is *not* the ability to measure the difference in distance between two separate objects when the two objects are angularly resolved. Some authors have mistakenly used the term resolution to describe this type of performance, which is more closely related to precision.

Although resolution is typically defined with respect to the minimum separation of equal amplitude point sources, real targets are not point sources and they seldom have equal amplitudes. As a result, this simplistic definition may have little to do with actual system performance under field conditions. A detailed examination of the ambiguity function and its sidelobes may be required to determine if one target is detectable in the presence of one or several other targets.

1.10.1.2 Accuracy. Accuracy is defined as the deviation about the expectation or true value. It is a statistical quantity that indicates how much an individual measurement can be expected to differ from the true value. It characterizes the probable error. If a laser radar has small systematic errors, it is often referred to as being very accurate.

Accuracy is defined by

$$\sigma = \left[\frac{\sum_{i=1}^{N} (x_{mi} - x)^2}{N} \right]^{1/2} \tag{1.41}$$

where

σ = measurement accuracy
x_m = measured value
x = actual (or expectation) value
N = number of measurements made
i = summation index.

1.10.1.3 Precision. Precision is defined as the deviation about the mean value. It is a statistical quantity that characterizes the repeatability of measurements. If the mean and expectation values are equal, then the precision and accuracy must also be equal. If a laser radar has small random error, it is often referred to as being very precise.

Precision is defined by

$$\varepsilon = \left[\frac{\sum_{i=1}^{N} (x_{mi} - \bar{x})^2}{N - 1} \right]^{1/2} \tag{1.42}$$

where ε = measurement precision, $\bar{x} = (\sum_{i=1}^{N} x_{mi})/N$, and all other terms are as previously defined.

Precision is closely related to the ability to measure the difference in two quantities and is frequently confused with resolution and accuracy. It is equivalent to the concept of significant digits in mathematics. For example, in a laser radar that measures range with a precision of 1 m, no significance can be ascribed to two measurements that differ by only 1 cm. However, for this laser radar, a difference of 5 m between two measurements indicates, with high confidence, that the two objects are at different distances.

1.10.1.4 Error. Error is the difference between the measured value and the expectation or true value. It is not a statistical quantity.

1.10.1.5 Ambiguity. Ambiguity arises in laser radar systems when the returns from well-separated objects provide identical returns. This is easily illustrated by the example of a pulsed laser radar with a pulse repetition frequency of 100 kHz. Such a laser radar transmits a pulse every 10 μs; 10 μs is also the round-trip time for a laser radar signal reflected by an object located at a range of approximately 1500 m. The laser radar receiver cannot distinguish between the arrival of a received signal for a target at any particular range and the arrival of previous pulses reflected off other targets located multiples of 1500 m farther away than the first. All repetitive waveforms experience some form of ambiguity. Its effect is greatly reduced by selecting a range ambiguity longer than the maximum anticipated operating range.

1.10.2 System Error Sources

The performance of a laser radar system is dependent on the performance of the entire system. Performance evaluation and prediction for the laser radars often emphasize the performance of the transmitter and receiver. Although these elements are important and provide a fundamental limitation on the performance of the system, other components are frequently the practical limitations on the overall system performance.

Error budgets must be established for the complete laser radar system. For example, the formula for range accuracy of a pulsed laser radar shown in Sec. 1.2 represents the ability of a well-designed receiver to measure (estimate) the actual time of arrival of the center of the incoming pulse. This formula presumes that a "perfect" clock is used to measure time. Real clocks, such as those normally used in the receiver electronics of laser radars, have timing errors. Such errors must be included in the overall error budget.

Neglecting systematic error sources (including clock limitations), the accuracy to which a laser radar return signal from a point target can be measured

is dependent on the SNR of the return and the bandwidth of the transmitted signal.

The measurement accuracy of the amplitude of the received signal is given by

$$\sigma_A \geqslant \frac{A}{(2\mathrm{SNR})^{1/2}} , \tag{1.43}$$

where σ_A is the accuracy of amplitude measurement and A is the signal amplitude. The measurement accuracy of the range to a point target is given by

$$\sigma_r = \frac{c}{2\pi\beta(2\mathrm{SNR})^{1/2}} , \tag{1.44}$$

where

σ_r = accuracy of range measurement in meters
c = speed of light in a vacuum
n = mean index of refraction in the propagation medium
β = effective bandwidth of the receiver in hertz.

The measurement accuracy of the radial velocity (the component of velocity along the line of sight) when measured by determining the Doppler shift of the reflected radiation is given by

$$\sigma_v = \frac{\lambda}{2} \frac{1}{2\pi\tau(2\mathrm{SNR})^{1/2}} , \tag{1.45}$$

where σ_v is the accuracy of the velocity measurement in meters per second, τ is the pulse length or coherence length of the transmitted signal or the integration time of the receiver (whichever is shorter) in seconds, and all other terms are as previously defined.

1.10.3 Solved Problems

Problem 1.18. Calculate the SNR required for a pulsed 10.6-μm system to have a velocity accuracy of 1 m/s. The pulse length is 1 μs.

Solution 1.18. Velocity accuracy is given by Eq. (1.45):

$$\sigma_v = \frac{\lambda}{2} \frac{1}{2\pi\tau(2\mathrm{SNR})^{1/2}} .$$

Solving for the SNR yields

$$\mathrm{SNR} = \frac{1}{4}\left(\frac{\lambda}{\pi\tau\sigma_v}\right)^2 .$$

In this problem,

$\tau = 1 \times 10^{-6}$ s
$\lambda = 10.6 \times 10^{-6}$ m
$\sigma_v = 1$ m/s.

Substitution yields

$$\text{SNR} = 268{,}502 \ ,$$

$$\therefore \text{SNR} = 54 \text{ dB} \ .$$

Problem 1.19. Calculate the range accuracy of a laser radar which transmits a Gaussian pulse with a full width at half maximum of 10 ns when the SNR is 8 dB; for an SNR of 40 dB.

Solution 1.19. The range measurement accuracy is given by Eq. (1.44) as

$$\sigma_r = \frac{c}{2\pi\beta(2\text{SNR})^{1/2}} \ .$$

The effective bandwidth (see Sec. 1.11.6) of a Gaussian pulse is

$$\beta = \frac{\tau}{0.441} \ .$$

Substituting for the effective bandwidth in Eq. (1.44) yields

$$\sigma_r = \frac{c\tau}{2 \times 1.18\pi(2\text{SNR})^{1/2}} \ .$$

An SNR of 8 dB is a factor of 6.31; 40 dB is a factor of 10,000. This yields

$$\sigma_r = \frac{2.998 \times 10^8 \times 10 \times 10^{-9}}{2 \times 0.59 \times 3.141 \times (2\text{SNR})^{1/2}} \ .$$

For an SNR of 6.31, the theoretical range measurement accuracy is 0.27 m. Increasing the SNR to 10,000 reduces the theoretical range measurement accuracy to only 0.0067 m or 6.7 mm. Achieving a true range accuracy this small also requires that the effect of systematic error sources (e.g., timing clock stability) and uncorrectable random error sources (e.g., random fluctuations in the index of refraction of the atmosphere) be much smaller than the measurement accuracy.

1.11 TRANSMITTER MODULATION TECHNIQUES

The laser that generates the transmitted radiation is critically important to the performance of the entire laser radar system. It is the existence of this one component that separates active sensors from all passive sensors. The function of the transmitter is to generate a modulated optical signal of sufficient power that the receiver can determine specific information about a target from the reflected radiation. The key points are modulation, power, and information. The type of information required determines the type, characteristics, and quality of the modulation required. The power transmitted must be sufficient that the losses associated with range, target cross section, and atmospheric propagation do not prevent detection of the reflected signal.

The radiation transmitted by laser radar must be modulated so that the round-trip time and, hence, the range can be determined. The most common modulation formats used in laser radars are pulse modulation, single- or multiple-tone sinusoidal amplitude modulation, and linear frequency modulation. Furthermore, the modulated radiation must be of sufficient power that the received signal can be detected. Selection of a modulation format that is achievable within the state of the art for the wavelength and power levels required for a specific application is one of the most challenging tasks facing the laser radar designer.

1.11.1 Pulse Modulation

Conceptually, pulse modulation is the simplest approach to range measurement. Laser radar transmits a single pulse of light and measures the round-trip time of that pulse to determine range. A sample of the transmitted light may be directed to a photodetector that triggers the start of a counter that is stopped by the detection of the received signal.

Although the modulation format and measurement technique used in pulsed systems appear simple, the techniques used to generate pulsed laser signals tend to be somewhat complex. The most common techniques employed to generate pulsed laser output are gain switching, Q switching, cavity dumping, Q switching with cavity dumping, and mode locking.

1.11.1.1 Gain Switching. Gain switching occurs when the gain in the laser medium is rapidly increased. Typically, this is accomplished by pulse pumping the gain media. Pulse pumping can cause inhomogeneities in the gain medium that reduce the frequency stability and coherence length of the laser output. Laser pulse lengths produced by this technique are of the order of the relaxation time of the excited state.

1.11.1.2 Q Switching. By rapidly reducing the losses internal to the laser cavity, the laser output can be pulsed. The gain medium is often continuously pumped, but it can also be pulse pumped just prior to Q switching. Continuous pumping requires that the interpulse period be shorter than the relaxation time of the excited state to ensure efficient operation. Because pump energy cannot be stored in the gain medium longer than the relaxation time, a longer interpulse period results in decreased laser efficiency, increased heat dissipation, and lower average power output. A wide variety of techniques can be employed to alter the cavity losses. These include acousto-optic, electro-optic, saturable absorbers, thin film absorbers, and rotating mirrors.[38] The pulse lengths produced by these techniques are typically limited by the bandwidth of the modulating device.

1.11.1.3 Cavity Dumping. Whereas Q switching begins with a high-loss optical cavity that is rapidly converted to low loss, cavity dumping does exactly the opposite. A low-loss cavity that is continuously or nearly continuously pumped is switched to a high loss by the activation of an optical switch. The optical switch acts as a reflective shutter that diverts the internally circulating radiation out of the laser cavity. The length of the pulse generated by this technique is limited by the physical length of the laser resonator.

1.11.1.4 *Q* Switching with Cavity Dumping. In a low-loss optical cavity, the pumping process and stimulated emission compete to populate and depopulate the excited state in the gain medium. The excited state can be made to reach a higher density in a high-loss cavity since stimulated emission is suppressed. In a Q switched and cavity-dumped laser, the gain medium is initially excited to a high level while the laser cavity is in a high-loss mode. The Q is then increased by reducing the losses and the circulating power in the cavity builds rapidly as the excited stated is depopulated by stimulated emission. While the intracavity flux is near its peak, the cavity is dumped. This results in higher peak powers than cavity dumping alone. As with cavity dumping, the pulse length is controlled by the physical length of the resonator cavity.

1.11.1.5 Mode Locking. It is well known that the length of a laser cavity must be an integer multiple of the laser wavelength. If the laser gain line is broad or if the laser cavity is long, then more than one wavelength may have an integer multiple equal to the cavity length. In this case, each wavelength may circulate through the cavity and experience gain. The laser output consists of a superposition of closely and uniformly spaced laser wavelengths, which have zero phase difference at the output coupler of the laser.

These multiple frequencies beat against each other to produce a time-varying signal. If a large enough number of frequencies are present, then the output appears as a train of extremely narrow, evenly spaced pulses that modulate a single carrier frequency. From a laser radar operations standpoint, this permits simultaneous high resolution range and velocity measurement.

1.11.2 Amplitude Modulation

If the transmitted laser signal is modulated sinusoidally in amplitude, then the range to the target can be measured by comparing the phase of the amplitude modulated received signal to the phase of the currently transmitted signal. The range to the target is proportional to the phase difference of the amplitude modulated envelope and is given by

$$r = \frac{c}{4}\frac{\phi_{AM}}{F_{AM}}, \tag{1.46}$$

where

r = range to the target in meters
ϕ_{AM} = phase difference between transmitted signal and received signal in radians
F_{AM} = amplitude modulation frequency in hertz,

and all other terms are as previously defined.

A single-tone AM system provides ambiguous range measurements unless the ambiguity range exceeds the maximum range of the laser radar. The ambiguity of the range measurement is a result of the identicality of the individual cycles of a sinusoid. Relative phase differences can be measured but absolute phase difference cannot. The ambiguity range of an amplitude modulated laser radar is given by

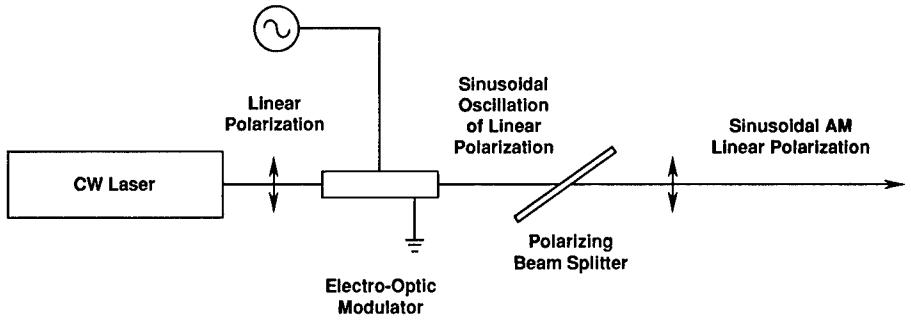


Fig. 1.31 Block diagram of a typical external amplitude modulator.

$$\Delta r = \frac{\lambda_{AM}}{2} = \frac{c}{2F_{AM}} , \tag{1.47}$$

where Δr is the ambiguity range in meters, λ_{AM} is the amplitude modulation wavelength in meters, and all other terms are as previously defined.

Multiple-tone sinusoidal modulation can be used to resolve the range ambiguities. Each tone has its own ambiguity range. The lowest frequency tone should have an ambiguity range greater than the maximum range of the system. In a multitone system, range errors associated with any of the lower frequency tones can result in significant range errors. The distribution of errors tends to be multimodal.

Sinusoidal AM can be implemented either internal or external to the cavity. When generated internal to the cavity, the gain of the medium can be modulated (e.g., by varying the drive current supplied to a semiconductor laser). The bandwidth of the modulation generated in this way is limited by the lifetime of the excited state in the gain medium and by the round-trip time of the laser cavity.

When generated externally, an element that introduces a time-varying optical loss is often inserted into the path of a cw laser beam. A common form of external amplitude modulator (Fig. 1.31) consists of an electro-optical modulator (EOM) followed by a polarizing beamsplitter (PBS). The EOM rotates the plane of polarization of the linearly polarized light as a function of the drive voltage applied to the EOM. The PBS permits only one polarization to pass through it and reflects the other. If an appropriate drive signal is applied to the EOM, the amplitude of the polarization components can be made to vary sinusoidally. In addition to being orthogonally polarized, the two modulated beams are 180 deg out of phase.

The primary disadvantage of external AM is that at least half of the laser power generated is often lost, i.e., only half of the power is contained in either of the polarized beams and only one is usually transmitted to the target. The other beam is routinely directed into a beam dump. This results in an immediate 3-dB loss in system efficiency and performance. Because the PBS simply separates the two polarization components, both are still usable. However, systems that transmit two, separate AM beams and detect them separately are significantly more complex than single-beam systems.

1.11.3 Frequency Modulation

Pulsed modulation and amplitude modulation can use either coherent or direct detection receivers. Frequency modulation (FM) requires a coherent receiver for demodulation because all of the information is encoded in the carrier frequency. FM can be implemented internal to the gain cavity (intracavity) or external to the gain cavity (extracavity).

1.11.4 Hybrid Modulation

Requirements to measure range and velocity to a high level of precision or resolution can require longer coherence times and higher modulation bandwidths than can be economically achieved with single modulation formats. Hybrid or complex waveforms have been developed to overcome this problem. The two most common hybrid formats are AM/FM and pulse burst. AM/FM is simply the amplitude modulation of a frequency modulated carrier. A pulse burst is generated when a series (train) of short, mode-locked laser pulses is transmitted together and the carrier remains coherent throughout the transmission.

1.11.5 Laser Amplifiers

Each of the modulation techniques previously discussed is limited in the total signal power that can be generated. The most damage-sensitive element of the laser (e.g., output coupler, electro-optic modulator, etc.) determines the maximum power output. If the specific application requires greater power than can be generated directly, an optical amplifier may, in some cases, be employed.

The laser architecture in which a modulated laser signal is amplified to a higher power level is referred to as a *master oscillator-power amplifier* (MOPA). The master oscillator (MO) signal can be internally or externally modulated. If internally modulated, it is common for the MO to be frequency locked to a reference oscillator (RO). The RO is a separate laser with very high frequency stability, which permits its use as a frequency standard. The laser power amplifier (PA) must retain the coherence and bandwidth of the signal fed into it to prevent signal distortion. An isolator is often inserted between the MO and the PA to prevent feedback from the PA entering the MO (and possibly the RO) and corrupting the signals generated by it.

1.11.6 Effective Transmitter Bandwidth and Pulse Length

The range and velocity accuracy of laser radar is limited by the effective transmitter modulation bandwidth and the effective pulse length. The effective modulation bandwidth is not simply related to the pulse length and may be completely unrelated to the noise bandwidth of the receiver (providing that the noise bandwidth is wider). The effective bandwidth of several common laser radar modulation types is shown in Table 1.4.

1.11.7 Laser Efficiency and Power

The ratio of the prime power used to pump the transmitter laser to the optical energy generated is the laser efficiency. In many applications, the optical power

Table 1.4 Effective Bandwidth of Common Modulation Types

Modulation Type	Effective Bandwidth	Notes
Unmodulated (CW)	B	B = receiver noise bandwidth (Hz)
Gaussian Pulse	$\frac{0.441}{\tau}$	
Bandwidth-Limited Rectangular Pulse	$\sqrt{\frac{2B}{\tau}}$	
Linear Frequency Modulation	$\dot{f}\,\tau$	$\dot{f}$ = linear frequency ramp rate
Pulse Burst	$\frac{0.441}{\tau_{micro}}$	τ_{micro} = width of an individual pulse from within burst

that can be generated by the transmitter laser source is limited by the laser efficiency and the prime electrical power available within the size and weight constraints of the specific application. In other applications, the maximum usable optical power may be limited by the damage threshold of the most sensitive element in the laser or beam relay optics.

Because the process of pumping a laser is not 100% efficient, the energy left over usually ends up as heat in the laser gain medium. Thermal gradients, turbulence (in the case of fluid laser gain medium), and expansion reduce the coherence length of the laser. Because long coherence lengths (i.e., narrow bandwidths) are required for coherent detection, this loss of coherence may also limit the maximum usable output power. This may occur even though the transmitter laser may be capable of much higher total output power if multimode operation or short coherence length operation is considered.

1.12 RECEIVER DEMODULATION TECHNIQUES

After the reflected laser signal has been detected by the optical receiver and converted into an electronic signal, the information contained in that signal must still be extracted. The specific information content of the electronic signal depends on the transmitter modulation format, atmospheric propagation and target characteristics, and the optical receiver design and performance. Electronic signal detection and signal filtering is, in itself, a large subject. Once the received signal has been converted into an electronic signal, the measurement techniques for laser radars are identical to those used in conventional radars and are a well-established art. A few examples of some of the most common demodulation techniques are described briefly here.

1.12.1 Pulse Demodulation

If the transmitter in a coherent detection radar uses a pulse modulation format, then the optical receiver must use heterodyne (instead of homodyne) detection

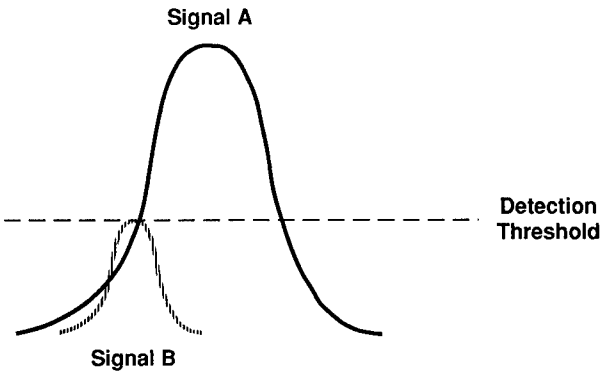


Fig. 1.32 Threshold-crossing pulse detection.

(Fig. 1.1). Homodyne type receivers are not generally an option because both the local oscillator and received signals must be present on the detector at the same time for optical mixing to occur, and synchronization of the transmitter and received signal is not possible because the time of arrival, which the system is to measure, is the unknown.

After the optical signal has been converted into an electronic signal, the primary remaining issue is the criterion to be used to determine the presence of a received signal.

1.12.1.1 Threshold Crossing. One criterion to determine the arrival of a pulse is a threshold crossing (Fig. 1.32). If the electronic signal exceeds a predetermined amplitude (as usually determined by a voltage comparator), then a counter is stopped and the value of the counter is a measure of the round-trip time to the target and back and, hence, the range. If no signal is detected before a preset time, the counter is reset for the next pulse. This approach suffers from several deficiencies. First, it records as the range only the first pulse to exceed threshold. An early pulse may be generated by noise or by close-in clutter. The second major disadvantage is that pulse-to-pulse variations in amplitude result in time-of-arrival jitter. In Fig. 1.32, signal A and signal B have significantly different arrival times. However, because they also have different amplitudes, they cross the detection threshold at the same time. This jitter may be up to one half of the total pulse width. The primary advantage of this technique is that it is simple and inexpensive.

1.12.1.2 Peak Sample and Hold. Another strategy that overcomes some of the limitations of the threshold-crossing detector is the peak-sample-and-hold (PSAH) circuit (Fig. 1.33). In this approach, the receiver holds the highest amplitude signal that occurs within the measurement period. A digital counter is commonly used to measure the time of arrival of the pulse but, unlike the threshold-crossing technique, the counter is not stopped when a peak is detected. Instead, the hold latch of the PSAH is used to freeze the time (count) in a digital register. If, at a later time, another signal of greater amplitude is detected, the PSAH hold latch and the digital register are released until the new peak is detected. This approach reduces the jitter associated with amplitude variations and reduces the detection of noise and clutter.

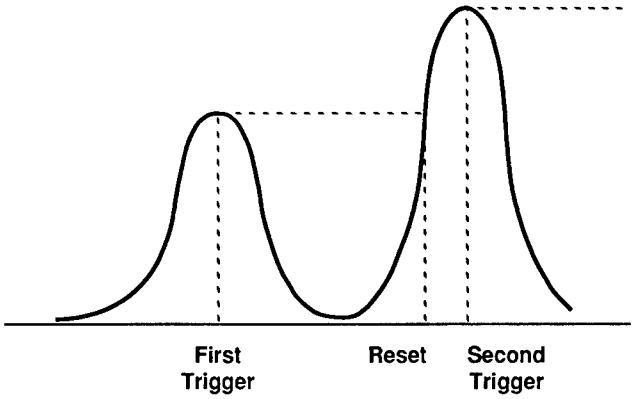


Fig. 1.33 Peak-sample-and-hold pulse detection.

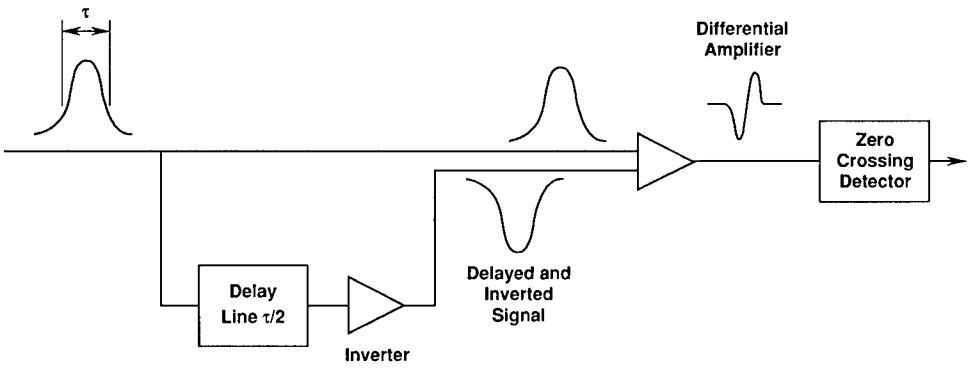


Fig. 1.34 Constant fraction discrimination peak detection.

1.12.1.3 Constant Fraction Discrimination. The jitter in the time-of-arrival measurement can be further reduced by using constant fraction discrimination (CFD). CFD is accomplished by splitting the incoming signal into two channels, delaying one channel by one half of a pulse width, and subtracting the delayed channel from the original (Fig. 1.34). This results in a positive and then negative signal with a characteristic S-shaped profile. It has been shown that the zero crossing of this derived signal is very insensitive to amplitude fluctuations. It is, however, somewhat sensitive to pulse length variations.

Use of the different techniques listed here is not exclusive. They can be combined for specific applications and achieve increased resistance to noise and clutter.

1.12.2 Amplitude Demodulation

An amplitude modulated (AM) system may use a heterodyne, conventional homodyne, or frequency offset homodyne optical receiver to convert the received optical signal into an electronic signal. The electronic signal consists of an AM carrier with the carrier frequency at the sum of the Doppler shift

and offset frequencies (for heterodyne or frequency offset homodyne receivers). Measurement of range with sinusoidal AM is usually accomplished with a phase-locked loop, which is used to make a phase comparison between the envelope of the currently transmitted signal and the received signal. The phase-locked loop generates, as one of its outputs, a voltage proportional to the difference in phase between the two signals. This approach requires that the carrier frequency be significantly larger than the AM frequency. This requirement often precludes the use of homodyne optical receivers because the carrier frequency is solely the result of the Doppler shift.

The primary limitation of this technique is the accuracy of the phase measurement. Accuracy is dependent on SNR, but for moderate SNR levels, 3 to 10 deg of phase is typical. Another limitation of this technique is its inability to resolve multiple targets (or clutter). Multiple targets at different ranges generate signals that are additive in the receiver. The demodulation electronics provide a measure of the centroid of these signals.

Another limitation of a phase-locked loop is that the phase difference output is provided continuously, even when no signal is present. Under no-signal or low-signal conditions, the phase difference signal out of the phase-locked loop is random. This technique is often used in conjunction with a threshold-crossing detector to provide a "data valid" signal. In imaging systems or other systems where successive signals are highly correlated in range, the random output of the phase-locked loop appears as clutter. Image-processing techniques can be used in these systems to postprocess the data and reduce the number of invalid data points.

1.12.3 Frequency Demodulation

Linear frequency modulation (LFM) is used to measure range in a manner similar to AM. LFM is often referred to as a *chirp,* allegedly because the sound made by crickets is also a nearly linear acoustic frequency ramp. In LFM, the currently transmitted frequency is compared to the received frequency. The ratio of the difference in frequency to the frequency ramp rate is equal to the round-trip time, provided that the target is not in relative motion to the laser radar (no Doppler shift). A Doppler shift generates a frequency difference that may be misinterpreted as a range in an FM system. This phenomenon is known as range-Doppler coupling or range-Doppler ambiguity. Generally, multiple LFM signals are transmitted, at different ramp rates or in different ramp directions (an up-down sweep), to resolve this ambiguity.

1.12.3.1 FM Homodyne. By mixing the received signal with the currently transmitted LFM signal, as in a homodyne optical receiver, the electronic heterodyne signal is at a constant frequency, which is the difference between the currently transmitted signal and the received signal. The same effect can be created using a heterodyne optical receiver and a constant frequency local oscillator if the electronic heterodyne signal is mixed with a replica (electronic local oscillator) of the currently transmitted signal. Sidebands are then created that are at a constant frequency. In either technique, the frequency of the signal is a characteristic of the range and the Doppler shift. Even though both signals (received and local oscillator) are at constantly changing frequencies,

their frequency difference is constant. This constant frequency can then be measured with a variety of spectral analysis techniques.

1.12.3.2 Pulse Compression. Pulse compression requires that heterodyne detection be used in the optical receiver and that the local oscillator be at a constant frequency. Pulse compression is a technique in which an LFM signal is passed through a dispersive delay line and converted into a pulse-modulated signal. The output signal is than processed with any of the discussed pulse demodulation techniques. A dispersive delay line is a filter that provides a time delay of the signal that depends on the frequency of the signal. To effect pulse compression, the delay characteristics of the filter must be equal but opposite to the LFM ramp. This causes the various frequencies to become superimposed at the output of the filter. The amplitude of the signal output is additive and results in a single, high amplitude pulse.

Acknowledgments

This chapter would not have been possible without the efforts of many individuals. The author would like to thank Gary E. Eberhart and Karl McKechnie of Teledyne Brown Engineering for their support during the preparation of this manuscript. The author would also like to thank Dr. Albert Reynolds of McDonnell Douglas Missile Systems Company for reviewing the manuscript and correcting the many mistakes left in the original. The author would also like to acknowledge the assistance of Dr. Mike Newchurch in the preparation and review of atmospheric effects and of Dr. Philip Smith and Lauren Dickerson in the preparation and review of laser target signature effects.

References

1. Merrill I. Skolnik, *Radar Handbook,* pp. 37-2–37-67, McGraw Hill, New York (1970).
2. B. Yuhas, "Laser gun ammunition measurement equipment," U.S. Patent 5,159,396.
3. Max Born and Emil Wolf, *Principles of Optics,* p. 396, Pergamon Press, Oxford (1980).
4. Warren J. Smith, *Modern Optical Engineering,* p. 311, McGraw Hill, New York (1966).
5. *ACCOS V Users Manual,* Scientific Calculations, Fisher, NY.
6. Adolph S. Jursa, Ed., *Handbook of Geophysics and the Space Environment,* Air Force Geophysics Laboratory, Air Force Systems Command, United States Air Force (1985).
7. P. Shettle and R. W. Fenn, "Models for the aerosols of the lower atmosphere and the effects of humidity variations on their optical properties," Air Force Geophysics Laboratory TR-79-0214 (1979).
8. "Final report to AFWAL/AART-2 on sensor blending for terrain and obstacle avoidance," Boeing Corporation (1984).
9. V. Chimelis, "Extinction of CO_2 laser radiation under adverse weather conditions," Air Force Wright Aeronautical Laboratory TR-81-1280 (1981).
10. P. Chylek, "Extinction and liquid water content of fogs and clouds," *Journal of Atmospheric Science* **35,** 296 (1978).
11. R. G. Pinnick, S. G. Jennings, P. Chylek, and H. J. Auvermann, "Verification of a linear relationship between IR extinction, absorption and liquid water content of fogs," *Journal of Atmospheric Science* **36,** 1577–1586 (1979).
12. D. G. Rensch and R. K. Long, "Comparative studies of extinction and backscattering by aerosols, fog, and rain at 10.6 μm and 0.63 μm," *Applied Optics* **9**(7), 1563–1573 (1970).
13. Warren L. Hatch, "Selective guide to climatic data sources," Key to Meteorological Records Documentation No. 4.11, National Climatic Center, Asheville, NC, U.S. Department of Commerce, National Oceanic and Atmospheric Administration, Nation Environmental Satellite, Data and Information Service.

14. C. C. Chen, "Attenuation of electromagnetic radiation by haze, fog, clouds, and rain," The Rand Corporation, R-1694-PR (Apr. 1975).
15. D. Duncan, R. Shirkey, and M. Richardson, "EOSAEL 82 volume 2: transmission through gasses and natural aerosols," Atmospheric Science Laboratory TR-0122 (1982).
16. M. A. Seagraves and J. F. Ebersole, "Visible and infrared transmission through snow," Atmospheric Sciences Laboratory TR-0119 (1982).
17. V. P. Bisyarin, I. P. Bisyarina, V. K. Rubash, and A. V. Sokolov, "Attenuation of 10.6 μm and 0.63 μm laser radiation in atmospheric precipitation," *Radio Engineering and Electron Physics* **16,** 1594–1597 (1971).
18. S. G. Jennings, R. G. Pinnick, and H. J. Auvermann, "Effects of particulate complex refractive index and particle size variations on atmospheric extinction for visible through middle IR wavelengths," *Applied Optics* **17,** 3922–3939 (1978).
19. Brent F. Walker and Roger T. Winn, "Climatological statistics on meteorological factors affecting electro-optics in Germany," AD-TR-80-61, U.S. Air Force Armament Division, Eglin AFB, Florida (Aug. 1980).
20. J. H. McCoy, D. B. Rencsh, and R. K. Long, "Water vapor continuum absorption of carbon dioxide laser radiation near 10 μm," *Applied Otpics* **8,** 1471–1478 (1969).
21. F. X. Kneizys, "Users guide to LOWTRAN 7," Air Force Geophysics Laboratory, AFGL-TR-88-0177, Environmental Research Papers, No. 1010 (16 Aug. 1988).
22. H. J. P. Smith et al., "FASCODE-fast atmospheric signature code (spectral transmittance and radiance)," Air Force Geophysics Laboratory, AFGL-TR-78-0081, Contract No. F19628-77-C-0041 (16 Jan. 1978).
23. "Guide to reference and standard atmospheric models," American Institute of Aeronautics and Astronautics, American National Standard, ANSI/AIAA, G-003-1990 (6 Aug. 1990).
24. Merrill I. Skolnik, *Radar Handbook,* pp. 27-11–27-21, McGraw Hill, New York (1970).
25. "Laser signature handbook," Vol. I, Ballistic Missile Defense Advance Technology Center Contract DASG60-75-C-0042 Report No. 90-02-CR (1976).
26. J. L. Gilbert and S. Weiner, "Laser signature handbook, range-Doppler cross-section predictions," Vol. II, Ballistic Missile Defense Advance Technology Center Contract DASG60-75-C-0042, Report No. 90-03-CR, DTIC No. ADC007630 (1976).
27. Paul J. Kramer, "Laser signature handbook, speckle effects," Vol. III, Ballistic Missile Defense Advance Technology Center Contract DASG60-75-C-0042, Report No. 90-05-CR, DTIC No. ADC009545 (1977).
28. *DELTAS® Users Guide,* Nichols Research Corporation, Huntsville, AL.
29. Merrill I. Skolnik, *Radar Handbook,* pp. 2-18–2-23, McGraw Hill, New York (1970).
30. Merrill I. Skolnik, *Introduction to Radar Systems,* pp. 23–29, McGraw Hill, New York (1980).
31. David K. Barton, *Radar System Analysis,* p. 15, Artech House, Dedham, MA (1976).
32. Merrill I. Skolnik, *Introduction to Radar Systems,* pp. 388–390, McGraw Hill, New York (1980).
33. Stephen C. Cohen, "Heterodyne detection: phase front alignment, beam spot size, and detector uniformity," *Applied Optics* **14**(8), 1953–1959 (1975).
34. Walter L. Tucker and John L. Barnett, "Heterodyne efficiency of quadrant photodetector," *Applied Optics* **28**(5), 892–896 (1989).
35. S. Fowler, G. W. Kamerman, and G. Lawson, "Analysis of heterodyne efficiency for coherent laser radars," to be presented at SPIE Conf. on Applied Laser Radar Technology, Orlando, FL, April 1993.
36. W. W. Montgomery, D. J. Wilson, W. D. Crowe, et al., "Applications of coherent laser Doppler velocimetry—phase III," Air Force Armament Laboratory Report AFATL-TR-89-33 (Sep. 1989).
37. J. Bilbro, G. Fichtl, D. Fitzjarrald, M. Krause, and R. Lee, "Airborne Doppler lidar wind field measurements," *Bulletin of the American Meteorological Society* **65**(4), 348–359 (1984).
38. Anthony E. Siegman, *Lasers,* University Science Books, Mill Valley, CA (1986).
39. Albert B. Jelalian, *Laser Radar Systems,* p. 3, Artech House, Boston (1992).
40. A. L. Kachelmyer, "Range-Doppler imaging with a laser radar," *The Lincoln Laboratory Journal* **3**(1), 87–118.
41. C. C. Chen, "A correction for Middleton's visible- and infrared-radiation extinction coefficients due to rain," The Rand Corporation, R-1523-PR (Aug. 1974).

Bibliography

Born, M. and E. Wolf, *Principles of Optics,* Pergamon Press, New York (1980).
Siegman, Anthony E., *Lasers,* University Science Books, Mill Valley, CA (1986).
Skolnik, Merrill I., *Introduction to Radar Systems,* 2nd ed., McGraw Hill, New York (1980).
Skolnik, Merrill I., *Radar Handbook,* McGraw Hill, New York (1970).
Smith, Warren J., *Modern Optical Engineering,* McGraw Hill, New York (1966).

CHAPTER 2

Laser Rangefinders

Robert W. Byren
Hughes Aircraft Company
El Segundo, California

CONTENTS

2.1 INTRODUCTION

A laser rangefinder (LRF) is a device that calculates the distance to a target by measuring the time it takes the laser light to travel a round-trip path from the laser transceiver to the target and back to the transceiver. This chapter discusses pulsed laser rangefinders that use direct (incoherent) detection receivers. Modulated continuous wave (cw) laser rangefinders that use heterodyne or coherent detection are typically classified as laser radars and the theory of operation of these devices is discussed in Chapter 1.

The neodymium-doped yttrium aluminum garnet (Nd:YAG) laser has been the workhorse of armed forces worldwide for laser rangefinder applications since replacing the ruby laser in the late 1960s. Modern battle tanks, such as the U.S. M1 Abrams tank and the German Leopard II tank, use Nd:YAG laser rangefinders as part of an integrated fire control system to improve the accuracy of the main weapon. Nd:YAG laser rangefinders are used extensively in air defense systems to improve tracking accuracy against high-velocity aircraft and maneuvering helicopters, and in hand-held sets for field commanders and forward observers.

Carbon dioxide (CO_2), Raman-shifted Nd:YAG, and erbium-doped glass (Er:glass) lasers were developed in the late 1970s to mid-1980s as eyesafe alternatives to the Nd:YAG laser. These lasers also provide improved performance in poor weather and smoke and are expected to supplant the Nd:YAG LRF for most single-function tactical rangefinding applications.[1]

Lasers are also used in military systems to illuminate remote targets to "designate" those targets for precision-guided weapons, such as bombs, missiles, or artillery rounds. Guidance is performed by an electro-optical seeker within the nose of the weapon, which provides guidance commands to home on the laser spot and hit the target. The earliest operational laser target designators used Nd:YAG lasers operating at 1.06 μm. A family of precision laser-guided weapons was developed in conjunction with these designators and produced in large quantities. As a result, the Nd:YAG laser remains the only laser in use today for designator applications and has also become the basis for all dual-function rangefinder-designator applications, with Raman shifting providing a compatible eyesafe mode for rangefinding and training.[2]

This chapter discusses laser rangefinder operation theory, the requirements for generic LRF applications, the performance equation used to design and analyze LRF systems, and the characteristics of typical laser rangefinders. Performance data are provided for CO_2, Raman-shifted Nd:YAG, and Er:glass laser rangefinders, as well as the conventional Nd:YAG and ruby LRFs.

2.2 THEORY OF OPERATION

Laser rangefinders considered in this chapter operate on the pulse-echo principle, in which target range is determined by measuring the time it takes a laser pulse, traveling at the speed of light, to traverse the round-trip path from the laser transceiver to the target and back to the transceiver. The relationship between the measured time of flight and the range is given by

$$R = \frac{cT}{2\langle n \rangle}, \tag{2.1}$$

where

R = range to target (m)
T = time of flight of laser pulse (s)
c = speed of light in vacuum (2.998×10^8 m/s)
$\langle n \rangle$ = path-averaged index of refraction.

The block diagram in Fig. 2.1 describes the operation of most pulse-echo laser rangefinders. A "laser fire" command from an operator or fire control computer triggers the laser to fire a temporally narrow laser pulse toward the target. An electrical signal is also generated when the laser is fired, which starts a digital range counter. The transmitted laser beam makes one pass through the atmosphere to the target, and a small portion of the beam is reflected or scattered by the target in the direction of the laser transceiver. This reflected portion experiences a second pass through the atmosphere, after which it is collected by the rangefinder optics and focused on a photodetector.

A passive sensor such as a visual sight, television camera, or thermal imager is often used in conjunction with a servo-controlled pointing and stabilization system to direct the line of sight of the laser rangefinder to the target. Rangefinder optics are used to adjust the divergence of the laser beam and receiver field of view. Selection of the optimal laser beam divergence and receiver field of view depends on many factors, including target size; maximum target range; line-of-sight pointing error resulting from imperfect tracking and stabilization; boresight error between the laser, receiver, and tracking sensor; clutter in the foreground or background of the target; and physical constraints on the size of the rangefinder optics. A narrow-band optical filter is normally used to enhance sensitivity by removing much of the unwanted background radiation from the spectrally narrow (monochromatic) laser signal.

A fast photodetector, such as a semiconductor photodiode or photomultiplier tube, converts the optical signal into an electrical signal, preserving to a large degree the temporal shape of the pulse. This electrical signal is then amplified and electronically filtered to suppress the high-frequency noise components introduced in the detection and preamplification processes. The peak signal level is next compared with an electrical threshold value (set sufficiently above the noise level to limit false alarms), and the digital output of the comparator is used to stop the range counter. The target range is finally read as a digital output from the range counter.

Most tactical laser rangefinders provide a multiple target indication and a means for selecting which target range is displayed (e.g., first-last reply). First-reply mode is normally selected when the target does not fill the receiver field of view to avoid ranging to the background clutter. Last-reply mode is normally selected in fog or smoke to avoid ranging on the aerosol backscatter signal. The operator can also use the first-last-reply mode control to resolve the range ambiguity when multiple targets or terrain features appear within the receiver field of view.

The described pulse-echo laser rangefinding technique offers considerable advantages over earlier forms of passive rangefinding, such as stadiametric and stereoscopic ranging.[3] Stadiametric ranging determines range by comparing the angular subtense of the target as seen through a passive sight with known target size (usually vertical height). This is usually done manually with

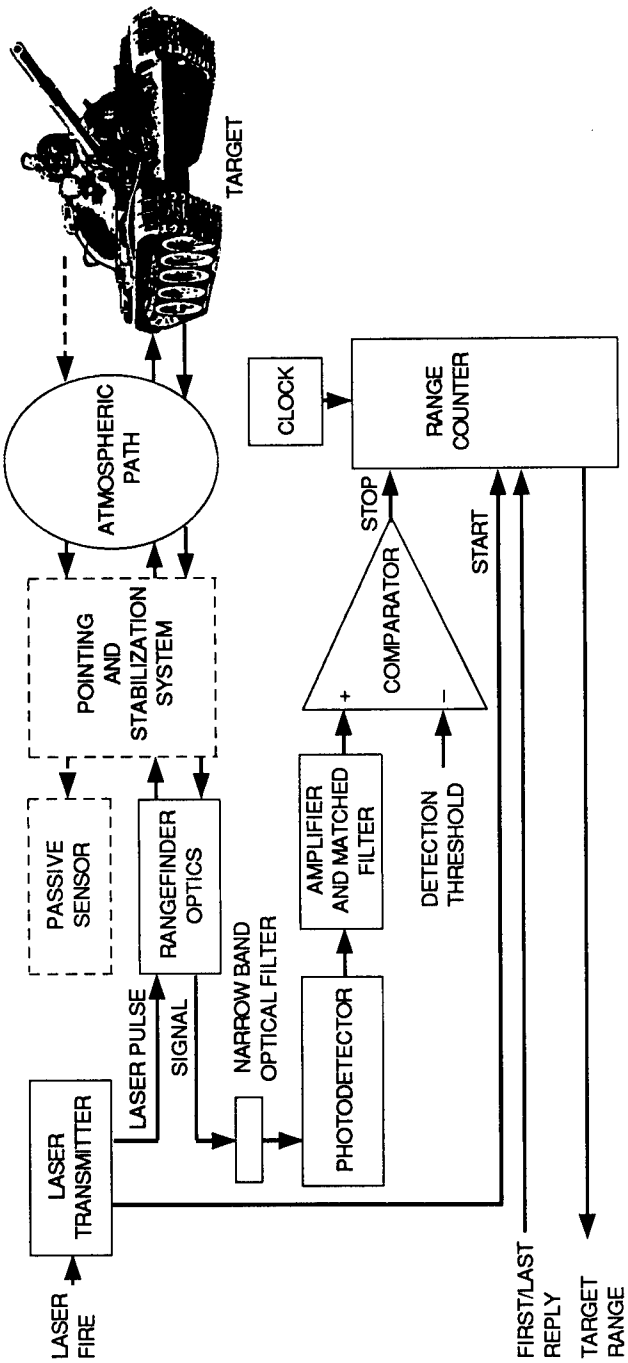


Fig. 2.1 Laser rangefinder block diagram.

a wedge-shaped pattern on the sighting reticle but can be done automatically using the size of a computer-generated tracking gate. Manual operation is time consuming, which adversely affects crew work load and limits the approach to applications that require infrequent range measurements (low ranging rates). The accuracy of stadiametric ranging is dependent on *a priori* knowledge of the target size and the ability of the sight to accurately resolve target dimensions, which degrade at longer ranges. Target obscuration, camouflage, and poor visibility also limit the accuracy of the stadiametric ranging technique.

Stereoscopic or optical coincidence ranging determines range by measuring the parallax angle between two precisely boresighted sights that are separated by a "baseline" distance. This is usually done manually, but can be automated with an image correlator. Stereoscopic ranging does not depend on *a priori* target size information and is less affected by obscuration and camouflage; however, the ranging accuracy is also range dependent and is limited by the practical length of the triangulation baseline (usually 1 to 2 m for combat vehicles) and the boresight accuracy of the two sights. The stereoscopic approach is not practical for man-portable systems because of size constraints and is also limited to low ranging rate applications.

Pulse-echo laser rangefinders avoid these inherent limitations in passive ranging and provide improved range accuracy, longer range performance, and higher ranging rates. Laser rangefinders, however, do have basic limitations that generally involve trade-offs between rangefinder performance, eye safety, size, weight, power consumption, and cost for a specific application.

2.3 LASER RANGEFINDER APPLICATIONS

Table 2.1 contrasts typical rangefinder requirements for eight representative military and commercial applications. These applications are grouped according to the host platform and operating environment.

The man-portable applications for laser rangefinders include hand-held devices for infantry and artillery observers (see Fig. 2.2) and dual-mode rangefinder-designators for forward observers and forward air controllers (FACs). The system requirements for these applications are driven by the need for portability, light weight, battery operation, and low unit production cost. Eye safety is a critical concern because of the close proximity with other friendly forces in combat and the need to train extensively with unprotected troops and noncooperative targets (i.e., no retroreflectors).[3,4]

Land vehicle applications include tank and infantry fighting vehicle (IFV) fire control and air defense gun-missile fire control, as well as vehicular variants of ground rangefinder-designators. The purpose of the rangefinder for tank fire control (see Fig. 2.3) is to provide superelevation correction for the ballistic trajectory and azimuth correction for cross-wind and target motion.[5] Infantry fighting vehicles use rangefinders primarily to verify whether the target is within antitank missile range, and secondarily for gun fire control and target prioritization. Both the tank and IFV require eye safety for the discussed reasons, although attenuation filters are presently used with noneye-safe (Nd:YAG) lasers to achieve eye safety in training against cooperative targets. To be totally effective, the laser rangefinder should be able to range to any target that can be detected and recognized by the passive fire control

Table 2.1 Typical Laser Rangefinder Requirements for Different Applications

	Manportable	Land Vehicle		Airborne		Ship	Space-based	Airfield
Requirement	Infantry, FAC, Artillery Observation[3,4]	Tank/IFV Fire Control[1,5]	Air Defense Gun/ Missile[7]	Helicopter Missile Command Guidance	Fixed Wing Targeting	Carrier Aircraft Landing Aids	Long Range Tracking	Ceilometry [8,9]
Maximum Range	4-10 km	4-10 km	10-20 km	4-10 km	10-20 km	5 km	>1000 km	2-3 km
Pulse Energy	5-40 mJ	5-40 mJ	25-150 mJ	5-40 mJ	25-150 mJ	25-100 mJ	0.1-10 J	5-40 mJ
Beam Divergence	1-2 mrad	0.4-1 mrad	0.5-2.5 mrad	0.4-1 mrad	0.1-0.5 mrad	0.25-0.75 mrad	<20 μrad	1-5 mrad
Beam Quality	Multi mode	Multi Mode	Multi Mode	Multi Mode	Multi Mode	Multi Mode	Near Diffraction Limited	Multi Mode
Pulse Rate	Single Shot	0.1 to 1 Hz	6-20 Hz	4 Hz	5-20 Hz	5-20 Hz	10-100 Hz	Single Shot
Duty Factor	>10 s between pulses	Burst 10-100 pulses	100%	25-100%	100%	100%	25-100%	>10 s between pulses
Range Accuracy	±10 m	±5-10 m	±2.5-5 m	±5-10 m	±1-10 m	±5-10 m	<±5-10m	±100 m
Target Discrimination	50-100 m	20 m	Not issue	20 m	20 m	Not issue	Not issue	100 m
Atmospheric Penetration	Compatible with Direct View Optics (DVO)	Compatible with DVO, TV, and FLIR	Clear to hazy weather	Clear to hazy weather	Clear to hazy weather	Clear to hazy weather	Vacuum propagation	Clear (under ceiling)
Eyesafety	Required	Required	Desired	Desired	Desired	Required	Not issue	Required
Shared Functions	Some ground designators	No	Possible uplink/ beamrider	Some airborne designator	Most airborne designator	No	Active tracker	No
Weight/Size	Critical (3-5 lb)	Important (10-20 lb)	Important (35-50 lb)	Very important (8-15 lb)	Very important (15-35 lb)	Not issue	Critical	Not issue
Electrical/Thermal Efficiency	Very important (Battery)	Important	Important	Important	Important	Not issue	Critical	Not issue
On-Time Fraction	Very low	Low	Medium	Low	Low	High	Very low	Medium
Reliability	Medium	Medium	High	Medium	High	High	Very high	Low
Maintenance	Infrequent	Infrequent	Infrequent	Frequent	Frequent	Frequent	Unattended	Frequent
Unit Cost	Critical	Very important	Important	Important	Important	Less important	Not issue	Important

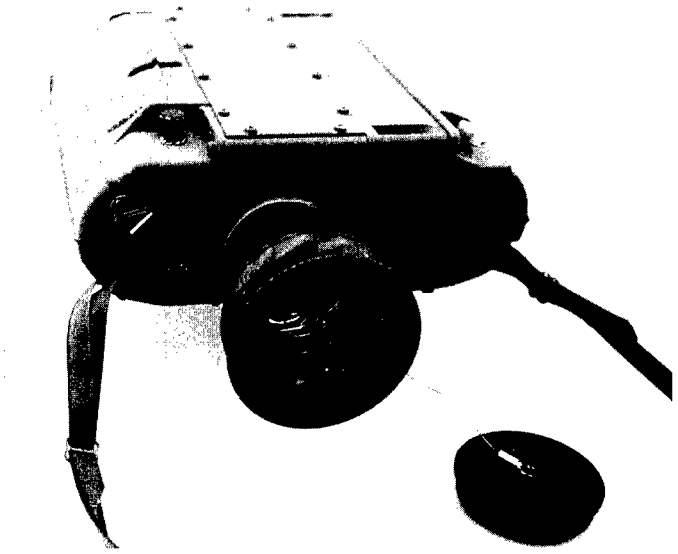


Fig. 2.2 Optic-Electronic Corporation mini eyesafe laser infrared observation set (MELIOS) utilizing Er:glass.

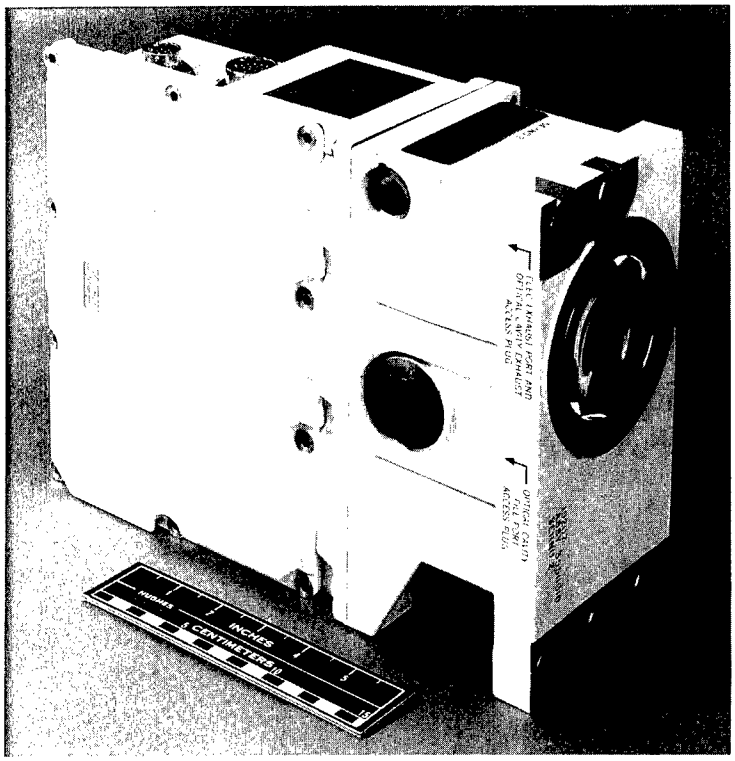


Fig. 2.3 Hughes Aircraft Company M1 Abrams tank laser rangefinder utilizing Nd:YAG.

sensors, which may include direct-view optics, television sensors, or thermal imagers (FLIRs). This is a very stressing requirement that cannot be satisfied completely by any single-wavelength laser rangefinder.[1]

Air defense rangefinders (see Fig. 2.4) and to some extent tank-IFV rangefinders used for self-protection against helicopter threats[6] are driven by a requirement to provide a stable track on a maneuvering target in range and range rate. This requires a high data rate (high laser pulse rate) and good range accuracy. The engagement ranges can be reasonably long (up to 20 km); however, these long-range encounters typically occur under good weather conditions without obscurants, thus requiring only slightly higher sensitivity than tank rangefinders.[7] High humidity in certain seasons and climatic regions limits the maximum range capability of long-wavelength laser rangefinders (CO_2 at 10.6 μm) because of strong H_2O molecular absorption.

Airborne applications include missile command guidance for attack helicopters and electro-optical targeting for fixed-wing close air support and interdiction aircraft (see Fig. 2.5). These applications typically require a 1.06-μm designation capability; therefore a baseline Nd:YAG laser designator may be used with Raman shifting for eyesafe ranging and training modes.[2] Maintenance for airborne systems is performed on a routine basis. System features, such as liquid cooling, which tend to require higher maintenance, are therefore

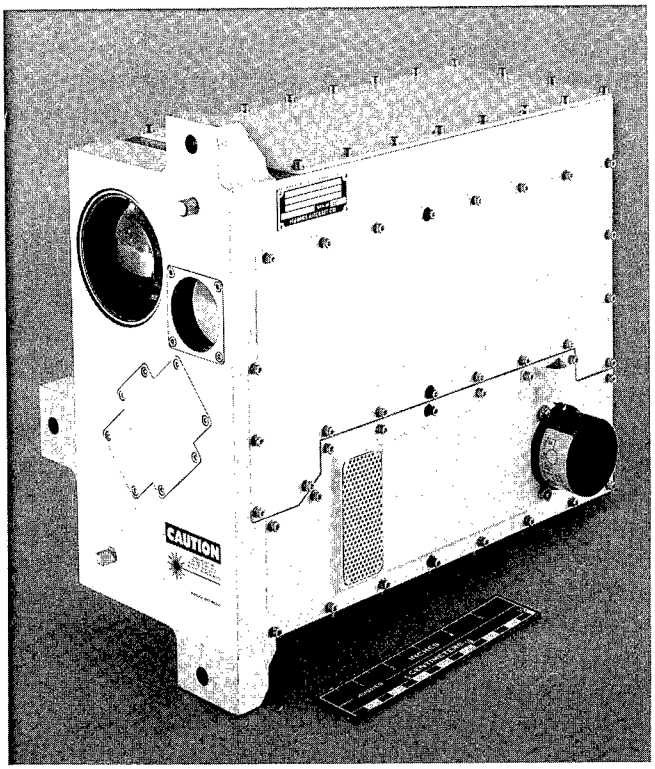


Fig. 2.4 Hughes Aircraft Company electro-optical tracking sensor (EOTS) laser rangefinder utilizing Raman-shifted Nd:YAG.

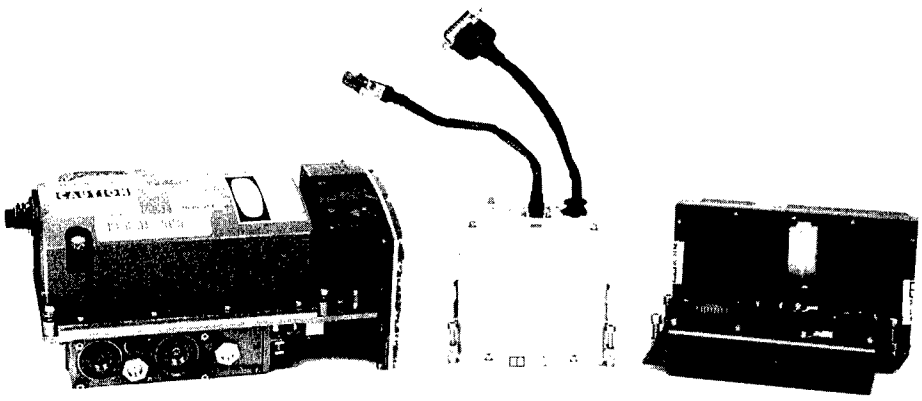


Fig. 2.5 Litton Laser Systems LANTIRN laser designator ranger utilizing Raman-shifted Nd:YAG.

accepted for airborne applications, whereas alternate approaches (pressurized gas or hybrid cooling) are preferred for man-portable and vehicular applications.

Laser rangefinders are also used on board naval vessels for conventional fire control and air defense applications; unique shipboard applications, such as covert (radar-silent) carrier aircraft recovery, are beginning to emerge. Although these systems are similar in performance to the tactical air defense systems, the requirements on size, weight, electrical efficiency, maintainability, and cost are generally less severe.

Space-based laser rangefinders for a variety of applications are still in the planning stages, although experimental systems have been flown. The stressing requirements for these applications are light weight (because cost is driven primarily by launch weight), high efficiency (because prime power is a premium and cooling is difficult in space), and long unattended lifetime. Space-based rangefinders used for long-range tracking in strategic defense will rarely be operated (occasional built-in tests and operational checks) but must be extremely reliable if ever called to action. Because aperture size is a major driver in launch cost, short-wavelength diffraction-limited lasers are preferred. Ultraviolet wavelengths are attractive for certain functions because of the extremely low earth background at wavelengths below about 300 nm. To save weight, the laser used for rangefinding is often used for other functions such as active tracking, thus requiring a laser wavelength that is in the optical band of the tracking sensor. A frequency-doubled Nd:YAG laser, for example, may be used as an illumination source for an intensified CCD active imager. Other potential shared functions, such as Doppler spectrum analysis and communications, require coherent or cw waveforms and are beyond the scope of this chapter.

Cloud-height indicators, or ceilometers, use laser rangefinders to measure the vertical height of cloud ceilings. These laser systems are reasonably covert, making them attractive for forward military air bases, and are relatively inexpensive, making them attractive for both large and small commercial airports. Eye safety is essential, because the laser beams are fired through high-traffic, unrestricted air space. It is of historical interest that the first eyesafe laser rangefinders were developed for cloud-height measurement.[8,9]

2.4 LASER RANGE EQUATION

The range performance of pulse-echo laser rangefinders is generally defined as the range at which the laser transceiver can detect the laser pulse reflected from the target with a certain probability. This subsection presents a general formulation of the signal-to-noise range equation for the laser rangefinder and reduces this equation to a specific form for tactical fire control applications. The following subsection describes how the probability of ranging is calculated from the signal-to-noise ratio (SNR) for a given false alarm rate.

The general range equation for laser rangefinder SNR is

$$\mathrm{SNR} = \frac{\int_{\Omega} I_L \, d\omega}{\mathrm{NEI}} f \eta_T \frac{\exp[-2(kR + \alpha\langle \mathrm{CL} \rangle)]}{R^2} , \tag{2.2}$$

where

SNR = signal-to-noise ratio
I_L = radiant intensity of the laser beam (W/sr)
Ω = solid angle subtended by target that also lies within the receiver field of view (sr)
$d\omega$ = differential solid angle from transceiver (sr)
NEI = noise-equivalent irradiance of the receiver (W/m^2)
f = retroreflectance of target (sr^{-1})
η_T = efficiency factor accounting for pulse stretching by the target
k = total atmospheric extinction coefficient (m^{-1})
R = slant range to target (m)
α = mass extinction coefficient of obscurant (m^2/g)
C = average mass concentration of obscurant cloud (g/m^3)
L = path length through obscurant cloud along line of sight (m)
$\langle \mathrm{CL} \rangle$ = path-averaged concentration–length product of obscurant (g/m^2).

The terms within this range equation are discussed in the following.

2.4.1 Power Integral

The integral term in the SNR range equation represents the total peak power transmitted by the laser within a solid angle that is bounded by the target angular subtense or the receiver field of view, whichever is smaller. For most applications, the target will be in the far field of the laser beam, and the far-

field radiant intensity pattern of the laser beam can be used in evaluating the integral.

Figure 2.6 illustrates three classical cases for the evaluation of the power integral. Figure 2.6(a) represents the normal tank fire control situation in which the target subtends a much larger angle than the receiver field of view (FOV) or the laser beam. The power integral in this case is simply the total transmitted power within the receiver field of view and is given by

$$\int_{\Omega} I_L \, d\omega = \int_{\text{FOV}} I_L \, d\omega = P_L \eta_B \quad \text{[see Fig. 2.6(a)]} \ , \tag{2.3}$$

where

FOV = solid angle of receiver field of view (sr)
P_L = total peak power of laser (W)
η_B = fraction of laser energy within receiver field of view.

Note that the power integral [Eq. (2.3)] is independent of range, therefore the geometric range dependence (no atmospheric attenuation) in the SNR is determined by the $1/R^2$ range term in Eq. (2.2).

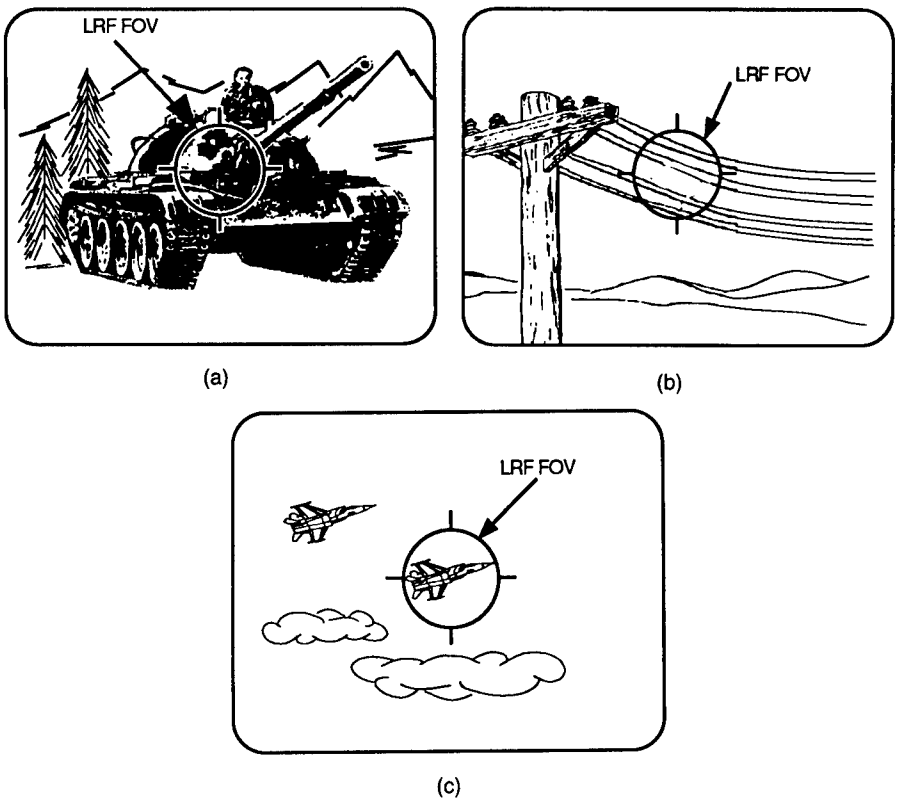


Fig. 2.6 Classical rangefinder situations: (a) target larger than receiver FOV, (b) one-dimensional target larger than receiver FOV, and (c) target smaller than receiver FOV.

Figure 2.6(b) illustrates the case where the target, such as a wire or pole, fills the receiver field of view in one dimension but is smaller than the receiver field of view in the other dimension. The power integral for this case is given by

$$\int_{\Omega} I_L \, d\omega = \int_{\mathrm{FOV}_x} \int_{\Theta_y} I_L \, d\theta_x \, d\theta_y \ , \tag{2.4}$$

where

FOV_x = receiver horizontal field of view (rad)
Θ_y = vertical angle subtended by feature (rad)
$d\theta_x, d\theta_y$ = differential horizontal and vertical angles (rad).

If the target is very thin and the radiant intensity of the beam does not vary appreciably across the thin dimension, the power integral can be further simplified to

$$\int_{\Omega} I_L \, d\omega = \left(\int_{\mathrm{FOV}_x} I_L|_y \, d\theta_x \right) \cdot \frac{T_T}{R} \quad \text{[see Fig. 2.6(b)]} \ , \tag{2.5}$$

where T_T is the projected thickness of the target across the thin dimension in meters. Note that the $1/R$ range term in the power integral [Eq. (2.5)] combined with the $1/R^2$ range term in Eq. (2.2) gives a net $1/R^3$ geometric range dependence in the SNR. For multiple targets at essentially the same range [e.g., multiple wires in Fig. 2.6(b)], the total power integral is calculated by summing the contribution from each target (wire).

Figure 2.6(c) illustrates the case where the target, such as an aircraft, satellite, or reentry vehicle at long range, is smaller than the receiver field of view. In this case, the power integral is taken over the solid angle subtended by the target. At sufficiently long range, targets will appear essentially as point sources, and a single value of the laser radiant intensity may be used. The power integral for this case is given by

$$\int_{\Omega} I_L \, d\omega = I_L|_T \cdot \Omega_T = I_L|_T \cdot \frac{A_T}{R^2} \quad \text{[see Fig. 2.6(c)]} \ , \tag{2.6}$$

where Ω_T is the target angular subtense (sr) and A_T is the projected area of the target (m^2). Retroflective targets, such as corner cubes or bicycle reflectors, also appear as point sources, even at relatively short range, and Eq. (2.6) can therefore be used. Note that the $1/R^2$ range term in the power integral [Eq. (2.6)] combined with the $1/R^2$ range term in Eq. (2.2) gives a net $1/R^4$ geometric range dependence in the SNR.

It is important to understand that a particular operational scenario may involve more than one of the described classical rangefinder situations. Consider, for example, a laser rangefinder associated with an air defense system that is tracking an inbound aircraft. At long range, the aircraft will be much smaller than the rangefinder field of view, as depicted in Fig. 2.6(c), and the geometric range dependence will follow a $1/R^4$ relation. At some range, the fuselage and/or wings of the aircraft will just fill the rangefinder field of view

in one dimension, and the geometric range dependence will transition smoothly from $1/R^4$ to $1/R^3$. At some shorter range, the fuselage will completely fill the rangefinder field of view, and the geometric range dependence will transition smoothly from $1/R^3$ to $1/R^2$. Plotting the SNR versus range for this case would yield a continuous curve with a change in slope at the transition points.

2.4.2 Noise-Equivalent Irradiance

Optical and electrical noise enters the signal-to-noise range equation [Eq. (2.2)] through the noise-equivalent irradiance (NEI) term. The NEI of the laser receiver is defined as the optical signal irradiance measured at the entrance aperture of the receiver that would give an SNR of unity. The NEI is calculated as the root-sum-square of the equivalent optical signals from all optical and electrical sources that degrade the laser signal during the detection, amplification, and bandwidth filtering processes.

The primary sources of noise for fast photodiode detectors are (1) shot noise caused by random fluctuations in the direct or reflected solar and thermal backgrounds (background noise), (2) shot noise in the detector caused by random generation and recombination of carriers (detector noise), and (3) Johnson noise in the preamplifier caused by thermal agitation of electrons within the dissipative elements (preamplifier noise). Other sources of noise, such as shot noise in the signal and $1/f$ noise in the detection circuit are also present but can often be ignored because of the short pulse durations and the wide receiver bandwidths. For most terrestrial applications, the received signal contains many photons, and quantum noise need not be considered. For space applications where the backgrounds are small and photomultipliers are used to boost sensitivity, the quantum nature of the signal and background photon flux must be considered. This quantum noise is discussed in Sec. 2.5.

The noise-equivalent irradiance for photodetectors that have internal amplification (gain), such as avalanche photodiodes and photomultiplier tubes, is given by

$$\mathrm{NEI} = \frac{\left[(\overline{I_{\mathrm{BN}}^2} + \overline{I_{\mathrm{DN}}^2})M^2F + \overline{I_{\mathrm{PN}}^2}\right]^{1/2}}{R_D M A_R T_O \eta_e} , \tag{2.7}$$

where

$(\overline{I_{\mathrm{BN}}^2})^{1/2}$ = unity-gain rms background noise current (A)

$(\overline{I_{\mathrm{DN}}^2})^{1/2}$ = unity-gain rms detector noise current resulting from bulk leakage (A)

$(\overline{I_{\mathrm{PN}}^2})^{1/2}$ = rms preamplifier noise current (A)

M = photodetector gain

F = excess noise factor

R_D = unity-gain responsivity of the detector (A/W)

A_R = effective area of receiver collecting aperture (m^2)

T_O = receiver optical transmission

η_e = signal processing efficiency factor.

The unity-gain rms noise current terms are discussed in the following paragraphs. The photodetector gain M is the enhancement in responsivity that results from avalanche multiplication of carriers in a high-field region near the junction of an avalanche photodiode or from electron acceleration and enhanced secondary emission from the cascaded dynodes or microchannel plates of a photomultiplier tube. This internal amplification is not noise free, and the excess noise factor F accounts for excessive noise caused by amplification. Excess noise results from the statistical nature of the amplification process in which individual carriers experience a gain that varies around the average gain value M.

The signal-to-noise advantage afforded by internal amplification devices can be appreciated by studying Eq. (2.7). Typically, the unity-gain values of background and detector noise are much smaller than the preamplifier noise, hence at unity gain the receiver will be preamplifier noise limited. Because the preamplifier noise is unaffected by the internal amplification within the detector, the signal will increase linearly with detector gain, whereas the noise measured at the output of the preamplifier remains essentially constant (linear SNR increase with gain for low gain values). This situation continues until the amplified values of background and/or detector noise become significant relative to the unamplified preamplifier noise. At this point, the increase in SNR at the preamplifier output levels off, and the receiver becomes background or detector noise limited. As the gain is increased further, the background and detector noise continue to increase but at a faster rate than the signal as a result of excess noise in the amplification process. Eventually, the SNR reaches a peak value (optimum gain) then decreases gradually with increased gain because of this excess noise.

2.4.3 Excess Noise Factor

For avalanche photodiodes, the excess noise factor is a function of the avalanche gain and is dependent on material properties (relative ionization rates of holes and electrons) and the physical construction of the device (the junction type, the width of the ionizing region, and the peak field). For n^+p silicon avalanche photodiodes, the excess noise is dominated by electron amplification noise. This excess noise factor has been predicted by McIntyre[10] to be

$$F_e \approx k_{\text{eff}} M_e + \left(2 - \frac{1}{M_e}\right)(1 - k_{\text{eff}}) , \tag{2.8}$$

where

F_e = excess noise for injected electrons
M_e = electron amplification
k_{eff} = weighted ratio of ionization coefficients of holes and electrons.

Figure 2.7 plots the predicted excess noise resulting from electron amplification versus gain for different values of k_{eff}. This plot is in good agreement with measured values for silicon, with $k_{\text{eff}} = 0.02$ (Ref. 11), and InGaAs, with $k_{\text{eff}} = 0.45$ (Ref. 12).

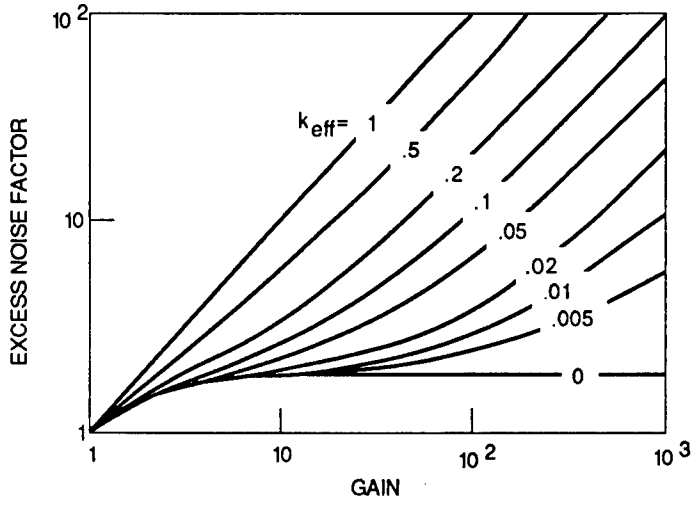


Fig. 2.7 Excess noise factor caused by electrons versus gain.[11]

2.4.4 Background Noise

The background shot noise current taken at the output of the detector (assuming unity gain) is given by

$$(\overline{I_{BN}^2})^{1/2} = (2qi_B\Delta f_n)^{1/2} , \tag{2.9}$$

where

q = electronic charge (1.6×10^{-19} C)
i_B = total dc background current from all sources (A)
Δf_n = receiver noise bandwidth (Hz).

For visible and near-infrared laser rangefinders, the dc background is dominated by solar sources, which include solar radiation reflected by the target and nearby terrain, solar radiation scattered by the atmosphere, and the sun itself, depending on which of these sources is within the receiver field of view. For long-wavelength infrared laser rangefinders operating within the 8- to 12-μm thermal band (e.g., CO_2), the dc background is dominated by thermal emission from the target, the terrain, the sky, and/or the sun. Mid-infrared laser rangefinders operating within the 3- to 5-μm band may have contributions from both solar and thermal sources.

For the "normal" tank fire control situation, the target fills the receiver field of view [see Fig. 2.6(a)], and the only contribution to the dc background is the solar energy reflected by the target and/or thermal emission from the target. In this case, the dc background current, ignoring atmospheric losses between the target and the receiver, is given by

$$i_B = R_D A_R T_O \Delta\lambda \Omega_{FOV}[E_\lambda f + M_\lambda/\pi] , \tag{2.10}$$

where

$\Delta\lambda$ = full width at half maximum (FWHM) optical bandpass of the receiver (μm)
Ω_{FOV} = solid angle of receiver field of view (sr)
E_λ = spectral irradiance of sun (W/m^2 μm)
f = retroreflectance of the target (sr^{-1})
M_λ = spectral radiant exitance of target (W/m^2 μm).

The detector responsivity is a parameter that is generally supplied in a specification sheet by the detector manufacturer. For detectors with internal amplification, such as avalanche photodiodes, the manufacturer may specify the responsivity at some gain value. In this case, the unity gain value is calculated by dividing the specified responsivity by the specified gain. The receiver optical bandpass is generally determined by the passband characteristics of a single narrow-band optical filter, which is used to limit the amount of background energy collected by the detector. For spectral irradiance values of the sun as a function of wavelength, see Volume 1, Chapter 3; for the spectral radiant exitance values for an ideal blackbody target as a function of temperature and wavelength, see Volume 1, Chapter 1. The target retroreflectance is discussed in a subsequent paragraph.

2.4.5 Detector Noise

The rms shot noise that results from random thermal generation and recombination of carriers within a semiconductor photodiode or the random thermal emission of electrons from the photocathode surfaces in a photomultiplier tube under unity-gain conditions is given by

$$(\overline{I_{DN}^2})^{1/2} = (2qi_D\Delta f_n)^{1/2} , \tag{2.11}$$

where i_d is the unity-gain detector dark current (A).

Manufacturers often specify a value of noise-equivalent power (NEP) at a given wavelength and gain value for the detector, which is a direct measure of detector shot noise. The unity-gain rms detector noise current can be calculated as follows, where the detector NEP and responsivity are given at the same wavelength and gain value. Note that the rms detector noise current is independent of operating wavelength.

$$(\overline{I_{DN}^2})^{1/2} = \frac{\text{NEP}|_{\lambda_R,M_R} R_D|_{\lambda_R,M_R} (\Delta f_n)^{1/2}}{M_R[F(M_R)]^{1/2}} , \tag{2.12}$$

where

λ_R = reference wavelength (μm)
M_R = reference gain
$\text{NEP}|_{\lambda_R,M_R}$ = detector rms noise-equivalent power measured at λ_R and M_R (W/$\sqrt{\text{Hz}}$)
$R_D|_{\lambda_R,M_R}$ = detector responsivity measured at λ_R and M_R (A/W)
$F(M_R)$ = excess noise factor of detector calculated at M_R.

Equation (2.12) assumes that all of the dark current in the detector experiences the same avalanche multiplication, which is true only for the bulk leakage component, not the surface leakage component. Equation (2.12) also assumes a flat noise spectrum such that the NEP is independent of frequency. Within the limits of these assumptions, Eq. (2.12) is a reasonably good approximation.

2.4.6 Preamplifier Noise

Johnson noise caused by thermal agitation of electrons in the resistive preamplifier load (as well as other noise sources in the preamplifier electronics) may be specified by the manufacturer but often must be measured. Manufacturers of integrated photodiode-preamplifier modules often specify a system noise-equivalent power at a particular wavelength under certain operating conditions, which includes both detector and preamplifier noise. The rms detector and preamplifier noise current for these modules can be calculated using Eq. (2.12).

2.4.7 Receiver Bandwidth

The electrical bandwidth of the receiver is determined by both the frequency response characteristics of the detector and the bandwidth shaping or filtering characteristics of the preamplifier circuit. For most applications, the designer selects a detector with adequate response (rise time) and tailors the bandwidth of the preamplifier-filter circuit to provide the best overall signal-to-noise performance. A wide receiver bandwidth preserves the amplitude of the electrical signal but allows too much of the noise spectrum to enter the detection process. Conversely, a narrow receiver bandwidth limits the detectable noise, but also degrades the amplitude and shape of the signal pulse prior to detection, reducing both the SNR and the ability of the rangefinder to resolve targets separated in range. The optimum signal bandwidth for a sine-squared pulse shape and a single-pole filter characteristic is given by

$$\Delta f_s = 0.189/\tau \ , \tag{2.13}$$

where Δf_s is the receiver signal bandwidth (Hz). The equivalent noise bandwidth is found by multiplying the signal bandwidth by $\pi/2$.

2.4.8 Signal-Processing Efficiency

The signal-processing efficiency factor in Eq. (2.7) is the factor by which the peak electrical signal is reduced because of band limiting or frequency response shaping within the receiver electronics. The signal processing efficiency factor for a sine-squared pulse shape and an optimized (matched) single-pole filter characteristic is given by

$$\eta_e = 0.595 \ . \tag{2.14}$$

2.4.9 Retroreflectance

The retroreflectance of a target is defined as the ratio of the radiant intensity of the laser radiation reflected by the target in the direction of the receiver to

Table 2.2 Hemispheric Reflectance of Standard U.S. Military Coatings (from Ref. 13)

Laser Wavelength (μm)	Hemispherical Reflectance (without specular component) Federal Color Number of Coating				
	30219 Light Brown	34079 Olive Drab	34102 Olive Drab	34159 Green Drab	36118 Gray
0.53	0.191	0.078	0.092	0.185	0.125
0.69*	0.32	0.09	0.08	0.16	0.11
0.80 – 0.85	0.281	0.097	0.107	0.140	0.087
1.06	0.257	0.118	0.118	0.121	0.074
1.3	0.226	0.122	0.105	0.106	0.065
1.54*	0.21	0.15	0.10	0.10	0.06
2.8	0.055	0.034	0.016	0.039	0.046
3.3	0.039	0.020	0.009	0.034	0.044
5.3	0.073	0.056	0.019	0.039	0.036
10.6	0.018	no signal detected	no signal detected	0.016	0.029

*Data at 0.69 and 1.54 μm estimated from hemispherical reflectances versus wavelength curves.[13]

the laser power incident on the target. The retroreflectance for Lambertian targets is given by

$$f = \frac{\rho}{\pi}, \tag{2.15}$$

where ρ is the hemispherical reflectance of the target. Hemispheric reflectance values of some military coatings are given in Table 2.2 for convenience. These reflectance values are for new (unweathered) coating samples measured under clean, dry surface conditions.

For very small targets, such as corner-cube retroreflectors, the reflectance is often given as a cross-section or effective-aperture value, which can be defined in a variety of ways (equivalent size of perfect reflecting sphere, equivalent size of perfect Lambertian reflector, ratio of retroreflected radiant intensity to incident irradiance, etc.). Note that the retroreflectance parameter used in Eq. (2.2) must be consistent with the particular cross-section definition. Section 12 of the *RCA Electro-Optics Handbook*[14] presents the theory of corner-cube retroreflection and provides cross-section values for perfect corner cubes and several commercial retroreflective array materials.

2.4.10 Pulse-Stretching Efficiency Factor

Targets that present a flat surface normal to the beam line of sight reflect the beam without changing its temporal characteristics. Most real targets, however, are not perfectly flat or may not necessarily be oriented at normal incidence. For these cases, some portion of the beam reflects from the target sooner (at closer range) than another portion of the beam, effectively stretching the temporal pulse shape of the reflected laser signal. If the target depth is

small compared with the physical length of the pulse, this effect is negligible. If the target depth is large compared with the length of the pulse, an efficiency factor must be calculated to account for the reduction in peak signal power caused by pulse stretching. For many situations, the efficiency factor is a unique function of the target shape, the small-scale target reflectivity, the ranging geometry, the laser temporal pulse shape, and the spatial characteristics of the laser beam. Therefore, a simple formulation cannot be given. For cases in which the target can be approximated by a tilted plane with uniform reflectance and the laser beam is uniform with a rectangular temporal pulse shape, the pulse-stretching efficiency factor is given by

$$\eta_T = \begin{cases} 1 & \text{for } \tau \geq 2(D/c) \\ \dfrac{\tau c}{2D} & \text{for } \tau \leq 2(D/c) \end{cases}, \tag{2.16}$$

where

η_T = efficiency factor accounting for pulse stretching by the target
τ = transmitted laser pulse width (s)
D = depth of target within receiver FOV (m).

2.4.11 Atmospheric Absorption and Scattering

The rate of signal attenuation resulting from the natural atmospheric absorption and scattering is assumed to be constant over the two-way path, therefore the atmospheric attenuation expression within Eq. (2.2) takes the form of a decaying exponential function of range. Values for the total atmospheric extinction coefficient k, which is the rate of signal attenuation, are typically calculated using a high spectral resolution atmospheric computer code such as the AFGL FASCOD2 (Ref. 15) or the ASL EOSAEL-87 (Ref. 16) models. Atmospheric extinction coefficients for specific laser wavelength and model atmospheres are also tabulated in references such as the JTCG/ME *Smoke and Natural Aerosol Parameters (SNAP) Manual.*[17] Values for atmospheric extinction derived from low spectral resolution computer codes, such as LOWTRAN, or from broadband measurements should not be used, because these will not properly account for narrow-band molecular absorption at specific laser wavelengths. Atmospheric attenuation may also be specified in units of decibels per meter, which is converted to an extinction coefficient by multiplying the decibels-per-meter values by 0.2303. Occasionally, atmospheric attenuation is characterized by an *absorption depth*, which is defined as the distance over which the signal is reduced to $1/e$ of its initial value. Absorption depth is the reciprocal of the extinction coefficient.

2.4.12 Attenuation by Obscurants

The rate of signal attenuation resulting from obscurants such as smoke or dust is generally not constant over the entire beam path. The rate of attenuation over a one-way path can be found using the Beer-Lambert relationship,[18,19] which is given by

$$\frac{dI}{dx} = -\alpha c_x I \ , \tag{2.17}$$

where

I = signal intensity (units of power or irradiance)
x = distance along beam path (m)
α = mass extinction coefficient of obscurant (m^2/g)
c_x = local mass concentration of obscurant at x (g/m^3).

Solving this differential equation (integrating over one-way path):

$$\ln\left(\frac{I_L}{I_o}\right) = -\alpha \int_L c_x \, dx = -\alpha \langle \mathrm{CL} \rangle \ , \tag{2.18}$$

where I_L/I_o is the transmission over the one-way path, $\langle \mathrm{CL} \rangle$ is the effective concentration-length product of the obscurant (g/m^2), and

$$\frac{I_L}{I_o} = \exp(-\alpha \langle \mathrm{CL} \rangle) \ , \tag{2.19}$$

which is the same form as the expression in Eq. (2.2).

The primary attenuation mechanism in battlefield obscurants is scattering by the aerosol droplets and smoke/dust particles. The scattering efficiency is determined by the relationship between the distribution of particle sizes and the wavelength of the laser radiation, therefore the mass extinction coefficient is a strong function of wavelength. The highest scattering efficiency occurs when the particle radius and the laser wavelength are nearly equal (assuming spherical particle).[19] Table 2.3 summarizes the mass extinction coefficients for common obscurants at specific laser wavelengths.

Example 2.1. Table 2.4 presents performance characteristics for typical laser rangefinders used for tank fire control. The parameters listed in the first column have been defined previously in this chapter. The derived values are based on the equations given or referenced in the source column. We assume for this example that the target is a military vehicle with a uniform diffuse Lambertian reflectance characteristic, that the depth of the target is insignificant, and that

Table 2.3 Mass Extinction Coefficients for Common Battlefield Obscurants (from Refs. 17, 18, and 20)

Laser Wavelength (μm)	Mass Extinction Coefficient (m^2/g) 50% Relative Humidity, 10°C					
	White Phosphorous	Hexachloro-Ethane*	Fog Oil	Diesel Fuel Fog	High Explosive Dust	
					Small	Large
0.53	3.7	—	—	—	0.32	0.035
0.694	3.24	3.89	6.86	5.56	0.32	0.035
0.80-0.85	2.6	3.3	5.8	4.9	0.29	0.036
1.06	1.41	2.30	3.50	3.25	0.26	0.037
1.54	0.57	1.16	1.66	1.63	0.26	0.037
10.6	0.39	0.14	0.01	0.03	0.24	0.036

* 85% relative humidity

Table 2.4 Performance Characteristics for Typical Tank Laser Rangefinders

Parameter	Symbol	Units	Source	Laser Rangefinder Type		
				Nd:YAG	Raman-shifted Nd:YAG	Carbon Dioxide
wavelength	λ	μm	Table 2-5	1.064	1.543	10.59
laser pulse energy	E_L	mJ	given	40	15	36
laser pulsewidth	τ	ns	given	8	6.5	60
laser peak power	P_L	W	E_L/τ	5.0E+06	2.3E+06	6.0E+05
energy fraction in FOV	η_B	—	given	0.8	0.8	0.8
detector type	—	—	given	EG&G C30872 Si APD	EG&G C30662 InGaAs APD	SBRC100μm PV HgCdTe
reference wavelength	λ_R	μm	mfr	1.06	1.54	10.59
reference gain	M_R	—	mfr	60	10	1
detector NEP (at λ_R, M_R)	$NEP\|_{\lambda_R,M_R}$	$W/\sqrt{Hz}$	mfr	1.0E-13	1.3E-13	5.OE-14
responsivity (at λ_R, M_R)	$R_D\|_{\lambda_R,M_R}$	A/W	mfr	9.0	9.3	7.1
unity-gain responsivity	R_D	A/W	$R_D\|_{\lambda_R,M_R}/M_R$	0.15	0.93	7.1
carrier ionization ratio	k_{eff}	—	mfr	0.02	0.45	N/A
photodetector gain	M	—	given	100	10	1
excess noise factor	F_e	—	Eq. 2-8	3.95	5.55	N/A
receiver signal bandwidth	Δf_s	MHz	Eq. 2-13	23.7	29.1	3.2
receiver noise bandwidth	Δf_n	MHz	$\frac{\pi}{2}\Delta f_s$	37.2	45.7	5.0
electrical efficiency	η_e	—	Eq. 2-14	0.60	0.60	0.60
receiver area	A_R	m^2	given	0.004	0.004	0.004
receiver optical transmission	T_O	—	given	0.5	0.5	0.5
receiver optical bandpass	$\Delta\lambda$	μm	given	0.025	0.060	0.10
receiver field-of-view	Θ_{FOV}	mrad	given	0.5	0.5	0.5
receiver FOV solid angle	Ω_{FOV}	sr	$\frac{\pi}{4}(\Theta_{FOV})^2$	1.96E-07	1.96E-07	1.96E-07
maximum range of counter	R_{max}	m	given	8000	8000	8000
target type	—	—	given	vehicle painted "green drab"		
target hemisphere reflectance	ρ	—	Table 2-2	0.121	0.100	0.016
retroreflectance	f	sr^{-1}	Eq. 2-15	0.039	0.032	0.005
solar spectral irradiance	E_λ	$W/m^2\cdot\mu m$	Volume 1, Chapter 3	607	246	N/A
spectral radiant exitance of 300 K target	M_λ	$W/m^2\cdot\mu m$	Volume 1, Chapter 1	N/A	N/A	30.0
unity gain dc background current	i_B	A	Eq. 2-10	3.44E-11	1.72E-11	2.66E-9
unity gain background noise current	$\left[\overline{I_{BN}^2}\right]^{1/2}$	A	Eq. 2-9	2.02E-11	5.01E-11	6.50E-11
unity gain detector noise current	$\left[\overline{I_{DN}^2}\right]^{1/2}$	A	Eq. 2-12	5.16E-11	3.47E-10	7.90E-10
preamp noise current	$\left[\overline{I_{PN}^2}\right]^{1/2}$	A	given	1.22E-8	1.35E-8	4.45E-9
receiver noise-equivalent irradiance	NEI	W/m^2	Eq. 2-7	9.20E-7	1.43E-6	5.35E-7
atmospheric extinction:	k	1/km	FASCOD2			
clear (23.5 km VMR, midlatitude summer)				0.077	0.055	0.279
thin fog (1.5 km VMR, midlatitude summer)				4.092	2.937	0.827
mass extinction of smoke: white phosphorous (50% RH)	α	m^2/g	Table 2-3	1.41	0.57	0.39
concentration-length	<CL>	g/m^2	given	5	5	5
signal-to-noise ratio	SNR	—	Eq. 2-20	See Figure 2-8		

the beam is always smaller than the target. Given these assumptions, Eq. (2.2) can be simplified to the following equation, which is useful for a variety of fire control applications:

$$\mathrm{SNR} = \frac{P_L \eta_B}{\mathrm{NEI}} \frac{\rho}{\pi} \frac{\exp[-2(kR + \alpha\langle \mathrm{CL}\rangle)]}{R^2} \cdot \square \tag{2.20}$$

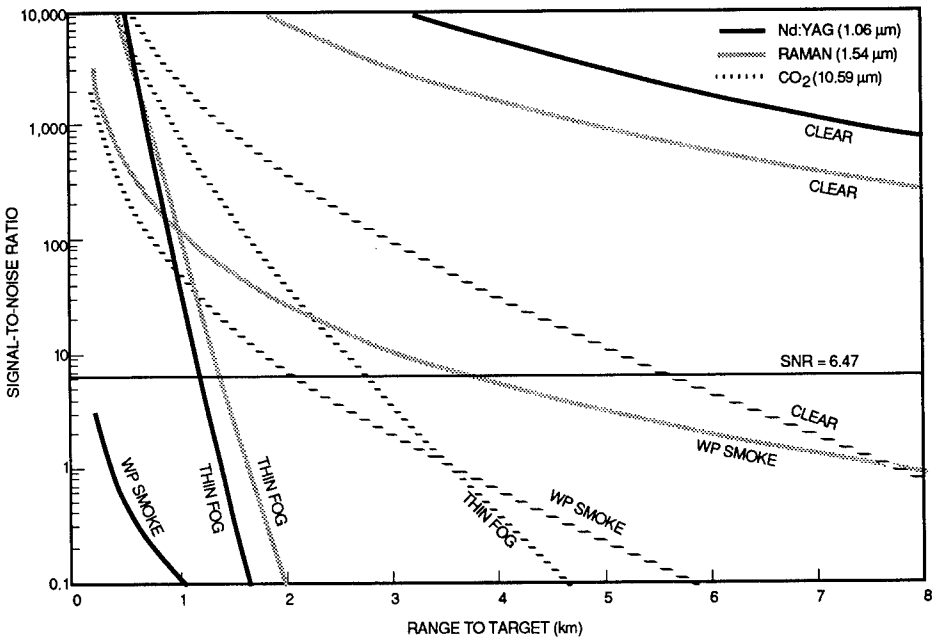


Fig. 2.8 Signal-to-noise ratio versus range.

Figure 2.8 plots the calculated value of the SNR versus range for three atmospheric conditions, clear (23.5-km visibility), thin fog (1.5-km visibility), and white phosphorous (WP) smoke (concentration length product of 5 g/m^2).

2.5 PROBABILITY OF RANGING AND FALSE ALARM RATE

This subsection discusses the relationship between SNR, the probability of ranging, and the false alarm rate. The effects of atmospheric turbulence are discussed as well as the benefits of multiple-pulse detection. Finally, an empirical formulation of the range equation based on the measured extinction ratio is presented.

2.5.1 Pulse Detection in Gaussian Noise

The probability of ranging to a target is essentially the probability that the peak electrical signal plus noise will exceed the detection threshold.[21] When background, detector, and/or preamplifier noise sources dominate the detection process, a Gaussian probability density provides a reasonable approximation to the noise distribution, and the probability of detection is given by

$$P_d = \frac{1}{2} + \frac{1}{2}\text{erf}\left(\frac{\text{SNR} - \text{TNR}}{\sqrt{2}}\right), \tag{2.21}$$

where

P_d = probability of detection
SNR = ratio of peak signal to rms noise
TNR = ratio of detection threshold to rms noise
erf(x) = single-sided error function,

and

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2}\, du \ . \tag{2.22}$$

2.5.2 Error Function

Equation (2.22) is the mathematical definition of the single-sided error function most commonly used.[22] The area under the normal (Gaussian) probability curve, which is tabulated in a variety of probability and statistics textbooks and mathematical handbooks,[22,23] can also be used, however, a change of variables and integration limits may be required to convert the values given in these tabulations to the error function defined in Eq. (2.22).

2.5.3 Laser Pulse Amplitude Jitter

The foregoing analysis assumed that the amplitude of the transmitted laser pulse is time invariant. In general, the laser output energy and beam shape vary from pulse to pulse as a result of many factors including variations in (1) the charge on the pulse forming network; (2) Q-switch timing and hold-off threshold; (3) the gain of the laser medium, which changes with temperature; (4) thermal lensing, wedging, and birefringence in the laser medium; and (5) thermal distortion and structural compliance in the resonator optical bench. Assuming that the distribution of laser radiant intensity is Gaussian, the rms jitter value can be added in quadrature to the rms noise value in calculating the probability of detection. Equation (2.21) therefore becomes

$$P_d = \frac{1}{2} + \frac{1}{2}\text{erf}\left[\frac{\text{SNR} - \text{TNR}}{\sqrt{2}(1 + \sigma_A^2 \cdot \text{SNR}^2)^{1/2}}\right] , \tag{2.23}$$

where σ_A is the rms jitter in the laser radiant intensity (far field) normalized to the average laser radiant intensity.

2.5.4 Threshold-to-Noise Ratio

Laser rangefinders generally detect the signal pulse by comparing the amplitude of the amplified and filtered signal with a threshold setting. The threshold-to-noise ratio (TNR) used in Eq. (2.21) is defined as the ratio of this electrical threshold to the total rms noise at the comparator. The electrical threshold may be fixed, time variable (time-programmable threshold), or adaptive (activity threshold). The time-programmable threshold and activity threshold techniques are discussed later in this subsection.

The value of the TNR is set by the requirement to limit noise-induced false alarms, and is approximated by Rice[24] as

$$\text{TNR} \approx \left[2\ln\left(\frac{2R_{\max}\Delta f_n}{\sqrt{3}cP_{\text{fa}}}\right)\right]^{1/2}, \tag{2.24}$$

where

$R_{\max}$ = maximum range of range counter (m)
Δf_n = noise bandwidth of the receiver (Hz)
P_{fa} = probability of a false alarm during a single pulse.

Equation (2.24) assumes a flat noise spectrum (white noise) and a matched receiver, in which the signal and noise bandwidth are approximately equal to the reciprocal of twice the laser pulse width. The value of $R_{\max}$ used in Eq. (2.24) is the maximum range at which a valid target will be accepted by the range logic. The value of $R_{\max}$ is thus determined by the capacity of the electronic range counter or by the size of the electronic range gate and not by the maximum sensitivity of the rangefinder system.

Repetitively pulsed laser rangefinders used in range- and range-rate-tracking applications (e.g., air defense) are often specified in terms of a long-term average system false alarm rate. In this situation, the single-pulse probability of false alarm is given by

$$P_{\text{fa}} = \frac{\overline{\text{FAR}}}{\overline{\text{PRF}}}, \tag{2.25}$$

where $\overline{\text{FAR}}$ is the long-term average system false alarm rate (false alarms per second) and PRF is the laser pulse repetition frequency (pulses per second, Hz). The long-term average system false alarm rate is not to be confused with the average false alarm rate during a single ranging interval.

Example 2.2. Using the parameters for the Nd:YAG LRF characterized in Table 2.4 and assuming a maximum false alarm rate of 1% (0.01), the required TNR is found from Eq. (2.24) to be 4.83. The probability of detection can now be calculated from Eq. (2.21) for any given SNR. Alternatively, Eq. (2.21) can be solved for the minimum SNR required for any given probability of detection and TNR. The minimum SNR for a 95% (0.95) probability of detection is 6.47. Referring to Fig. 2.8, the probability of detection is at least 95% for all ranges corresponding to SNR values of 6.47 or higher. □

2.5.5 Time-Programmable Threshold and Activity Threshold

The time-programmable threshold technique (sometimes called time-programmable gain) is used in most tactical laser rangefinders to reduce the probability of false alarm caused by (1) electromagnetic interference from the high-voltage flashlamp trigger pulse, (2) optical backscatter from optical elements that are shared by the transmitter and receiver, and (3) optical backscatter from atmospheric aerosols in the region where the transmitter and receiver fields of view overlap. The electrical signals from these false alarm sources decrease rapidly with time after the laser pulse is fired. Because the target return signal also decreases as a function of range (time), as shown in Fig. 2.8, a detection threshold that decreases as a function of time can be

programmed in the receiver electronics to discriminate against these sources of false alarm while maintaining a high probability of detecting both short- and long-range targets.

The activity threshold technique allows the rangefinder to adapt to the average noise environment, providing maximum sensitivity under the prevailing background, backscatter, and electrical noise conditions. The activity threshold is typically implemented as an initial offset to the time-programmable threshold waveform and controlled in a feedback circuit, which maintains a constant false alarm rate at the output of the threshold comparator.

2.5.6 Pulse Detection in Quantum Noise

The foregoing analysis used a Gaussian probability distribution to represent noise in the detection process. This is an accurate approximation only when the signal and noise involve the detection of many photons. When the signal and noise are small, as in the case of a space-based laser rangefinder with a high-gain photomultiplier detector operating against a very low background, the Gaussian approximation is no longer valid, and a Poisson distribution should be used.[21] The probability of ranging, which is the probability that the number of signal-plus-noise photons exceeds a threshold number in a given time interval, is given by

$$P(n \geq n_t) = \sum_{r=n_t}^{\infty} \frac{e^{-\bar{n}}(\bar{n})^r}{r!} = 1 - \sum_{r=0}^{n_t-1} \frac{e^{-\bar{n}}(\bar{n})^r}{r!} , \tag{2.26}$$

where

$P(n \geq n_t)$ = probability of detection
n = number of signal-plus-noise photons within time interval
n_t = threshold number of photons
$\bar{n}$ = average number of noise photons within time interval.

The summation of terms of the Poisson distribution [Eq. (2.26)] is tabulated in a variety of probability and statistics textbooks and mathematical handbooks.[25] Also, the probability of detection is plotted as a function of threshold number and average noise count in Sec. 8 of the *RCA Electro-Optics Handbook*.[21]

Example 2.3. A space-based tracking sensor uses a laser rangefinder to provide fine-range measurements to distant orbiting objects. Passive ranging techniques are used to provide a coarse ranging resolution of ± 1000 m (i.e., at handover, target range is known to be within a specific 2000-m range interval, and range measurements falling outside this interval are presumed to result from false alarms or false targets). The LRF is required to have a range resolution of 1 m. Using Eq. (2.1), this 1-m range resolution corresponds to a time resolution of 6.67 ns, which, for photon counting applications, is the time interval over which signal and noise events are integrated or counted. The average number of events detected by the LRF in the absence of signal during the 6.67-ns time interval is determined to be 3.5 (based on the average number of background photons received, the average number of electronic noise events, and the number of radiation-induced events). The problem is to determine the number of signal photons required to provide a 95% probability of detection with a maximum false alarm probability of 1% per ranging measurement.

The first step is to calculate the probability of false alarm within a single 6.67-ns integration time interval. The 2000-m handover range interval corresponds to a handover time period of 13.3 μs [Eq. (2.1)]. Because there are 2000 of the 6.67-ns integration time intervals within the 13.3-μs handover time period, the probability of false alarm within a single integration interval is $0.01 \times 1/2000 = 5 \times 10^{-6}$.

The second step is to determine the threshold number that gives a probability of exceeding threshold $P(n \geq n_t)$ that is less than or equal to 5×10^{-6} for an average event count $\bar{n}$ of 3.5. Using Eq. (2.26), the threshold number n_t is calculated to be greater than 14 but less than 15. Because the counting circuit can only count integral numbers of events, the next higher integer, which is 15, is selected.

The third step is to determine the number of signal-plus-noise events required to give a probability of detection of 0.95 for a threshold number n_t of 15. Using Eq. (2.26), the average signal-plus-noise events $\bar{n}$ is 21.9.

The final step is to subtract the average number of noise-only events from the signal-plus-noise events to give a required number of signal photons, which is $21.9 - 3.5 = 18.4$. □

2.5.7 Atmospheric Turbulence

The foregoing analysis treated the atmospheric propagation process as time invariant when, in fact, turbulence in the atmosphere can cause significant spatial and temporal fluctuations (scintillation) in the propagated signal. These signal fluctuations result from optical refraction along the propagation path because of thermally induced variations in the refractive index of the air mass. For most tactical laser rangefinder applications, the illuminated portion of the target is much larger than the lateral correlation distance, and scintillation over the transmit path is reduced considerably through spatial averaging. This is not the case, however, for small retroreflective targets such as corner cubes, retroreflective tape, or bicycle-type reflectors (Stimsonite) or where the target reflectance characteristic is dominated by specular reflections caused by glints and trihedral retroreflections. For these targets, scintillation over the transmit path may be significant.

Spatial averaging across the receiver aperture also reduces scintillation over the receive path but to a lesser extent because the receiver aperture may not be large compared with the lateral correlation distance. For typical multimode laser rangefinders, beam-broadening and beam-wander effects resulting from atmospheric turbulence are small compared to the laser beam divergence and can be ignored.

Atmospheric turbulence can be treated as a requirement to increase the SNR of the rangefinder to achieve a given probability of detection. In the weak-scattering regime (log-normal probability distribution), the signal-to-noise enhancement factor required for a specified detection probability is given by Knepp and Valley[26] as

$$\bar{I}/I = \exp\left[\sqrt{2}\sigma_{\ln I/\bar{I}} \operatorname{erf}^{-1}(2P_d - 1) + \frac{1}{2}\sigma^2_{\ln I/\bar{I}}\right], \tag{2.27}$$

where

$\bar{I}/I$ = signal-to-noise enhancement factor
$\sigma^2_{\ln I/\bar{I}}$ = variance of log intensity ratio
P_d = required probability of detection in turbulence
$\text{erf}^{-1}(x)$ = argument of the error function defined in Eq. (2.22).

Example 2.4. The Nd:YAG LRF in example 2.2 is operated across a turbulent atmospheric path. The scintillation index, which is defined as the standard deviation of the log intensity ratio $\sigma_{\ln I/\bar{I}}$, is found to be 0.6 at maximum target range. By what factor must the SNR (calculated for no turbulence) be increased to achieve a 99% probability of detection?

The minimum SNR with no turbulence is calculated using Eq. (2.21), as in example 2.2. For $P_d = 0.99$ and TNR = 4.83, the minimum SNR is found to be 7.15.

The SNR enhancement factor is calculated using Eq. (2.27). For $\sigma_{\ln I/\bar{I}} = 0.6$ and $P_d = 0.99$, $\bar{I}/I$ is found to be 4.84. The minimum SNR of 7.15 must therefore be increased to 34.6 to provide the same ranging probability. □

2.5.8 Cumulative Probability of Detection

For laser rangefinder applications in which the propagation properties of the atmosphere or the reflective properties of the target are expected to fluctuate significantly (e.g., strong atmospheric turbulence, inhomogeneous smoke screens, or specular reflections) the probability of detection can be improved more dramatically by firing multiple pulses than by firing a single pulse with higher energy. The cumulative probability of detecting at least m-out-of-n pulses, assuming random independent events, is given by the binomial equation as

$$P(x \geq m) = \sum_{k=m}^{n} \binom{n}{k} (P_d)^k (1 - P_d)^{n-k}$$

$$= 1 - \sum_{k=0}^{m-1} \binom{n}{k} (P_d)^k (1 - P_d)^{n-k} , \tag{2.28}$$

where $P(x \geq m)$ is the probability of detecting at least m pulses and P_d is the single-shot probability of detection. For one-out-of-n logic, Eq. (2.26) simplifies to

$$P(x \geq 1) = 1 - (1 - P_d)^n . \tag{2.29}$$

Example 2.5. For a high-pulse-rate Nd:YAG LRF operating against the same maximum range target and across the same turbulent atmosphere given in example 2.4, would it be more advantageous in terms of SNR to operate the LRF at twice the pulse rate and half the pulse energy, maintaining the same average laser power?

Operating the LRF at twice the pulse rate allows the processor to use one-out-of-two detection logic to achieve a cumulative probability of detection of 99% without affecting the average ranging rate. The required single-shot probability of detection is calculated by solving Eq. (2.29) for P_d. For $P(X \geq 1) = 0.99$ and $n = 2$, P_d is calculated to be 0.90.

The TNR calculated in example 2.2 was based on a single-shot false alarm probability of 1%, which is no longer valid at twice the pulse rate. Equation (2.25) states that to maintain a given long-term average system false alarm rate, $\overline{\text{FAR}}$, the single-shot false alarm probability P_{fa} must be reduced in inverse proportion to the increase in the laser pulse repetition frequency, PRF. Doubling the PRF necessitates a factor of 2 reduction in P_{fa} from 0.01 to 0.005. Using Eq. (2.24) with this P_{fa} and the other parameters specified in Table 2.4, the required TNR is calculated to be 4.97.

The SNR required for a 90% detection probability in no turbulence is found from Eq. (2.21) to be 6.25. The SNR enhancement factor for a scintillation index of 0.6 is found from Eq. (2.27) to be 2.58. The required SNR with turbulence is therefore $6.25 \times 2.58 = 16.15$. Because this required SNR is less than half the SNR requirement calculated in example 2.4, it is more advantageous in this application to operate the LRF at twice the pulse rate with half the pulse energy. □

2.5.9 Extinction Ratio

The sensitivity of a laser rangefinder is often specified empirically, in terms of an extinction ratio. The extinction ratio (ER) is defined as the amount of optical attenuation placed in the receiver path of the rangefinder that gives a 50% probability of detection when operating the rangefinder over a specified range to a large, normally incident target board of specified reflectance. If the receiver uses a time-programmable threshold, this circuit is defeated for the extinction ratio measurement and the threshold is set for maximum sensitivity. The SNR for an arbitrary set of range, atmospheric attenuation, and target parameters is given in terms of this extinction ratio by

$$\text{SNR} = \frac{\text{TNR}}{\text{ER}} \frac{\rho}{\rho_0} \left(\frac{R_0}{R}\right)^2 \exp[-2(kR + \alpha\langle \text{CL} \rangle)] \ , \tag{2.30}$$

where

ER = transmission of extinction filters
ρ_0 = hemispherical reflectivity of reference target board
R_0 = range to reference target board (m).

Atmospheric transmission loss across the specified range is generally included in the extinction ratio specification and therefore does not appear explicitly in the SNR calculation [Eq. (2.30)].

Example 2.6. An extinction ratio measurement is made to characterize the sensitivity of a Nd:YAG laser rangefinder prior to shipment from the factory. The detection threshold is set at the factory to provide a TNR of 4.83 (with time-programmable threshold circuit disabled). The laser range used for this measurement has a large target board with a hemispherical reflectance of 85% at 1.06 μm, which is located precisely 500 m from the LRF test station. Attenuation filters are placed within the receiver path of the LRF such that the attenuation of the filter stack corresponds with the two-way attenuation in the beam path. With 44.5-dB attenuation in the filter stack, half the range measurements provided valid ranges with the remaining measurements failing

to register a return (P_d = 50%). What will the SNR of this LRF be against a large 10% reflectance target at 3 km when the atmospheric attenuation is 3 dB/km at 1.06 μm (assume no turbulence).

The first step is to convert atmospheric attenuation to an atmospheric extinction coefficient. Multiplying the 3×10^{-3} dB/m value by the conversion factor 0.2303 gives $k = 6.91 \times 10^{-4}$ m^{-1}. The second step is to convert the 44.5-dB attenuation of the filter stack to an extinction ratio (transmission of filters), ER = 3.55×10^{-5}. Finally, the SNR is calculated using Eq. (2.30) to be 6.95. □

2.6 RANGING ACCURACY

2.6.1 Range Errors

The range to the target measured by a pulse-echo laser rangefinder may differ from the actual range as a result of the following four error sources: (1) quantization error in the range counter, (2) timing errors in the logic circuitry, (3) detection error resulting from the finite rise time of the laser pulse, and (4) clock frequency error in the range counter. The first three error sources are random and independent, therefore the total rms range error can be found by adding these errors in quadrature (root-sum-square). The clock frequency error is not random and is treated as a bias or calibration error.

Quantization error results from the fact that digital circuits are used to measure the time interval between transmission of the laser pulse and detection of the reflected pulse from the target. Most digital range counters use a free-running oscillator (or clock) to establish the time base, and the clock has a finite time resolution set by its oscillation frequency. The least significant bit within the digital range counter corresponds to the period of oscillation of the clock (or half the clock period if the counter triggers on both positive and negative transitions in the clock waveform). Laser pulsing is typically asynchronous with the clock, therefore the first recorded range count will be the next clock transition that occurs after the laser is fired. This introduces a timing error that can assume any value between zero and the clock period with equal probability. The rms range error resulting from start-time quantization is given by

$$\sigma_{RQ(start)} = \frac{c}{4\sqrt{3}f_c}, \tag{2.31}$$

where $\sigma_{RQ(start)}$ is the rms range error resulting from start-time quantization (m) and f_c is the frequency of the oscillator in the range counter (Hz).

A similar quantization error arises when the pulse is detected. The last range count will be the clock transition that occurs just prior to detecting the reflected pulse, and the range error resulting from stop-time quantization is given by

$$\sigma_{RQ(stop)} = \frac{c}{4\sqrt{3}f_c}, \tag{2.32}$$

where $\sigma_{RQ(stop)}$ is the rms range error resulting from stop-time quantization (m).

Timing errors result from switching-time jitter in the logic gates that are used to turn the range counter on and off. Timing error is therefore a function of the type of logic gates used in the range logic circuitry. An rms timing error of 1 ns per event, which is an upper limit for most tactical rangefinder logic circuits, corresponds to an rms range error of 0.15 m for either the turn-on or turn-off event.

Detection error in simple threshold detection circuits results from the finite rise time of the laser pulse (and pulse stretching by the target). The finite rise time causes pulses of higher amplitude to cross the detection threshold before those of lower amplitude, thus indicating a shorter range to the target. The range error is therefore a deterministic function of the SNR (signal amplitude), the threshold value, and the temporal shape of the laser pulse. However, because pulse amplitude is not measured by the rangefinder, it cannot be biased out and must be handled as a random error source. Assuming a linear rise-time characteristic and uniform probability of threshold crossing along this rising pulse shape, the range error resulting from detection error is given by

$$\sigma_{RD} = \frac{ct_R}{4\sqrt{3}} , \tag{2.33}$$

where σ_{RD} is the rms range error due to detection error (m) and t_R is the rise time of the laser pulse (s).

Advanced signal processing techniques are used in some LRF systems to reduce detection error. Constant fraction discrimination, for example, subtracts a delayed replica of the return pulse from itself, which provides a zero-crossing-time reference that is independent of pulse amplitude and, to a first order, pulse shape. Such techniques can be very effective but are more complex and costly to implement.

Clock frequency error (bias) results from frequency drift within the oscillator circuit of the digital clock. Because the timing error resulting from clock frequency error increases linearly with time, the range error will be a linear function of target range. Clock frequency error is normally specified as a fraction of the absolute clock frequency. The range error (bias) resulting from clock frequency error is given by

$$b_{RC} = Rb_C , \tag{2.34}$$

where

b_{RC} = range error (bias) resulting from clock frequency error (m)
R = range to target (worst case is maximum range) (m)
b_C = clock frequency error (fraction of clock frequency).

2.6.2 Range-Rate Errors

Tracking applications often require range rate (radial velocity) as well as range information. For direct detection laser rangefinders, range rate is calculated

within the fire control computer from differential range measurements to the target. The range rate (for a one-sided derivative) is given by

$$\dot{R}_n = (R_n - R_{n-1})\text{PRF} \ , \tag{2.35}$$

where

$\dot{R}_i$ = i'th value of range rate (m/s)
R_i = i'th value of range to target (m)
PRF = laser pulse repetition frequency (Hz).

For the general case in which the laser pulse rate is not synchronous with the range counter clock, the distribution of range-rate error resulting from start- and stop-time quantization error is found through simulation to approximate a Gaussian distribution with a mean of zero and a standard deviation (rms value) given by

$$\sigma_V = 0.291 \frac{c\text{PRF}}{f_c} \ , \tag{2.36}$$

where σ_V is the rms range-rate error (m/s).

When the laser pulse rate is synchronous with the range counter clock (which is difficult to achieve because of laser-trigger-timing errors), the form of the range-rate error distribution is triangular with a mean of zero and a standard deviation given by

$$\sigma_V = \frac{1}{2\sqrt{3}} \frac{c\text{PRF}}{f_c} \ . \tag{2.37}$$

The range-rate data derived from differential range measurements are often smoothed using weighted-average filtering techniques prior to tracking to reduce noise in the tracking process. The penalty for filtering is a lag in the response of the tracker to changes in target radial velocity. This tracking lag can be reduced by increasing the sample rate, which is equivalent to increasing the laser pulse repetition frequency. A laser rangefinder with improved range accuracy requires less filtering and therefore can operate at a lower pulse repetition frequency.

Example 2.7. An air defense laser rangefinder operates at a pulse rate of 10 pulses/s, which is asynchronous with the internal range counter clock operating at 40 MHz. What are the range and range-rate errors, assuming the only error source is start- and stop-time quantization?

The start- and stop-time rms range errors are identical and are calculated using Eqs. (2.31) or (2.32) to be $\sigma_{\text{RQ(start/stop)}} = 1.08$ m. The total rms range error is the root-sum-square of the start- and stop-time errors, which is 1.53 m. The instantaneous rms range-rate error (no filtering) is calculated using Eq. (2.36) to be $\sigma_V = 21.8$ m/s. □

2.7 CHARACTERISTICS OF COMMON LASER RANGEFINDERS

Table 2.5 provides a range of typical parameters for ruby, Nd:YAG, CO_2, Raman-shifted Nd:YAG, and Er:glass laser rangefinders.

The advantages of the Nd:YAG laser rangefinder over the ruby laser are its covertness, its improved electrical efficiency, and its ability to operate at elevated pulse repetition frequencies (an essential requirement for the air defense application and a desirable feature in tank fire control for helicopter self-defense). The major drawback of the Nd:YAG laser lies in its wavelength of operation. The 1.06-μm wavelength of the Nd:YAG laser is not strictly compatible with that of the 8- to 12-μm thermal imaging systems used in modern fire control systems under conditions of natural haze, fog, and battlefield smoke. Also, the 1.06-μm wavelength is transmitted and focused by the human eye onto the retina where, at sufficiently short ranges, it can permanently blind an unprotected observer.[1]

Three solutions have been developed that address the drawbacks of the 1.06-μm Nd:YAG LRF; these are, in order of development chronology, (1) carbon dioxide (CO_2), (2) Raman-shifted Nd:YAG, and (3) erbium:glass. The CO_2 laser operates at 10.59 μm, which is within the 8- to 12-μm wavelength band of typical thermal imaging sights. This allows the CO_2 LRF to range to most of the targets that the thermal imager can detect, which is its chief advantage. The CO_2 wavelength is attenuated by water vapor in the atmosphere, thereby limiting maximum range performance under clear, humid conditions. Tactical targets provide a lower reflectivity at the CO_2 wavelength than at shorter wavelengths, a problem that is compounded when the targets are wet or snow covered, because under these conditions the reflection is more specular (ge-

Table 2.5 Characteristics of Common Laser Rangefinders

Characteristic	Ruby	Nd:YAG	CO_2	Raman-Shifted Nd:YAG	Er:Glass
Wavelength (μm)	0.6943	1.064	10.59	1.543	1.54
Laser Type	Solid-state (Crystal)	Solid-state (Crystal)	Gas	Solid-state W/Conversion	Solid-state
Excitation	Flashlamp	Flashlamp or Laser Diode	Electric Discharge	Flashlamp or Laser Diode	Flashlamp or Laser Diode
Pulse Energy (J)	0.01-50	0.01-10	0.01-100	0.005-0.3	0.01-0.1
Pulsewidth (ns)	2-50	8-20	20-100	6-15	20-30
Pulse Repetition Frequency (Hz):					
Uncooled	< 0.1	< 1	< 1	< 1	< 0.3
Cooled	≤ 10	> 100	> 100	> 100	< 10
Typical Eyesafety Classification	Class IV Non-eyesafe	Class IIIb to Class IV Non-eyesafe	Essentially Class IIIa Eyesafe	Class I to Class IIIa Eyesafe	Class I to Class IIIa Eyesafe
Detector Type	Si APD	Si APD	Cooled PV HgCdTe	InGaAs APD or PIN	InGaAs APD or PIN
NEP (nW)	~ 1	~ 2	~ 1	2-15	2-15

ometry dependent). The CO_2 wavelength may be severely attenuated by falling snow conditions, which do not significantly degrade the target detection and fire control capability of the direct-view and TV sights or significantly reduce the range performance of the shorter wavelength LRFs (Ref. 27).

Functional Raman-shifted Nd:YAG laser rangefinders were developed for military applications beginning in the late 1970s to provide an inexpensive (relative to CO_2) eyesafe laser ranging capability for infantry, artillery observers, air defense, and tank fire control.[4] Stimulated Raman scattering (SRS) in methane gas shifts the 1.06-μm Nd:YAG wavelength to 1.54 μm, which is absorbed and dissipated in the aqueous portion of the eye before reaching the retina.

The erbium-doped glass (Er:glass) laser is also used in laser rangefinders to provide the 1.54-μm eyesafe output. Rather than shifting a 1.06-μm laser with SRS, this laser uses a flashlamp-pumped Er:glass rod that lases directly at 1.54 μm (Ref. 28). The pulse repetition frequency capability and flashlamp lifetime of the Er:glass laser are limited, thus making it unsuitable for high PRF air defense applications.[2]

References

1. R. M. Ogorkiewicz, "Eye-safe neodymium lasers," *International Defense Review* **2,** 174–175 (1990).
2. S. Guch, "Laser range finders find new applications," *Laser Focus World,* pp. 130–133 (June 1990).
3. R. C. Renairi and A. M. Johnson, "MELIOS—status report of the U.S. Army's eyesafe laser rangefinder program," *Proceedings of the SPIE* **1207,** 112–119 (1990).
4. R. W. Nichols and W. K. Ng, "Raman-shifted ND:YAG class I eye-safe laser development," *Proceedings of the SPIE* **610,** 92 (1986).
5. S. Feinstein, "The evolution of the tank fire control system," *Armor,* pp. 21–24 (Jan./Feb. 1973).
6. R. Fletcher, "Tank guns against the attack helicopter," *International Defense Review* **6,** 637–640 (July 1988).
7. E. Gergor, D. E. Nieuwsma, and R. D. Stultz, "20 Hz eyesafe laser rangefinder for air defense," *Proceedings of the SPIE* **1207,** 124 (1990).
8. D. G. Bruns, H. W. Bruesselbach, and D. A. Rockwell, "A Nd:YAG-based eyesafe Raman laser," *Proceedings of the International Conference on Lasers '80,* p. 406, STS Press, McLean, VA (1981).
9. M. A. Woodall, J. R. Minch, J. Nunez, H. S. Keeter, and A. M. Johnson, "Eyesafe laser cloud mapper," *Proceedings of the SPIE* **1207,** 100–111 (1990).
10. R. J. McIntyre, "Multiplication noise in uniform avalanche diodes," *IEEE Transactions on Electron Devices* **ED-13**(1), 164–168 (1966).
11. P. P. Webb, R. J. McIntyre, and J. Conradi, "Properties of avalanche photodiodes," *RCA Review* **35**(2), 234–278 (1974).
12. *Planar InGaAs APD C30662 Series, Data Sheet,* EG&G Canada Ltd., Optoelectronics Division, Vaudreuil, Quebec, Canada (1988).
13. G. A. Helgeson and R. T. Mousel, *Reflectance Data Reference Manual,* General Image Engineering Corp. Report No. GIE-TR-82-002, USAF Wright Aeronautical Labs, Wright-Patterson AFB, Ohio (Nov. 1982).
14. *Electro-Optics Handbook,* pp. 217–224, RCA Corp., Lancaster, PA (1974).
15. S. A. Clough, F. X. Kneizys, E. P. Shettle, and G. P. Anderson, "Atmospheric radiance and transmittance: FASCOD2," in *Sixth Conference on Atmospheric Radiation,* pp. 141–144, American Meteorological Society, Boston (1986).
16. R. E. Davis, "EOSAEL 87, Vol. 5, Laser transmittance model (LZTRAN)," Report No. TR-0021-5,

U.S. Army Atmospheric Sciences Laboratory, White Sands Missile Range, NM (Oct. 1987). This document is classified unclassified/limited.

17. *Smoke and Natural Aerosol Parameters (SNAP) Manual,* p. 2-2, Report No. 61 JTCG/ME-85-2, Joint Technical Coordinating Group for Munitions Effectiveness, Smoke and Aerosol Working Group, U.S. Army Material Systems Analysis Activity, Aberdeen Proving Ground, MD (1985). This document is classified unclassified/limited.
18. D. W. Hoock, R. A. Sutherland, and D. Clayton, "EOSAEL 87, Vol. 11, Combined obscuration model for battlefield-induced contaminants (COMBIC)," p. 16, Report No. ASL TR-0221-11, U.S. Army Atmospheric Sciences Laboratory, White Sands Missile Range, NM (Oct. 1987). This document is classified unclassified/limited.
19. "Quantitative description of obscuration factors for electro-optical and millimeter wave systems," pp. 2-5–2-6, Military Handbook DOD-HDBK-178(ER) (July 25, 1986).
20. M. E. Milham and D. H. Anderson, "Obscuration sciences smoke data compendium: standard smokes," Government Accession No. AD-B077078, Report No. ARCSL-SP-82024, U.S. Army Chemical Systems Laboratory, Aberdeen Proving Ground, MD (1983). This document is classified unclassified/limited.
21. *Electro-Optics Handbook,* pp. 109–116, RCA Corp., Lancaster, PA (1974).
22. M. R. Spiegel, *Schaum's Outline Series: Mathematical Handbook of Formulas and Tables,* pp. 183, 257, McGraw-Hill, New York (1968).
23. W. H. Beyer, *CRC Standard Mathematical Tables,* 25th ed., p. 524, CRC Press, West Palm Beach, FL (1978).
24. S. O. Rice, "Mathematical analysis of random noise," *Bell System Technical Journal* **23,** 282–332 (1944); **24,** 46–156 (1945).
25. W. H. Beyer, *CRC Standard Mathematical Tables,* 25th ed., p. 530, CRC Press, West Palm Beach, FL (1978).
26. D. L. Knepp and G. C. Valley, "Properties of joint Gaussian statistics," *Radio Science* **13,** 59–68 (1978).
27. B. W. Kennedy, B. A. Locke, W. M. Farmer, W. G. Klimek, and F. E. Perron, "Joint U.S.-Canadian obscuration analysis for smokes in snow (Smoke Week XI), Final Report: Part I," Report No. SMK-001-90, Project Manager, Smoke/Obscurants, Aberdeen Proving Ground, MD (April 1990). This document is classified unclassified/limited.
28. R. Pengelley, "OEC's eye-safe laser option," *International Defense Review* **2,** 176–177 (1990).

Bibliography

Laser Rangefinder Systems

Cooke, C., J. Cernius, and A. J. LaRocca, "Ranging, communications, and simulation systems," Chap. 23, *The Infrared Handbook,* pp. 23-3–23-23, Environmental Research Institute of Michigan, Ann Arbor, MI (1978).

Electro-Optics Handbook, RCA Corporation, Lancaster, PA (1974).

Torberg, H. E., W. J. Rowan, and J. R. Vyce, "Optical instruments for metrology," Sec. 16 in *Handbook of Optics,* pp. 16-49–16-50, Optical Society of America, McGraw-Hill, New York (1978).

Target Reflectivity

Barr, D. N., and M. Savan, "Laser radar for wire detection," in *Proceedings of the 1988 Meeting of the IRIS Specialty Group on Active Systems,* Vol. 2, pp. 11–42, Report No. WRDC-TR-89-1117, Defense Technical Information Center, Alexandria, VA (June 1989).

Bartell, F. O., E. L. Dereniak, and W. L. Wolfe, "The theory and measurement of bidirectional reflectance distribution function (BRDF) and bidirectional transmittance distribution function (BTDF)," *Proceedings of the SPIE,* **257,** 154–160 (1980).

Barthel, K., and W. Büchtemann, "Reflexionseigenschaften natürlicher und künstlicher Materialien bei $\lambda = 10.6$ μm," Report FfO 1980/52, Forschungsinstitut für Optik, Forschungsgesellschaft für Angewandte Naturwissenschaften e.V., Tübingen, West Germany (1980). This document is classified NATO restricted.

Durand, W. W., and P. B. Zimmer, "E-O camouflage coatings," Technical Report No. AFWAL-TR-80-1142, USAF Contract No. F33615-77-C-1144, U.S. Air Force Avionics Laboratory, Wright-Patterson AFB, OH (Apr. 1980). This document is classified confidential.

Hayes, C. L., and R. A. Brandewie, "Reflection coefficients for wires and cable at 10.6 μm," *Applied Optics,* **12**(7), 1564–1569 (1973).

Helgeson, G. A., and R. T. Mousel, "Reflectance data reference manual," General Image Engineering Corp. Report No. GIE-TR-82-002, U.S. Air Force Wright Aeronautical Laboratories, Wright-Patterson AFB, OH (Nov. 1982).

Huffman, P. J., "Measurement of retroreflectance from paint coatings," Report No. AFAL-TR-78-181, U.S. Air Force Avionics Laboratory, Wright-Patterson AFB, OH (June 1978). This document is classified confidential.

Reynolds, G. O., "An evaluation of laser reflectance from dusty and wetted surfaces," Contract No. DAAK70-79-D-0036, Arthur D. Little, Inc. (July 1981).

"Target signature analysis center: data compilation eleventh supplement, volume 1, bidirectional reflectance: definition, discussion, and utilization," Report No. AFAL-TR-72-226, U.S. Air Force Avionics Laboratory, Wright-Patterson AFB, OH (Oct. 1972). This report is classified confidential.

"Target signature analysis center: data compilation eleventh supplement, volume 2, bidirectional reflectance: graphic data," Report No. AFAL-TR-72-226, U.S. Air Force Avionics Laboratory, Wright-Patterson AFB, OH (Oct. 1972). This report is classified confidential.

"Target signatures study interim report, volume V: catalog of spectral reflectance data," Report No. 5698-22-T(V), The University of Michigan, Ann Arbor, MI (Oct. 1964).

Venable, W. H., Jr., H. F. Stephenson, and H. Terstiege, "Factors affecting the metrology of retroreflecting materials," *Applied Optics* **19**(8), 1236–1241 (1980).

Laser Propagation and Obscurants

Burlbaw, E. J., Y. K. Behl, and A. Deepak, "Analysis of contributions of soil-derived dust to mass extinction in the visible and infrared (U)," Government Accession No. AD-B076601, Technical Report No. CR-83-0100-1, U.S. Army Atmospheric Sciences Laboratory, White Sands Missile Range, NM (May 1983). This document is classified unclassified/limited.

Carlon, H. R., "Contributions of particle absorption to mass extinction coefficients (0.55–14 μm) of soil-derived atmospheric dusts," Government Accession No. AD-A203377, Technical Report CRDEC-TR-005, U.S. Army Chemical Research, Development & Engineering Center, Aberdeen Proving Ground, MD (Dec. 1988).

Duncan, L. D., R. C. Shirkey, and M. B. Richardson, Eds., "EOSAEL 82 volume III—Transmission through battlefield aerosols," Report No. ASL-TR-0122, U.S. Army Atmospheric Sciences Laboratory, White Sands Missile Range, NM (1982).

Englehardt, R. E., and G. W. Knebel, "Characteristics of the dust environment in the vicinity of military activities," Government Accession No. AD-665439, Report No. AR-642, U.S. Army Mobility Equipment Research and Development Center, Ft. Belvoir, VA (Jan. 1968).

Evans, A. J., and G. Roy, "Determination of the mass concentration and extinction efficiency of some aerosols in the DREV aerosol test facility (U)," Government Accession No. AD-C952024, Report No. DREV-4279/82, Defense Research Establishment Valcartier, Quebec, Canada (Nov. 1982). This document is classified Canada restricted.

Farmer, W. M., B. W. Kennedy, B. A. Locke, and W. G. Klimek, "Evaluation of electro-optical system and model performance in smoke/obscurant environments (Smoke Week X), final report: Part I," Report No. AMCPM-SMK-CT-004-89, Project Manager, Smoke/Obscurants, Aberdeen Proving Ground, MD (Nov. 1989).

Flanigan, D. F., and H. P. DeLong, "Spectral absorption characteristics of the major components of dust clouds," *Applied Optics* **10**(1), 51–57 (1971).

Gomez, R. B., D. W. Hoock, and R. A. Sutherland, "Objective summary of the US Army electro-optical modeling and field testing in an obscuring environment," *Optical Engineering* **22**(1), 2–19 (1983).

Heyl, M. J., "Obscuration sciences smoke data compendium: candidate smokes (U)," Government Accession No. AD-C031896, Report No. ARCSL-SP-82017, U.S. Army Chemical Systems Laboratory, Aberdeen Proving Ground, MD (June 1983). This document is classified confidential.

Heyl, M. J., "Obscuration sciences smoke data compendium: experimental smokes (U)," Government Accession No. AD-C033835, Report No. ARCSL-SP-82020, U.S. Army Chemical Systems Laboratory, Aberdeen Proving Ground, MD (1983). This document is classified confidential.

Hoock, D. W., R. A. Sutherland, and D. Clayton, "EOSAEL 87, volume 11—combined obscuration model for battlefield-induced contaminants, COMBIC," Report No. ASL-TR-0221-11, U.S. Army Atmospheric Sciences Laboratory, White Sands Missile Range, NM (Oct. 1987).

Hoock, D. W., and R. A. Sutherland, "Modeling the complex dirty battlefield," U.S. Army Atmospheric Sciences Laboratory, White Sands Missile Range, NM (1983).

Johnson, M., "Joint technical coordinating group for munition expenditures report," U.S. Army Munitions Command, Operations Research Group, Edgewood Arsenal, MD (1972).

Kayes, M. D., M. A. Seagraves, H. H. Monahan, and R. A. Sutherland, "Qualitative description of obscuration factors in Central Europe," ASL Monograph No. 4, U.S. Army Atmospheric Sciences Laboratory, White Sands Missile Range, NM (1980).

Kennedy, B. W., B. A. Locke, W. M. Farmer, W. G. Klimek, and F. E. Perron, Jr., "Joint U.S.-Canadian obscuration analysis for smokes in snow (Smoke Week XI), final report: Part I," Report No. SMK-001-90, Project Manager, Smoke/Obscurants, Aberdeen Proving Ground, MD (Apr. 1990). This document is classified unclassified/limited.

Kennedy, B. W., B. A. Locke, W. G. Klimek, and R. Laughman, "Large-area and self-screening smokes and obscurants (LASSO) (Smoke Week XII) quick look report," STC Report No. 4056, Science and Technology Corp., 101 Research Drive, Hampton, VA (Aug. 1990).

Klett, J. D., and R. A. Sutherland, "Effects of atmospheric moisture in determining obscurant properties from field data," in *Proceedings of the Ninth Annual EOSAEL/TWI Conference,* pp. 523–532, Physical Science Laboratory, Las Cruses, NM (Dec. 1987).

Locke, B., W. M. Farmer, B. W. Kennedy, and W. G. Klimek, "Joint United States-Canadian obscuration analysis for smokes in snow (Smoke Week XI) quick look report," STC Report No. 3053, Science and Technology Corp., 101 Research Drive, Hampton, VA (1989).

Maddix, M. V., R. E. Alongi, B. L. Williams, W. Greenleaf, and S. Brazelton, "A survey of extinction coefficients available for various transient aerosols and selection of extinction coefficients for BELDWSS modelling of transient aerosols," Special Report No. RG-82-1, U.S. Army Missile Command, Redstone Arsenal, AL (1982).

Matise, B. K., S. T. Hanley, B. L. Bean, W. R. Watkins, and K. O. White, "BICT-III HE dust data reduction data report," Report No. ASL-DR-84-0001, U.S. Army Atmospheric Sciences Laboratory, White Sands Missile Range, NM (1983).

Matise, B. K., "Effects of a cold environment and snow on obscuration," Interim Contractor Report, Optometrics, Inc., U.S. Army Atmospheric Sciences Laboratory, White Sands Missile Range, NM (1984).

Milham, M. E., and D. H. Anderson, "Obscuration sciences smoke data compendium: standard smokes," Government Accession No. AD-B077078, Report No. ARCSL-SP-82024, U.S. Army Chemical Systems Laboratory, Aberdeen Proving Ground, MD (1983). This document is classified unclassified/limited.

Obert, L. P., J. T. Wood, and C. J. Nash, *Visionics Handbook of EO Sensor Performance, Volume 1, Natural European Environments* (U), U.S. Army Night Vision and Electro-Optics Laboratory, Ft. Belvoir, VA (June 1981). This document is classified confidential.

Pena, R., S. L. Cohn, and F. V. Hansen, "The KWIK smoke munition expenditures model," Government Accession No. AD-B105187, U.S. Army Atmospheric Sciences Laboratory, White Sands Missile Range, NM (July 1986).

Quantitative Description of Obscuration Factors for Electro-Optical and Millimeter Wave Systems, Military Handbook DOD-HDBK-178(ER) (July 25, 1986).

Smith, M. D., and W. W. Jones, "Characteristics of filtered smoke data with associated empirical relationships," in *Proceedings of the Smoke/Obscurants Symposium VII,* Vol. 1, Project Manager, Smoke/Obscurants, Aberdeen Proving Ground, MD (1983).

Smoke and Natural Aerosol Parameters (SNAP) Manual, Report No. 61 JTCG/ME-85-2, Joint Technical Coordinating Group for Munitions Effectiveness, Smoke and Aerosol Working Group, U.S. Army Material Systems Analysis Activity, Aberdeen Proving Ground, MD (1985).

Sutherland, R. A., "Determination and use of the Hanel growth factors for modeling hygroscopic aerosols," ASL Internal Report, U.S. Army Atmospheric Sciences Laboratory, White Sands Missile Range, NM (Jan. 1982).

Turetsky, A. L., D. Anderson, M. Milham, and E. Peterson, "Obscuration sciences smoke data compendium: foreign smokes (U)," Report No. ARCSL-SP-82023, U.S. Army Chemical Systems Laboratory, Aberdeen Proving Ground, MD (1983). This document is classified confidential.

Turner, R. E., P. G. Eitner, C. D. Leonard, and D. G. S. Snyder, *Battlefield Environment Obscuration Handbook, Volume I,* Government Accession No. AD-A102822, U.S. Army Materiel Systems Analysis Activity, Aberdeen Proving Ground, MD (Dec. 1980).

Turner, R. E., P. G. Eitner, C. D. Leonard, and D. G. S. Snyder, *Battlefield Environment Obscuration Handbook, Volume II (U),* Government Accession No. AD-C025742, U.S. Army Materiel Systems Analysis Activity, Aberdeen Proving Ground, MD (Dec. 1980). This document is classified confidential.

Vervier, J., "Smoke obscurant technology (U)," in *Proceedings of the Smoke Symposium III,* Program Manager, Smoke/Obscurants, Aberdeen Proving Ground, MD (1979). This document is classified confidential.

Laser Safety

ANSI, Safe Use of Lasers, ANSI Z-136.1-1986, American National Standards Institute, Laser Institute of America, Toledo, OH (1986).

"Carbon-dioxide laser rangefinder for the M-1 tank," Nonionizing Radiation Protection Survey No. 25-42-0350-83, U.S. Army Environmental Hygiene Agency, Aberdeen Proving Ground, MD (May 1983).

"Control of hazards to health from laser radiation" U.S. Army Technical Bulletin TB-MED-524, Headquarters, Dept. of Army (June 1985).

"Control of health hazards from lasers and other high intensity optical sources," AR 40–46 (Feb. 6, 1974) (with Change 1, dated Nov. 15, 1978).

"Exemption for military laser products," Department of Defense Instruction 6050.6 (May 1, 1978).

"Evaluation and control of laser hazards, standardization agreement," Document No. STANAG 3606 LAS (Edition No. 4), North Atlantic Treaty Organization, Military Agency for Standardization (Feb. 10, 1983).

"Laser products; amendments to performance standard; final rule, *Federal Register* **50**(161) (1985).

Lund, D. J., M. B. Landers, G. H. Bresnick, J. O. Powell, J. E. Chester, and C. Carver, "Ocular hazards of the *Q*-switched erbium laser," *Investigative Ophthalmology* **9**(6), 463–470 (1970).

Performance Standards for Light-Emitting Products, Title 21, Code of Federal Regulations (CFR), Part 1040 (1982 revision).

"Potential eye and skin hazards with the Hughes eye-safe air defense (HEAD) laser," Nonionizing Radiation Protection Study No. 25-42-0332-90, U.S. Army Environmental Hygiene Agency, Aberdeen Proving Ground, MD (Mar. 1990).

"Potential ocular hazards associated with the use of the KEI Corporation modified LH-83 handheld laser rangefinder," Nonionizing Radiation Protection Survey No. 25-42-0301-87, U.S. Army Environmental Hygiene Agency, Aberdeen Proving Ground, MD (Mar. 1987).

"Radiation control for Health and Safety Act of 1968," Public Law 90-602 (Oct. 18, 1968).

"Radiation safety of laser products, equipment classification, and user's guide," Document No. WS 825, International Electrotechnical Commission, Geneva, Switzerland (1984).

"Regulations for the administration and enforcement of the Radiation Control for Health and Safety Act of 1968," U.S. Department of Health and Human Services, 21 CFR Subchapter J, revised (Apr. 1985).

"Safety design requirements for military lasers and associated support equipment," MIL-STD-1425 (Dec. 13, 1985).

"Safety requirements for military lasers," AR 385-9 (Apr. 1, 1982).

Sliney, D. H., "Radiation safety. The maximum permissible exposure levels: our knowledge of the hazards," *Optics and Laser Technology* **21**(4) (1989).

CHAPTER 3

Millimeter-Wave Radar

Elmer L. Johansen
Environmental Research Institute of Michigan
Ann Arbor, Michigan

CONTENTS

3.1 INTRODUCTION

Although the IEEE standard defines the millimeter-wave (MMW) band as 40 to 300 GHz, we consider the millimeter-wave band as the frequency spectrum between 30 and 300 GHz, in accordance with most observers. Thus, the millimeter-wave region is above the microwave region and below the submillimeter-wave region. In U.S. usage,[1] millimeter-wave bands include K_a band, 33 to 36 GHz; Q band, 36 to 46 GHz; V band, 46 to 56 GHz; and W band 56 to 100 GHz.

The development of better and less costly components has led to increased interest in millimeter-wave radar in recent years.[1–7] Millimeter-wave radars have advantages and disadvantages when compared with microwave radars or with sensors operating in the infrared and visible portions of the frequency spectrum. For a fixed antenna size, a millimeter-wave antenna has a much narrower beamwidth than a microwave antenna, because the beamwidth is proportional to the wavelength for a fixed aperture size. Millimeter-wave radar thus has better angular resolution and higher antenna gain, and the narrower antenna beam makes jamming more difficult. Range resolution is inversely proportional to the transmitted bandwidth, and the wider bandwidths available at millimeter wavelengths can thus provide fine range resolution.

Components available at millimeter wavelengths are smaller in size and weight than microwave components, and millimeter-wave systems are smaller and lighter. Disadvantages of millimeter-wave operation include transmitter power that decreases with frequency, and component losses and noise figure that increase with frequency; millimeter-wave receivers are less sensitive than microwave receivers. Fewer components are available at millimeter wavelengths, and existing components are more expensive. Most millimeter-wave radar systems developed in the past have operated between 35 and 100 GHz, partly because of component availability and cost.

Propagation effects are quite important for millimeter-wave systems. Clear-air attenuation and attenuation and backscatter from rain, clouds, fog, or snow are much higher than at microwave frequencies. There are atmospheric attenuation minima near 35, 94, 140, and 200 GHz, and most millimeter-wave radars operate near one of those frequencies. Millimeter-wave radar can, however, operate in the high-attenuation band near 60 GHz when there is a desire for covert operation. Although millimeter-wave radar performance in adverse weather is much worse than that in the microwave region, millimeter-wave radar has much better performance than a sensor operating at infrared or visible wavelengths: millimeter-wave radar can operate with limited performance through rain, fog, smoke, or dust when operation at infrared or visible wavelengths is impossible.

Millimeter-wave radar applications include situations where the narrower antenna beamwidth, increased antenna gain, and smaller physical size of the radar are advantageous and where increased propagation losses compared with microwave frequencies are tolerable. Millimeter-wave radars operate at shorter ranges than those in the microwave region. Military applications include missile seekers, terminal guidance, airborne ground mappers, ground-to-ground surveillance, terrain clearance, and target tracking at low grazing angles where the narrow antenna beam minimizes multipath and clutter. Other applications

include operation in space, where propagation losses are not a problem, and automotive warning and braking systems.

3.2 RADAR FUNDAMENTALS

3.2.1 Basics and Key Parameters

RADAR is an acronym for *ra*dio *d*etection *a*nd *r*anging. Radar is a device that radiates electromagnetic energy and detects the energy scattered from objects in the path of the radiation.[8,9] The basic principle of operation is the same whether radar operates in a microwave, millimeter-wave, or optical frequency band, although the components may be quite different. The time delay Δt between the transmission and reception of the scattered energy determines the target range R:

$$R = \frac{c\Delta t}{2} \,. \tag{3.1}$$

Here c is the velocity of propagation in the medium, 2.9979×10^8 m/s in free space, and the factor of 2 accounts for the round-trip path. The round-trip time delay for a nautical mile is 12.36 μs. Because the radar antenna normally has directivity, radar can determine the target location in azimuth and elevation from the antenna orientation of the peak response. Radars can also find the radial velocity of moving targets by measuring the difference in range between pulses, or, if the radar is coherent, by the Doppler frequency shift between the transmitted and received signals.

With monostatic radar, the transmitter and receiver are at the same location; bistatic radar has its transmitter and receiver at separate locations. Nearly all millimeter-wave radars are monostatic. A number of different monostatic radar configurations are possible; Fig. 3.1 shows a simple monostatic

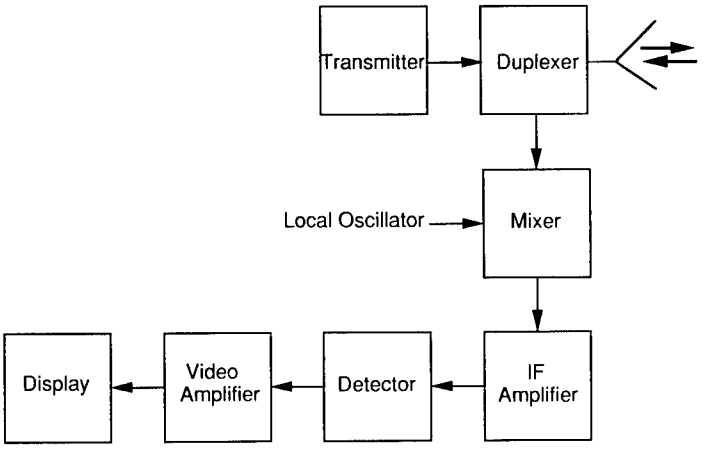


Fig. 3.1 Simple monostatic radar with superheterodyne receiver.

radar with a superheterodyne receiver. The advantage of the superheterodyne receiver is that it is easier to process signals at the intermediate frequency (IF) than at much higher transmitted frequencies. The transmitter generates a pulsed signal at a carrier frequency, and this signal passes through the duplexer, a device that separates the transmitted and received signals before the antenna. The signal travels to the target, and the target backscatters a portion of the incident energy to the radar antenna. The echo travels through the duplexer to a mixer whose other input is the local oscillator signal. The mixer is a nonlinear device that generates a number of signal components; the IF amplifier following the mixer amplifies the signal components at the difference frequency between the local oscillator and the signal. The detector following the IF amplifier recovers the amplitude information of the return, and the amplitude-detected signal passes through the video amplifier to the display for viewing by an observer.

Most millimeter-wave radars operate with pulsed waveforms, and the important transmitter parameters are the peak power P_t (averaged over the rf cycle), the average power P_{av}, and the duty cycle Du. The duty cycle is the product of the transmitted pulse width τ and the pulse repetition frequency (PRF) f_r:

$$\text{Du} = \tau f_r \ , \tag{3.2}$$

and

$$P_{av} = \text{Du} P_t \ . \tag{3.3}$$

Figures 3.2 and 3.3 plot these relationships. The energy in a single pulse is

$$E_1 = \tau P_t \ . \tag{3.4}$$

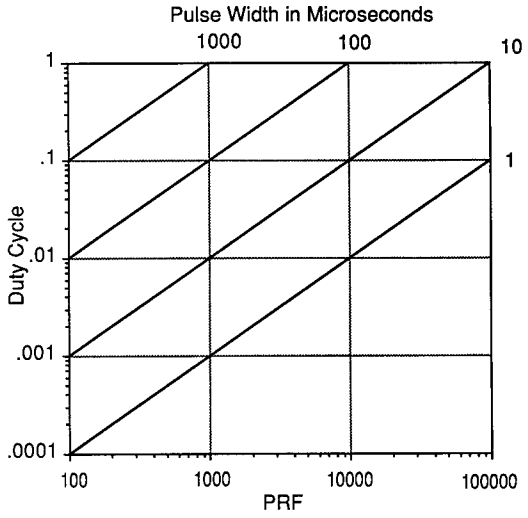


Fig. 3.2 Duty cycle versus PRF.

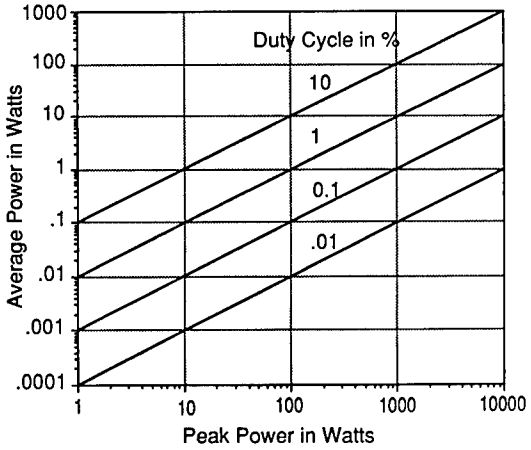


Fig. 3.3 Average power versus peak power.

The pulse length of millimeter-wave radar is typically of the order of microseconds; transmitter source restrictions may limit the maximum possible pulse length.

The range resolution ρ_r(−3-dB width) depends on the transmitted bandwidth ΔB:

$$\rho_r \cong \frac{c}{2\Delta B} \ . \tag{3.5}$$

For a short pulse,

$$\tau \cong 1/\Delta B \tag{3.6}$$

and

$$\rho_r \cong c\tau/2 \ . \tag{3.7}$$

When the peak power of the transmitter limits the operating range, pulse compression techniques provide a means of increasing the average transmitted power, while providing range resolution better than that corresponding to the transmitted pulse length. A common pulse compression technique is the use of a linear FM or chirp waveform in which the frequency of the transmitted signal varies linearly with time during the pulse to increase the bandwidth. The improvement in range resolution after range compression in the receiver equals the time-bandwidth product $\tau\Delta B$. Another pulse compression technique[10] often used in millimeter-wave radars is a frequency-stepped waveform of N pulses that are separated by equal frequency increments Δf to provide an effective bandwidth of $(N - 1)\Delta f$. Phase-coded pulse compression waveforms and other techniques are also available.[11] Figure 3.4 shows a plot of range resolution versus bandwidth.

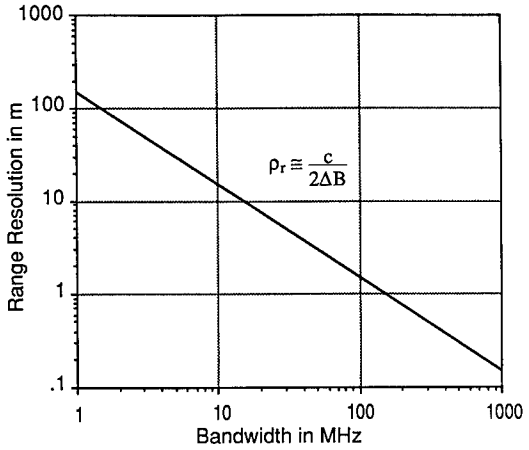


Fig. 3.4 Range resolution versus bandwidth.

The cross-range (angular) resolution ρ_a depends on the range R to the target and the antenna -3-dB beamwidth $\Delta\theta$:

$$\rho_a = R\Delta\theta \ . \tag{3.8}$$

From antenna theory, the beamwidth is proportional to the wavelength and inversely proportional to the antenna diameter D, so that

$$\rho_a = \frac{k\lambda R}{D} \ . \tag{3.9}$$

Here k is a constant near unity. Thus, good cross-range resolution requires a large antenna aperture and a short range; the cross-range resolution is inversely proportional to the frequency. Figure 3.5 plots cross-range resolution versus range as a function of the antenna diameter at 94 GHz.

The Doppler frequency shift f_d resulting from target motion depends on the radial velocity v_r of the target and the wavelength λ:

$$f_d = \frac{2v_r}{\lambda} \ . \tag{3.10}$$

Figure 3.6 shows the Doppler frequency shift as a function of the radial velocity at 35, 94, and 140 GHz. A pulsed radar samples f_d at the PRF f_r, and from the Nyquist sampling theorem, radar can measure f_d unambiguously when $f_d < f_r/2$. At higher Doppler frequencies, the Doppler spectrum is folded about $f_r/2$. Whenever f_d is an integer multiple of f_r, the sampled Doppler spectrum appears at zero frequency, and it cannot be distinguished from stationary clutter centered at zero frequency. These target velocities v_b are called *blind speeds*:

$$v_b = \frac{n\lambda f_r}{2} \ . \tag{3.11}$$

Figure 3.7 shows the first five blind speeds as a function of the PRF at 94 GHz.

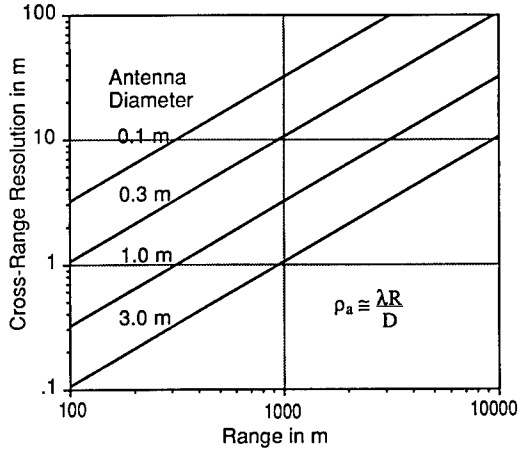


Fig. 3.5 Cross-range resolution versus range at 94 GHz.

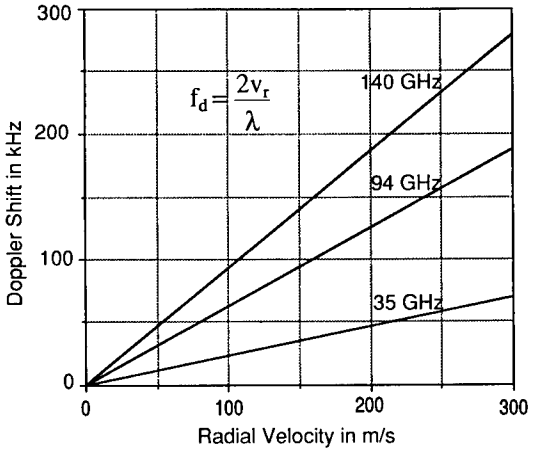


Fig. 3.6 Doppler shift versus radial velocity.

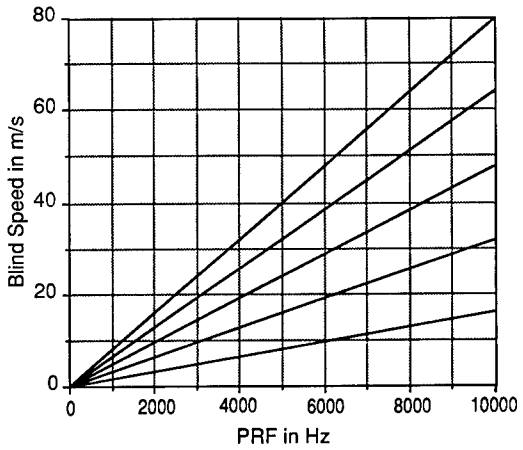


Fig. 3.7 First five blind speeds at 94 GHz.

3.2.2 Radar Range Prediction

3.2.2.1 SNR Computation. Return signals received by radar are always embedded in a background of thermal noise and, possibly, undesired clutter from unwanted targets in the antenna beam. An optimum radar design provides the highest possible ratio of peak signal to average noise power or peak signal to average clutter power at the receiver output. It can be shown that a receiver that acts as a matched filter optimizes the signal-to-noise ratio (SNR) for a single pulse embedded in thermal noise.[8–15] A matched filter has an impulse response that equals the input signal waveform with the time axis reversed. In the frequency domain, the voltage transfer function of the matched filter is the complex conjugate of the signal amplitude spectrum.

The maximum SNR at the output of a matched filter is

$$\mathrm{SNR}_{\max} = \frac{E}{N_0} , \tag{3.12}$$

where E is the received signal energy and N_0 is the noise spectral density in the receiver, which is assumed to be flat,

$$N_0 = kT_s . \tag{3.13}$$

Here k is Boltzmann's constant (1.38×10^{-23} J/K) and T_S is the system noise temperature. Practical radars are always less efficient than a matched filter; consequently, for practical radars,

$$\mathrm{SNR} = \frac{E}{N_0 L_M} , \tag{3.14}$$

where the loss factor $L_M > 1$. The matched filter concept is appropriate because many practical radars approach matched filter performance with loss factors of only 1 or 2 dB. In these equations, the signal power is the power averaged over the rf cycle.

The free-space radar transmission equation provides the energy received by the radar; for monostatic radar,[8,9]

$$E = \frac{P_{\mathrm{av}} t_0 G_t G_r \lambda^2 \sigma}{(4\pi)^3 R^4} . \tag{3.15}$$

Here P_{av} is the average power, t_0 is the observation time, G_t and G_r are the gains of the transmitting and receiving antennas in the target direction, σ is the target radar cross section (RCS), λ is the transmitted wavelength, and R is the slant range to the target. The RCS, which has dimensions of square meters, is a measure of the amount of the incident power backscattered from the target to the radar; the backscattered power is proportional to the RCS. Section 3.3.1 defines and discusses RCS.

To account for cases where free-space conditions do not exist and for the effects of the antenna pattern, we must add a pattern propagation factor F to

the right side of Eq. (3.15). If the same antenna is used for both transmission and reception, and there are losses, Eq. (3.15) becomes

$$E = \frac{P_{av} t_0 G^2 \lambda^2 \sigma F^4}{(4\pi)^3 R^4 L} \quad . \tag{3.16}$$

The propagation factor F is the ratio of the electric field intensity actually present at the target to the electric field intensity that would exist under free-space conditions at the antenna maximum gain condition. Now G is the maximum antenna gain, not the gain in the target direction, and F is raised to the fourth power because the antenna gain in the target direction is the product of the peak gain and the square of the antenna factor that accounts for the relative gain at direction angles θ and ϕ. The loss term L accounts for energy dissipated as heat and for other factors that lower the effective SNR.

Dividing Eq. (3.16) by Eq. (3.13) provides the SNR,

$$\frac{E}{N} = \text{SNR} = \frac{P_{av} t_0 G^2 \lambda^2 \sigma F^4}{(4\pi)^3 R^4 L k T_s} \quad . \tag{3.17}$$

Following Refs. 8 and 13, we select a detectability factor D to meet system requirements: D is the required SNR at the input of the envelope detector that provides the required probability of detection for a given false alarm rate. SNR differs from D in that SNR is the SNR computed with the radar equation, whereas D is the required value of SNR for a given probability of detection. Inserting D for SNR and solving Eq. (3.17) for R yields the maximum range R_{max}:

$$R_{max} = \left[\frac{P_{av} t_0 G^2 \lambda^2 \sigma F^4}{(4\pi)^3 L k T_s D} \right]^{1/4} \quad . \tag{3.18}$$

At R_{max}, the radar has a specified probability of detecting a particular target. Section 3.2.5 discusses D.

An alternative form of the radar equation relating SNR to the peak transmitter power P_t and the system noise bandwidth B_n is

$$\text{SNR} = \frac{P_t G^2 \lambda^2 \sigma F^4}{(4\pi)^3 R^4 L k T_s B_n} \quad . \tag{3.19}$$

The receiver half-power bandwidth is usually a good approximation to B_n.

3.2.2.2 System Noise Temperature. SNR computations require a knowledge of the system noise temperature T_s. Following Blake,[13,14] T_s is the sum of three contributions:

$$T_s = T_a + T_r + L_r T_e \quad [\text{K}] \ , \tag{3.20}$$

where T_a is the antenna noise temperature, T_r is the noise temperature of the transmission line between the antenna and the receiver input, L_r is the loss in the transmission line, and T_e is the noise temperature of the receiver. Many

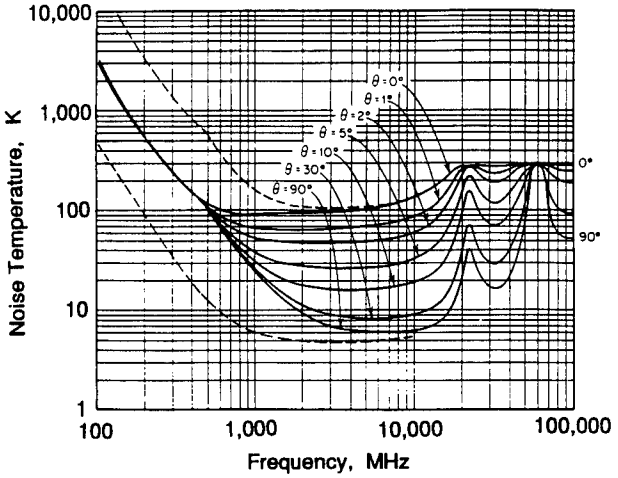


Fig. 3.8 Noise temperature of an antenna aimed at the sky. The antenna is lossless with no sidelobes directed at the earth.[14]

millimeter-wave radars have receiver noise temperatures that are so high that the antenna and transmission line noise temperatures can be ignored.

The antenna temperature is the temperature of natural radiating sources in the beam. If the antenna points at the ground, T_a is assumed equal to a reference temperature of 290 K. When the antenna is pointing at the sky, its temperature may be considerably less than 290 K. Typically, 88% of the total power received is from the sky and the rest from the ground through the antenna sidelobes, and[8,13]

$$T_a = \frac{0.88T_a' - 254}{L_a} + 290 \ , \tag{3.21}$$

where T_a' is the apparent sky temperature and L_a is the ohmic loss of the antenna. Blake computed the noise temperature of a lossless antenna aimed at the sky, and T_a' can be obtained from Fig. 3.8 for an elevation angle θ. Figure 3.8 assumes an antenna pattern with no sidelobes viewing the ground. The solid curves in Fig. 3.8 are for the geometric mean galactic temperature, and the dashed curves are for the maximum and minimum galactic noise. The noise temperature T_r of the transmission line is

$$T_r = T_{\text{tr}}(L_r - 1) \ , \tag{3.22}$$

where T_{tr} is the physical temperature of the transmission line (often assumed to be 290 K) and L_r is the loss in the transmission line.

The receiver noise temperature T_e is

$$T_e = T_0(F_n - 1) \ . \tag{3.23}$$

Here T_0 is 290 K, the reference temperature, and F_n is the receiver noise figure. The noise power at the output of a receiver equals the product of

F_n ($F_n > 1$) and the thermal noise power from an ideal receiver, which adds no noise beyond that which can be accounted for by thermal noise alone.

3.2.2.3 System Losses. The loss term L in the radar equation is a key factor for predicting radar performance; loss terms differ with the type of radar. Some losses can be calculated, whereas others involving operator performance and field degradation are more difficult to quantify. The total loss is the product of the individual losses or the sum of the decibel values when the losses are given in decibels. References 8, 9, and 13 have more detail on losses.

Propagation Loss. The two-way attenuation through the atmosphere is usually significant at millimeter wavelengths; Sec. 3.3.2 discusses atmospheric losses. Because the range prediction involves atmospheric attenuation, which is range dependent, an iterative procedure is necessary to determine the range when atmospheric loss is significant.

Antenna Loss. The ohmic antenna loss L_a is needed for the computation of the system noise temperature, but this loss may not be specified for the antenna, and it is often hard to obtain.

Matched Filter Loss. This loss accounts for a decrease in the efficiency of actual radar compared with a perfect matched filter; in some cases this loss is only 1 or 2 dB.

Plumbing Loss. The transmitter power is normally measured at the transmitter output terminal, and the plumbing loss between the transmitter and the input to the antenna accounts for the reduction in the power actually radiated. Plumbing losses between the receiver and the antenna are usually included in the computation of the system noise temperature.

Beamshape Loss. With search radar, the power received from an antenna scanning past a target varies because the antenna gain changes during the scan. Blake estimates this loss to be about 1.6 dB for a one-dimensional scan and about 3.2 dB for a two-dimensional scan.[13]

Integration Loss. This loss represents the decrease in efficiency for radar that integrates pulses noncoherently rather than coherently; Sec. 3.2.5 discusses integration losses.

Collapsing Loss. The collapsing loss is the loss resulting from combining extra noise samples with the desired signal plus noise samples. If, for example, the display resolution is less than the range resolution, more than one range cell will be collapsed into a single display resolution cell.

Fluctuation Loss. The fluctuation loss accounts for the additional SNR required to detect a target whose RCS varies; Sec. 3.2.5 discusses this loss.

Processing Loss. Losses that occur during processing may require a compensating increase in the SNR. Imaging radar may compress digital output data in both range and azimuth by taking fast Fourier transforms. When aperture weighting is added to reduce sidelobes, there is a reduction in the peak target response. Losses also occur when signals straddle range gates or Doppler filters.

CFAR Loss. When clutter is the dominant background interference, the clutter level must be estimated with special techniques to set the detection thresh-

old to maintain a constant false alarm rate. The constant false alarm rate (CFAR) loss accounts for the decrease in performance associated with the measurement of the clutter level and the establishment of the detection threshold.

3.2.3 Sample Range Computation

For a sample range computation, we assume scanning radar that must detect a 1-m^2 target in receiver noise. Table 3.1(A) lists radar and target parameters for the sample computation; the radar is assumed to use pulse compression so that the range resolution will be much narrower than that corresponding to a 10-μs pulse width. The first step in the computation is finding the system noise temperature T_s. The 8.6-mm radar wavelength corresponds to a frequency of 35 GHz, and the apparent sky temperature T_a' is 70 K from Fig. 3.8 when a 10-deg elevation angle is assumed.

From Eq. (3.21), the antenna noise temperature is

$$T_a = \frac{0.88 \times 70 - 254}{1.12} + 290 = 118 \text{ K} .$$

From Eq. (3.22), the noise temperature of the transmission line to the receiver is

$$T_r = 290(1.258 - 1) = 75 \text{ K} .$$

From Eq. (3.23), the receiver noise temperature is

$$T_e = 290(1.995 - 1) = 289 \text{ K} .$$

From Eq. (3.20), the system noise temperature is

$$T_s = 118 + 75 + 1.258 \times 289 = 557 \text{ K} .$$

The losses are

L_t	0.9 dB
L_m	2.0 dB
L_s	1.6 dB
L_{mis}	2.0 db
total loss L	6.5 dB .

The calculation in Table 3.1(B), using Eq. (3.18), provides the maximum range; Table 3.1(B) uses the logarithmic-decibel radar equation form, which is frequently used for range and SNR computations because many radar parameters are expressed in decibels. The initial $R_{\max}$ calculation in Table 3.1(B) assumes no atmospheric loss, because the atmospheric loss, which depends on the range, is not yet known. The final value of $R_{\max}$ is found by iteration. The first iteration for the atmospheric loss assumes that the range equals the value computed with no atmospheric attenuation, 26.3 km from Table 3.1(B). With a one-way loss of 0.12 dB/km, the round-trip atmospheric loss L_{at} is

$$L_{\text{at}} \cong 0.12 \times 2 \times 26.3 = 6.3 \text{ dB} .$$

Table 3.1 Range Computation Example

(A) Radar and Target Parameters		
Peak power	P_t	5000 W
Pulse width	τ	10 μs
PRF	f_r	750 Hz
Wavelength	λ	8.6 mm
Antenna gain	G	40 dB
Propagation factor	F	1
Elevation angle	θ	10 deg
Antenna ohmic loss	L_a	0.5 dB
Transmission line loss	L_r	1.0 dB
Atmospheric attenuation	γ_T	0.12 dB/km
Noise figure	F_n	3.0 dB
Transmitter loss	L_t	0.9 dB
Mismatch loss	L_m	2.0 dB
Scanning pattern loss	L_s	1.6 dB
Miscellaneous losses	L_{mis}	2.0 dB
Number of pulses integrated	n	31
Target RCS	σ	1 m^2
Probability of detection	P_d	0.9
Probability of false alarm	P_{fa}	10^{-8}
Target model		Swerling case 1
Detectability factor	$D_1(31)$	10.5 dB (from Sec. 3.2.5)

(B) $R_{\max}$ Calculation

$$R_{\max} = \left[\frac{P_{\text{av}} t_0 G^2 \lambda^2 \sigma F^4}{(4\pi)^3 L k T_s D_1(31)}\right]^{1/4}$$

		+dB	−dB
P_{av}	10 log(5000)	37.	
$\tau = t_0$	10 log(10^{-5})		50.
G^2	2 × 40 dB	80.	
λ^2	20 log(8.6×10^{-3})		41.3
σ	10 log(1)	0.	
F	10 log(1)	0.	
$(4\pi)^3$	30 log(4π)		33.0
L	6.5 dB		6.5
k	10 log(1.38×10^{-23})	228.6	
T_s	10 log(557)		27.5
$D_1(31)$	10.5 dB		10.5
		Total = $\overline{345.6\text{ dB}}$	− $\overline{168.8\text{ dB}}$ = 176.8 dB

$R_{\max}$ = 176.8 / 4 = 44.2 dB, $R_{\max} = 10^{4.42}$ = 26,303 m
(no atmospheric loss included)

The 176.8-dB value computed in Table 3.1(B) for R_{max} in decibels must be reduced an additional 6.3 dB to account for the atmospheric loss. Thus,

$$R_{max} \cong (176.8 - 6.3)/4 = 42.63 \text{ dB} \quad \text{or } 18{,}323 \text{ m} .$$

The second iteration assumes a range of 18.3 km, and

$$L_{at} \cong 0.12 \times 2 \times 18.3 = 4.4 \text{ dB} ,$$

$$R_{max} \cong (176.8 - 4.4)/4 = 43.1 \text{ dB} \quad \text{or} \quad 20{,}417 \text{ m} .$$

If the loss at 20.4 km, 4.9 dB, is used to compute R_{max} in a third iteration, R_{max} equals 19.8 km. No further iterations are necessary because the loss at 19.8 km, 4.8 dB, is very close to the loss assumed. Reference 13 discusses graphic solutions for determining R_{max} for situations for which the iterative procedure does not converge.

3.2.4 Signal-to-Clutter Computations

Many situations require the detection of targets in a background of distributed clutter. Detection predictions in clutter are less accurate than predictions in noise, because clutter reflectivity is more difficult to model than thermal noise. Distributed clutter may be in a volume or on a surface. The signal-to-clutter ratio (SCR) is the ratio of the target radar cross section (RCS) σ to the clutter RCS σ_c in a resolution cell:

$$\text{SCR} = \frac{\sigma}{\sigma_c} . \tag{3.24}$$

Section 3.3.1 discusses RCS.

3.2.4.1 Volume Clutter. Volume clutter results from the presence of rain, snow, hail, fog, or chaff in a volume. The total clutter RCS of the volume is the product of the reflectivity per unit volume η and the volume V:

$$\sigma_c = \eta V . \tag{3.25}$$

The units of η are square meters per cubic meter. The volume V is

$$V = \frac{\pi}{4} \theta_a \theta_e R^2 \rho_r , \tag{3.26}$$

as shown in Fig. 3.9, where R is the range and ρ_r is the range resolution. The azimuth and elevation beamwidths, θ_a and θ_e, to use in Eq. (3.26) are approximately 75% of the -3-dB one-way beamwidths.[13] Section 3.3.3 discusses the volume backscatter coefficient η.

3.2.4.2 Surface Clutter. The parameter σ_0 (sometimes written σ^0) is the average reflectivity of distributed clutter on a surface. The clutter RCS σ_c for a surface is the product of σ_0 and the illuminated area A:

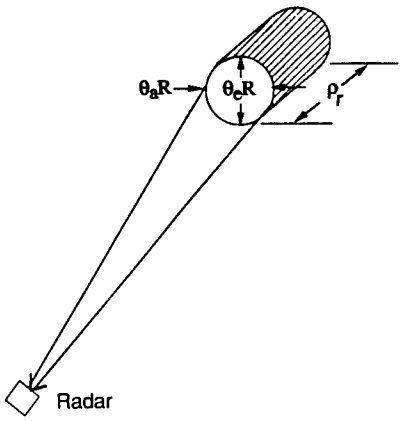


Fig. 3.9 Volume clutter cell size.

$$\sigma_c = \sigma_0 A \ . \tag{3.27}$$

When the antenna beamwidth determines the range resolution, the antenna spot is elliptic with an area

$$A = \frac{\pi}{4}(\theta_a R)(\theta_e R \csc\psi) \ , \tag{3.28}$$

where θ_a and θ_e are the azimuth and elevation beamwidths, R is the range, and ψ is the grazing angle. The values of θ_a and θ_e to use are approximately 75% of the one-way −3-dB beamwidth.[13]

If the range resolution from the antenna beamwidth is greater than the range resolution ρ_r resulting from the radar bandwidth, i.e.,

$$\theta_e R \csc\psi > \rho_r \sec\psi \ , \tag{3.29}$$

then,

$$A = \theta_a R \rho_r \sec\psi \ . \tag{3.30}$$

In the case of short-pulse radar,

$$\rho_r = \frac{c\tau}{2} \ , \tag{3.31}$$

where c is the velocity of light and τ is the pulse width. Figure 3.10 shows the areas for the two cases.

For synthetic aperture radar (see Sec. 3.2.7),

$$A = \rho_r \rho_a \sec\psi \ , \tag{3.32}$$

where ρ_r and ρ_a are the −3-dB-range and azimuth (cross-range) resolutions.

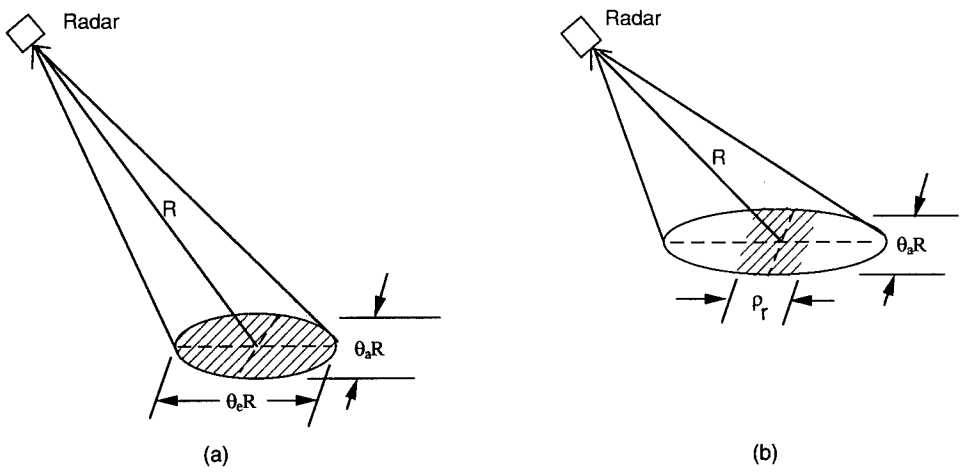


Fig. 3.10 Surface clutter area illuminated: (a) area limited by beamwidth and (b) area limited by range resolution.

When the clutter power is much higher than the noise power, the signal-to-clutter ratio (SCR) determines the probability of detection. When both clutter and noise are present, the total background power determines the probability of detection, and the SNR and SCR can be combined to find the signal-to-interference ratio (SIR)

$$\text{SIR} = \frac{1}{1/(\text{SNR}) + 1/(\text{SCR})} . \tag{3.33}$$

The quantity SIR is useful whenever the clutter amplitude has a Rayleigh distribution like receiver noise so that the total interference power also has a Rayleigh distribution. In this case, standard probability of detection curves are applicable. At millimeter wavelengths, however, rain clutter has a log-normal distribution, and surface clutter often has a distribution with higher tails than a Rayleigh distribution; consequently, probability of detection computations based on Rayleigh statistics are optimistic. For a clutter distribution with higher tails, the detection threshold must be raised to maintain a constant false alarm rate, and the probability of detection is therefore lower. Section 3.3.3 discusses clutter statistics in more detail.

3.2.5 Target Detection

3.2.5.1 Single-Pulse Detection. To determine the radar operating range, the radar designer must select the value of the SNR that provides the desired performance. Figure 3.1 is a block diagram of a typical pulsed radar for detecting pulsed signals in a background of Gaussian noise whose spectral density is constant over the receiver passband. Assuming that the bandwidth of the IF amplifier is small compared with the center frequency, the IF noise is composed of in-phase (I) and quadrature (Q) components, each of which is statistically independent and has a Gaussian probability distribution. It can

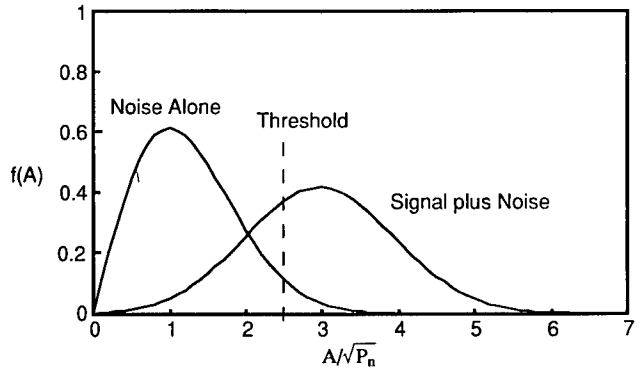


Fig. 3.11 Probability density functions of noise and signal plus noise.

be shown that the envelope of the IF amplifier noise amplitude A has a Rayleigh probability distribution function[8,9,13,15]:

$$f_n(A) = \frac{A}{P_n} \exp\left(-\frac{A^2}{2P_n}\right) , \quad A > 0 , \tag{3.34}$$

where P_n is the mean noise power. When a sinusoidal signal of amplitude a is present, the envelope of the signal plus noise has a Rician distribution,

$$f_{sn}(A) = \frac{A}{P_n} \exp\left[-\frac{(A^2 + a^2)}{2P_n}\right] I_0\left(\frac{aA}{P_n}\right) . \tag{3.35}$$

Here I_0 is the zero-order modified Bessel function.

Figure 3.11 is a plot of these two functions, which are present at the input of the envelope detector. To determine if a signal is present, a threshold Th is established, and a threshold detector senses whenever the voltage exceeds the threshold level. A false alarm occurs whenever the noise alone exceeds the threshold, and the probability of false alarm P_{fa} is

$$P_{\text{fa}} = \int_{\text{Th}}^{\infty} \frac{A}{P_n} \exp\left(\frac{-A^2}{2P_n}\right) dA = \exp\left(\frac{-\text{Th}^2}{2P_n}\right) . \tag{3.36}$$

A detection occurs whenever the signal plus noise exceeds the threshold; the probability of detection P_d is

$$P_d = \int_{\text{Th}}^{\infty} \frac{A}{P_n} \exp\left[-\frac{(A^2 + a^2)}{2P_n}\right] I_0\left(\frac{aA}{P_n}\right) dA . \tag{3.37}$$

Thus, the threshold setting affects both the probability of detection and the probability of false alarm: raising the threshold decreases both the probability of detection and the probability of false alarm. The probability of detection depends only on the SNR for a fixed threshold.

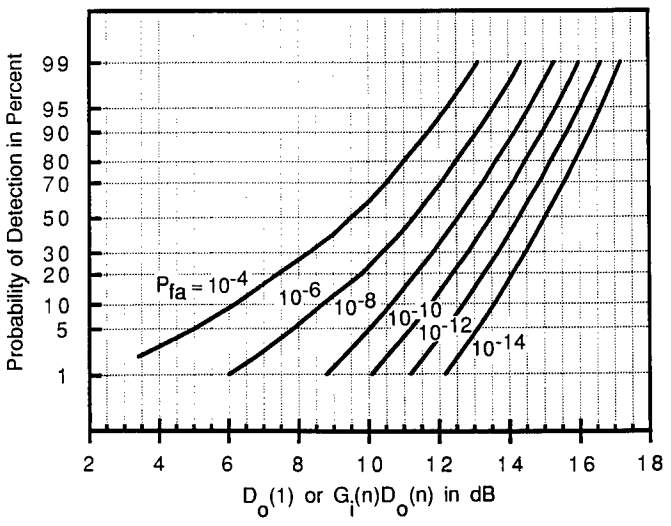


Fig. 3.12 Detectability factor for nonfluctuating target of unknown phase.

Marcum and others have calculated the probability of detection versus the SNR for single-pulse detection with the probability of false alarm as a parameter using the incomplete Toronto function.[16] Figure 3.12 plots the probability of detection for a nonfluctuating target of unknown phase for false alarm probabilities between 10^{-4} and 10^{-14}; the horizontal scale is $D_0(1)$ or $G_i(n)D_0(n)$. The detectability factor $D_0(1)$ is the required single-pulse SNR at the input of the envelope detector for a given probability of detection for the false alarm probabilities shown on the curves. The quantity $G_i(n)$ is the gain when n pulses are integrated noncoherently, and $D_0(n)$ is the required SNR when n pulses from a nonfluctuating target are integrated noncoherently. Note that a relatively small difference in $D_0(1)$ results in a large change in the probability of detection. Although Marcum used a square-law detector for his calculations, the detectability factor for a linear detector is almost the same.[15]

If noise is present continuously, independent noise samples occur at a rate equal to the noise bandwidth B_n, resulting in a false alarm rate of $B_n P_{fa}$. The average time between noise threshold crossings is the false alarm time T_{fa}, which equals the reciprocal of the false alarm rate:

$$T_{fa} = \frac{1}{B_n P_{fa}} . \tag{3.38}$$

Some authors prefer the false alarm number n_f, where

$$n_f = 1/P_{fa} . \tag{3.39}$$

There will be an average of n_f decisions between false alarms. Marcum and others have defined the false alarm number n_f' differently[15]:

$$n_f' \cong \frac{0.693}{P_{\text{fa}}} \quad . \tag{3.40}$$

Example. Suppose that the required probability of detection is 0.90 for a false alarm probability of 10^{-8}; from Fig. 3.12, $D_0(1) = 14.2$ dB. If $B_n = 2$ MHz and $T_{\text{fa}} = 50$ s, the false alarm rate $= 0.02\ \text{s}^{-1}$, and $n_f = 10^8$.

3.2.5.2 Pulse Integration. Most radars integrate a number of pulses to increase the probability of detection over that which can be obtained from single-pulse detection. With scanning radar, for example, the number of pulses received when the beam scans past the target is[9]

$$n = \frac{\theta_B f_r}{6\omega_m} \quad , \tag{3.41}$$

where θ_B is the antenna -3-dB beamwidth in degrees, f_r is the PRF in hertz, and ω_m is the number of revolutions the antenna makes per minute.

Two types of integration are possible: coherent and noncoherent. With coherent integration, the phase relationship between pulses is preserved, and the pulse amplitudes add in phase, providing an output power proportional to the square of the number of pulses. The noise, however, has a random phase from pulse to pulse, and the integrated noise power is proportional to the number of pulses. The detectability is proportional to the number n of the pulses coherently integrated and

$$D_c(n) = D_c(1)/n \quad . \tag{3.42}$$

Here $D_c(n)$ is the detectability factor with n pulses coherently integrated, and $D_c(1)$ is the single-pulse coherent detectability factor. These factors are the required SNRs at the detector input. With coherent integration, the probability of detection can be found from Fig. 3.13 by using $D_c(1)$ or $nD_c(n)$ as the input; Fig. 3.13 shows detection plots for a nonfluctuating target for an ideal radar system where the phase of the return is known.[8,15] Because coherent integration requires that phase information be preserved from pulse to pulse, it must be done at the intermediate frequency before envelope detection, or synchronous detectors must be used to provide in-phase and quadrature video signals to preserve phase information.

A coherent detector responds only to the in-phase noise component, not to the quadrature noise component. With an envelope detector, both in-phase and quadrature noise compete with the signal at low SNRs, but at high SNRs, the quadrature noise component has little effect. Thus, there are additional losses with envelope detection compared with coherent detection. It is, however, much easier to construct instrumentation to perform noncoherent integration after envelope detection.

Barton has derived an empirical equation for the loss $C_x(1)$ with envelope detection compared with coherent detection[17]:

$$C_x(1) = \frac{D_0(1)}{D_c(1)} \cong \frac{D_0(1) + 2.3}{D_0(1)} = \frac{\text{SNR} + 2.3}{\text{SNR}} \quad . \tag{3.43}$$

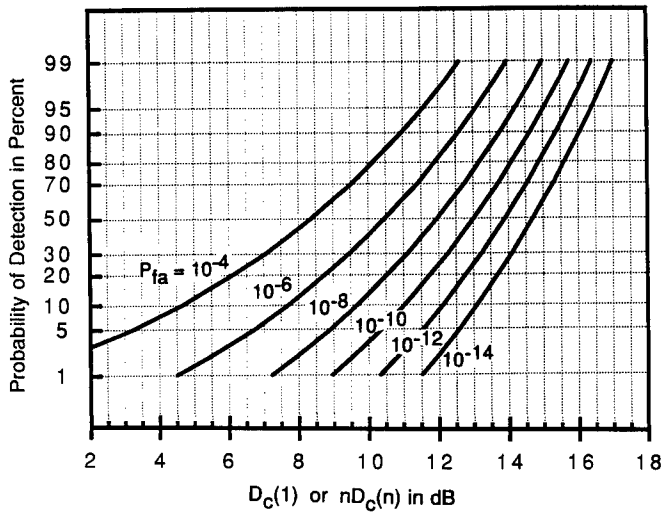


Fig. 3.13 Detectability factor for coherent detection and a nonfluctuating target.

Probability of detection computations made with this empirical formula are within 0.2 dB of those made with more accurate formulas for nonfluctuating targets; consequently, Barton's empirical approach is adequate for most engineering purposes. When there is integration after envelope detection, the required SNR per pulse decreases, but there is an increase in the detector loss, and the full benefit of the increased total energy ratio is not obtained. The additional loss occurring during detection is called the *integration loss* $L_i(n)$, which represents the decrease in efficiency of noncoherent integration compared with coherent integration:

$$L_i(n) = \frac{nD_0(n)}{D^0(1)} = \frac{n}{G^i(n)} \,. \tag{3.44}$$

Figure 3.14 plots $L_i(n)$ versus the number of pulses integrated with $D_0(1)$ as a parameter. Figures 3.12 and 3.14 provide sufficient information to permit the radar designer to calculate the required single-pulse detectability factor when n pulses are integrated noncoherently. Barton's equations also provide the probability of detection when the SNR is known. The coherent detectability factor can be expressed as a function of SNR using Eqs. (3.42) and (3.43),

$$D_c(1) = \frac{n(\text{SNR})^2}{\text{SNR} + 2.3} \,. \tag{3.45}$$

The probability of detection can then be found from Fig. 3.13. For instance, if the single-pulse SNR = 3.16 (5 dB), and 20 pulses are integrated,

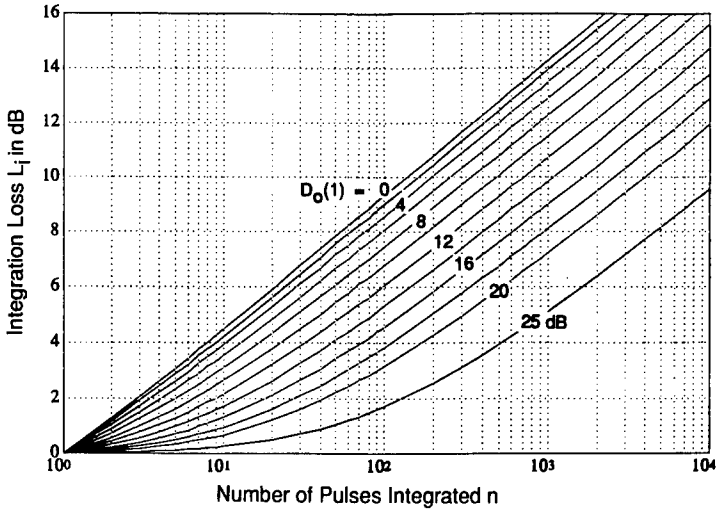


Fig. 3.14 Integration loss versus number of pulses integrated [$D_0(1)$ = effective single-pulse SNR].

$$D_c(1) = \frac{20 \times (3.16)^2}{3.16 + 2.3} = 36.6 \quad \text{or } 15.6 \text{ dB} \ .$$

From Fig. 3.13, $P_d = 0.93$ for $P_{\text{fa}} = 10^{-12}$.

3.2.5.3 Targets with a Fluctuating RCS. The detection computations have thus far assumed a nonfluctuating RCS; however, most targets have an RCS that fluctuates. Man-made targets are complex structures with many component parts, and the RCS changes rapidly with aspect angle, particularly at millimeter wavelengths. It is therefore necessary to describe the target RCS with a probability density function and correlation properties with time and frequency. Table 3.2 lists target models often employed in target detection predictions.[9,15,18–20]

Modeling is not a precise procedure because the model probability density functions seldom match the target probability density functions perfectly, or the target probability density function may be unknown. The four Swerling models are the most commonly used models for predicting radar performance, because the models provide approximate fits to experimental data, and extensive computational results are available.[20] The probability density functions for the Swerling models, which are special cases of chi-squared distributions, are not based on any physical scattering mechanisms.

Swerling Case 1. A Swerling case 1 target has an RCS with an exponential probability density function; this model is applicable to a target that has a number of scatterers with no single dominant scatterer. All the pulses integrated in a pulse group are correlated, but there is no correlation from pulse group to pulse group. A Swerling case 1 target RCS fluctuates slowly, and for a scanning radar, all pulses during a scan are correlated, but there is no correlation from scan to scan.

Table 3.2 Target RCS Models for Detection Prediction

Model	Probability Density Function	Fluctuation Model	Applicability
Steady or Marcum model	Constant	—	Sphere or object that does not rotate
Swerling case 1	$\frac{1}{\bar{\sigma}} \exp\left(-\frac{\sigma}{\bar{\sigma}}\right), \quad \sigma \geq 0$	Slow, constant during scan	Several reflectors, no one dominant
Swerling case 2	$\frac{1}{\bar{\sigma}} \exp\left(-\frac{\sigma}{\bar{\sigma}}\right)$	Fast, independent pulse to pulse	Several reflectors, no one dominant
Swerling case 3	$\frac{4\sigma}{\bar{\sigma}^2} \exp\left(-\frac{2\sigma}{\bar{\sigma}}\right)$	Slow, constant during scan	One large, several small
Swerling case 4	$\frac{4\sigma}{\bar{\sigma}^2} \exp\left(-\frac{2\sigma}{\bar{\sigma}}\right)$	Fast, independent pulse to pulse	One large, several small
Log-normal	$\frac{1}{\sigma s\sqrt{2\pi}} \exp\left\{-\frac{[\ln(\sigma/\sigma_m)]^2}{2s^2}\right\}$ s = standard dev. of $\ln(\sigma/\sigma_m)$ σ_m = median RCS	—	Target with large mean-to-median ratio

Swerling Case 2. The Swerling case 2 target has the same exponential probability density function as the case 1 target, but there is no correlation from pulse to pulse. Thus, the Swerling case 2 target uses a fast fluctuation model, which is applicable, for example, to a radar with frequency diversity and no pulse-to-pulse correlation.

Swerling Case 3. The probability density function for a Swerling case 3 target approximates the distribution for a target with one large scatterer and several smaller scatterers. The RCS fluctuates slowly, and all the pulses within a pulse group are correlated, but there is no correlation from scan to scan.

Swerling Case 4. The probability density function for a case 4 target is the same as for case 3, but there is no pulse-to-pulse correlation. A case 4 target is applicable, for example, to a radar with frequency diversity and no pulse-to-pulse correlation.

There are a number of other RCS prediction models.[19] The Weinstock cases correspond to the RCS of cylinders and other simple shapes whose RCS can be approximated with a chi-squared distribution with degrees of freedom between 0.6 and 4.0. The Rice distribution is exact for a target consisting of a steady reflector whose return is significant compared with the returns from an assembly of other reflectors. The probability density function for Swerling cases 3 and 4, however, has only small differences from the Rice distribution for the case where the power from a steady reflector equals the power from all the other reflectors. The log-normal distributions fit the RCS distributions of ships, missiles, and satellites; these distributions do not match any of the chi-squared models. Reference 21 provides log-normal detection statistics.

References 15 and 20 include a large number of accurate probability of detection curves for the four Swerling models, but the concept of fluctuation loss provides a method of solving detection problems with sufficient accuracy

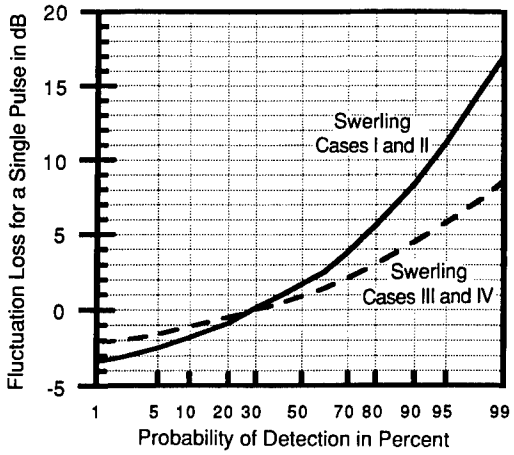


Fig. 3.15 Fluctuation loss versus probability of detection for $P_{fa} = 10^{-8}$.

for most engineering purposes. The fluctuation loss L_f is the ratio of the detectability factor D_f for a fluctuating target to the detectability factor D_0 for a steady target, which gives the same probability of detection for a given probability of false alarm:

$$L_f = \frac{D_f}{D_0} \ . \tag{3.46}$$

Figure 3.15 shows plots of the fluctuation loss for the Swerling models for $P_{fa} = 10^{-8}$; the fluctuation loss is not very sensitive to the probability of false alarm, and the value for $P_{fa} = 10^{-8}$ is accurate enough in most cases. For probabilities of detection less than 0.3, the fluctuation loss is negative, and the fluctuation loss is, therefore, actually a fluctuation gain. For probabilities of detection greater than 0.3, the detectability factor must increase to make up for target fading. With slow fluctuation, the effect of the noncoherent integration is the same as for a steady target, and the fluctuation loss should be added to D_0, not to the detectability factors for Swerling case 1 and case 3 targets, D_1 or D_3.

Reference 15 includes detection curves for rapidly fluctuating targets with noncoherent integration, but such curves are not necessary because there is a simple empirical relationship between the fluctuation loss and the number of independent samples n_e during noncoherent integration,[17]

$$L_f(n_e) = [L_f(1)]^{1/n_e} \ . \tag{3.47}$$

Thus, the fluctuation loss for a Swerling case 2 target when n pulses are integrated is the fluctuation loss for a case 1 target to the $1/n$'th power. This equation is also useful for fluctuation rates between cases 1 and 2 or cases 3 and 4. If n pulses are integrated, but there are only n_e blocks of pulses that are uncorrelated, Eq. (3.47) provides the fluctuation loss.

At the shorter millimeter wavelengths, the integration time t_0 is more likely to exceed the correlation time t_c than at microwave frequencies. The number of independent samples n_e is[8]

$$n_e = 1 + \frac{t_0}{t_c} \cong 1 + \frac{\lambda}{2\omega_a L_x} , \tag{3.48}$$

where λ is the wavelength, ω_a is the rotation rate of the target about the line of sight, and L_x is the width of the target normal to the line of sight. If the radar employs frequency diversity,

$$n_e = 1 + \frac{\Delta f}{f_c} \cong \frac{2\Delta f L_r}{c} . \tag{3.49}$$

Here Δf is the transmitted bandwidth, f_c is the correlation frequency of the target, and L_r is the radial extent of the target.

Example. Consider radar with a 1.5-deg −3-dB beamwidth turning at 6 rpm with a 750-Hz PRF, a nonfluctuating target, and $P_d = 0.90$ and $P_{fa} = 10^{-8}$. What is $D_0(n)$, the required SNR at the input to the envelope detector for a nonfluctuating target? From Eq. (3.41), the number of pulses integrated is

$$n = 1.5 \times 750/(6 \times 6) = 31 \text{ pulses} ,$$

from Fig. 3.12,	$D_0(1)$	=	14.2 dB
from Fig. 3.14,	$L_i(31)$	=	3.2 dB
from Eq. (3.44),	$nD_0(31)$	=	17.4 dB
	$-10 \log(31)$	=	−14.9 dB
	$D_0(31)$	=	2.5 dB .

For a Swerling case 1 target, the fluctuation loss is about 8 dB (6.3) from Fig. 3.15, and the detectability factor $D_1(31)$ is

$D_0(31)$	=	2.5 dB
$+L_f(1)$	=	8.0 dB
$D_1(31)$	=	10.5 dB .

For a case 2 target from Eq. (3.47),

$$L_f(n_e)_{dB} = L_f(1)_{dB}/n_e = 8.0/31 = 0.26 \text{ dB} .$$

The detectability factor $D_2(31)$ for the case 2 target is thus

$$D_0(1) = 2.50 \text{ dB}$$
$$L_f(31) = \underline{0.26 \text{ dB}}$$
$$D_2(31) = 2.76 \text{ dB} .$$

If 31 pulses are integrated, but there are only three uncorrelated pulse blocks, $n_e = 3$. The fluctuation loss from Eq. (3.47) becomes

$$L_f(3) = L_f(1)_{\text{dB}}/3 = 8/3 = 2.7 \text{ dB} .$$

The detectability factor D for this target, which is between case 1 and case 2, is

$$D_0(1) = 2.5 \text{ dB}$$
$$L_f(3) = \underline{2.7 \text{ dB}}$$
$$D = 5.2 \text{ dB} .$$

3.2.5.4 Detection Processing. A detection occurs when the output signal from either a coherent or envelope detector exceeds a threshold level. With automatic detection, electronic circuitry compares the output signal to the threshold level and triggers an alarm when the output signal exceeds the threshold. Although automatic detection has wide application, there are other detection methods. From the early days of radar, radar operators have detected targets on cathode ray tube displays. The persistence of the tube results in noncoherent pulse integration within the beamwidth of a scanning antenna. A useful approximation to the performance of an operator is the use of the automatic detection curves presented in the previous subsection with an operator integration time[13] between 1 and 10 s.

Binary integration is another form of detection sometimes used. With binary integration, a threshold detector and an accumulator follow the envelope detector, and the accumulator counts m threshold crossings and triggers an alarm when there are m crossings out of n pulses. The optimum setting for m is approximately $1.5\ n^{1/2}$, but the setting has a very broad optimum, and it is not critical. Performance with binary integration is within about 1.5 dB of ideal video integration.[8,13]

Radar may require constant false alarm rate (CFAR) circuitry when the background noise power, jamming power, or clutter power varies in some unknown manner, and the radar must maintain a constant false alarm rate.[8,13] A CFAR processor samples reference cells in range, Doppler, or angle; obtains an estimate of the average background level; and sets the detection threshold so that the false alarm rate remains constant. In some cases when there is a clutter background with non-Rayleigh statistics, the CFAR circuitry also measures higher order statistics of the probability density function to make a better estimate of the required detection threshold. Because there will be an error in the estimate of the background level, it is necessary to increase the detection threshold above the value that the background level would normally require, and the probability of detection decreases. The CFAR loss is the increase in the SNR required to restore the desired probability of detection.

3.2.6 Radar Measurement Accuracy

There is often a requirement for the extraction of target information after radar has detected a signal in a noise background. Noise limits the accuracy with which radar can extract information on the fundamental radar parameters normally of interest: range, Doppler frequency shift, and the angle of arrival.

Various authors have demonstrated that the theoretical rms error δM of a radar measurement M is[8,9,12,22,23]

$$\frac{\delta M}{M} = \frac{k}{(2E/N_0)^{1/2}} , \tag{3.50}$$

where k is a constant close to unity, E is the received signal energy, and N_0 is the noise spectral density. As discussed earlier, E/N_0 is the maximum output SNR for a matched filter. This formula is valid for E/N_0 greater than about 4; below 4, there are additional noise terms to account for.[8] When radar is less efficient than a matched filter, there is a matching loss L_m, where $L_m > 1$, and

$$\frac{\delta M}{M} = \frac{k\sqrt{L_m}}{(2E/N_0)^{1/2}} . \tag{3.51}$$

3.2.6.1 Range Measurement Accuracy. Radar measures range by determining the time delay of the return pulse. The theoretical rms time delay error δT is

$$\delta T = \frac{1}{\beta(2E/N_0)^{1/2}} , \tag{3.52}$$

and the rms range error δR is

$$\delta R = \frac{c\delta T}{2} . \tag{3.53}$$

Here β is the effective bandwidth of the transmitted signal, i.e., the normalized second moment of the transmitted signal spectrum about the mean. Table 3.3 lists effective noise bandwidths for pulse shapes and for pulse spectra.[8,9]

Equation (3.52) is valid for a single-pulse or a coherent-pulse train for a matched filter. When the radar is not a perfectly matched filter, the mismatch loss reduces the SNR and the measurement efficiency. For a train of n noncoherent pulses,

$$\delta T \cong \frac{1}{\beta^*[2n(\mathrm{SNR})_1]^{1/2}} , \tag{3.54}$$

where $(\mathrm{SNR})_1$ is the single-pulse SNR achieved with the mismatch loss, and β^* is slightly less than β.

Table 3.3 Effective Noise Bandwidths for Range Measurements

(A) For Pulse Shapes (from Ref. 9)	
Pulse Shape	Effective Bandwidth β
Rectangular of width τ passed through filter of bandwidth B	$(2B/\tau)^{1/2}$
Trapezoidal of width $2T_1$, rise and fall times T_2	$[3/(T_2^2 + 3T_1T_2)]^{1/2}$
Gaussian with half-power bandwidth B	$\pi B/1.18$

(B) For Signal Spectra (from Ref. 8)		
Signal Spectra	$A(f)$	β/B
Rectangular	$1, \lvert f\rvert < 0.5B$	1.81
Triangular	$1 - \lvert 2f/B\rvert$	0.99
Parabolic	$1 - (2f/B)^2$	1.19
Cosine	$\cos(\pi f/B)$	1.14
Cosine squared	$[\cos(\pi f/B)]^2$	0.89
Gaussian, $B = 6\sigma_f$	$\exp(-f^2/2\sigma_f^2)$	0.74

As an example, assume a Gaussian pulse with a half-power bandwidth of 2 MHz so that $\beta = 5.3 \times 10^6$ [from Table 3.3(A)], $n = 30$, and $(\text{SNR})_1 = 3.16$ (5 dB). Then,

$$\delta R = 1.5 \times 10^8 \times \frac{1}{5.3 \times 10^6 (2 \times 30 \times 3.16)^{1/2}} = 2.1 \text{ m} .$$

3.2.6.2 Doppler Frequency Measurement Accuracy. The theoretical minimum rms error in the measurement of Doppler frequency during a single pulse is

$$\delta f = \frac{1}{\alpha(2E/N_0)^{1/2}} . \tag{3.55}$$

Here α is the rms time duration of the signal; i.e., $(\alpha/2\pi)^{1/2}$ is the normalized second moment of the signal squared about the mean. As an example, if $\alpha = 2$ s and $E/N_0 = 10$, the rms Doppler frequency measurement error is

$$\delta f = \frac{1}{2(2 \times 10)^{1/2}} = 0.11 \text{ Hz} .$$

When measurements are made on a single pulse, the pulse is often rectangular, and it can be shown that for the rectangular pulse,[8]

$$\alpha = \left(\frac{\pi}{\sqrt{3}}\right)\tau = 1.81\tau = \frac{1.61}{B_3} , \tag{3.56}$$

where τ is the pulse width and B_3 is the -3-dB spectrum bandwidth. For a Gaussian-shaped pulse,

$$\alpha = \frac{1.18}{B_3} , \tag{3.57}$$

where B_3 is the half-power width of the spectrum.

When measurements are made with a noncoherent pulse train of n pulses, and there is a mismatch loss

$$\delta f = \frac{1}{\alpha^*[2n(\mathrm{SNR})_1]^{1/2}} , \tag{3.58}$$

where α^* is slightly less than α because of the reduced measurement precision, and $(\mathrm{SNR})_1$ is the single-pulse SNR achieved. For measurements on a coherent pulse train, the effective signal duration for use in Eq. (3.55) is

$$\alpha = 1.81 t_0 , \tag{3.59}$$

where t_0 is the length of the pulse train.

3.2.6.3 Angular Measurement Accuracy. It can be shown that the minimum rms angular error $\delta\theta$ in the presence of noise is

$$\delta\theta = \frac{K^*\theta_B}{(2E/N_0)^{1/2}} , \tag{3.60}$$

where K^* is a constant that depends on the aperture distribution and θ_B is the -3-dB antenna beamwidth. Table 3.4 lists values of K^* for a number of aperture distributions. The rms angular error for a tracking scan radar is[23]

$$\delta\theta = \frac{k_1\theta_B}{k_s[B\tau(\mathrm{SNR})(f_r/\beta_s)]^{1/2}} , \tag{3.61}$$

where

k_1 = a constant
k_s = angular error detection slope
SNR = SNR per pulse
f_r = pulse repetition frequency

Table 3.4 Angular Measurement Constants

Aperture Distribution $\lvert x\rvert < D/2$		K^*
Uniform	$A(x) = 1$	0.628
Parabolic	$A(x) = 1 - 4x^2/D^2$	0.936
Cosine	$A(x) = \cos(\pi x/D)$	0.734
Triangular	$A(x) = 1 - 2\lvert x\rvert/D$	0.793

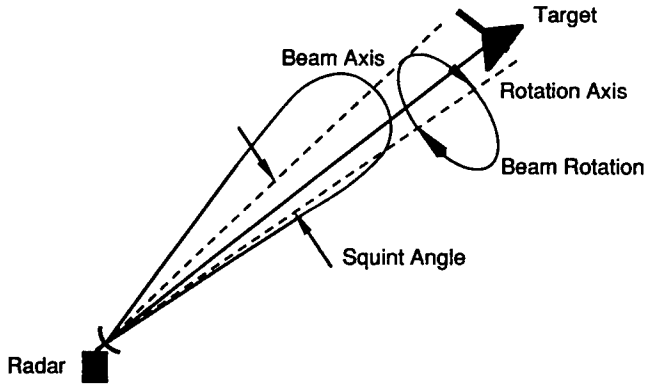


Fig. 3.16 Conical scan angular tracking.

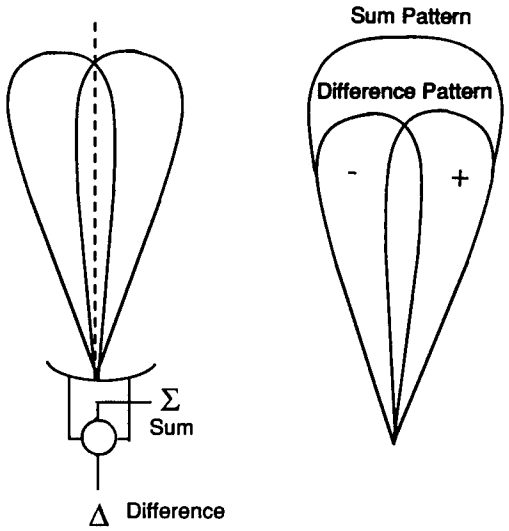


Fig. 3.17 Single-channel amplitude comparison monopulse.

β_s = servo bandwidth
B = receiver bandwidth
τ = pulse width.

Conical scan radar is one type of angular tracking radar in common use. With conical scan radar, the antenna rotates about an axis, and the peak return from the target during the rotation determines the target angle; see Fig. 3.16. Conical scan radar has the advantage of simplicity, but target cross-section scintillations limit performance at short range. For conical scan radar, k_1 = 1.4 and k_s = 1.5 for the optimum offset angle.

Another common angular tracking radar is monopulse radar, which finds the angular location of the target by processing the sum (Σ) and difference (Δ) returns from two or more beams simultaneously. Figure 3.17 shows the an-

tenna beams for one channel of an amplitude comparison monopulse; tracking radar will normally have both azimuth and elevation channels. A typical value for k_m is 1.57. Monopulse radar is more complicated than conical scan radar, but the measurement efficiency and data rate are higher, and target scintillation does not affect the measurements. Antenna crossover losses with a conical scan result in an rms error 1.8 times higher than for monopulse, and the range performance of monopulse radar is a factor of 1.3 longer.[23] For a monopulse radar, $k_1 = 1$ in Eq. (3.61), and a typical value for k_s is 1.57.

As an example, consider monopulse radar with the parameters: $\theta_B = 0.5$ deg, $k_s = 1.57$, $B = 1.2$ MHz, $\tau = 1.0$ μs, SNR = 1, $f_r = 1000$, and $\beta_n = 2.5$ Hz. The rms angular error is

$$\delta\theta = \frac{0.5}{1.57(1.2 \times 10^6 \times 10^{-6} \times 1 \times 1000/2.5)^{1/2}} = 0.0145 \text{ deg} .$$

3.2.7 Synthetic Aperture and Inverse Synthetic Aperture Radars

3.2.7.1 ISAR Principles. Synthetic aperture radar (SAR) or inverse synthetic aperture radar (ISAR) provide cross-range resolution that is much finer than that obtainable with the azimuth beamwidth of the radar antenna.[24–30] SAR and ISAR differ in that with SAR, the radar motion provides the change in aspect angle required for fine resolution, whereas target motion causes the corresponding change in aspect angle for ISAR. Millimeter-wave instrumentation radars often have an ISAR mode to form range-Doppler images of a complex target; applications include imaging airborne targets, tumbling space objects, or ships at sea with roll, pitch, and yaw motions. SARs and ISARs are coherent because phase information must be preserved to generate the synthetic aperture.

Consider the geometry sketched in Fig. 3.18, where a target scene is rotating at a constant angular velocity Ω. Assuming that the rotating target is at long range, the lines of constant range and constant Doppler shift form an approximately rectangular grid covering the target. The range r to a point x_0,y_0 in the rotating field is

$$r = r_a + x_0 \sin\Omega t + y_0 \cos\Omega t , \tag{3.62}$$

where r_a is the range to the center of rotation. The Doppler frequency for x_0,y_0 is

$$f_d = \frac{2}{\lambda}\frac{dr}{dt} = \frac{2}{\lambda}\left(\frac{dr_a}{dt} + x_0\Omega \cos\Omega t - y_0\Omega \sin\Omega t\right) . \tag{3.63}$$

For a small value of t,

$$f_d \cong \frac{2}{\lambda}\frac{dr_a}{dt} + \frac{2\Omega x_0}{\lambda} . \tag{3.64}$$

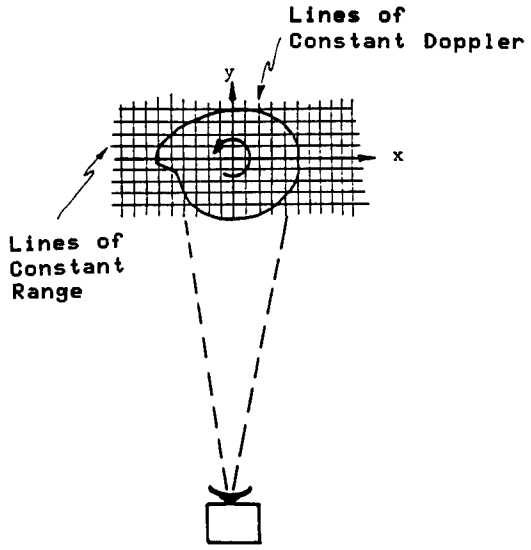


Fig. 3.18 Lines of constant Doppler and constant range for a rotating scene.

Assume a stationary platform or that the first term on the right side of Eq. (3.64) is removed by the radar. The minimum resolvable Doppler frequency Δf_d is related to the minimum resolvable change in x_0, Δx_0, by

$$\Delta f_d = \frac{2\Omega \Delta x_0}{\lambda} = \frac{2\Omega \rho_a}{\lambda} \ . \tag{3.65}$$

Here ρ_a is the azimuth or cross-range resolution, which equals Δx_0. Because the minimum resolvable Doppler frequency is the inverse of the integration time T,

$$\rho_a = \frac{\lambda}{2\Omega T} = \frac{\lambda}{2\Delta\theta} \ , \tag{3.66}$$

where $\Delta\theta$ is the integration angle.

Let r_0 be the radius from the center of rotation to x_0, y_0. The range resolution must satisfy

$$\rho_r > r_0 \Delta\theta = \frac{r_0 \lambda}{2\rho_a} \ . \tag{3.67}$$

Thus, the motion of the point x_0, y_0 through range and Doppler cells limits the product of the range and azimuth resolution

$$\rho_r \rho_a > \frac{\lambda r_0}{2} \ . \tag{3.68}$$

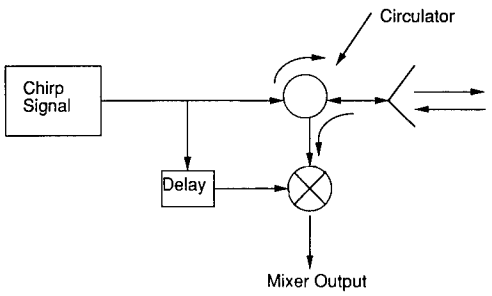


Fig. 3.19 Simplified schematic for an ISAR radar.

To obtain fine range resolution, ISAR must transmit a pulse with a wide bandwidth such as a chirp or stepped frequency waveform. Figure 3.19 is a simplified schematic of ISAR using a chirp waveform; the received echo mixes with the transmitted signal to provide a difference frequency proportional to the product of the FM rate of the chirp signal and the target range. Figure 3.20(a) is a sketch of the mixer output. Walker[25] showed that recording the data in a polar format, as shown in Fig. 3.20(b), eliminates the problem of motion through range and Doppler cells and the resulting resolution restriction of Eq. (3.68). In the polar recording format, individual pulses are recorded as radial line segments on a recording medium, which rotates in synchronism with the target, and the recording radius is proportional to the transmitted frequency. With polar format recording, a point target forms a grating of parallel lines as indicated in Fig. 3.20(c), and taking a two-dimensional Fourier transform of all or a portion of the grating provides an image of the target scene.

Figure 3.21 is an ISAR image of a jeep collected at a center frequency near 35 GHz with a resolution of about 5 cm. The recording medium for this image was photographic film, and the image is the result of the noncoherent summation of contributions from a 360-deg target rotation in an optical processor. Most ISAR processing is now digital.

3.2.7.2 SAR Principles. SAR uses radar motion to generate the radar images. The original SAR mode was the strip-map mode: in the strip-map mode, the SAR moves along a straight line, as shown in Fig. 3.22, while transmitting pulses and recording the amplitude and phase of the returns along the line.[24] If D is the horizontal aperture of the antenna, and λ is the wavelength, the beamwidth β of the physical antenna is approximately

$$\beta = \frac{\lambda}{D} . \tag{3.69}$$

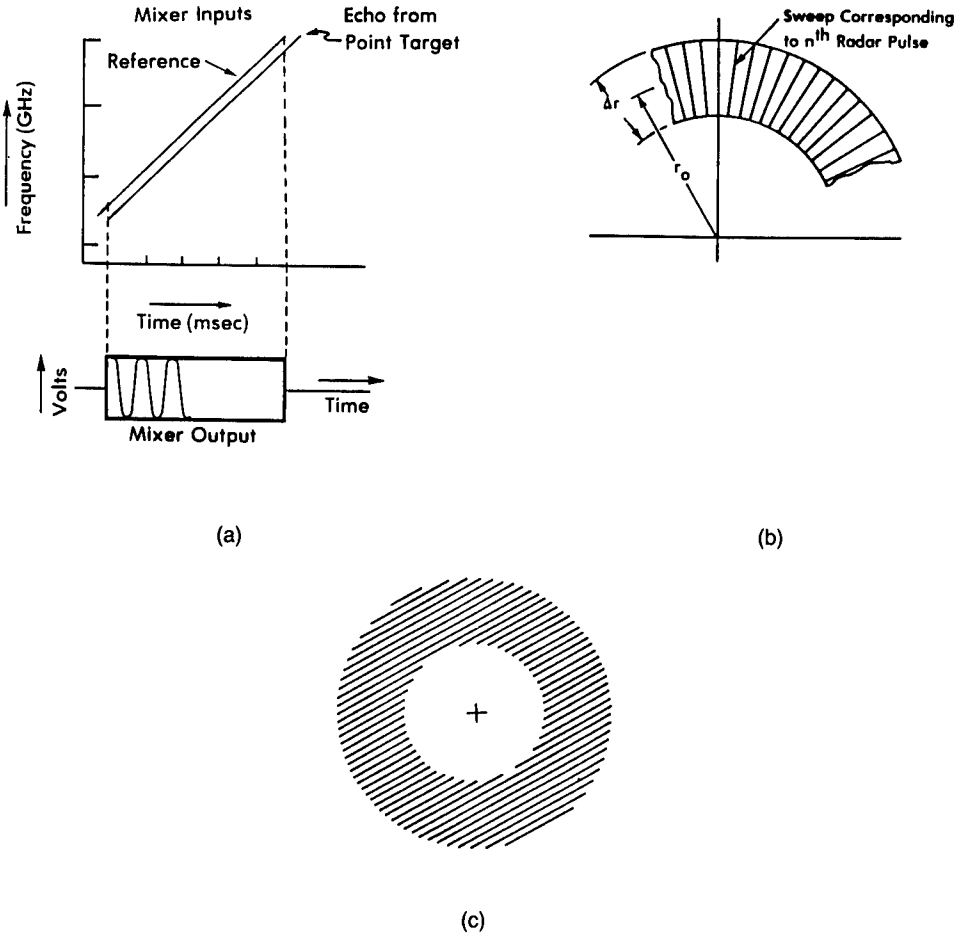


Fig. 3.20 Polar format recording: (a) signal format with chirp signal, (b) polar format, and (c) point target grating.

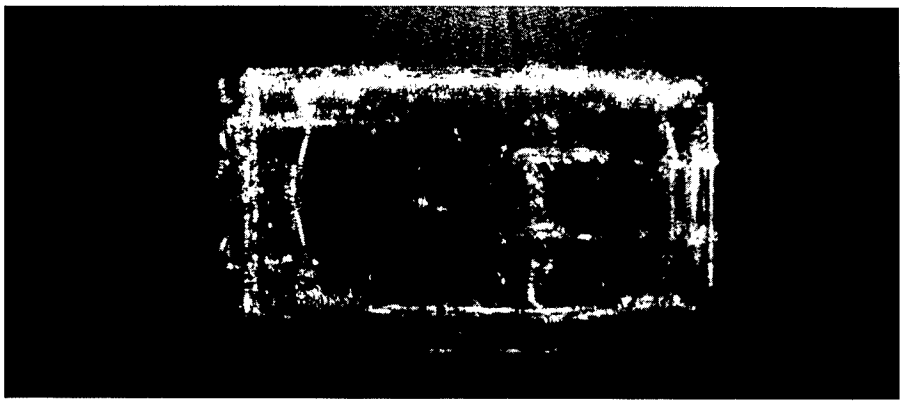


Fig. 3.21 A 35-GHz image of a jeep; vertical polarization, 4.5-deg elevation angle. (Courtesy of ERIM)

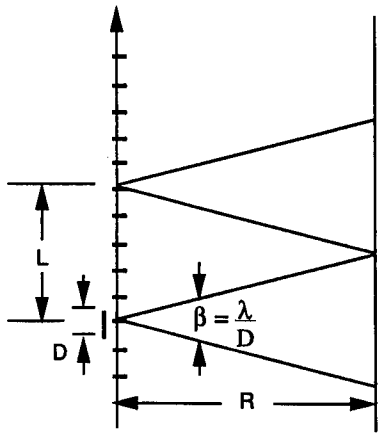


Fig. 3.22 Generation of a synthetic array.

Thus, the effective length of the synthetic array L_{eff} is

$$L_{\text{eff}} = \frac{R\lambda}{D} \ . \tag{3.70}$$

Because of the round-trip phase shift along the synthetic array, the effective beamwidth β_{eff} of the synthetic array is half as wide as that of a physical antenna or $\lambda/2L_{\text{eff}}$, and the cross-range resolution ρ_a is

$$\rho_a = \frac{\lambda R}{2L_{\text{eff}}} = \frac{D}{2} \ . \tag{3.71}$$

The azimuth resolution is therefore independent of both the range and the wavelength, and a smaller physical antenna provides finer resolution because of its wider physical beamwidth.

This derivation assumes that the target is in the far field of the synthetic array; however, the array is normally long enough so that there is a quadratic phase error that increases from the center to the edges of the array. Coherently summing the returns from the synthetic array without compensating for this phase error generates an unfocused array with resolution greater than $D/2$. Nearly all SAR systems are focused: i.e., the processor compensates for the quadratic phase error along the aperture during image formation to obtain a resolution of $D/2$. It can be shown that Eq. (3.71) is a special case of Eq. (3.66) derived for an ISAR.

There are three SAR modes, as sketched in Fig. 3.23. In the strip-map mode just discussed, the side-looking antenna points at a fixed angle from the ground track, usually 90 deg, and illuminates a strip of terrain parallel to the ground track. In the spotlight mode, the antenna illuminates a fixed point on the ground as it flies by to obtain improved azimuth resolution at the sacrifice of area coverage. The scan mode can be thought of as a universal mode: in this mode, the center of the antenna beam illuminates a scan path with arbitrary

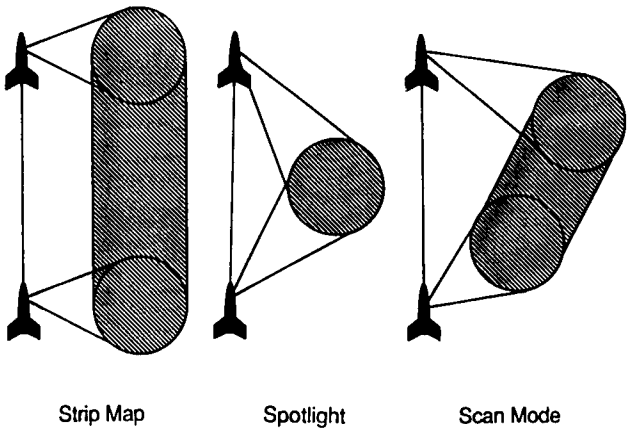


Fig. 3.23 SAR modes.

orientation with respect to the flight path to trade off resolution and coverage. The strip-map and spotlight modes are really special cases of the scan mode. References 24 to 30 discuss SAR systems in more detail; a key factor in collecting high-quality imagery is compensation for the deviation of the aircraft motion from a straight-line path.

3.3 MILLIMETER-WAVE PHENOMENOLOGY

3.3.1 Radar Targets

3.3.1.1 Radar Cross Section. The radar cross section (RCS) characterizes the radar scattering properties of a target.[31–34] The RCS of a target is the area that intercepts an amount of power that, when scattered isotropically, produces an echo equal to that received from the target. From the definition, the RCS of a target that scatters uniformly in all directions equals its projected area. A metal sphere whose radius a is large compared with the wavelength is an isotropic scatterer, and its RCS equals its projected area, πa^2. The RCS σ is defined as

$$\sigma = \lim_{R \to \infty} 4\pi R^2 \left| \frac{E_s}{E_i} \right|^2 , \tag{3.72}$$

where R is the distance from the point of observation, E_s is the magnitude of the scattered electric field, and E_i is the magnitude of the incident electric field. The RCS is normally defined when the incident field is a plane wave, which implies that the illuminating source is far from the scattering body. The RCS depends on the transmitted wavelength, the polarization, the target orientation, and the dielectric or conductive properties of the target. RCS is normally specified in square meters or in decibels relative to 1 m^2 (dBsm).

3.3.1.2 Frequency Regimes and Millimeter-Wave Scattering. The sphere is a three-dimensional target whose RCS can be computed exactly at all frequencies from equations derived by Mie in 1908; the scattering properties of a sphere indicate the general wavelength dependence of radar targets. Figure 3.24 is a plot of the normalized RCS of a conducting sphere showing the three scattering frequency regions: the Rayleigh region, the resonance region, and the optics region. Asymptotic computational methods are valid in the Rayleigh and optics regions.

Rayleigh Scattering. In the Rayleigh region, the wavelength λ is much larger than the dimensions of the body, and the RCS is proportional to the inverse fourth power of the wavelength. At millimeter wavelengths, scattering from hydrometeors often falls within the Rayleigh region. The RCS is proportional to the square of the volume with some modification for particle shape.

Resonance Scattering. In the resonance region, the wavelength is of the order of the dimensions of the body, and the backscattered field oscillates with wavelength changes. The backscattered power can be found from integral equation methods, which are limited to body sizes less than about 10 wavelengths.

High-Frequency or Optics Scattering. In the optics region, the body dimensions are much larger than the wavelength. The RCS of a sphere approaches a constant value of πa^2. A complex body can be treated as a collection of scattering centers that act independently of one another. The RCS is computed with geometrical optics, physical optics, the geometric theory of diffraction, the physical theory of diffraction, and extensions of the latter two techniques.

Scattering at millimeter-wave frequencies is similar to scattering at microwave frequencies with the exceptions that the target is more likely to be in

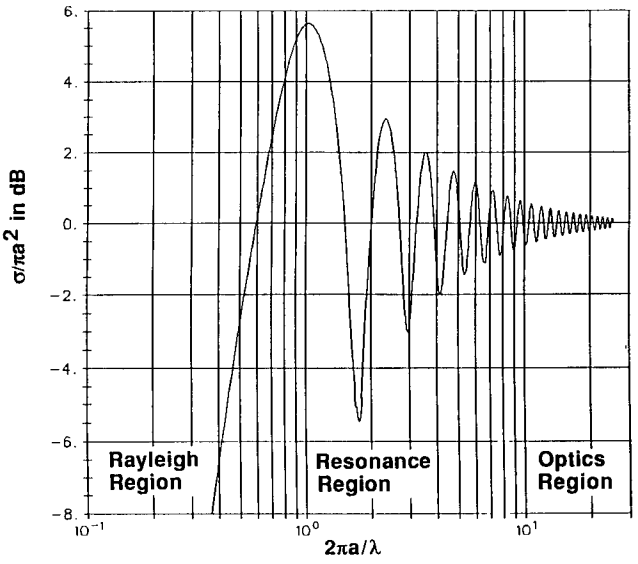


Fig. 3.24 Normalized RCS of a conducting sphere.

the near field and that the surface roughness is a larger fraction of the wavelength.[35] The usual far-field criterion is that the incident field be within $\pi/8$ rad of planar, which requires

$$R > \frac{2D^2}{\lambda} , \tag{3.73}$$

where R is the range from the source to the target and D is the largest target dimension. Because λ is so small, millimeter-wavelength scattering often occurs in the near field. Except at very close range, a complex target may be in the near field when considered as a whole, but the individual scattering centers, which act independently of one another, can still be in the far field. In this case, the average RCS is the same as it would be in the far field, but there are changes in the RCS pattern such as shifts in the locations and magnitudes of peaks and nulls. At aspects where a long scatterer such as a flat plate or a cylinder is dominant, the scattering from the simple shape will change as the range decreases: the nulls fill in, the sidelobes increase, and the peak RCS decreases.

At microwave frequencies, the assumption that a metallic target is perfectly smooth is usually good; however, the surface roughness becomes a higher fraction of the wavelength as the wavelength decreases. The roughness decreases the RCS compared with what would be expected for smooth, metallic shapes. The calculated RCS must be modified to account for the roughness and the Fresnel reflection coefficient when the surface roughness is somewhat greater than a thousandth of a wavelength.[35]

3.3.1.3 Simple Shapes. This subsection discusses the RCS of some smooth and perfectly conducting simple shapes when the wavelength is small compared to the body dimensions so that high-frequency techniques are applicable. Figure 3.25 shows sketches of most of these simple shapes.

Convex Surface. Geometric optics techniques show that the RCS of a convex surface with principal radii of curvature a_1 and a_2 is[34]

$$\sigma = \pi a_1 a_2 . \tag{3.74}$$

A special case is the sphere where $a_1 = a_2$.

Flat Plates. Physical optics provides the RCS of a flat plate.[32] The maximum RCS at normal incidence is

$$\sigma_m = \frac{4\pi A^2}{\lambda^2} , \tag{3.75}$$

where A is the area of the plate and λ is the wavelength. Neglecting edge effects, the RCS at an angle θ from the normal to a circular flat plate of radius a is

$$\sigma = \frac{\pi a^2}{(\tan\theta)^2}\left[J_1\left(\frac{4\pi a \sin\theta}{\lambda}\right)\right]^2 . \tag{3.76}$$

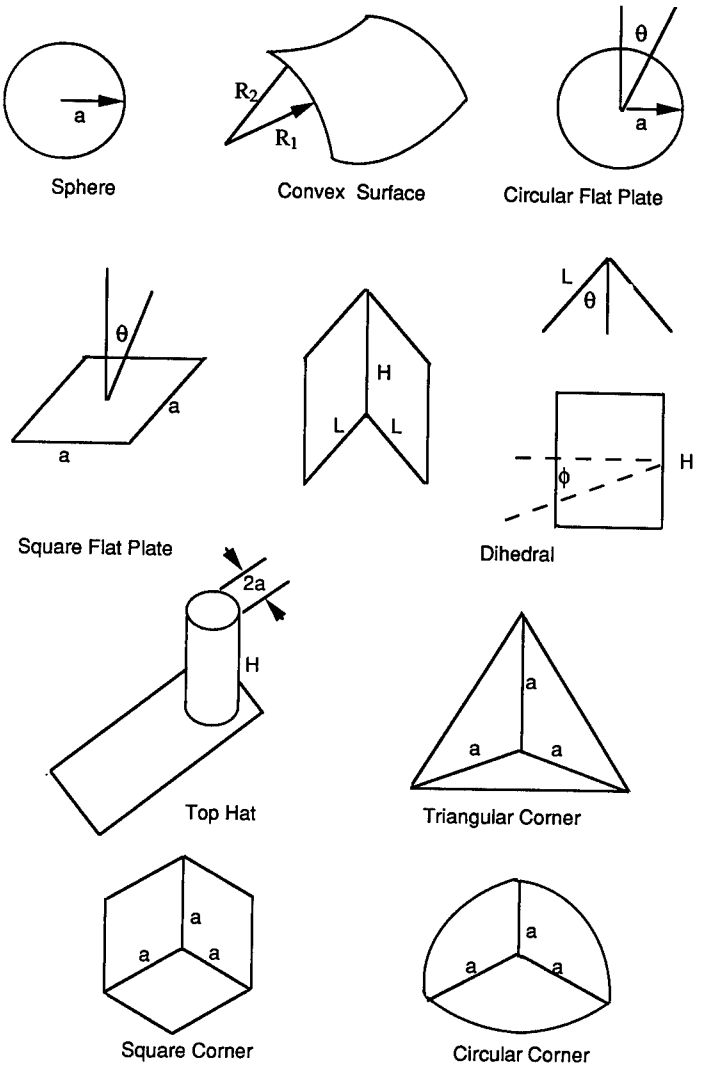


Fig. 3.25 Reflectors with simple shapes.

For a square flat plate with sides of length a parallel to the x and y axes,

$$\sigma = \frac{4\pi a^4}{\lambda^2}\left[\frac{\sin(ka\,\sin\theta)}{ka\,\sin\theta}\right]^2(\cos\theta)^2, \quad k = \frac{2\pi}{\lambda}, \tag{3.77}$$

where θ is the angle from the normal (the z axis) in either the xz or yz planes.

Circular Cylinder. From physical optics, the RCS of a cylinder of length L and radius a is[31]

$$\sigma = \frac{2\pi a L^2}{\lambda}\left[\frac{\sin(kL\,\sin\theta)}{kL\,\sin\theta}\right]^2(\cos\theta)^2, \tag{3.78}$$

where θ is the angle between the normal to the cylinder side and the point of observation; θ is in a plane containing the cylinder axis and the normal.

Dihedral. Refer to Fig. 3.25 for the angular coordinate system. For two-bounce reflections when the viewing angle is normal to the crease ($\phi = 0$),

$$\sigma(\theta) = \frac{8\pi L^2 H^2(1 - \sin 2|\theta|)}{\lambda^2} \quad . \tag{3.79}$$

If θ is now held fixed, and ϕ is varied, the reflector behaves like a flat plate whose normal is fixed in the θ direction:

$$\sigma(\theta,\phi) = \sigma(\theta)\left[\frac{\sin(kH\sin\phi)}{kH\sin\phi}\right]^2 (\cos\phi)^2 \quad . \tag{3.80}$$

Top Hat Reflector. Figure 3.25 shows a sketch of the top hat reflector whose height is h and whose cylinder radius is a. In the two-bounce reflection region where the cylinder is completely imaged in the ground plane, and the incident polarization is parallel or perpendicular to the cylinder axis, the RCS is[36]

$$\sigma \cong \frac{8\pi a h^2 \cos\psi}{\lambda} \quad , \tag{3.81}$$

where ψ is the grazing angle above the ground plane. The ground plane length must be greater than $h \cot\psi$ for complete imaging in the ground plane.

Corner Reflectors. Corner reflectors are frequently used for RCS calibration.[32–34] The maximum RCS and the angular coverage for the trihedral corner reflector shapes in Fig. 3.25 are

Type	Maximum RCS	Angular Coverage (−3 dB)
Triangular	$4\pi a^4/(3\lambda^2)$	40 deg
Circular	$15.6a^4/\lambda^2$	32 deg
Square	$12\pi a^4/\lambda^2$	23 deg

The peak return occurs at an elevation angle of 35.3 deg when the bottom side of the corner is horizontal. Figure 3.26 shows the relative responses of the three corners with $a^2/\lambda = 1.545$.

3.3.1.4 Scattering Matrix. Conventional radar transmits and receives the same linear polarization; but radar systems are now being built that exploit the polarization properties of a target to improve target discrimination or detection.[37–39] The scattering matrix describes the polarization properties of a target. Assuming that the radar operates with a linear polarization base, we can represent the incident electric field with horizontally and vertically polarized orthogonal components E_H^T and E_V^T and the backscattered field with horizontally and vertically polarized orthogonal components[40] E_H^S and E_V^S. These components will be in phase with a linearly polarized field and out of phase

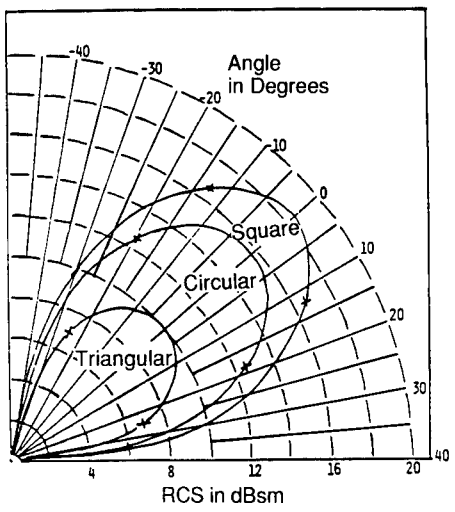


Fig. 3.26 Trihedral corner reflector response.

with an elliptically polarized field. It can be shown that the scattered field is related to the transmitted field by matrix multiplication

$$\begin{vmatrix} E_H^S \\ E_V^S \end{vmatrix} = \begin{vmatrix} a_{HH} & a_{VH} \\ a_{HV} & a_{VV} \end{vmatrix} \begin{vmatrix} E_H^T \\ E_V^T \end{vmatrix} . \tag{3.82}$$

In the elements of the scattering matrix, a_{HH}, a_{VH}, a_{HV}, and a_{VV}, the first subscript denotes the transmitted field, and the second subscript, the scattered field. (This radar notation referring to the element in the matrix is opposite to the practice customary with matrices.) The elements a_{VH} and a_{HV} are equal for a monostatic radar.

The scattering matrix contains all the polarization properties of the target. The matrix elements are proportional to the square root of the RCS; a target has parallel- and cross-polarized cross sections. A scattering matrix also exists for a circular polarization base, and the scattering matrix for linear polarization can be converted to a circular base and vice versa.[41] The scattering matrix for an odd-bounce scatterer, such as a convex surface, a flat plate, a cylinder, or a corner reflector, is

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$$

for the H-V polarization base. For an even-bounce scatterer such as the dihedral or the top hat, the scattering matrix is

$$\begin{vmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{vmatrix} .$$

Here the dihedral crease is assumed to be horizontal. The axis of the top-hat cylinder is assumed to be vertical, and α is the angle the electric field makes with the horizontal.

3.3.1.5 Complex Targets. The total backscattered field for a target at short wavelengths is the vector summation of the contributions of the individual scattering centers that act independently of one another.[32,34] With the relative phase method for finding the RCS, the total RCS is the vector summation of the contributions of the individual scattering centers

$$\sigma_p = \left| \sum_{i=1}^{n} (\sigma_i)^{1/2} \exp(j\phi_i) \right|^2 . \tag{3.83}$$

Here σ_i is the RCS of the i'th scattering center and ϕ_i is the phase of the i'th scattering center. Each of the individual scattering centers is, of course, aspect dependent.

The phase angles of the individual scattering centers are usually not known or only known approximately. As a consequence, it is not possible to compute the total RCS with a vector summation, and it is necessary to use statistical techniques. We assume that ϕ_i has a uniform distribution from 0 to 2π. On averaging over the ϕ_i, we obtain an average RCS σ,

$$\sigma = \sum_{i=1}^{n} \sigma_i , \tag{3.84}$$

and an rms spread S,

$$S^2 = \left(\sum_{i=1}^{n} \sigma_i \right)^2 - \sum_{i=1}^{n} \sigma_i^2 . \tag{3.85}$$

An examination of RCS data has shown that about 60% of the time, the measured RCS pattern is between $\sigma + S$ and $\sigma - S$.

Figure 3.27 is an example of a RCS plot for a complex target. This plot shows[42] the relative RCS of a 1/25 scale model of a Chieftain tank at a measurement frequency of 100 GHz. The measurements provide RCS data for a full-size target at a frequency of 1/25 of the measurement frequency, or 4 GHz. Note the rapid RCS changes with small azimuth angle changes and the large specular returns, which are typical of vehicular targets, at the front, rear, and sides of the model.

Measured RCS values for tactical targets remain fairly constant over the frequency range from 10 to 100 GHz. As the frequency increases, the surface roughness and curvature increase relative to the wavelength, which offsets the expected increase in the scattering center RCS values resulting from the smaller wavelength. Table 3.5 has typical RCS values for tactical targets taken from Refs. 43 and 44 and other sources. Figure 3.28 shows[43] the measured RCS distribution for numerous vehicles at 35 GHz.

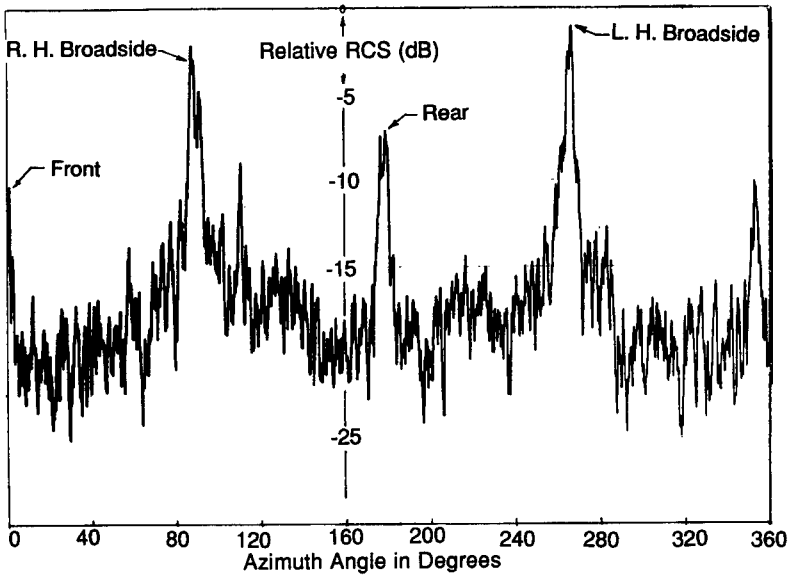


Fig. 3.27 RCS plot of a 1/25 scale model of a Chieftain tank[42]; measurement frequency of 100 GHz.

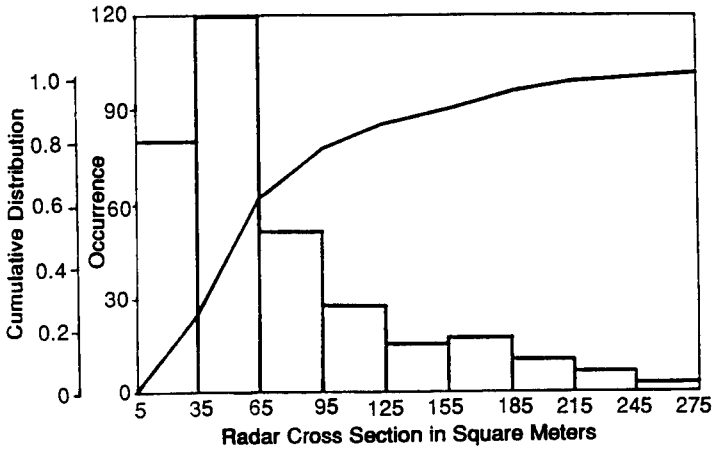


Fig. 3.28 Measured RCS distribution of numerous vehicular targets.[43] (© 1981 Academic Press, reprinted with permission)

Table 3.5 Typical RCS Values for Tactical Targets (in dBsm)

Large truck	18
Tank	15
APC	10
Scout car	7
Man	1

3.3.2 Millimeter-Wave Propagation

3.3.2.1 Introduction. The propagation characteristics at millimeter-wave frequencies, 30 to 300 GHz, lie between those at microwave and optical frequencies. Although atmospheric attenuation in adverse weather at millimeter-wave frequencies is much higher than at microwave frequencies, the attenuation is much less than at optical frequencies, and a millimeter-wave sensor provides much better adverse-weather range performance than an optical sensor.[5,45–48] Figure 3.29 shows the atmospheric attenuation in the millimeter, submillimeter, infrared, and visible frequency bands.

The attenuation and backscatter along the path from a radar to the target have a significant effect on performance at millimeter wavelengths. The attenuation lowers the SNR, and the backscatter increases the clutter background around the target. There is considerable variability and uncertainty

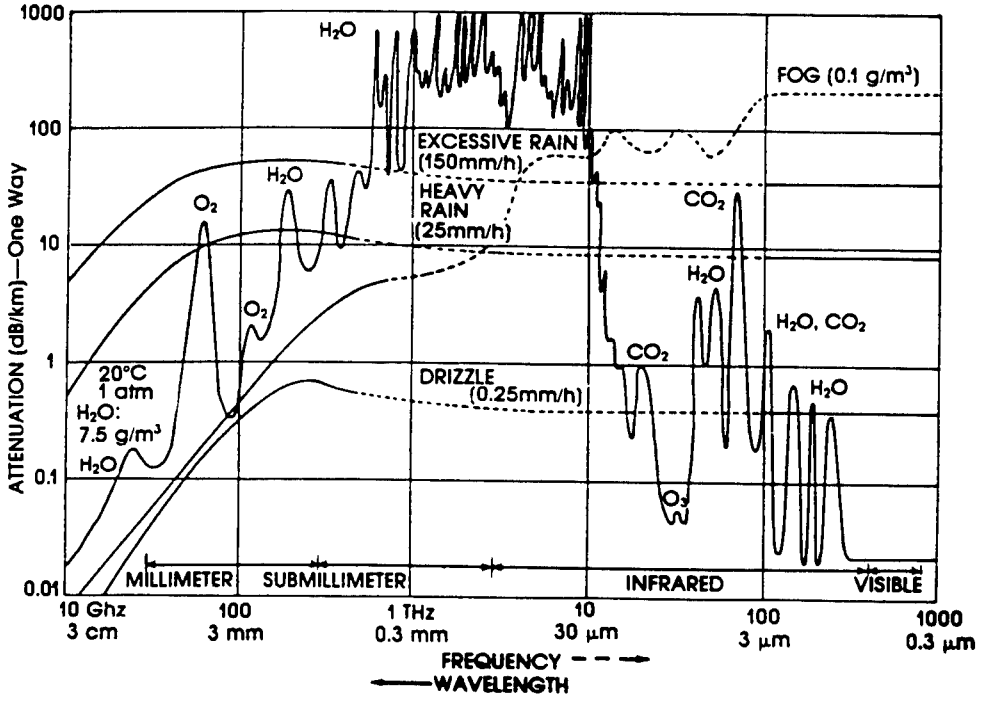


Fig. 3.29 Attenuation by atmospheric gases, rain, and fog.[45] [The original version of this material was first published by the Advisory Group for Aerospace Research and Development, North Atlantic Treaty Organization (AGARD/NATO) in Conference Proceedings CP 245, "Millimeter wave and submillimeter wave propagation and circuits," in February 1979.]

in existing millimeter-wave data bases, and accurate predictions of propagation effects are difficult to make[5]; a millimeter-wave propagation model is available for finding millimeter-wave attenuation with computer algorithms.[48]

The following subsections discuss atmospheric refractivity; millimeter-wave attenuation in clear air; attenuation from hydrometeors, dust, and smoke; volume backscatter; and multipath effects.

3.3.2.2 Refractivity. Because the refractive index n of air is slightly greater than unity, the velocity of propagation in air is slightly slower than in a vacuum.[49] At sea level, n is approximately 1.0003, and although n is very close to unity, the difference from unity results in two effects. First, the normal decrease in refractivity with height causes a downward curvature of a ray in the troposphere. Second, the apparent range to a target is greater than in free space because of the slower velocity of propagation. These factors cause radar tracking errors that increase as the target elevation angle decreases.

Because the refractive index n is so close to unity, the refractivity N is usually a more convenient quantity for numerical computation,

$$N = (n - 1) \times 10^6 \ . \tag{3.86}$$

The range error ΔR between propagation in the atmosphere and free space is

$$\Delta R = NR \times 10^{-6} \ , \tag{3.87}$$

where R is the range. If $N = 313$ N-units, a typical value for the continental United States, and R is 10 km, $\Delta R = 313 \times 10{,}000 \times 10^{-6} = 3.13$ m.

The equation for the refractivity is[50]

$$N = \frac{77.6P}{T} + \frac{3.73 \times 10^5 e}{T^2} \ , \tag{3.88}$$

where T is the atmospheric temperature in kelvins, P is the atmospheric pressure in millibars, and e is the partial pressure of water vapor in millibars. Equation (3.88) is accurate to about 0.5% at frequencies[51] below 100 GHz, but above 100 GHz, O_2 and H_2O absorption lines cause some dispersion. Curves of dispersive delay, however, indicate that the dispersion is small enough so that Eq. (3.88) is still accurate to within about 1.5% at frequencies[52] up to 300 GHz. At sea level, values of N are usually between 300 and 350. Above sea level, N decreases approximately as an exponential

$$N(h) \cong N_0 \exp(-h/7.0) \ , \tag{3.89}$$

where N_0 is the refractivity at sea level and h is the height above sea level in kilometers. Reference 49 includes more accurate models for the variation of N with height.

At altitudes below about 4 km, the bending of radio waves is approximately uniform, and ray propagation relative to the earth can be represented by straight lines, if the earth's radius is modified to maintain a constant relative curvature between the ray and the earth. The effective earth radius is the ratio of the modified earth radius to the actual earth radius; a ratio of 1.33 is typical of temperate climates, but values between 5.0 and 0.6 can be found for different atmospheric conditions.[53]

Irregular propagation conditions, which are quite frequent in some parts of the world, can cause anomalous propagation. When a cool, moist air mass is located below a warmer, dryer air mass, there may be a steep gradient in the refractive index, which causes increased downward bending. Rays entering the transition layer between the two air masses at a shallow grazing angle may be trapped in the transition layer, which acts as a duct.[49] Ducting may cause holes in radar coverage.

Turbulence in the troposphere causes random variations in the temperature, pressure, and water vapor content, which induce random fluctuations in the refractive index. Turbulence causes amplitude scintillation, angle of arrival fluctuations, and depolarization. Studies of the effect show that scintillation is not likely to affect millimeter-wave radar, unless the radar is operating at a performance limit, but that angle of arrival fluctuations can approach[54] 350 μrad.

3.3.2.3 Clear Air Attenuation. Atmospheric gases attenuate radio waves in clear air by molecular absorption.[45–49,51,52,55] A quantum level change in the rotational energy results in radio wave absorption at a resonant frequency or over a narrow band of frequencies. Oxygen and water vapor are the only gases that have a significant effect on atmospheric propagation. These gases have resonances at microwave or millimeter-wave frequencies: oxygen has a large number of closely spaced resonances near 60 GHz and a single resonance at 118.7 GHz, whereas water vapor has resonances at 22.2, 183, and 325 GHz.

The total atmospheric attenuation γ_T in decibels is

$$\gamma_T = \int_0^L (\gamma_o + \gamma_w)\, dL \quad [\text{dB}] \; . \tag{3.90}$$

Here γ_o and γ_w are the specific attenuations for oxygen and water vapor in decibels per kilometer, one way, and L is the path length in kilometers; the two-way loss in decibels for radar is twice the one-way loss. Figure 3.30 has plots of the one-way specific attenuations γ_o and γ_w for oxygen and water vapor at the average sea-level pressure of 1013 mbars, a temperature of 15°C, and a water-vapor concentration ρ of 7.5 g/m^3; these plots were computed with the algorithms in Ref. 56. The water-vapor algorithm accounts for the quadratic dependence on the water-vapor density, because quadratic effects are significant at high densities; a high-water-vapor density greatly increases the attenuation. The temperature dependence is approximately -0.6% for water vapor and -1% for oxygen per degree Celsius over the range of -20 to $+40$°C; hence, attenuation increases with decreasing temperature, as well as with increasing water-vapor density. The curves are accurate to about 15% between 963 and 1063 mbars; the curves cannot be scaled to other pressures because

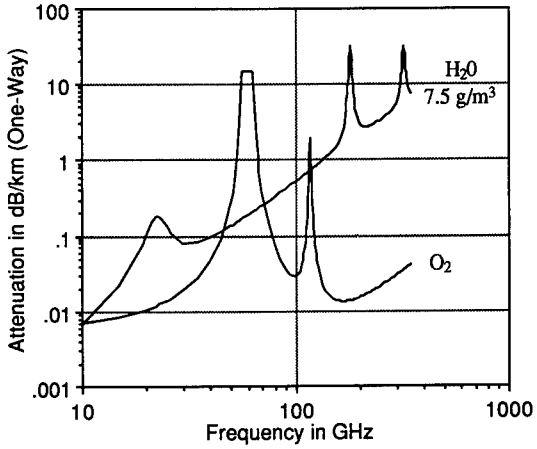


Fig. 3.30 Oxygen and water vapor attenuation: 1013 mbar, 15°C.

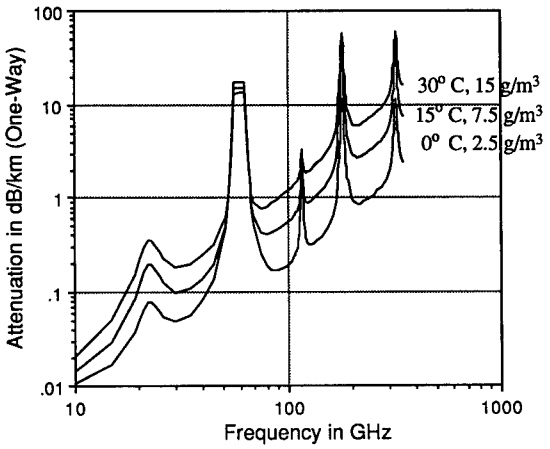


Fig. 3.31 Clear air attenuation at 1013 mbar.

the shapes of the absorption lines change with pressure, and more detailed computational methods are necessary.[52] Figure 3.31 plots the total one-way atmospheric attenuation for three different weather conditions; these plots show that there can be considerable variation in attenuation. Clear air attenuation is high above 100 GHz.

The decrease in clear air attenuation with height is approximately exponential;

$$\gamma = A \exp(-BH) \ . \tag{3.91}$$

Here A and B are constants and H is the height in kilometers. Table 3.6 shows these constants, which were taken from Ref. 46.

Table 3.6 Constants for the Decrease in Clear Air Attenuation with Height (data from Ref. 46)

Frequency	Relative Humidity	A	B	Limits
35 GHz	100%	0.15	0.3816	$H < 7$ km
	50	0.09	0.311	$\gamma = 0, H > 7$ km
	5	0.04	0.1984	
95 GHz	100	0.70	0.4846	$H < 10$ km
	50	0.36	0.4181	$\gamma = 0, H > 10$ km
	5	0.08	0.2677	

3.3.2.4 Hydrometeor Attenuation. Hydrometeor attenuation, which is much higher at millimeter wavelengths than at microwave wavelengths, can cause significant performance degradation for millimeter-wave radars. The term *hydrometeor* refers to condensed water-vapor products in the atmosphere: rain, clouds, fog, hail, ice, or snow. Rain, clouds, or fog normally have the most significant effect on radar operation.

Rain Attenuation. The classical rain attenuation approach assumes an exponential decay of wave intensity during propagation, spherical rain droplets, and that the scattering from a droplet is additive and independent of the scattering from other droplets.[55] Rain attenuation depends on the size of the rain droplets, the distribution of the droplet size, the rain rate, and the droplet attenuation cross section, which is a function of the frequency and the temperature. Theoretical calculations of the scattering from rain droplets depend on Mie's 1908 theory for the scattering of a plane wave from a dielectric sphere. Unfortunately, calculations using Mie's theory are difficult to make, but considerable simplification results when the drop radius r is small compared with the wavelength, $2\pi r \ll \lambda$. In this case, the Rayleigh approximation is valid, and only one term of the series expansion for the droplet attenuation cross section is needed. The Rayleigh approximation is valid[55] up to 40 to 100 GHz.

It has been observed that the specific attenuation γ in decibels per kilometer has the form

$$\gamma = aR_r^b \quad [\text{dB/km}] \ , \tag{3.92}$$

where R_r is the rain rate in millimeters per hour, and a and b are constants that depend on the frequency and temperature. The aR_r^b approximation is in good agreement with Mie calculations.[51] The coefficients a and b have been calculated for rain drops that are not spherical using techniques similar to Mie's classical approach, with the result that the attenuations for vertical and horizontal polarization differ slightly. However, in practice the uncertainties in knowledge of the rain rate, of rain thickness, and in the drop size distribution cause a much greater uncertainty than the polarization difference. Figure 3.32 plots rain attenuation at 20°C for vertical or horizontal polarization for the commonly used Laws and Parsons drop size distribution; these plots were calculated from CCIR data reported in Ref. 51. Rain attenuation data from

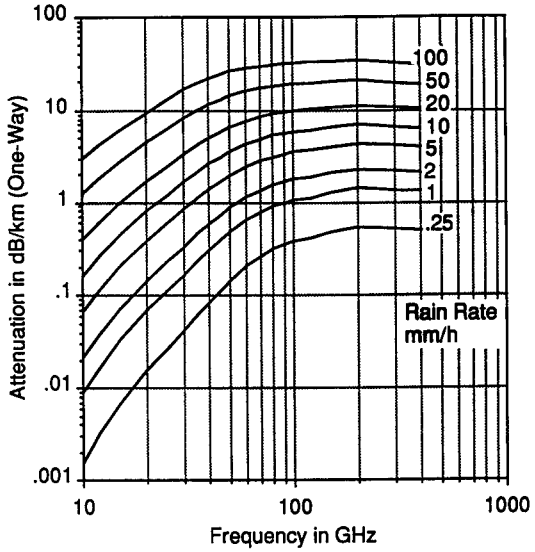


Fig. 3.32 Rain attenuation at 20°C; Laws and Parsons drop size distribution.

other sources will be slightly different. The total attenuation (in decibels) is the sum of the rain attenuation and the clear air attenuation from oxygen and water vapor.

Fog and Cloud Attenuation. Although rain is the hydrometeor having the highest attenuation at millimeter wavelengths, cloud and fog attenuation can also be significant.[57,58] It is difficult to predict cloud and fog attenuation accurately because of the uncertainty of the water content, cloud temperature and thickness, and the lack of experimental data. Cloud and fog attenuation depend on the size of the water droplets, their distribution, their temperature, the index of refraction of water, the extent of the clouds or fog, and the radar wavelength.

Cloud or fog water droplets are usually less than 100 μm in diameter; consequently, the Rayleigh approximation is valid up to about 100 GHz. Absorption, not scattering, is the primary cause of attenuation. Altshuler[57] has developed an expression for cloud or fog attenuation that is valid for wavelengths between 3 and 20 mm, except for the water vapor resonance near 22 GHz and the oxygen resonance near 60 GHz, at a nominal temperature of 10°C:

$$\gamma = -1.2 + 0.0371\lambda + 19.96/\lambda^{1.15} \quad [\text{dB km}^{-1}\,\text{g}^{-1}\,\text{m}^{-3}] \; . \tag{3.93}$$

Here λ is the wavelength in millimeters. The attenuation changes -0.022 dB $\text{km}^{-1}\,\text{g}^{-1}\,\text{m}^{-3}$ per degree of temperature change. Figure 3.33 plots cloud or fog attenuation for values of liquid water content ρ between 0.05 and 2 g/m^3 for frequencies between 20 and 100 GHz.

Liebe has also developed a model to predict the cloud or fog attenuation.[52,58] Figure 3.34 shows plots of cloud or fog attenuation computed with the Liebe model at 20, 10, and 0°C at frequencies up to 400 GHz. Note that the attenuation

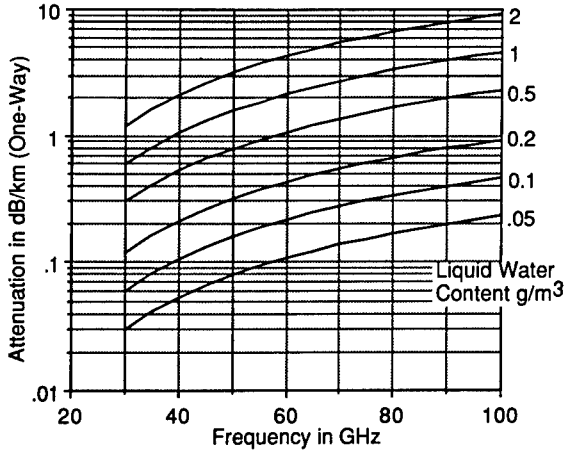


Fig. 3.33 Cloud or fog attenuation at 10°C (Altshuler model).

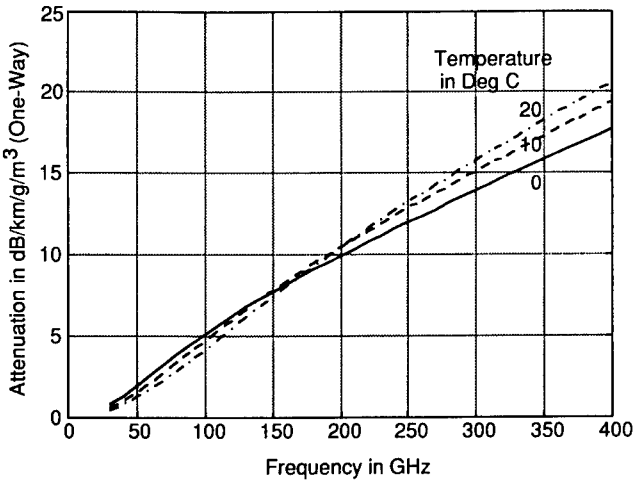


Fig. 3.34 Cloud or fog attenuation (Liebe model).

is higher at 0°C than at 20°C at the lower frequencies, but the reverse is true at the higher frequencies.

The liquid water content ρ of clouds varies from 0.05 to more than 2 g/m^3. Although ρ can be as high as 5 g/m^3 in heavy cumulus clouds, ρ is usually less than 1 g/m^3 in fair weather clouds.[51] The thickness of heavy clouds (≥ 1 g/m^3) is typically 1 to 2 km, whereas the thickness of medium to light clouds can vary from 0.5 to 1 km. For heavy fog, ρ varies from 0.2 to 0.4 g/m^3; ρ is typically 0.02 to 0.06 g/m^3 for medium fog. Typical fog layers extend[51] 50 to 100 m above the ground and seldom more than 150 m.

The total attenuation is the sum of the cloud or fog attenuation and the oxygen and water-vapor attenuation from Fig. 3.31. The water-vapor content

in the cloud or fog layer will be that corresponding to 100% relative humidity (10 g/m^3 at 10°C). For example, the approximate attenuation for a fog with a liquid water content $\rho = 0.2$ g/m^3 is 0.16 dB/km at 35 GHz from Fig. 3.33. From Fig. 3.31, the clear-air attenuation is 0.1 dB/km at 15°C with a water-vapor content $\rho = 7.5$ g/m^3. The attenuation at 10°C increases about 5% because of the lower temperature and about 33% because of the higher water-vapor content for a total increase of approximately 38%. Thus, the total attenuation is approximately 0.16 + 0.14 = 0.30 dB/km, one way; the two-way attenuation for a radar is approximately 0.60 dB/km.

Snow and Ice. Compared with rain attenuation data, attenuation data from snow and hail is scant at millimeter wavelengths,[51,59] and the variation in the amount of moisture present in snow makes snow attenuation difficult to quantify. The attenuation of dry snow is 10 dB less than the attenuation of rain with the same rainfall rate at microwave frequencies; but the attenuation of very wet snow can exceed the attenuation of rain with the same rainfall rate. The attenuation of ice clouds is one to two orders of magnitude less than the attenuation of water clouds or fog, because ice spheres are less lossy than water spheres at the same frequency.[55]

3.3.2.5 Obscurants

Smoke and Dust. Measurements have shown[59–62] that smokes including hexachloroethane and phosphorous cause no significant attenuation of millimeter-wave signals up to 140 GHz. Dust is the only obscurant that had a significant effect on millimeter-wave propagation. Like water droplets, dust particles attenuate and backscatter signals. Measurements showed that dust and debris blown in the air from exploding artillery rounds caused attenuations of 10 to 30 dB with a recovery time of about 10 s. Backscatter from soil and debris was also observed when the attenuation was significant.[61]

Dust storm attenuation is of interest in some regions.[63] Dust attenuation and backscatter depend on the wavelength, particle size and distribution, particle density, particle composition, and moisture content. The particle size in dust storms tends to be small; hence, the attenuation tends to be small. An upper bound on the attenuation at 37 GHz is 0.7 dB/km for a sand storm with 10-m visibility. The attenuation is less at lower frequencies and greater at higher frequencies.

Foliage. Measurements documented in Ref. 59 provide an empirical expression for the two-way attenuation of dry foliage:

$$\gamma = 1.10 + 1.48 \log_{10} f \quad \text{[dB/m two-way]} , \tag{3.94}$$

where f is the frequency in gigahertz. The attenuation increases when the foliage is wet, there is little difference between vertical and horizontal polarization, and the attenuation increases with frequency, as indicated by the equation.

3.3.2.6 Multipath Effects. Radar propagation effects also include reflections from the surface of the earth.[59,64–66] Multipath reflections are, in general, less troublesome for millimeter-wave radars than for lower frequency radars:

millimeter-wave antennas are usually more directive, minimizing ground reflections, and the ground reflections are more diffuse because of the shorter wavelength. When the elevation angle is two beamwidths or more, the signal reflected from the ground passes through the antenna sidelobes, and it is usually far below the direct signal. Thus, multipath is a problem only at low grazing angles.

Consider the case when the radar and target are situated above a smooth, flat earth, as shown in Fig. 3.35. Assume that the radar is far enough from the target that the incident field is a plane wave. The field at the target is the vector sum of the direct signal and the signal reflected off the ground; the return signal to the radar has the same two contributions. The angle of incidence equals the angle of reflection, and there is a radar image below the ground plane. The path difference δ between the direct and the reflected ray is

$$\delta = 2H \sin\psi \ . \tag{3.95}$$

Let F represent the ratio of the field strength in the presence of the ground plane to the free-space field strength. Then, assuming equal antenna gain for both paths,

$$F = 1 + \rho_s \exp\left(\frac{j4\pi H \sin\psi}{\lambda}\right) \ . \tag{3.96}$$

Here ψ is the grazing angle and ρ_s is the complex reflection coefficient of the surface. When the surface is smooth, $\rho_s = \Gamma$, the reflection coefficient for a smooth surface, which can be computed with Fresnel's equations.[40]

At very low grazing angles, $\Gamma = -1$ for both vertical and horizontal polarizations, and

$$F = 1 - \exp\left(\frac{j4\pi H \sin\psi}{\lambda}\right) \ . \tag{3.97}$$

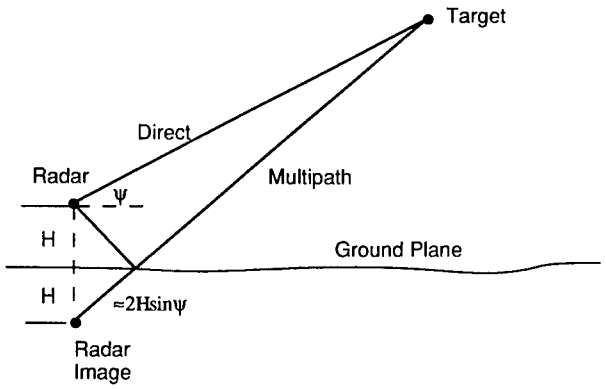


Fig. 3.35 Multipath geometry.

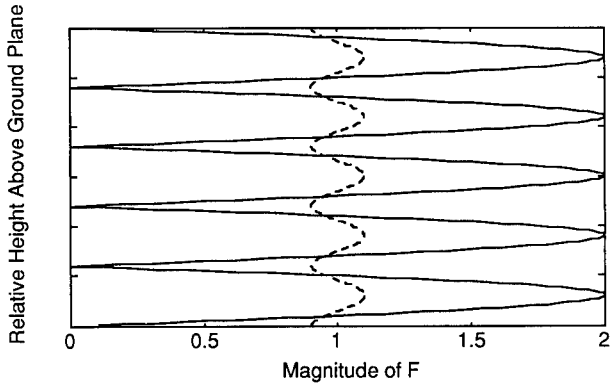


Fig. 3.36 Lobe structure above ground plane: solid, $\rho_s = -1$ and dashed, $\rho_s = -0.1$.

Thus,

$$|F|^2 = 4 \sin\left(\frac{2\pi H \sin\psi}{\lambda}\right)^2 . \tag{3.98}$$

The ratio F is zero whenever the quantity in the parentheses is zero, or an integer multiple of π, and a maximum whenever this quantity is an integer multiple of $\pi/2$, as sketched in Fig. 3.36. The null-to-null lobe width w is

$$w = \frac{\lambda}{2 \sin\psi} , \tag{3.99}$$

which is small at millimeter wavelengths.

When a surface is rough, the scattered field is composed of specular and diffuse components. The specular component adds to the direct field and produces a field that varies periodically with the height because the phase is correlated. The diffuse component, on the other hand, is the result of the combination of signals having random phases, and the field strength of diffuse scattering is small compared with the direct signal.[40] With rough terrain, the specular reflection coefficient becomes

$$\rho_s = R_s \Gamma , \tag{3.100}$$

where

$$\langle R_s^2 \rangle = \exp[-(\Delta\Phi)^2] . \tag{3.101}$$

Here $\Delta\Phi$ is a measure of the surface roughness,

$$\Delta\Phi = \frac{4\pi \Delta h \sin\psi}{\lambda} , \tag{3.102}$$

and Δh is the standard deviation of the assumed Gaussian distribution of heights measured from the plane surface.

3.3.3 Millimeter-Wave Clutter

3.3.3.1 Volume Clutter. Rain backscatter is appreciable at millimeter wavelengths, and it limits radar performance much more than at longer wavelengths. Precipitation is the principal source of volumetric backscatter, although backscatter from chaff, birds, or insects can also occur. Backscatter from water droplets depends on the droplet diameter, the density of the droplets, the dielectric constant of water, and the wavelength.[55,66,67] When $\lambda > 4\pi rn$, where r is the drop radius and n is the index of refraction, Rayleigh scattering occurs (see Fig. 3.24), and the backscatter is proportional to the inverse fourth power of the wavelength. As a consequence, the backscatter increases rapidly as the wavelength decreases. When $\lambda = 4\pi rn$, the scattering is in the Mie or resonance region, and the backscattered power oscillates rapidly about an average value as the wavelength changes, and the backscattered power saturates.

Measured values for rain backscatter have a rather large spread, but an empirical model for the rain reflectivity η has been developed that is similar in form to the expression for attenuation from rain[66]

$$\eta = aR_r^b \quad [\mathrm{m^2/m^3}] \ , \tag{3.103}$$

where R_r is the rain rate in millimeters per hour and the constants a and b are determined from measured data at a particular frequency. From the Georgia Tech rain backscatter model,[66]

$$\begin{aligned} \eta &= 1.2 \times 10^{-6} R_r^{1.6} \quad \text{at 35 GHz} \\ &= 4.2 \times 10^{-5} R_r^{1.1} \quad \text{at 70 GHz} \\ &= 1.5 \times 10^{-5} R_r \quad \text{at 95 GHz} \ . \end{aligned} \tag{3.104}$$

Figure 3.37 plots the rain backscatter as a function of the rain rate. Note that the model predicts a higher backscatter coefficient at 70 GHz than at 95 GHz. Backscatter from rain increases with frequency up to 70 GHz, but there is an apparent reversal of the backscatter dependence on frequency above 70 GHz. This reversal may be related to scattering in the resonance region. The values of η computed at 35 GHz from Eq. (3.104) were lower than values computed with theoretical expressions in Ref. 55.

As an example, we compute the backscattered cross section σ of a resolution cell for a 35-GHz radar with a 1-deg (17.45-mrad) circular beamwidth and a 0.1-μs pulse width, when the range is 10,000 m, and the rain rate is 10 mm/h.

$$\eta = 1.2 \times 10^{-6} \times 10^{1.6} = 4.8 \times 10^{-5}\ \mathrm{m^{-1}} \ .$$

As discussed in Sec. 3.2, the effective beamwidth for volume clutter calculation is about 75% of the -3-dB beamwidth β; hence, the effective clutter volume V is

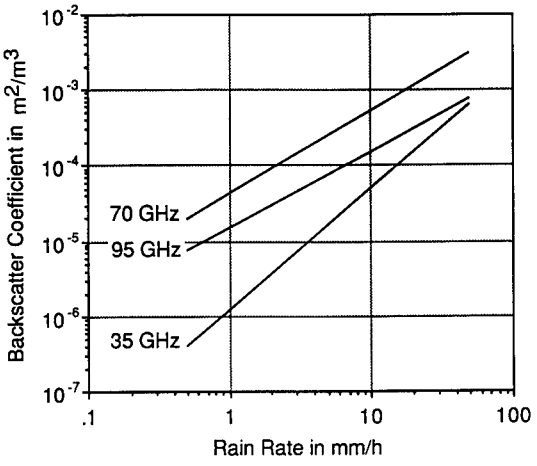


Fig. 3.37 Rain backscatter. (Model from Ref. 66)

$$V = \frac{\pi}{4}(0.75\beta R)^2\frac{c\tau}{2} = 2.02 \times 10^5 \text{ m}^3 \ .$$

Thus,

$$\sigma = \eta V = 4.8 \times 10^{-5} \times 2.02 \times 10^5 = 9.6 \text{ m}^2 \ .$$

The backscattered power from rain at millimeter wavelengths has a log-normal distribution with standard deviations between 2.5 and 6 dB. The decorrelation time decreases with increasing frequency and increasing rain rate. At 35 GHz, the decorrelation time varies from 8 ms at 5 mm/h to 3 ms at 100 mm/h, whereas at 95 GHz, the decorrelation time varies from 3.5 to 1.5 ms at the same rain rates.[66]

The backscatter from clouds or fog is much less than the backscatter from rain. The Rayleigh approximation is valid up to 300 GHz for fair weather clouds, which have a small drop size, and to lower frequencies for rain clouds, with larger drops. Figure 3.38 gives values of cloud or fog backscatter for frequencies of 95, 70, and 35 GHz computed from equations based on the Rayleigh approximation listed in Ref. 55.

3.3.3.2 Ground Return. A distributed ground target has a large number of scatterers randomly distributed over a homogeneous surface, and the overall return is the vector sum of the returns from the individual scatterers in the illuminated area. The quantity σ_0 (sometimes written σ^0) is the backscattered radar cross section (RCS) per unit area of a distributed target. The RCS σ of a distributed target with an illuminated area A is

$$\sigma = \sigma_0 A \ . \tag{3.105}$$

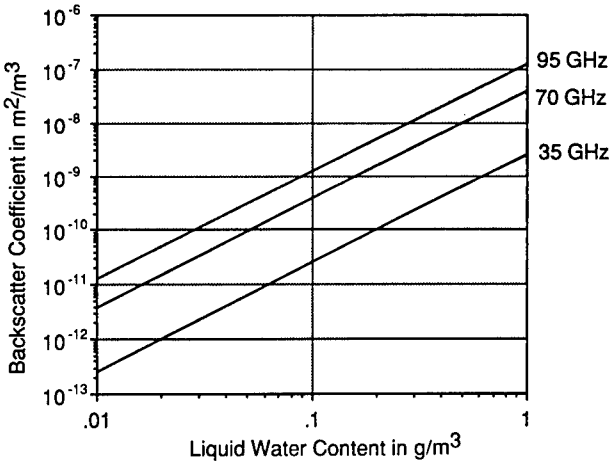


Fig. 3.38 Backscatter from clouds or fog at 20°C.

Some authors prefer to describe the backscatter with the parameter γ and the area A' normal to the antenna beam, so that

$$\sigma = \gamma A' \ . \tag{3.106}$$

If ψ is the grazing angle between the incident ray and the average plane of the surface,

$$A' = A \sin\psi \ , \tag{3.107}$$

and

$$\sigma_0 = \gamma \sin\psi \ . \tag{3.108}$$

The parameter γ varies less with changes in ψ than σ_0 does; both parameters are dimensionless.[40,66–71]

The magnitude of σ_0 depends on the radar wavelength and polarization, the dielectric and geometric properties of the surface, and the grazing angle. By the Rayleigh roughness criterion,[40] a surface is considered smooth when

$$\Delta h \sin\psi \leq \lambda/8 \ , \tag{3.109}$$

where Δh is the rms height of the surface irregularities and λ is the wavelength. Figure 3.39 shows the three radar scattering regions for a planar surface at microwave and millimeter-wave frequencies: the near-grazing incidence region, the plateau region, and the near-vertical incidence or specular region.[40] The properties of the regions depend on the wavelength, the polarization, and the surface. Near grazing incidence, σ_0 increases rapidly with increases in the grazing angle and radar frequency up to the critical angle ψ_c; σ_0 changes slowly with the grazing angle and transmitted frequency in the plateau region be-

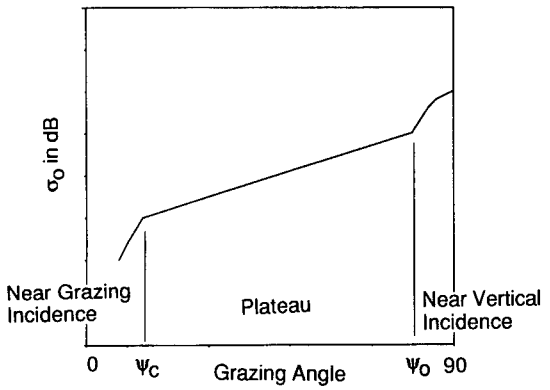


Fig. 3.39 Typical σ_0 grazing angle dependence.

tween ψ_c and ψ_0. In the near-grazing incidence and plateau regions, σ_0 increases as the surface roughness increases, but σ_0 decreases as the surface roughness increases near vertical incidence, and the dependence on the radar frequency is weaker. Resonant scatterers are sometimes present at millimeter wavelengths, and the critical angle is smaller than at microwave frequencies because the terrain appears rougher at the shorter wavelengths.

Measured values of σ_0 depend on the transmitted and received polarizations. Data have been collected much more frequently with linear than with circular polarization. With linear polarization the earth's surface serves as a convenient reference; the term *horizontal* refers to a wave whose electric field is parallel to the earth's surface, whereas *vertical* refers to a wave whose electric field has a component perpendicular to the earth's surface. Thus, σ_{0HH} represents parallel-polarized values measured with horizontal transmitter and receiver polarizations; σ_{0VV}, parallel-polarized values measured with vertical transmitter and receiver polarizations; and σ_{0VH}, cross-polarized values measured with vertical transmitter and horizontal receiver polarizations. With the monostatic radars we are considering, the cross-polarized components, σ_{0VH} and σ_{0HV}, are equal, and they are usually several decibels lower than σ_{0HH} and σ_{0VV}. Values of σ_0 measured with circular polarization are about 3 dB lower than the values measured with parallel-linear polarization.

Typical Values. Most past millimeter-wave clutter measurements[66–71] have been at 35 or 95 GHz, but data are now becoming available[72–74] at higher frequencies such as 140 and 215 GHz. In the millimeter-wave band where the surface roughness is a large number of wavelengths, the Georgia Tech clutter model has the simple form

$$\sigma_0 = A(\psi + C)^B \ , \tag{3.110}$$

where the grazing angle ψ is in radians and A, B, and C are empirical constants determined from measurements.[68] Other models are available; see, for example, Ref. 70, which has an excellent data base as well as an extensive bibliography on land clutter measurements.

The updated Georgia Tech clutter model provides typical σ_0 values at 35 and 95 GHz for parallel-linear polarizations; the model assumes the same value for *HH* and *VV* polarizations.[68] Figures 3.40 and 3.41 plot the average σ_0 values for grass and crops and for trees; the 90th percentile level for these terrains is 4 to 6 dB above the average values shown. Reflectivity measurements of grass, crops, and trees indicate that there is a large intrinsic variability in terrain backscatter. The reflectivity of vegetation changes with the season of the year: the reflectivity is higher with tall grass than with short grass, and the reflectivity of crops increases with increasing foliage. Deciduous trees are more reflective with leaves than when the trees are bare, and the reflectivity of wet terrain is approximately 5 dB higher than the reflectivity of dry terrain.[66]

The reflectivity of snow at millimeter wavelengths depends on the free water content, and the reflectivity can change rapidly with time.[67] Wet snow, which has a liquid water content greater than 1% by volume, is much less reflective than dry snow. Figures 3.42 and 3.43 show plots of the mean reflectivity of wet and dry snow at 35 and 95 GHz computed with the updated Georgia Tech

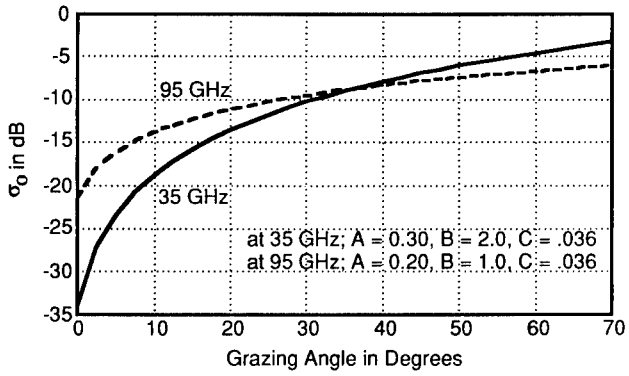


Fig. 3.40 Average σ_0 for grass and crops.

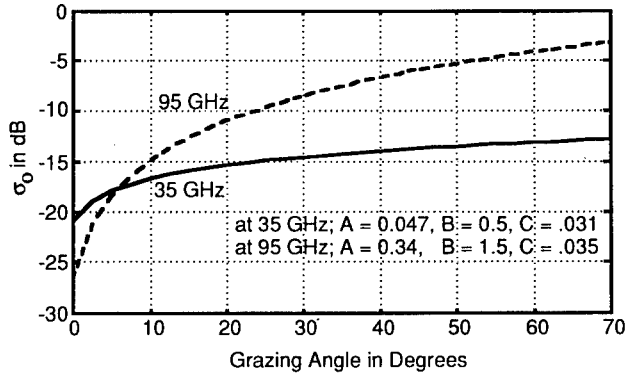


Fig. 3.41 Average σ_0 for trees.

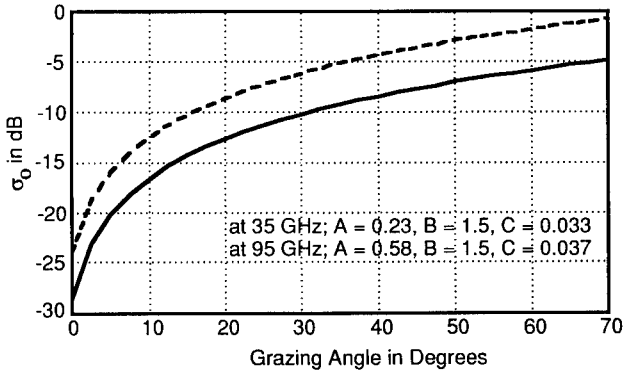


Fig. 3.42 Average σ_0 for wet snow.

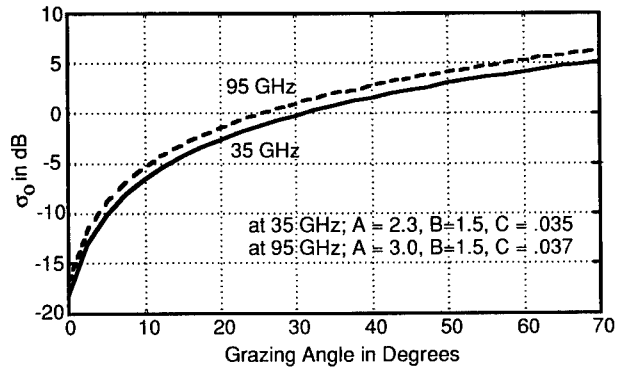


Fig. 3.43 Average σ_0 for dry snow.

model.[68] There is considerable variability in σ_0 with snow measurements, and the values in Figs. 3.42 and 3.43 are somewhat higher than similar data in Ref. 70.

Example. Find the average radar cross section of tree clutter for a 95-GHz radar with resolution cell area A of 10 m^2 at a grazing angle of 20 deg. From Fig. 3.41, the mean value of σ_0 for tree clutter at 95 GHz at $\psi = 20$ deg is -11 dB or 0.079. From Eq. (3.105),

$$\sigma = \sigma_0 A = 0.079 \times 10 = 0.79 \text{ m}^2 .$$

In terms of decibels relative to 1 m^2,

$$\sigma = 10 \log(0.79) = -1 \text{ dBsm} .$$

Spatial Clutter Variations. The classical approach to the analysis of distributed targets assumes that the backscattered field at the radar antenna (in

volts per meter) has a Rayleigh distribution, because the backscattered field is the vector sum of a large number of statistically independent, far-field returns that have about the same amplitude.[75] The Rayleigh density function, which describes the fluctuations (fading) of the voltage amplitude due to Rayleigh clutter (or thermal noise), is

$$p(E) = \frac{E}{\alpha} \exp\left(\frac{-E^2}{2\alpha}\right) , \quad E \geq 0 . \tag{3.111}$$

Here α is the mean-square voltage. The return power P is proportional to E^2, and the probability distribution for the return power from Rayleigh clutter is

$$p(P) = \frac{1}{P_{av}} \exp\left(-\frac{P}{P_{av}}\right) , \tag{3.112}$$

where P_{av} is the average power. Although this distribution is an exponential, it is sometimes called the Rayleigh probability density function for power.

The Rayleigh clutter model is usually valid for uniform clutter below 30 GHz at grazing angles above about 5 deg, but it is sometimes inadequate at higher frequencies because of longer tails on the probability density function. The Rayleigh distribution may be appropriate at millimeter wavelengths for a stationary radar system, or for scanning systems when the integration time is very short relative to the beam motion.[75] The probability of a large scatterer causing a high tail in the clutter is higher for a scanning radar, and log-normal or Weibull probability density functions are better clutter models. See Ref. 76 for a summary of clutter models.

With the log-normal distribution, a random variable Y is log-normally distributed if its logarithm has a normal (Gaussian) distribution[40]:

$$p(Y) = \frac{1}{Ys\sqrt{2\pi}} \exp\left[-\frac{1}{2s^2}\left(\ln\frac{Y}{Y_m}\right)^2\right] , \tag{3.113}$$

where Y is the log-normally distributed variable (σ_0 or σ), Y_m is the median value of Y, and s is the standard deviation of $\ln(Y/Y_m)$.

The Weibull probability density function lies between the exponential distribution for Rayleigh clutter and the log-normal distribution. The Weibull distribution for σ_0 or σ has the form[40]

$$p(\sigma) = \frac{b}{\sigma_{av}} \sigma^{(b-1)} \exp\left(-\frac{\sigma^b}{\sigma_{av}}\right) , \tag{3.114}$$

where the shape factor b is a constant and σ_{av} is the average value of σ_0 or σ. For $b = 1$, the Weibull distribution reduces to the exponential power distribution for Rayleigh clutter.

Probability of detection computations are simplified with Rayleigh clutter because the statistics are the same as for thermal noise, and detection curves computed for thermal noise are applicable. With other clutter backgrounds, Rayleigh statistics will predict too high a probability of detection, because the

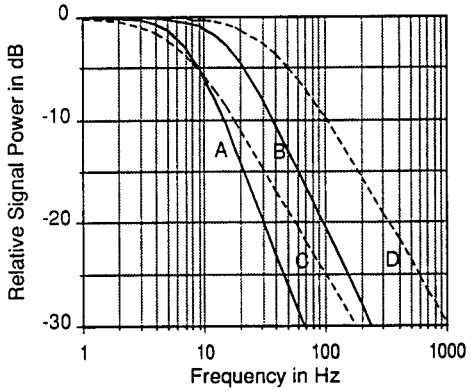


Fig. 3.44 Normalized frequency spectrum of deciduous trees with a linear receiver: A—35 GHz, wind 0 to 5 mph, $f_c = 7$, $n = 3$; B—35 GHz, wind 6 to 15 mph, $f_c = 16$, $n = 2.5$; C—95 GHz, wind 0 to 5 mph, $f_c = 6$, $n = 2$; and D—95 GHz, wind 6 to 15 mph, $f_c = 35$, $n = 2$. (Model from Ref. 67)

detection thresholds must be raised to maintain the same false alarm rate. Probability of detection computations with other clutter models may require numerical techniques.

Temporal Clutter Variations. The amplitude of the return power from a clutter cell that contains foliage or branches moving in the wind varies with time. The amplitude statistics and the frequency spectrum or the decorrelation time describe the temporal response. Using a noncoherent pulse radar, Georgia Tech measured data on the temporal characteristics of tree clutter.[66] At 35 GHz, a Rayleigh distribution was a close approximation to the amplitude characteristics, but at 95 GHz, a log-normal distribution was a better clutter model. The standard deviation of the tree clutter was about 4 or 5 dB at 35 GHz and about 5.5 to 6.5 dB at 95 GHz.

The frequency spectra of deciduous trees depended on the frequency and the wind speed; the spectra had the form

$$W(f) = \frac{A}{1 + (f/f_c)^n}, \tag{3.115}$$

where A is the dc value, f is the frequency, f_c is the corner frequency, and n is a constant. Figure 3.44 plots the Georgia Tech models of the frequency spectra of deciduous trees measured with a linear receiver; the frequency spectra measured with a log receiver were slightly different.

3.3.3.3 Sea Clutter. Investigators have studied radar backscatter from the sea since World War II, and they have collected a large amount of experimental data at frequencies below 30 GHz but relatively little data at millimeter wavelengths.[19,66,67,77,78] Sea backscatter is a complex phenomenon that involves a number of factors. Sea factors include the height of the waves, the wind speed, the length of time the wind has been blowing, the direction of the wind relative to the radar beam, the temperature differential between the air and the sea,

and the presence of swell. The radar factors of importance are the frequency, the polarization, the grazing angle, and the size of the radar resolution cell. In general, sea clutter varies more temporally than land clutter.

As with land clutter, there are three scattering regions that depend on the grazing angle: the quasispecular, the plateau, and the interference region near grazing incidence. The quasispecular region is the result of specular scatter at grazing angles near 90 deg. At very low grazing angles, σ_0 decreases very rapidly with decreasing angle, and the region is called the *interference region,* because the direct wave and the reflected wave are out of phase and produce destructive interference. In the plateau region between the quasispecular and interference regions, σ_0 changes less rapidly with the grazing angle.

The sea surface has ripples or capillary waves developed by the surface friction velocity of winds, and these ripples are superimposed on much longer gravity waves, which have a tilting effect. The ripples become increasingly important at millimeter wavelengths when the ripple spacing results in backscatter contributions being in phase (Bragg scatter). Backscatter is low at velocities less than 5 kn, except in the specular region, and backscatter increases with wind speed at higher velocities.

The magnitude of σ_0 depends on the radar look direction with respect to the wind: σ_0 is highest looking upwind (into the wind) and lowest looking crosswind, whereas σ_0 looking downwind is usually slightly less than looking upwind. There can be a 5- to 10-dB variation in σ_0 as the radar scans 360 deg in azimuth, and the variation of σ_0 with look angle resembles a sine wave.[79] The change in σ_0 with azimuth angle decreases as the wind speed increases. Because the capillary waves build up more rapidly with the wind than gravity waves, backscatter increases more rapidly with the wind at higher frequencies. The breaking of waves, spray, foam, and swell all affect backscatter.

In general, it is difficult to describe completely the sea and wind conditions under which experimental data are collected, and there can be a considerable variation in data collected by different organizations or by the same organization at different times. There are few published measurements of sea backscatter in the millimeter-wave band, especially above 35 GHz, but the data available indicate that backscatter behavior at millimeter wavelengths is much like the behavior at microwave frequencies.[66,67,77]

Figure 3.45 plots σ_0 for sea backscatter for vertical and horizontal polarizations for the frequency band covering 26.4 to 52 GHz for sea states 1, 3, and 5. The data for Fig. 3.45 are from Table 15.4.2 in Ref. 80; Petts produced the data for this table by interpolating and extrapolating data compiled by Nathanson from a number of sources.[19] The data on which Petts's model is based were collected at a time when investigators did not recognize the importance of the radar look angle relative to wind direction; as a consequence, it represents an average over all wind directions. As mentioned in Ref. 19, errors of 5 dB are likely for many of the data points. Reference 67 discusses the Georgia Tech sea clutter model for grazing angles below 10 deg and presents some low-angle 35- and 95-GHz measurement results. At grazing angles between 1 and 6 deg, σ_0 is about -30 dB looking upwind at both frequencies for a 0.6-m wave height; σ_0 is 12 to 18 dB less looking downwind.

Plots of σ_{0HH} and σ_{0VV} looking upwind versus grazing angle at 34.43 GHz are shown in Fig. 3.46; the plots are from scatterometer data measured by

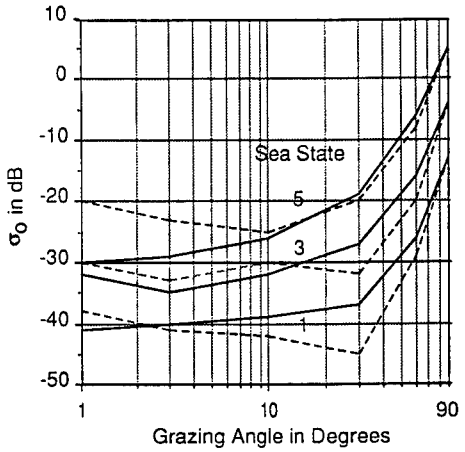


Fig. 3.45 Typical sea clutter reflectivity from 26.4 to 52 GHz; solid lines—*VV* and dashed lines—*HH*. (Petts's model reported by Morris.[80])

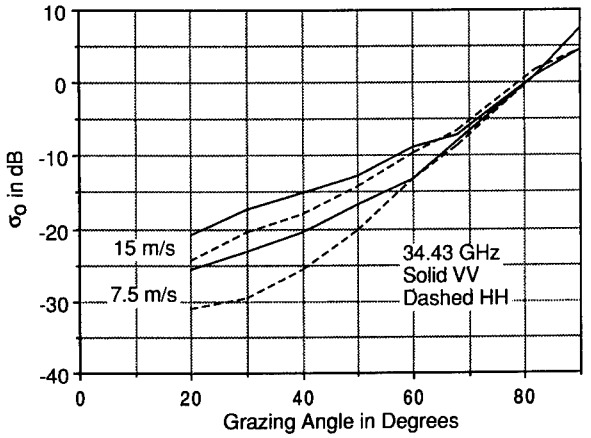


Fig. 3.46 Plot of σ_0 versus grazing angle looking upwind.[79]

Masuko et al.[79] The wind speeds for the plots are 15 and 7.5 m/s; a speed of 7.5 m/s equals 14.6 kn. Because the median wind speed over the world's oceans is about 15 kn, corresponding to sea state 3, the 7.5 m/s plot corresponds to sea state 3. At 34.43 GHz, grazing angles between 20 and 70 deg, and wind speeds between 3.2 and 17.2 m/s, Table 2 in Ref. 79 indicates that σ_0 looking upwind averages 2.6 and 1.8 dB higher than σ_0 looking downwind for *HH* and *VV* polarizations, respectively, and σ_0 looking upwind averages 4.8 and 5.7 dB higher than σ_0 looking crosswind for the same polarizations. Thus, subtracting the appropriate amount from the curves in Fig. 3.46 provides downwind and crosswind σ_0 values.

The illuminated area affects the nature of the backscatter.[77] When the radar resolution is fine enough, the return contains target echoes of short duration

that appear as spikes on a display. Spikes, which can last for several seconds, occur when whitecaps form, but they may also occur in the absence of whitecaps. Sea clutter is more spiky for horizontal than vertical polarization. The temporal and spatial statistics of sea clutter are essentially the same.[67] Statistical data on spatial statistics are sparse at millimeter wavelengths, but experimental data at lower frequencies show that the probability density function of sea clutter is between Rayleigh and log-normal.[66]

3.4 MILLIMETER-WAVE COMPONENTS

3.4.1 Antennas

3.4.1.1 Introduction. An antenna is a device that couples power from free space to a transmission line or from the transmission line to free space. Many millimeter-wave antennas are similar to those in use at lower frequencies, but in the millimeter-wave band, antennas are smaller in size, they have wider bandwidths, the losses are higher, construction tolerances are tighter, and scanning is more difficult because of the lack of rotary joints. Factors important for antenna performance include the beam shape, the sidelobe level, the polarization, scanning capabilities, the efficiency, and the bandwidth. Types of antennas in common use at millimeter-wave frequencies include horns, lenses, reflectors, and arrays.[55,81–84] Figure 3.47 is an example of a millimeter-wave reflector antenna.

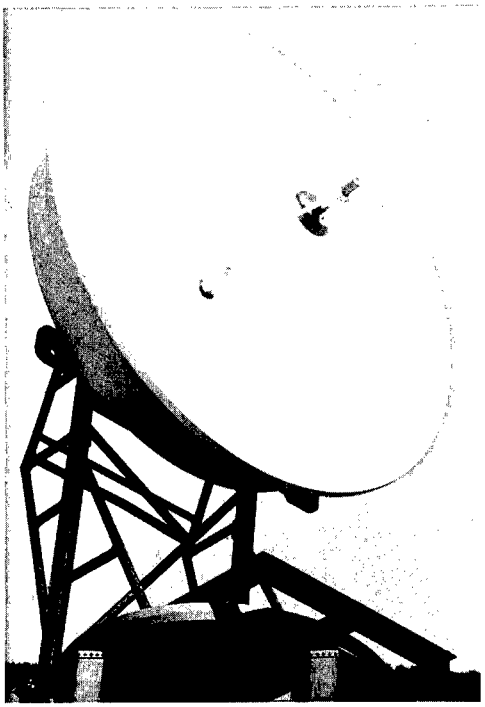


Fig. 3.47 A 35-GHz Cassegrain antenna. (Courtesy of Alpha Industries, Inc.)

The directivity and gain of an antenna determine how well the antenna concentrates energy in a given direction. The directivity $D(\theta,\phi)$ of an antenna is the ratio of the power density in a given direction to the power density that would have been present had the power been radiated isotropically; directivity is usually expressed in decibels above isotropic (dBi). If $\Phi(\theta,\phi)$ is the radiation intensity in watts per steradian, the directivity $D(\theta,\phi)$ in the angular direction denoted by θ and ϕ is

$$D(\theta,\phi) = \frac{\Phi(\theta,\phi)}{P_r/4\pi} , \tag{3.116}$$

where P_r is the radiated power. The radiation efficiency η is the ratio of the radiated power to the input power P_0 of the antenna

$$\eta = \frac{P_r}{P_0} . \tag{3.117}$$

The radiation efficiency is always less than 1 because of ohmic losses in the antenna. The antenna gain G is

$$G(\theta,\phi) = \eta D(\theta,\phi) = \frac{\eta\Phi(\theta,\phi)}{P_0/4\pi} . \tag{3.118}$$

Thus, the antenna gain is the ratio of the radiation intensity in the direction (θ,ϕ) to the radiation intensity that would have been present had all the input power to the antenna been radiated isotropically.

The power density $P(\theta,\phi,R)$ at range R in watts per square meter equals the antenna gain times the power radiated by an isotropic antenna

$$P(\theta,\phi,R) = G(\theta,\phi)\frac{P_0}{4\pi R^2} . \tag{3.119}$$

The incident power absorbed by an antenna is the product of the effective area of the antenna A_e and the incident power density. The effective area depends on the wavelength λ and the gain:

$$A_e(\theta,\phi) = \frac{\lambda^2}{4\pi}G(\theta,\phi) . \tag{3.120}$$

The antenna bandwidth is the range of frequencies within which the antenna performance conforms to a specified standard with respect to some characteristic such as a decrease of the radiated power by a factor of 2, or a specified change in the input impedance.

The field radiated from an antenna depends on the angular coordinates and the distance from the antenna.[83] Very close to the antenna, the field contains a reactive component that decays rapidly with distance. Farther away in the radiating near field, the amplitude and phase of the field depend on the distance

to the antenna. In the far field, the radiation pattern is essentially independent of the distance to the observation point. A common criterion for the far field is that

$$R \geq \frac{2D^2}{\lambda} , \tag{3.121}$$

where R is the range, D is the antenna diameter, and λ is the wavelength. Operation at shorter ranges in the near field is frequent at millimeter wavelengths, and there are some pattern changes compared to the far field.

High-gain antennas such as horn or reflector antennas are said to be aperture antennas because most of the radiated power passes through a plane in front of the antenna. It can be shown that the far-field pattern is approximately equal to the Fourier transform of the field illumination across the aperture, or in the case of a line source, the illumination across the line source. The −3-dB beamwidth BW_{3dB} of an antenna is

$$BW_{3dB} = \frac{k\lambda}{L} \quad \text{rectangular aperture} , \tag{3.122a}$$

$$BW_{3dB} = \frac{k\lambda}{D} \quad \text{circular aperture} , \tag{3.122b}$$

where k is a constant, L is the length of the rectangular aperture, and D is the diameter of the circular aperture. For uniform illumination and a rectangular aperture, k = 50.8 deg, and the first sidelobe is 13 dB down from the peak of the main lobe. With uniform illumination and a circular aperture, k = 58.9 deg, and the first sidelobe is 17.6 dB down. Normally, the antenna designer tapers the illumination from the center to the edges of the aperture to lower the sidelobe level at the expense of mainlobe broadening.[82] Figure 3.48 has plots of beamwidth versus frequency for a typical value of k, 70 deg, for antenna apertures between 7.5 and 180 cm.

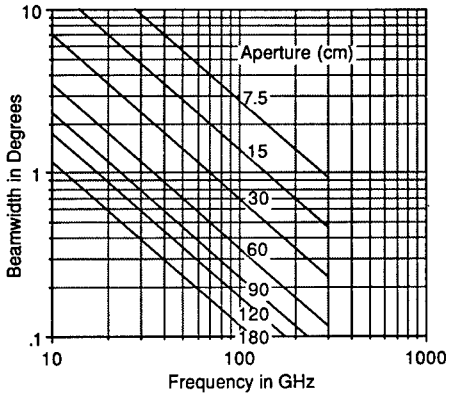


Fig. 3.48 Aperture antenna beamwidth versus frequency; k = 70 deg.

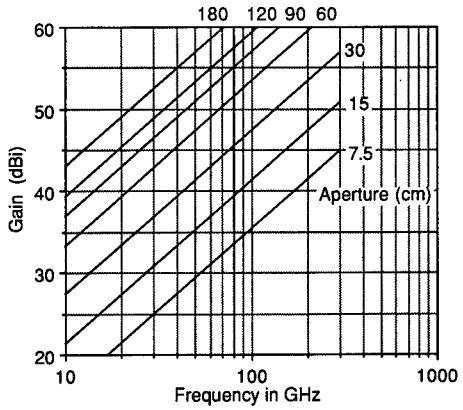


Fig. 3.49 Aperture antenna gain versus frequency for a circular aperture.

The antenna gain also depends on the aperture size, and for a circular aperture of diameter D,

$$G = 20 \log(\pi D/\lambda) + 10 \log\eta \ . \tag{3.123}$$

A typical value for the efficiency η at millimeter wavelengths is 0.5. Figure 3.49 plots antenna gain versus frequency for a circular aperture and a loss of 2.6 dB.

Sample Computation with Charts. An application requires an antenna beamwidth of 0.5 deg at a frequency of 100 GHz. In Fig. 3.48, the horizontal 0.5-deg beamwidth line intersects the vertical 100-GHz frequency line between the 30- and 60-cm apertures, and the aperture is about 42 cm. From Fig. 3.49, the gain for a 42-cm aperture at the 100-GHz line is about 50 dBi.

3.4.1.2 Horn and Horn Lens Antennas. Horn antennas, Fig. 3.50, are widely used at millimeter wavelengths to illuminate larger apertures and for test purposes when gains of 15 to 25 dB or beamwidths of more than about 10 deg are required.[81,84] Horn antennas are formed by gradually flaring the end of a waveguide to avoid the generation of higher order modes. The flare is in the direction of the electric field of the input waveguide for an E plane sectoral horn, and in the direction of the input magnetic field for the H plane sectoral horn. However, the pyramidal horn, which has flares in both the E and H planes, is much more commonly used for power monitoring, in test bench applications, and as a standard gain horn to calibrate other antennas. Because of its circular symmetry, the conical horn can handle linear polarization at any angle as well as circular polarization. Corrugations in a conical horn generate a hybrid mixture of propagating modes, which results in almost equal beamwidths in the E and H planes and low sidelobes and cross-polarization levels.

A pyramidal horn that produces maximum far-field gain for a given slant length in the E and H planes is called an *optimal gain horn*. If a is the width

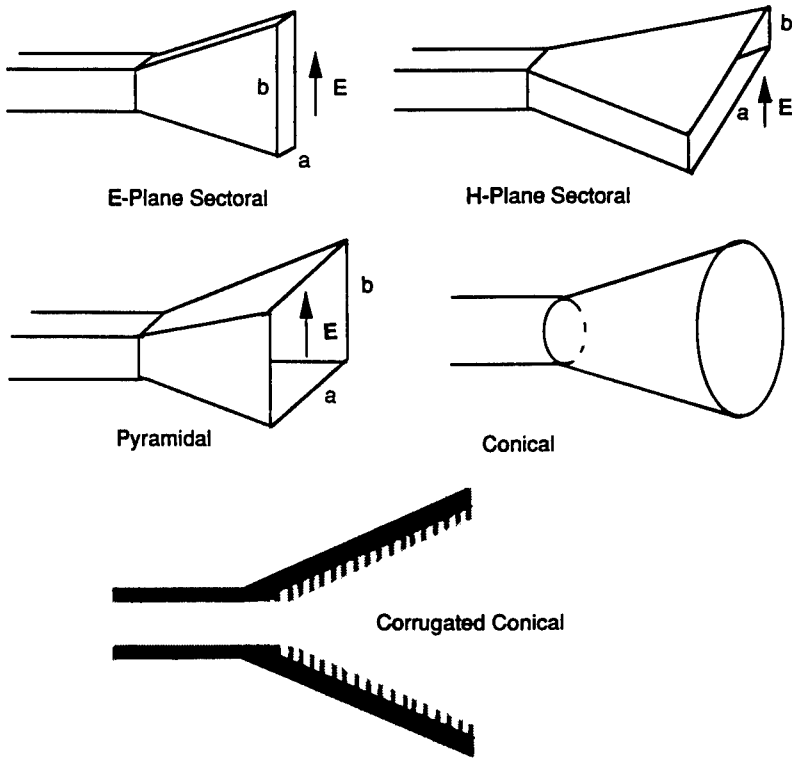


Fig. 3.50 Horn antennas.

of the horn mouth in the H plane and b is the width in the E plane, the approximate gain, beamwidth, and first-sidelobe levels of the optimal gain horn are[84]

$$G = 8.1 + 10 \log_{10}(ab/\lambda^2) \quad [\text{dB}] \; . \tag{3.124}$$

BW_{3dB}	Sidelobe Level	
$2 \sin^{-1}(0.45\lambda/b)$	-9 dB	E plane
$2 \sin^{-1}(0.70\lambda/a)$	-16 dB	H plane

When more gain and a narrower beamwidth are required than a horn antenna can provide, a conical horn or a corrugated conical horn can be combined with a dielectric lens to form a more compact structure. Typical horn lens antennas have gains of 25 to 47 dBi, beamwidths of 10 to 0.8 deg, and sidelobe levels of -18 to -30 dB. When constructed with circular components, horn lens antennas provide horizontal, vertical, or circular polarizations.[85] Figure 3.51 shows a horn lens antenna.

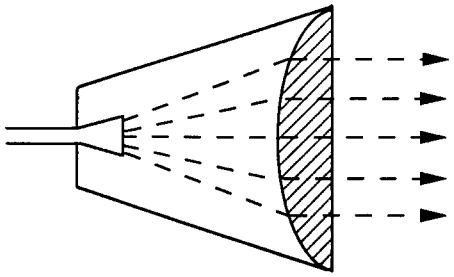


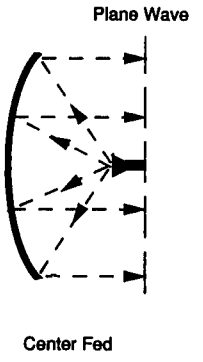
Fig. 3.51 Horn lens antenna.

3.4.1.3 Reflector Antennas. Parabolic and Cassegrain antennas find frequent application in the millimeter-wave band; these antennas are available with a wide range of reflector diameters.[81,85] Figure 3.52 shows center-fed and offset-fed parabolic reflector antennas, which consist of a parabolic reflector, a primary feed, and a feed support assembly. The feed is at the focal point of the parabola, and because the path lengths from the focal point to the aperture plane are equal, a collimated beam is formed at the aperture plane. A parabolic reflector is a broadband device that can operate over a wide frequency range with suitable feeds; however, the losses in the waveguide transmission line to the feed are high. Aperture blockage by the feed structure of the center-fed parabola lowers the gain and increases the sidelobe level; the offset feed eliminates aperture blockage.

Cassegrain antennas, Fig. 3.53, are often recommended for applications where low waveguide losses to the feed are desired, and when the required beamwidth is 1 deg or less. The feed horn, which is at the primary focal point, illuminates the hyperboloid subreflector, and the parabolic main reflector collimates the energy reflected from the subreflector. The subreflector must be large enough to act as an optical reflector and small enough to minimize aperture blockage. The proximity of the feed horn to the main reflector reduces the transmission line losses. The Cassegrain antenna has better low-noise performance than a single reflector antenna because the spillover radiation from the virtual feed is normally much smaller.

Parabolic and Cassegrain reflector antennas can provide horizontal, vertical, or circular polarizations with suitable feeds. A circular feed radiates both linear and circular polarizations, and with an orthomode transducer assembly, a feed can operate with both horizontal and vertical, or both right- and left-hand circular, polarizations simultaneously.[85]

Because construction tolerances for a given performance with a reflector antenna are in fractions of a wavelength, tolerances become tighter as the wavelength decreases. Surface deviations from the desired shape are statistical in nature, and an rms surface error of 0.03λ produces about a 0.5-dB loss in gain. Surface tolerances of 1/20 to 1/30 of a wavelength result in aperture efficiencies of 67 to 84%. Millimeter-wave reflector antennas have been built with diameters of 13.7 m (45 ft) for radio astronomy.[86]



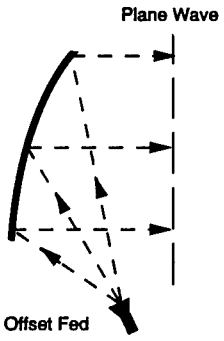
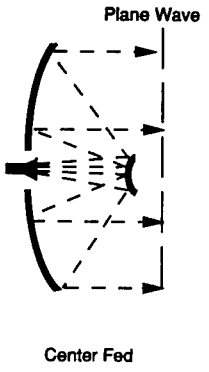


Fig. 3.52 Parabolic antennas.



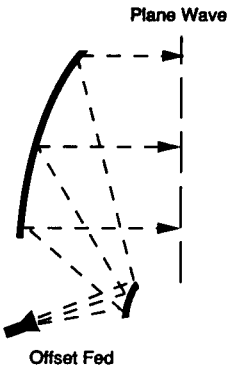


Fig. 3.53 Cassegrain antennas.

3.4.1.4 Other Antenna Types. There are many other types of millimeter-wave antennas in addition to horn and reflector antennas. Although reflector antennas have the best gain, sidelobe, and pattern performance, there are many applications where a reflector antenna is unsuitable because of its size and cost, or because a conformal structure is required, and it is necessary to use another type of antenna whose performance is not quite as good. As at lower frequencies, cutting slots in the narrow or broad wall of a waveguide forms a linear array, and slotted waveguides placed side by side form a planar array, as shown in Fig. 3.54.

With present technology, microstrip (patch) antennas are attractive for planar arrays up to 100 or 140 GHz. Microstrip antennas have a low profile, they are light, easy to construct, and rugged, and they can be built on a single-layer printed circuit board; disadvantages of microstrip antennas include narrow bandwidth, feed network losses, and tight construction tolerances. Figure 3.55

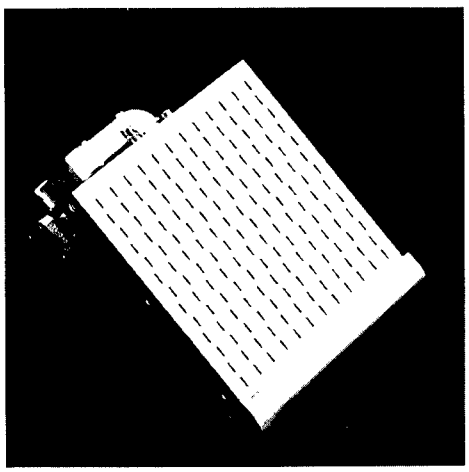


Fig. 3.54 A 35-GHz planar waveguide array. (Courtesy of Alpha Industries, Inc.)

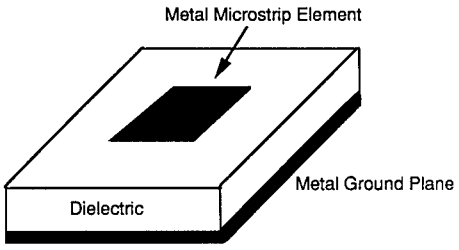


Fig. 3.55 Microstrip antenna element.

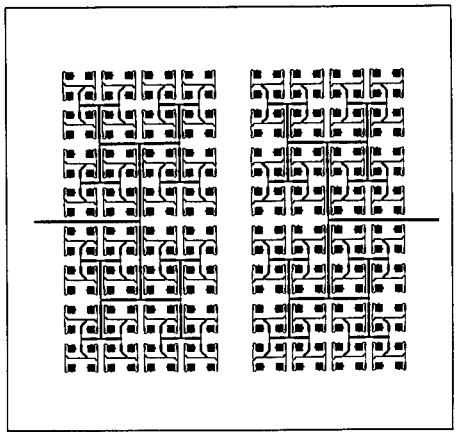


Fig. 3.56 Dual circularly polarized microstrip array.[89] (© 1988, Cardiff Publishing Company, reprinted with permission)

is a sketch of a single microstrip element,[88] and Fig. 3.56 shows a microstrip array. The gain of a microstrip array is[89]

$$G = 10 \log 4\pi L^2 - \alpha L - C \quad [\text{dB}] \; . \tag{3.125}$$

Here L is the length in wavelengths of one side of the aperture, α is the loss per unit wavelength, and C is a constant resulting from other loss contributors. At 35 GHz, a typical value of C is 1.7 dB, and losses are about 0.5 dB/in.

3.4.1.5 Radomes. A radome is an electromagnetic window that covers and protects enclosed antennas from the environment.[90,91] The properties of importance for a radome include high transmission, broad bandwidth, low boresight error, and low reflection. Radome construction is more difficult at millimeter wavelengths because a radome thickness that provides adequate structural strength is larger in terms of the wavelength. Millimeter-wave radomes are often constructed with a wall thickness that is a multiple of half a wavelength; desirable properties for radome materials include a low dielectric

constant and a low loss target. Ceramics and large molecule polymers are often used at millimeter wavelengths.

3.4.2 Millimeter-Wave Sources

The availability of suitable power sources is essential for the construction of a millimeter-wave radar system. Both vacuum tube and solid-state power sources are currently in use, but vacuum tube sources provide much higher output power. There are two approaches to output power generation in a radar system.[92] In the power oscillator approach, the output of an oscillator goes directly to an antenna, whereas in the power amplifier approach, an oscillator generates power at a low level, and a power amplifier or amplifier chain raises the level of the power sent to the antenna. The rf signal generated may be cw or pulsed, depending on the application, and a pulsed signal may have frequency coding for pulse compression. Tube types available at millimeter wavelengths include gyrotrons, magnetrons, klystrons, extended interaction oscillators and amplifiers, traveling-wave tubes, and backward-wave oscillators. Although Gunn diodes and IMPATT diodes are still the primary millimeter-wave solid-state sources, transistor amplifiers are now available at frequencies up to 40 to 60 GHz, and they will be available at higher frequencies in the future.

3.4.2.1 Vacuum Tube Sources

Gyrotrons. The gyrotron, also known as the cyclotron resonance maser (CRM), is a relatively recent development. In a gyrotron device, coupling between an electron beam and the microwave field occurs by means of the cyclotron resonance condition.[6,93,94] Because the strength of the magnetic field, not the dimensions of some resonant structure, determines the operating frequency, the beam and microwave structure can be large compared with the wavelength, increasing the power capability. High magnetic fields are required at millimeter wavelengths, and superconducting magnets are required above 60 GHz. In addition to oscillators, amplifier versions have been developed using resonant cavities and traveling-wave circuits; both pulsed and cw tubes have been constructed. Figure 3.57 is a drawing of a gyrotron.

Gyrotron oscillators commercially available in the United States have cw or average pulsed output powers of the order of 200 kW at frequencies between 28 and 70 GHz, and Varian has developed a 140-GHz oscillator with 100 kW of cw output. Megawatt power outputs have been obtained at millimeter wavelengths in the Soviet Union, and ongoing gyrotron research is expected to provide higher power in the future.

So far, the practical application of gyrotrons has been mainly for heating confined plasmas for fusion reactors but not for radar. Commercially available gyrotrons operating between 30 and 70 GHz are 1 to 3 m long, and they weigh 100 to 250 kg with the focusing magnet. The beam voltage is 80 kV at 8 A, which requires a large, high-voltage power supply. Gyrotrons are suitable only for relatively permanent ground-based operation, and there are also problems with mode purity and obtaining millimeter-wave components to handle the high power levels.

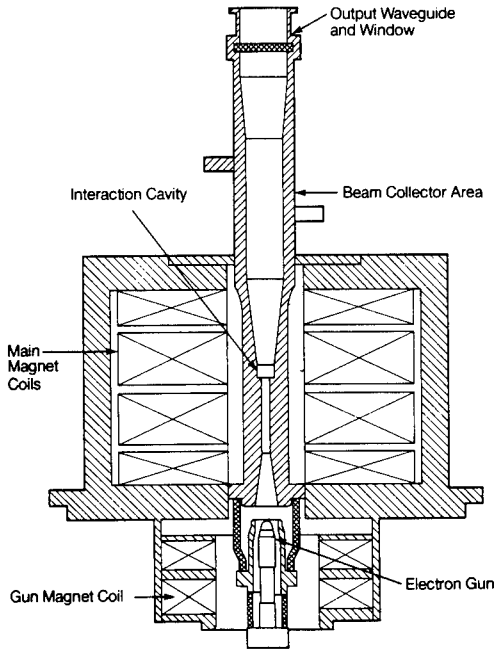


Fig. 3.57 Gyrotron in focusing magnet. (Courtesy of Varian Associates)

Magnetrons. The magnetron oscillator has been used in radar systems for about 50 yr because it is a very efficient and low-cost source of microwave or millimeter-wave power. Pulsed magnetrons are useful for simple radar systems where high output power is required.[6,92,94,95] Three types of magnetrons are in common use at millimeter wavelengths: rising sun, coaxial, and inverted.

A magnetron is a cross-field device; it consists of a concentric cathode and anode with resonant cavities and a magnet; Fig. 3.58 shows a rising sun type magnetron. The magnet applies a strong magnetic field perpendicular to the axis of symmetry, while a radial electric field is applied between the anode and the cathode. The dc voltage between the anode and the heated cathode draws electrons from the cathode, which, because of the transverse magnetic field, follow a curved path toward the anode. Near the anode the rf field either accelerates an electron so that its path becomes more curved, and it returns to the cathode, or the field decelerates an electron so that its path is less curved, and it reaches the anode. The decelerated electrons transfer part of their potential energy to the rf field. Because of this mechanism, electron clouds form spokes that rotate around the cathode at the proper velocity to transfer energy to the rf field and maintain oscillation. Table 3.7 lists the characteristics of some commercially available millimeter-wave magnetrons.

Klystrons, Extended Interaction Oscillators, and Extended Interaction Amplifiers. Klystrons are linear beam tubes in which the rf field bunches electrons traveling in a drift space, and the bunched electrons induce currents in the output circuit. The three types available include klystron amplifiers, multicavity klystron oscillators, and reflex klystron oscillators.[6,92,94,95]

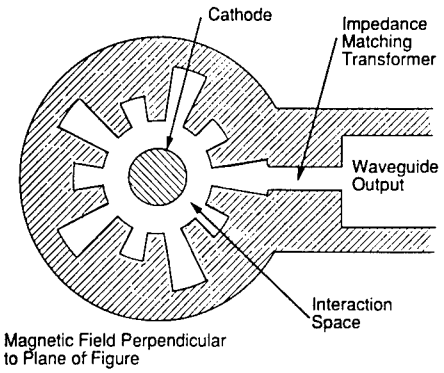


Fig. 3.58 Rising sun magnetron.

Klystron amplifiers have high gain and high peak power but narrow bandwidth. Figure 3.59 is a sketch of a three-cavity klystron amplifier; a magnetic field from a solenoid or a permanent magnet focuses the electron beam that passes through the drift tube. The anode accelerates the electrons that pass through the first cavity where they are bunched by the input rf field. After further bunching in the second cavity, the bunched electrons transfer energy to the rf field at the output cavity. Each of the cavities is resonant at a particular frequency, and the cavity tuning determines the gain and bandwidth. With proper feedback from the output cavity to the input cavity, the klystron acts as an oscillator. The VKA-7852A, a pulsed klystron amplifier in the 33- to 36-GHz frequency band, provides 2.5 kW of peak power with a 10% duty cycle, 130 MHz of bandwidth, and a 47-dB gain. A cw tube, the VA-928, has an output power of 1 kW.

A reflex klystron has a single resonant cavity. Electrons from the cathode are bunched while passing through the cavity toward the reflector on the other side, the reflector repels the electrons, and they travel back through the cavity a second time with the proper phase relationship to add energy to the rf field and sustain the oscillation. Commercially available reflex klystron oscillators

Table 3.7 Typical Parameters of Commercially Available MMW Magnetrons

Type	Center Frequency (GHz)	Peak Power (kW)	Peak Voltage (kV)	Peak Current (A)	Pulse Width	Duty Cycle
SFD-332	32.9–33.5	60	18	16	1–2 μs	0.002
M5154	33–37	1.3	4	1.5	400 ns	0.0016
MG5311	34.5–35.5	50	14.5	14.5	200 ns	0.0004
VMA-1616	34.6–35.4	25	12	9	1–2 μs	0.002
M5057	78–82	6.0	11	5	50 ns	0.0002
MG5294	94.2–96	2.5	10.5	6	50 ns	0.0005
DX252	120	2.5	10	11	—	0.0002

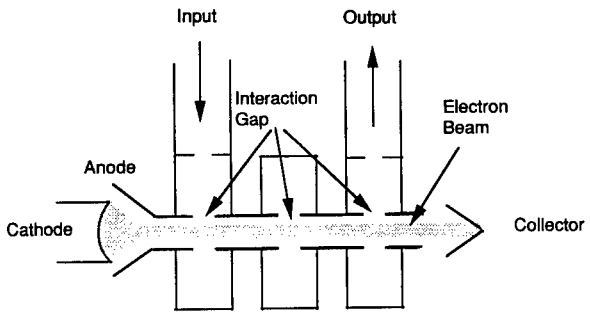


Fig. 3.59 Three-cavity klystron amplifier; magnetic field parallel to electron beam.

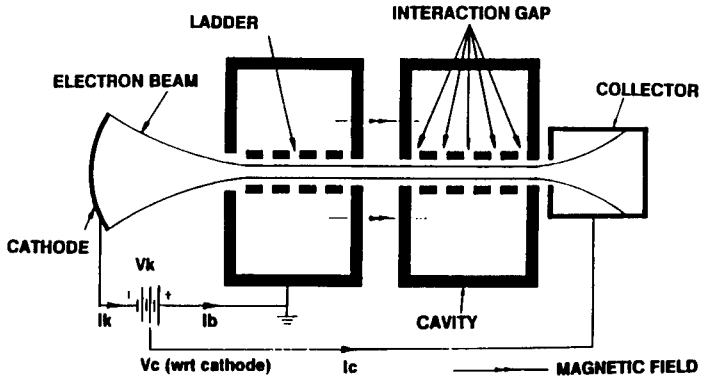
provide 500 mW of power for airborne radars or for pumping parametric amplifiers at frequencies up to about 95 GHz, 100 mW at 140 GHz, and 10 mW at 220 GHz. Solid-state sources are preferred for lower power applications.

Extended interaction oscillators (EIO) and amplifiers (EIA) provide hundreds of watts of average power and kilowatts of peak power at millimeter-wave frequencies from relatively small, lightweight packages suitable for radars.[6,94,96] An extended interaction klystron is a linear beam tube where the electrons interact with the rf field along several interaction gaps per cavity, as sketched in Fig. 3.60(a); a conventional klystron has only one gap per cavity. The extended interaction between the electron beam and the gaps provides more efficient electron bunching, a higher interaction efficiency, and a higher gain-bandwidth product. Performance at millimeter-wave frequencies is comparable to that of conventional klystrons at centimeter wavelengths in spite of higher losses.

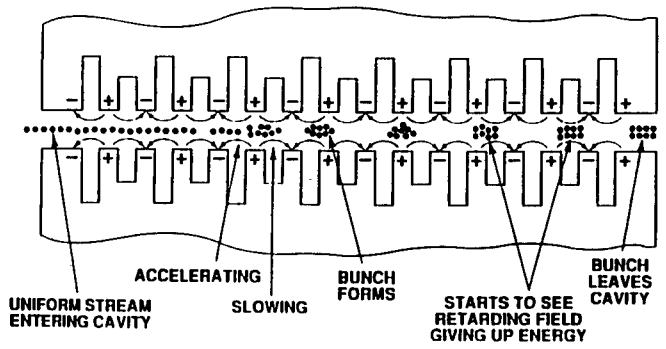
As shown in Fig. 3.60(b), the electron beam travels down a "ladderlike" circuit of alternating long and short slots in each rectangular cavity. Some electrons entering the ladder structure are influenced by accelerating fields and others by decelerating fields, and bunches are formed. Because the initial electron velocity is greater than the circuit phase velocity, the bunches drift into retarding fields and give up kinetic energy to the rf field. The collector at the far end of the drift space collects the electrons emerging from the final cavity. Figure 3.60(c) is a cross-sectional drawing of an EIO. With a 10-kV beam voltage, the electrons travel at about 20% of the speed of light. Because the electrons must reach successive slots in half an rf cycle, the slot spacing is 0.3 mm at a 95-GHz operating frequency.

Extended interaction devices are available for both cw and pulsed operation; cathode or grid modulation is available for the pulsed tubes. Table 3.8 lists the power outputs and operating frequencies for some selected extended interaction tubes. The amplifier bandwidths are several hundreds of megahertz.

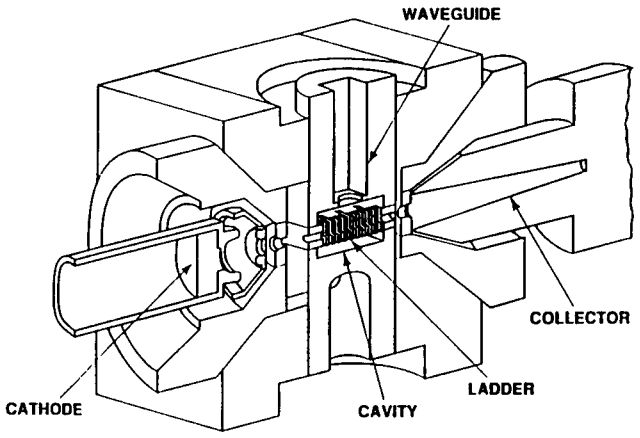
Traveling-Wave Tubes. Traveling-wave tubes (TWTs) are well suited for wide bandwidth applications.[92–95,97] A TWT, Fig. 3.61, contains an electron gun that sends a dense electron beam down the center of a slow-wave structure that supports a traveling wave that interacts with the electron beam. A collector at the end of the drift space collects the electrons after they have passed



(a)



(b)



(c)

Fig. 3.60 Extended interaction oscillator theory: (a) operation of extended interaction klystron, (b) electron and electric field at one instant in time, and (c) cross section of an EIO. (Courtesy of Varian Canada Microwave Products)

Table 3.8 Selected Extended Interaction Oscillators (EIO) and Amplifiers (EIA) (Varian Canada Microwave Products)

Device	Tube Type	Frequency (GHz)	Minimum Peak Output (kW)	dc Beam	
				Voltage (kV)	Current (A)
EIO	VKQ 2456T	30–50	1.00	21.0	0.70
EIO	VKE 2457T	50–80	1.00	21.0	0.70
EIO	VKB 2445T	80–110	1.00	21.0	0.70
EIO	VKT 2458T	110–140	0.20	21.0	0.25
EIO	VKT 2459T	140–170	0.10	21.0	0.25
EIO	VKY 2460T	170–220	0.06	21.0	0.15
EIA	VKB 2449T	94–95	1.00	21.0	0.70
EIA	VKQ 2471T	33–36	1.00	14.0	0.70

through the slow-wave structure parallel to a focusing magnetic field. The electrons in the beam travel at a velocity close to the velocity of wave propagation in the slow-wave structure, and velocity modulation occurs at the input of the slow-wave structure. The bunched electrons near the slow-wave structure output transfer energy to the rf field, and this energy is coupled out of the tube. The slow-wave structure is a helix for low-power tubes, but a helix-derived circuit, such as a ring bar or a ring loop, has a higher power capability than a helix. A typical millimeter-wave helix TWT has a bandwidth of 2 to 3 GHz near 35 GHz. A coupled-cavity (CC) slow-wave structure has a higher thermal dissipation capability than a helix, and the output power of coupled-cavity TWTs is an order of magnitude higher, but a coupled-cavity TWT has a narrower bandwidth, typically 1 GHz at 35 GHz.

The highest power TWTs use a solenoid to furnish the high magnetic field required to focus the beam, but lower power tubes use periodic permanent magnet (PPM) focusing. Both pulsed and cw TWTs are available; Fig. 3.62 shows[97] the peak output power for state-of-the-art pulsed U.S. TWTs between 10 and 100 GHz. The output power for cw tubes of similar construction is about an order of magnitude less than the values shown in Fig. 3.62. The upper frequency limit for helix TWTs is 50 or 60 GHz, but coupled-cavity TWTs operate above 300 GHz.

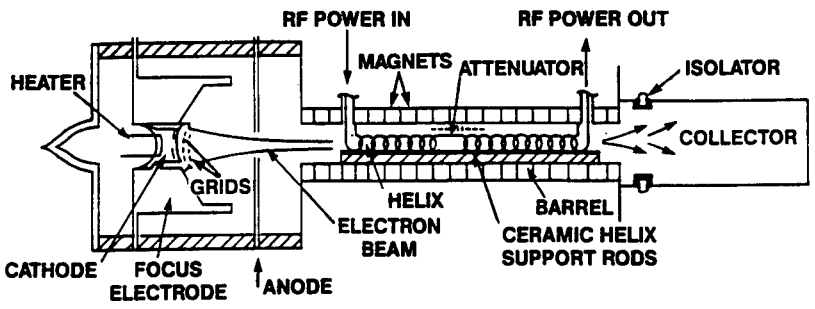


Fig. 3.61 Helix TWT.[97] (© 1989, *Microwave Journal,* reprinted with permission)

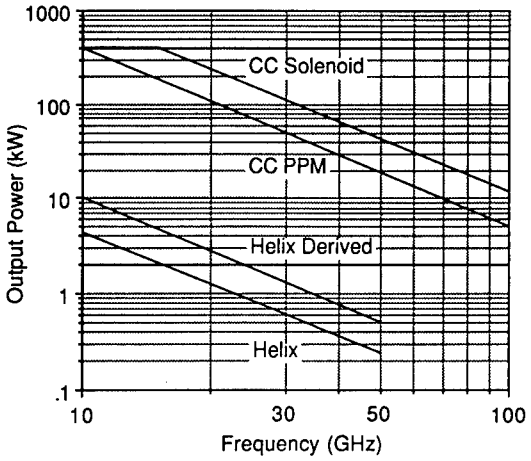


Fig. 3.62 Output power of pulsed U.S. TWTs. (Adapted from Figs. 13 and 14 of Ref. 97)

Backward-Wave Oscillators. A backward-wave oscillator (BWO) has a large electronic tuning range. BWOs, which are also known as carcinotrons, are similar to TWTs in that an electron beam interacts with a periodic slow-wave structure.[6] With a BWO, however, the electron beam and the electromagnetic wave propagate in opposite directions, and the output power is extracted near the electron gun. The helix voltage controls the frequency. Although BWOs can be built from microwave frequencies to over 1000 GHz, solid-state sources have replaced BWOs at the lower frequencies. Table 3.9 lists the frequency ranges and minimum output powers for some commercially available BWOs; the output power increases at the upper end of a tube's frequency band.

3.4.2.2 Solid-State Sources

The power outputs of millimeter-wave solid-state devices are much less than those of millimeter-wave tube sources, and there is no prospect of making up this deficiency with existing technology. The advantages of solid-state devices include small size, no requirement for a magnetic focusing field, and a low operating voltage; there are a number of short-range applications where only

Table 3.9 Typical MMW BWO Frequency Ranges and Powers

Tube Type	Frequency (GHz)	Output Power (Minimum) (mW)
VA-164M	26.5–40.0	10
F 4076/CO 40B	68–72	500
TH 4221	80–120	50
TH 4223	180–210	10
TH 4219	300–315	50

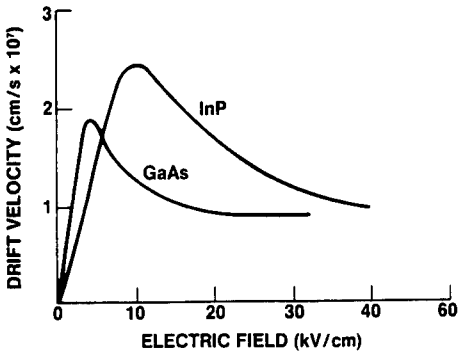


Fig. 3.63 Velocity field curves of GaAs and InP.[99] (© 1989, *Microwave Journal,* reprinted with permission)

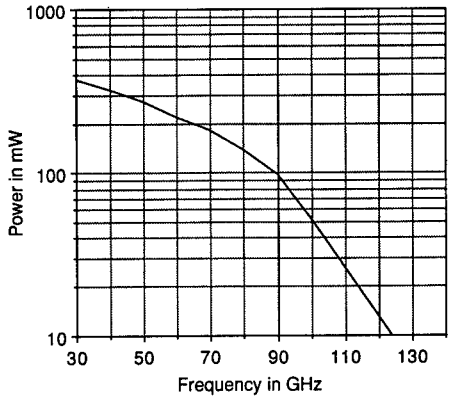


Fig 3.64 Output power from commercial InP Gunn oscillators.

solid-state sources are useful. Two-port solid-state sources include Gunn diodes and IMPATT diodes. Transistor amplifiers, which are three-port devices, are also available at the lower end of the band. Commercially available solid-state frequency multipliers provide output power up to 500 GHz.

Gunn Diodes. Gunn diodes, also called *transferred electron devices,* are semiconductors that generate rf energy because of the negative resistance that occurs when an electrical bias is present.[6,98–102] Because of their very low amplitude modulation (AM) and phase modulation (PM) noise properties, Gunn oscillators make good local oscillators for radar receivers. Gunn oscillators are also used as low-power sources for transmitters and test instrumentation and as drivers for frequency multipliers; Gunn oscillators operate at frequencies up to 140 GHz. Fixed frequency, mechanically tuned devices with bandwidths up to 6 GHz, and varactor-tuned devices with bandwidths up to 1 GHz are available; injection-locked and phase-locked sources can also be purchased. Gunn diode amplifiers, which have lower noise but less power than IMPATT amplifiers, operate[102] up to 100 GHz.

Gunn diodes are transit time devices whose operating frequency is approximately equal to the inverse of the time it takes the carrier to cross the active region.[100] If the bias across the semiconductor material is above a threshold value, the transfer of conduction band electrons from a low-energy, high-mobility valley to a higher energy valley with lower mobility causes a decrease in electron velocity with an increase in the electric field. This mobility decrease is equivalent to a negative resistance; see Fig. 3.63. With biasing in the negative resistance region, there are oscillations that depend on the doping of the semiconductor material, the diode contact, and the local circuit.

The most common Gunn diode semiconductor materials are GaAs and InP. The higher peak-to-valley ratio for InP results in higher efficiency than for GaAs. InP Gunn diodes have a higher output power than GaAs diodes above 40 GHz; the power of GaAs diodes in the fundamental mode drops rapidly above 70 GHz, although oscillation at higher frequencies is possible in a harmonic mode. Figure 3.64 plots the output power from typical commercially

available cw Gunn oscillators made of InP. The efficiency of InP diodes decreases from about 13% at 35 GHz to about 5% at 94 GHz.

IMPATT Diodes. IMPATT diodes are two-terminal devices that generate rf power because of the negative resistance that results when an electrical bias is present[99,100]; IMPATT devices are useful as pulsed or cw oscillators or amplifiers. The name IMPATT stands for *imp*act *a*valanche and *t*ransit *t*ime. Because IMPATT diodes have high AM noise, they do not make good local oscillators when the intermediate frequency is below 1 GHz. Although GaAs diodes have higher efficiency and power below 50 GHz, silicon diodes, which operate up to 300 GHz, produce more power at higher frequencies.

IMPATT diodes have either a single- or a double-drift structure. Figure 3.65 shows doping concentration for a double-drift structure that generates higher power than a single-drift structure because there are two drift zones and the junction area can be made larger. The diode consists of a heavily doped p region (p^+), moderately doped p and n regions, and a heavily doped n region (n^+); the heavily doped outer regions are connected to metal contacts. In operation, the p-n junction is reverse-biased to the breakdown voltage. If a high-frequency sinusoidal voltage is superimposed, impact ionization occurs during the positive voltage swing, and the current generates peaks near the end of the positive swing. The holes and electrons comprising this avalanche current move in opposite directions across their respective drift regions at the saturated drift velocity under the influence of the dc bias field. The drift regions have the proper length so that the current induced in the external rf circuit is 180 deg out of phase with the voltage; this causes a negative rf resistance and the generation of rf power.

IMPATT diodes are designed differently for cw and pulsed operation. The peak power output of the pulsed diodes is an order of magnitude higher than that of the cw diodes for pulse lengths that are short compared with the thermal time constant of the semiconductor. The frequency of a pulsed IMPATT changes during a pulse because of thermal heating unless the bias current is controlled. Figure 3.66 shows the state-of-the-art output power for single-diode cw and pulsed IMPATT oscillators. According to theory, the power varies inversely with the frequency below about 100 GHz because of thermal limitations; above 100 GHz, space charge limits the device current, and the power decreases inversely with the square of the frequency. Measured output powers vary approximately as the inverse third power of the frequency above 100 GHz. IMPATT oscillators can be tuned mechanically or by changing the bias. There are various techniques to combine the outputs of a number of IMPATT diodes to increase the overall output power to several times that available from a single diode.

IMPATT amplifiers operate in either the reflection mode or the injection-locked oscillator mode; both modes use a three-port circulator as an input device. In the reflection mode, the input power passes through one port of the circulator to the IMPATT diode. The IMPATT acts as a stable amplifier if the impedance of the circuit connected to the diode is greater than the negative resistance of the diode, and the diode output power then passes back through the circulator to the output port. For injection locking, a stable input signal passes through the circulator to an oscillating IMPATT diode, and the phase

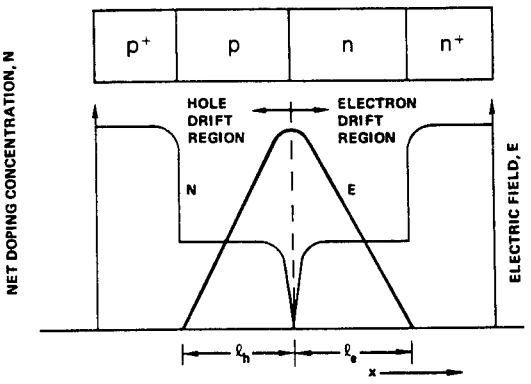


Fig. 3.65 Double-drift IMPATT diode showing field profile. (Courtesy of Hughes Aircraft Company)

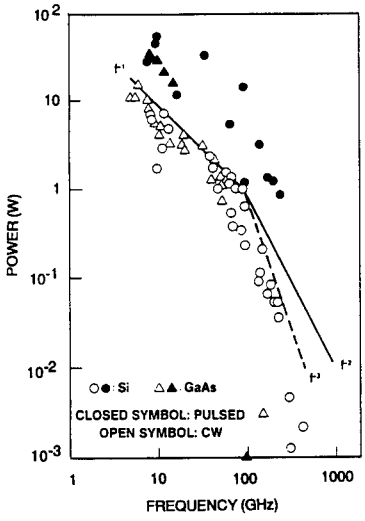


Fig. 3.66 State-of-the-art performance of IMPATT oscillators.[99] (© 1989, *Microwave Journal,* reprinted with permission)

of the oscillator locks to the phase of the input signal. Injection-locked amplifiers have nearly the same stability and noise characteristics as the locking signal and more gain but less bandwidth than the reflection amplifier. With injection locking, the output power does not go to zero with no input, unless the bias current is reduced.

A wide variety of cw and pulsed IMPATT oscillators and amplifiers is available. Up to 40 W of pulsed power at 35 and 95 GHz and 1 W at 140 GHz are available with an injection-locked amplifier over a 500-MHz bandwidth with a four-diode combiner. Injection-locked cw amplifiers provide[100] 25 W of power and 1.5 GHz of bandwidth at 30 GHz, 25 W and 2 GHz of bandwidth at 45 GHz, and 10 W and 2 GHz of bandwidth at 60 GHz. IMPATT diodes are still the only practical solid-state devices for higher power levels at millimeter wavelengths.

Millimeter-Wave Transistors. Recent developments in millimeter-wave three-terminal devices such as the GaAs metal-semiconductor field effect transistor (MESFET) and the high electron mobility transistor (HEMT) have resulted in the availability of amplifiers and oscillators in the millimeter-wave region[99] below 60 GHz. Although the power output is now lower, transistor amplifiers are likely to replace Gunn and IMPATT devices for many applications below 100 GHz in the future. Commercially available transistor amplifiers now have 2 W of output power at 35 GHz and 200 mW at 44 GHz. A noise figure of 4.3 dB has been obtained at 35 GHz.

Frequency Multipliers. Frequency multipliers utilize the nonlinear properties of semiconductor diodes to convert input power at one frequency to output power at a higher frequency.[85,98] Frequency multipliers are useful in low-power applications where it is easier or cheaper to generate power at a lower input frequency; the local oscillator for a mixer is a typical application. The output of a frequency multiplier is phase coherent with the input signal. A wide variety of frequency multipliers exist with different bandwidths and efficiencies; frequency multiplication factors of 2, 3, or 4 are typical. The output power from commercial frequency multipliers varies from tens of milliwatts at 35 GHz to 1 or 2 mW at 300 GHz.

3.4.3 Transmission Lines

There are a number of different millimeter-wave transmission lines: coaxial cable is in use at the low end of the band, waveguide, over the entire band, and quasioptical transmission line, at the upper end. Planar transmission lines find application in millimeter-wave integrated circuits and phased-array antennas.

3.4.3.1 Coaxial Cable

Coaxial cable is usable at frequencies where only the TEM mode propagates; consequently, the cable dimensions must be small to prevent higher order mode propagation. Families of millimeter-wave coaxial components are available[103] for operation up to 50 GHz, although cable losses are high: 1 dB/ft at 40 GHz and 0.8 dB/ft at 30 GHz. Connectors and cables are now becoming available for operation up to 65 GHz, and it is likely that the operating frequency of coaxial connectors will be raised[104] to 110 GHz.

3.4.3.2 Waveguide

Conventional Waveguide. Conventional waveguide continues to be in widespread use in millimeter-wave radar systems, and commercial vendors manufacture conventional waveguide components up to 140 to 220 GHz. Waveguide is a hollow tube of uniform cross section with conducting walls that confine the electromagnetic field. The solutions of Maxwell's equations determine the electromagnetic field within the waveguide, and there are two basic sets of solutions or modes.[6,105,106] The transverse electric, TE or H modes, have no longitudinal electric field component parallel to the direction of propagation along the length of the guide, whereas the transverse magnetic, TM or E modes, have no magnetic field component parallel to the direction of propagation. A TEM mode, where both the electric and magnetic field are transverse to the direction of propagation, is not possible in a hollow waveguide.

The dimensions of the waveguide and the frequency determine the propagating modes, and each mode has a cutoff frequency below which the mode will not propagate; hence, a waveguide acts as a high-pass filter. The TE_{10} mode has the lowest cutoff frequency f_c, where

$$f_c = \frac{c}{2a} \,. \tag{3.126}$$

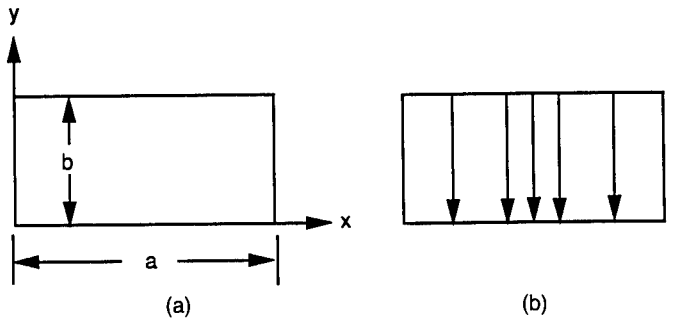


Fig. 3.67 Rectangular waveguide: (a) geometry and (b) transverse electric field for TE_{10} mode.

Here c is the velocity of light, and a is the guide width. A conventional waveguide, Fig. 3.67, is rectangular in shape with a height b that is half the width a. The normal frequency range of operation for conventional waveguide is somewhat above the cutoff frequency of the TE_{10} mode and below the cutoff frequency of the next higher propagating mode to avoid the conversion of power from the TE_{10} mode to unwanted higher order modes. Figure 3.67 also shows the electric field distribution for the TE_{10} mode: the field strength is constant in the y direction, and it varies sinusoidally in the x direction with a maximum at the center of the guide and zero intensity at the ends of the guide. There can be no component of the electric field parallel to the sides of the guide.

The wavelength in an air-filled guide is always greater than in free space. In an air-filled guide, the guide wavelength λ_g is

$$\lambda_g = \frac{\lambda}{[1 - (\lambda/\lambda_c)^2]^{1/2}} . \tag{3.127}$$

Here λ is the free-space wavelength and λ_c is the cutoff wavelength.

$$\lambda = \frac{c}{f} , \tag{3.128}$$

where f is the frequency and

$$\lambda_c = 2a . \tag{3.129}$$

The phase shift ϕ for a distance increment z along the guide is

$$\phi = \beta z , \tag{3.130}$$

where the phase constant β is

$$\beta = \frac{2\pi}{\lambda_g} = \frac{2\pi}{\lambda}\left[1 - \left(\frac{\lambda}{\lambda_c}\right)^2\right]^{1/2} . \tag{3.131}$$

Table 3.10 Parameters for Coin Silver Millimeter Waveguides for TE_{10} Mode (voltage breakdown = 15,000 V/cm)

Designator Jan & WR	Inner Dimensions (in.)	Normal Operating Range (GHz)	Cut-off Frequency (GHz)	Theoretical cw Breakdown, Low to High Freq. (kW)	Estimated Attenuation, Low to High Freq. (dB/ft)
RG-96/U WR-28	0.280 × 0.140	26.5–40.0	21.081	22–31	0.28–0.19
RG-97/U WR-22	0.224 × 0.112	33.0–50.0	26.342	14–20	0.39–0.26
WR-19	0.188 × 0.094	40.0–60.0	31.357	10–14	0.49–0.34
RG-98/U WR-15	0.148 × 0.074	50.0–75.0	39.863	6.3–9.0	0.71–0.49
RG-99/U WR-12	0.122 × 0.061	60.0–90.0	48.350	4.2–6.0	0.98–0.66
WR-10	0.100 × 0.050	75.0–110.0	59.010	3.0–4.3	1.28–0.89
RG-138/U WR-8	0.080 × 0.040	90.0–140.0	73.764	1.8–2.6	1.90–1.23
RG-136/U WR-7	0.065 × 0.0325	110.0–170.0	90.786	1.2–1.7	2.65–1.70
RG-135/U WR-5	0.051 × 0.0255	140.0–220.0	115.71	0.71–1.07	3.81–2.41

Most commercial millimeter waveguides are made of coin silver, which is 90% silver and 10% copper. Tischer[107] measured the attenuation of commercial coin silver waveguides, and he found that the attenuation was 5 to 15% above the theoretical coin silver values because of surface roughness. The excess attenuation of copper waveguide is much higher because of a room temperature anomaly of the skin effect. The attenuation α, in decibels per 100 ft, for commercial coin silver millimeter waveguide operating in the TE_{10} mode with a vacuum dielectric is approximately[106]

$$\alpha = \frac{1.34\left[0.5\left(\frac{a}{b}\right)\left(\frac{f}{f_c}\right)^{3/2} + \left(\frac{f}{f_c}\right)^{-1/2}\right]}{a^{3/2}[(f/f_c)^2 - 1]^{1/2}} \quad [\text{dB}/100\ \text{ft}] \ . \tag{3.132}$$

Figure 3.67 shows the waveguide dimensions a and b, which are in inches in this formula; f is the operating frequency, and f_c is the cutoff frequency.

Table 3.10 lists the important parameters for the commercial millimeter waveguide sizes listed in most catalogs. The theoretical cw breakdown values assume a 15,000 V/cm breakdown voltage for air, which results in a safety factor of about 2 for sea level.[106] The allowable temperature rise of the waveguide determines the average power which the waveguide can handle.[6] The attenuations in Table 3.10 for coin silver waveguide are 125% of the theoretical values for pure silver waveguide to account for the higher resistivity of coin silver and the surface roughness.

Oversize Waveguide. Oversize waveguides have lower attenuation and higher power handling capabilities than conventional waveguides in which only the TE_{10} mode propagates. A conventional waveguide is unsuitable for many millimeter-wave applications requiring low attenuation or high power. Oversize waveguides have dimensions much larger than the operating wavelength, and a large number of unwanted modes can propagate. In addition to wall attenuation, power loss occurs because of the conversion of energy from the desired TE_{10} mode to unwanted higher order propagating modes; energy conversion occurs at guide imperfections, discontinuities, and waveguide transitions to or from a smaller size guide. Oversize rectangular and circular waveguides are both in use.[6]

Oversize circular waveguide is applicable for situations where the attenuation of oversize rectangular waveguide is too large. The attenuation of TE_{0m} modes in circular waveguide, which have a circular electric field, decreases with increasing frequency, and the attenuation approaches zero at infinite frequency. Thus, oversize circular waveguide can transmit millimeter-wave power over long distances with very low loss.

3.4.3.3 Microstrip Transmission Line

Microstrip transmission line consists of a strip conductor on a dielectric substrate mounted on top of a conducting ground plane; Fig. 3.68 shows a sketch of microstrip geometry.[6,108–110] It can be shown that microstrip transmission lines employ a nontransverse electromagnetic (non-TEM) propagation mode. Because of the fringing electric and magnetic fields at the dielectric interface, longitudinal components of the electric and magnetic fields must exist. The fringing fields are, however, much smaller than the main field in the dielectric between the two conductors, and the deviation from TEM propagation is small, although the deviation increases with frequency.

One application for microstrip transmission lines at millimeter wavelengths is in microstrip patch arrays. Because of its open structure, discrete components can be mounted on microstrip or variations of microstrip such as inverted microstrip (also shown in Fig. 3.68) for the construction of hybrid integrated circuits. Microstrip transmission lines find application in hybrid or monolithic integrated circuits up to 140 GHz.

The characteristic impedance of microstrip is typically from 25 to 125 Ω; the measured attenuation of 50-Ω microstrip lines used for antennas is on the

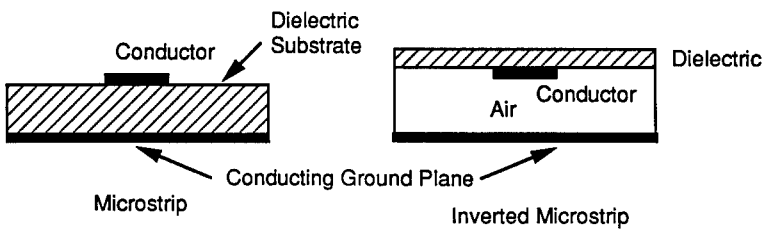


Fig. 3.68 Microstrip and inverted microstrip transmission lines.

order of 0.08 to 0.16 dB per guide wavelength. The microstrip wavelength λ_m depends[109] on the effective dielectric constant ε_{re}:

$$\lambda_m = \frac{\lambda}{\sqrt{\varepsilon_{re}}}, \tag{3.133}$$

where λ is the free-space wavelength. The effective dielectric constant is a function of the microstrip dimensions, the relative dielectric constant ε_r of the substrate, and the frequency; λ_m can be 30 to 70% of the free-space wavelength. The dimensions of microstrip lines decrease with increasing frequency, and factors such as higher mode excitation, losses, construction tolerances, and mechanical stability limit the maximum operating frequency.

3.4.3.4 Other Millimeter-Wave Transmission Lines

In addition to coax, waveguide, and microstrip, there are a number of other types of millimeter-wave transmission lines, and more than 15 types have been used for hybrid millimeter-wave integrated circuits. Brief descriptions of three other types of transmission lines follow.

Fin Line. Fin line is in wide use in millimeter-wave integrated circuits at frequencies up to 140 GHz; fin line is essentially a printed circuit ridged waveguide that has more bandwidth and a higher Q than microstrip.[6,108] Fin line consists of printed circuit fins on one or both sides of a dielectric substrate which is mounted across a rectangular waveguide parallel to the electric field, as shown in Fig. 3.69. Typical characteristic impedances range from 100 to 600 Ω, and the guide wavelength can vary from 0.9 to 1.6 times the free-space wavelength.

Coplanar Guide. Coplanar guide consists of a center strip between two ground planes parallel to and in the plane of the strip on top of a dielectric substrate (see Fig. 3.70).[6,110] Because of its planar geometry, coplanar guide is suitable for millimeter-wave integrated circuits, but it is not as popular as microstrip. Like microstrip, coplanar guide uses quasi-TEM propagation.

Image Line. Image line consists of a dielectric substrate with a rectangular cross section on a ground plane, as shown[6] in Fig. 3.71. Image line finds use in both passive and active integrated circuits, and active devices can be built directly on the dielectric substrate. Energy propagates in a transverse electric mode as in rectangular waveguide.

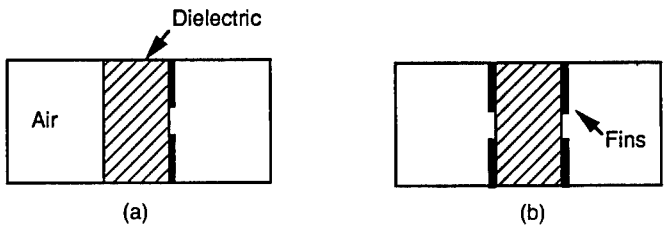


Fig. 3.69 Fin lines: (a) unilateral and (b) bilateral.

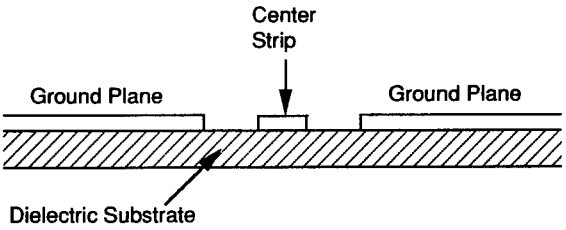


Fig. 3.70 Coplanar transmission line.

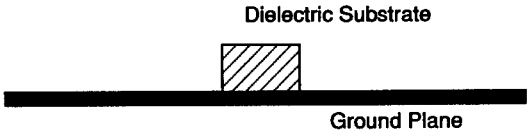


Fig. 3.71 Image line.

3.4.4 Mixers and Detectors

3.4.4.1 Schottky Diodes. Important receiver parameters include noise figure, bandwidth, dynamic range, and the level of the intermodulation products generated. Although low-noise preamplifiers are becoming increasingly popular at frequencies up to 50 GHz, Schottky barrier mixers are less expensive for applications below 50 GHz, and they are the most suitable front-end component at higher frequencies where low-noise preamplifiers are not yet available. In addition to the mixer contribution, the local oscillator and IF amplifier contributions to the overall receiver noise figure are also significant.[111]

A Schottky diode has a metal-to-semiconductor contact whose rectifying properties are similar to those of a conventional p-n junction, but the lower barrier potential results in higher sensitivity, an advantage for unbiased operation with a low local oscillator drive level. Because the current flowing through a Schottky diode is the result of majority carrier flow, there are no charge storage effects because of the absence of minority carriers.[6,100] The honeycomb and beam-lead Schottky diodes shown in Fig. 3.72 are both in use at millimeter wavelengths, and the two diode types have the same basic junction characteristics.

The top surface of the honeycomb structure is a protective silicon dioxide layer through which an array of 2- to 5-μm-diam circular holes is etched. The silicon dioxide layer is on top of an n-type epitaxial layer grown on a n^+ substrate; the semiconductor material is GaAs. There is an ohmic contact on the bottom of the n^+ substrate. The Schottky barrier junction is formed by depositing gold or another suitable conductor on the n-type epitaxial layer through the holes in the silicon dioxide. A whisker contact is made to one of the small diodes. The honeycomb diodes, which have a higher cutoff frequency than beam lead diodes, exhibit the best performance possible at millimeter wavelengths, but the whisker contact is unreliable in the presence of shock or vibration. Honeycomb diodes are poorly suited to low-cost volume production.

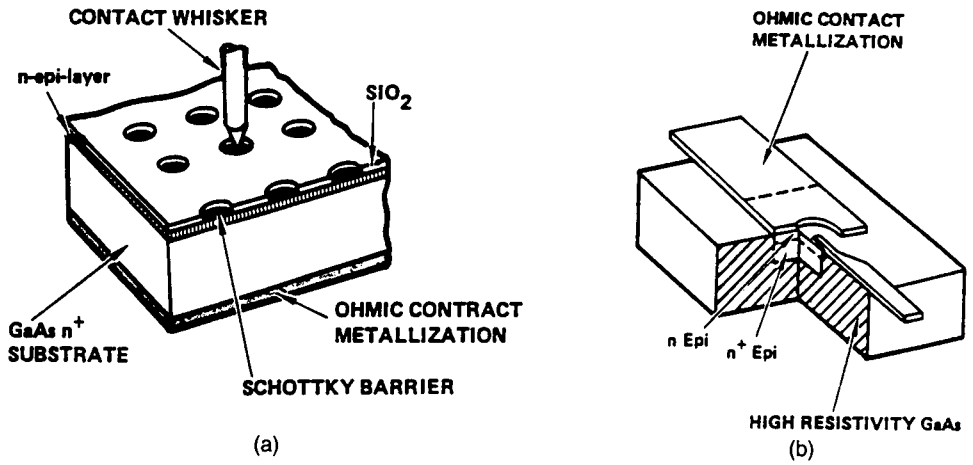


Fig. 3.72 Schottky barrier diodes: (a) honeycomb and (b) planar beam lead. (Courtesy of Hughes Aircraft Company)

A beam lead diode also contains an n-type epitaxial layer grown on a n^+ substrate. One lead to the diode makes an ohmic contact to the n^+ substrate, and the Schottky barrier is formed between the other lead and the n-type epitaxial layer. The beam lead diode is suitable for mass production, and it is mechanically reliable.

The current-voltage characteristic of a diode determines its nonlinear resistance.[6]

$$I = I_s\left[\exp\left(\frac{eV}{nkT}\right) - 1\right] . \tag{3.134}$$

Here I is the current, I_s is the saturation current, e is the electronic charge, V is the junction voltage, n is the ideality factor (about one for a Schottky diode), k is Boltzmann's constant, and T is the absolute temperature of the crystal. Schottky barrier diodes are now the only diodes in widespread use for millimeter-wave detectors, mixers, upconverters, and downconverters; Schottky barrier diodes are used in components built in waveguide, fin line, microstrip, other types of transmission line, and in monolithic integrated circuits.

3.4.4.2 Detectors. Single diodes that rectify and detect low-level rf signals are called video detectors.[101] The sensitivity of a diode depends on its noise properties and its equivalent circuit: diode noise contributors include thermal noise, shot noise, and flicker noise. The random motion of electrons in a resistance causes thermal noise, whereas shot noise is the result of the random crossing of electrical carriers through a junction. Thermal noise is constant from dc to 1 GHz, and shot noise is constant from dc to 100 MHz. Flicker noise, which is the dominant noise source in video detectors, is not well understood; it appears to be related to semiconductor surfaces in detectors. Because flicker noise power varies inversely with the frequency below a corner frequency, flicker noise is high at low frequencies.

Video detectors have a flat response over wide rf bandwidths; their sensitivity is proportional to the square root of the bandwidth of the amplifier following the detector. The tangential sensitivity (TSS), which describes the noise performance of a video detector, is the input signal level that gives an output SNR of 8 dB. For a bandwidth of 40 Hz at an output center frequency of 1 kHz, commercial detectors have TSS values[101] varying from −59 dBm at 60 GHz to −40 dBm at 300 GHz.

3.4.4.3 Mixers. A superheterodyne receiver has much better sensitivity than a video detector receiver, and amplification is much easier at an intermediate frequency than at the rf carrier frequency. In a superheterodyne receiver, the rf signal and the local oscillator signal are both applied to the mixer diode, and the nonlinear resistance of the diode results in a large number of modulation products. In normal operation, the signal frequency equals the local oscillator frequency plus or minus the intermediate frequency. The IF amplifier amplifies the difference frequency, and the higher frequency outputs of the mixer are filtered out. The IF center frequency is normally high enough to avoid the flicker noise of the mixer.

A balanced mixer has two diodes, and the voltage across one diode is the sum of the signal plus the local oscillator, whereas the voltage across the other diode is the signal minus the local oscillator. A balanced mixer has better performance than a single-ended (single-diode) mixer because of the reduction in spurious responses, and the 20- to 30-dB suppression of the local oscillator noise sidebands. Millimeter-wave balanced mixers are commercially available up to 140 GHz; single-ended mixers are commercially available through the entire millimeter-wave band.

Using the cascade noise figure equation, it can be shown that the noise figure F of a mixer is[6]

$$F = L_c(N_R + F_{\text{IF}} - 1) \ , \tag{3.135}$$

where L_c is the conversion loss between the signal and IF terminals, N_R is the noise ratio of the mixer, and F_{IF} is the noise figure of the IF amplifier. The noise ratio is the output noise power seen at the IF terminal divided by the output power of a resistor with the same IF impedance seen looking into the mixer when the local oscillator is present. Because N_R equals 0.85 to 1.05 for Schottky barrier diodes, N_R is approximately unity, and the noise figure of a mixer and an IF amplifier is

$$F \cong L_c + F_{\text{IF}} \quad [\text{dB}] \ , \tag{3.136}$$

when all quantities are expressed in decibels. A low conversion loss and a low IF noise figure are thus essential for a good overall noise figure. The IF noise figure increases as the IF bandwidth increases; IF bandwidths of up to 18 GHz are available with millimeter-wave mixers.

Because it is difficult to separate noise contributions from the upper and lower sidebands at the mixer input during noise figure measurements, millimeter-wave catalogs normally provide the double-sideband (DSB) noise temperature or noise figure rather than the single-sideband (SSB) values. The SSB noise

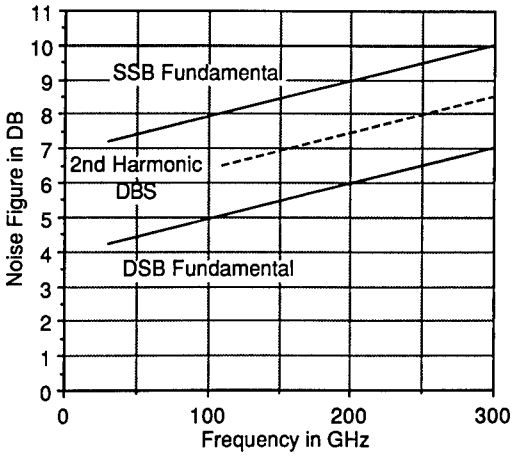


Fig. 3.73 Typical noise figure of commercial MMW mixers.

figure is applicable to a radar where a signal is present only at the signal sideband, but noise is received through both the signal and image sidebands. The DSB noise figure is appropriate for radiometers which utilize both mixer sidebands. With equal conversion loss at both sidebands, it can be shown that the SSB and DSB noise temperatures, T(SSB) and T(DSB), and noise figures, F(SSB) and F(DSB), are

$$T(\text{SSB}) = 2T(\text{DSB}) = 580[F(\text{DSB}) - 1] \ , \tag{3.137}$$

$$F(\text{SSB}) = 2F(\text{DSB}) \ . \tag{3.138}$$

Figure 3.73 shows plots of typical SSB and DSB noise figures versus frequency for commercial millimeter-wave mixers, assuming an IF noise figure of about 1.5 dB.

At the higher millimeter-wave frequencies, harmonic or subharmonic mixers are sometimes used to take advantage of cheaper sources at lower frequencies. With these mixers, the signal mixes with the second or higher harmonic of the local oscillator frequency. A harmonic mixer employs a single diode, but the subharmonic mixer has antiparallel diodes to suppress unwanted modulation products.[6] A typical noise figure for a subharmonic mixer using the second harmonic is only 1.5 dB higher than for a fundamental mixer. Harmonic mixers using a higher harmonic of the local oscillator degrade in performance as the harmonic number increases, and the noise figure is much higher than that of a fundamental mixer.

3.4.5 Waveguide, Ferrite, and Switching Components

3.4.5.1 Standard Components. Although some of the simpler waveguide components are commercially available at frequencies up to 220 GHz or even 300 GHz, waveguide losses increase, and many components are either unavailable or have poor performance[111] beyond 140 GHz. Components required for millimeter-wave radars are in common use from 30 to 140 GHz, but the

lower the frequency, the better the performance: losses increase and power handling capabilities decrease with increasing frequency. Waveguide attenuators, loads, directional couplers, tuners, tees, filters, etc., are catalog items up to 140 GHz or higher, and such devices can also be fabricated in microstrip and other transmission lines.

3.4.5.2 Ferrite Devices. Millimeter-wave ferrite devices include circulators, isolators, switches, phase shifters, and polarization switches for antennas[85,111,112]; circulators and isolators are the ferrite devices most widely used. As shown in Fig. 3.74, a three-port circulator can serve as a duplexer to separate a transmitter and receiver connected to the same antenna, or as an isolator to absorb unwanted reflections from a load when the third port is terminated in a matched load. Both fixed and switching circulators are sold; the switching versions serve as variable attenuators or for load switching.

Although the first circulators utilized Faraday rotation, most present devices are either differential phase shift or Y-junction circulators. Differential phase-shift circulators, which utilize the nonreciprocal phase shift properties of ferrites in an applied magnetic field, find application for very high peak or average powers. The Y-junction circulator, however, is more common; the operation of a Y-junction circulator depends on the resonance properties of ferrite material, which is usually in the form of a circular disk with an applied magnetic field normal to the disk.

The important properties of a circulator are its insertion loss, isolation, input voltage standing wave ratio, bandwidth, and peak and average power handling capabilities. The insertion loss for a differential phase shift circulator increases from about 0.5 dB at 30 GHz to 1.3 dB at 100 GHz with an isolation of about 20 dB. The peak power decreases from about 50 kW at 30 GHz to 4 kW at 100 GHz. Y-junction circulators have a peak power capability 7 to 10 dB below that of differential phase shift circulators. Low-power isolators and circulators are available at frequencies up to 170 or 220 GHz.

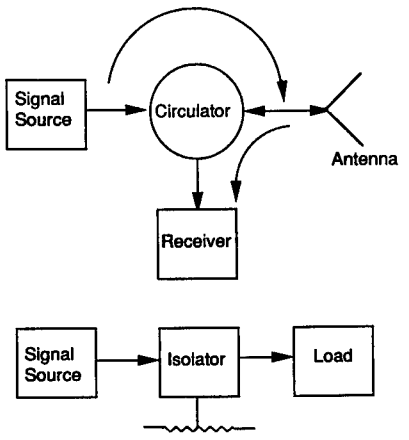


Fig. 3.74 Circulator and isolator operation.

Ferrite phase shifters vary the phase along a signal path without changing the physical path length[113]; their major use is in electronic antenna scanning. Phase shifters are two-port devices that have a low insertion loss and a low standing wave ratio. Reciprocal phase shifters have the same variable differential phase shift in both directions of propagation, whereas the differential phase shift depends on the direction of propagation for nonreciprocal phase shifters. With a latching phase shifter, a current drives the ferrite into saturation at one end of the hysteresis curve, the current is then removed, and the magnetic bias remains. A nonlatching phase shifter requires a continuous current to maintain the magnetic bias.

The two types of ferrite phase shifters in use at millimeter wavelengths are both latching phase shifters: the twin-toroid phaser is nonreciprocal, whereas the dual-mode phaser is reciprocal. The ferrite materials used in their construction have a hysteresis loop that is nearly square. The twin-toroid and dual-mode phase shifters cover the frequency range from about 2 to 94 GHz with an insertion loss that increases from about 0.8 dB at 30 GHz to 2.5 or 3.5 dB at 94 GHz. The phase shift can be continuously varied over a range of more than 360 deg in a typical unit. At millimeter frequencies, the switching time for a twin-toroid phase shifter is of the order of 1 μs, whereas a dual-mode phase shifter has a switching time of the order of tens of microseconds.

3.4.5.3 Semiconductor Switching Components. Millimeter-wave radar systems make extensive use of semiconductor control devices for switches, phase shifters, attenuators, limiters, and biphase modulators.[6,85,100,114] Semiconductor switching components include Schottky-barrier diodes, and field-effect transistors (FETs), but silicon *p-i-n* diodes are used most frequently. A *p-i-n* diode has a heavily doped *p* region and a heavily doped *n* region separated by a lightly doped intrinsic region with high resistivity. The *p-i-n* diode has very low impedance with forward bias and very high impedance with reverse bias.

Important properties of semiconductor switches include insertion loss, isolation, power handling capability, and switching time. Single-pole-single-throw (SPST) and single-pole-double-throw (DPDT) *p-i-n* switches are commercially available up to 110 GHz. A typical SPST switch covers a full waveguide bandwidth with an isolation of about 20 dB and an insertion loss that increases from about 1 dB at 30 GHz to 2 dB at 100 GHz. The peak and cw power-handling capacities are 10 and 1 W, respectively. The switching time is typically 250 ns for a broadband switch, but high-speed switches with narrower bandwidths have 1-ns switching times.[100]

3.4.6 Quasioptical Components

At the higher millimeter-wave frequencies, problems with waveguide components include relatively high losses, tight manufacturing tolerances, and difficulty in performing certain functions. Quasioptical techniques are becoming increasingly popular because of their relatively low loss, wide bandwidth, multipolarization capability, and versatility. Although quasioptical techniques are often applied at lower frequencies, they are particularly useful above 140 to 200 GHz, where waveguide components become unattractive. Quasioptical components include transmission lines, power combiners, power dividers, polarization processors, multiplexers, filters, mixers, and detectors.[115–117]

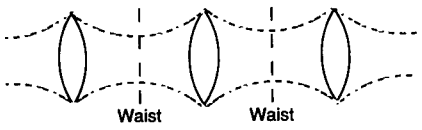


Fig. 3.75 Beam waveguide using lenses.

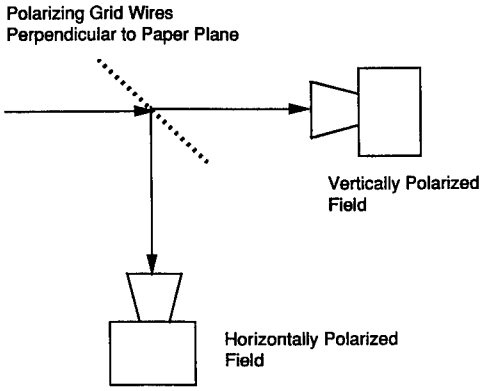


Fig. 3.76 Polarizing grid to separate two received signals.

Quasioptical techniques use Gaussian beams propagating through free space. Gaussian beams have an electric field distribution of the form

$$\frac{E(r)}{E(0)} = \exp[-(r/w)^2] \ , \tag{3.139}$$

perpendicular to the direction of propagation. Here r is the distance from the propagation axis and w is the beam radius where the electric field is 36.8% of its on-axis value. A Gaussian beam maintains its form as it propagates, and it can be focused to a minimum size called the *beam waist.* The wave fronts are collimated at the beam waist, but they become curved as the distance from the waist increases. Figure 3.75 shows a beam waveguide for power transmission; a series of dielectric lenses refocuses the Gaussian beam at periodic intervals to control beam spreading. A clear diameter of 3.5 to 4 times the waist is necessary to prevent excessive truncation at the beam edges.

The polarizing grid in Fig. 3.76 is an example of a commercially available quasioptical component for beam division. The closely spaced wires reflect nearly 100% of the incident field whose electric vector is parallel to the wires and transmit the electric field components perpendicular to the wire; hence, the polarizing grid separates signals at two different frequencies that have polarizations parallel and perpendicular to the grid wires.

3.4.7 Millimeter-Wave Integrated Circuits

3.4.7.1 Hybrid Integrated Circuits. A hybrid microwave integrated circuit (MIC) is a circuit where printed transmission lines are photoetched on a low-loss dielectric substrate, and then active and passive devices, such as resistors, chip capacitors, transistors, mixer diodes, or *p-i-n* diodes, are added with suitable bonding techniques to provide the required circuit function.[110,118–120] Circuit designers use a variety of media with a planar surface suitable for fabrication using photomasking and etching; some of the most popular transmission line types include microstrip, suspended stripline, and fin line. Millimeter-

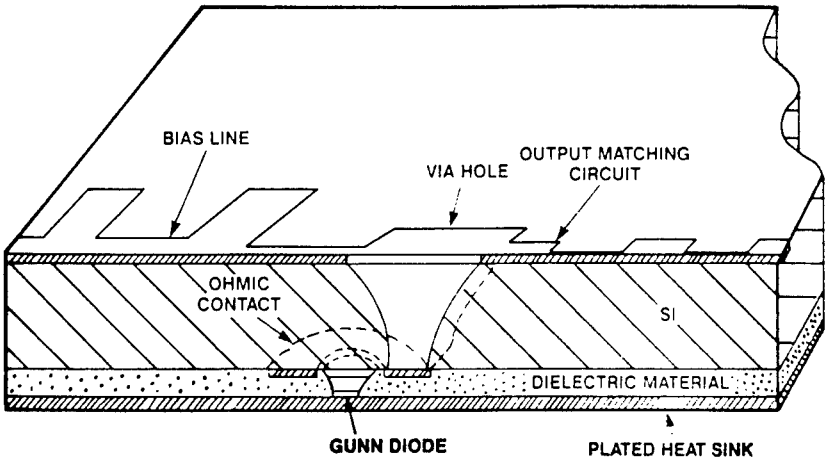


Fig. 3.77 Schematic diagram of a Gunn oscillator MMIC.[120] (© 1990, *Microwave Journal*, reprinted with permission)

wave hybrid microstrip components have been constructed up to 140 GHz. Hybrid MIC construction has the disadvantage of requiring many manual processing steps, which results in a reduction in yield and an increase in cost.

3.4.7.1 Monolithic Microwave Integrated Circuits (MMIC). Future military applications of millimeter-waves will require large numbers of inexpensive components for precision guided munitions, sensor fused munitions, and phased-array antennas. Waveguide components needed for these applications are large and bulky, and waveguide and hybrid MIC components are not well suited to mass production. Because MMICs are the key to future large-scale millimeter-wave system deployment, the Department of Defense has sponsored the Microwave and Millimeter Wave Integrated Circuits Program to improve the MMIC technology base.

Fabricating all the active and passive components of a MMIC on the same substrate eliminates the need for discrete component mounting for hybrid MICs, and only a small number of wire bonds are required for outside connections.[119] Figure 3.77 is a schematic diagram of a Gunn oscillator MMIC; the substrate material for a millimeter-wave MMIC is usually GaAs. Most past MMICs have performed only one function, but many functions can now be combined on a single chip.[120]

3.5 APPLICATIONS

3.5.1 Discussion

Millimeter-wave radars have advantages and disadvantages compared with radars operating in the microwave frequency band and with optical sensors.[1–7] Compared with the microwave band, a millimeter-wave radar has the advantages of a narrower antenna beamwidth for a given aperture size; consequently, the antenna gain is higher, and the angular resolution is better. The narrow

beamwidth is also an advantage for multipath reduction for tracking airborne targets at low angles and for reduced electronic countermeasures vulnerability. Wide bandwidths are available at millimeter waves for fine range resolution and low probability of intercept operation, and Doppler frequency shifts are larger for a given target velocity. Other advantages include the smaller size and weight of components compared with microwave radars.

One fundamental disadvantage of operation at millimeter-wave frequencies versus microwave frequencies is increased atmospheric attenuation. Although atmospheric losses in clear air, clouds, or rain are small at microwave frequencies, losses become significant in the millimeter-wave band and limit the operating range. Other disadvantages of millimeter-wave operation include higher receiver noise figures and lower transmitter power because of the smaller component size. Component availability is poorer at millimeter wavelengths, and components are more expensive. Most millimeter-wave radar systems have a maximum operating range of 10 to 20 km; microwave radars can operate at much longer ranges.

The fundamental advantage of millimeter-wave radar compared with optical sensors is much better performance in adverse weather conditions or in the presence of smoke or dust: millimeter-wave radar can operate under circumstances when an optical sensor is unusable. However, the angular resolution of an optical sensor is much better, Doppler shifts are larger, and the probability of intercept can be lower.

Section 3.5.2 presents a simple millimeter-wave design trade-off example, showing how information from Secs. 3.2, 3.3, and 3.4 can be applied to determine system performance. Most millimeter-wave radars operate in the atmospheric windows at 35 or 95 GHz where components are readily available, and the design example compares the performance of airborne mapping radars operating in these windows. The mapping radars are short-range systems that make strip maps of the terrain to either side of the flight path. Such radars, called *SLARs* (side-looking airborne radars), obtain cross-range resolution by using long antennas whose long dimension is parallel to the longitudinal axis of the airborne vehicle (see Fig. 3.78). References 121 and 122 have similar analyses.

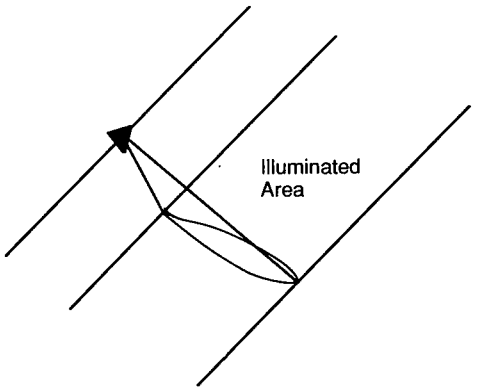


Fig. 3.78 Side-looking ground-mapping radar.

Typical millimeter-wave radar applications include

1. seekers and missile guidance
2. airborne radars for ground surveillance, fire control, or remote sensing
3. ground-based radars for surveillance and fire control for ground targets and tracking of airborne targets or space objects
4. instrumentation radars
5. automotive warning and braking systems.

Sections 3.5.3 to 3.5.7 review representative past, present, and possible future radar systems covering these applications.

3.5.2 Design Trade-Off Example

3.5.2.1 System Parameters. We assume that a small airborne platform such as a remotely piloted vehicle (RPV) carries the SLAR. The platform velocity is 250 m/s (486 kn), its altitude is 1 km, and the SLAR maps a swath extending from 1 to 10 km approximately perpendicular to the flight path; the forward motion of the platform scans the antenna. The range resolution ρ_r is 5 m, and space is available for a 3-m horizontal aperture. As discussed in Sec. 3.2, the transmitted bandwidth ΔB is

$$\Delta B = \frac{c}{2\rho_r} = \frac{3 \times 10^8}{2 \times 5} = 30 \text{ MHz} \ . \tag{3.140}$$

A 33.3-ns pulse width, therefore, provides the required bandwidth for 5-m range resolution. The cross-range (azimuth) resolution ρ_a equals the product of the horizontal antenna beamwidth β_a and the slant range. As discussed in Sec. 3.2,

$$\beta_a = \frac{k\lambda}{L} \ , \tag{3.141}$$

$$\rho_a = \beta_a R \ , \tag{3.142}$$

where λ is the wavelength, 0.00857 m at 35 GHz and 0.00316 m at 95 GHz; and L is the horizontal antenna aperture (3 m). The constant k depends on the aperture weighting for sidelobe reduction; a value of 1.2 is assumed. Figure 3.79 shows plots of the resolution versus the ground range at the two frequencies.

Assuming a flat earth, the grazing angle varies from 5.7 deg at a ground range of 10 km to 45 deg at 1 km. Elevation beam shaping, as sketched in Fig. 3.80, is necessary to provide return power that is relatively constant with ground range. A cosecant-squared elevation pattern is assumed to provide return power that is independent of slant range for a point target.[121] The elevation gain pattern has the form

$$G(\psi) = Ch^2(\csc\psi)^2 \ , \tag{3.143}$$

where C is a constant, h is the platform height, and ψ is the grazing angle.

The maximum antenna gain for an antenna aperture whose dimensions are large compared with the wavelength is[121]

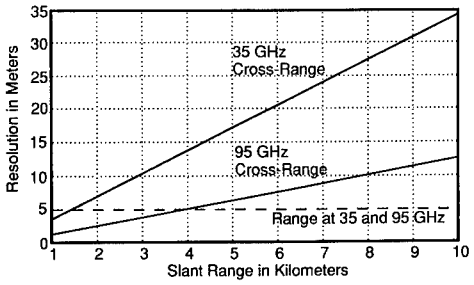


Fig. 3.79 Resolution versus slant range.

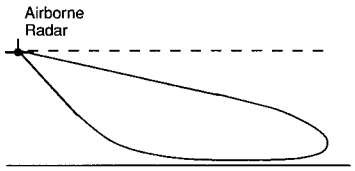


Fig. 3.80 Shaped elevation beam for surface search.

$$G_{\max} = \frac{41{,}000}{\beta_a \beta_e} \eta_a \eta_b \ , \tag{3.144}$$

where β_a and β_e are the azimuth and elevation beamwidths in degrees, η_a is the antenna efficiency (about 0.5), and η_b is the elevation beam shaping loss (0.6). From Eq. (3.141), the azimuth beamwidths are 0.196 deg at 35 GHz and 0.0724 deg at 95 GHz; the elevation beamwidth is approximately 4.6 deg with the cosecant-squared elevation pattern. With the peak response at a 5.7-deg depression angle, $G_{\max}$ = 13,642 (41.3 dB) at 35 GHz and 36,933 (45.7 dB) at 95 GHz. From Eq. (3.143), the gains as a function of grazing angle ψ become

$$G(\psi) = 134.6(\csc\psi)^2 \quad \text{at 35 GHz} \ , \tag{3.145}$$

$$G(\psi) = 369.3(\csc\psi)^2 \quad \text{at 95 GHz} \ . \tag{3.146}$$

Radar performance in both clear air and rain is of interest. For clear air, we assume a temperature of 15°C and a water-vapor density of 7.5 g/m^3; the middle curve in Fig. 3.31 gives an attenuation of 0.1 dB/km at 35 GHz and 0.5 dB/km at 95 GHz. A moderate rainfall rate of 1 mm/h is assumed, and from Fig. 3.32, the rain attenuation is 0.2 dB/km at 35 GHz and 1.0 dB/km at 95 GHz. The total attenuation with rain is the sum of the attenuations from the air with 100% relative humidity and the rain. At 15°C, the water-vapor density with 100% relative humidity is 12.8 g/m^3, and the clear air attenuation with 100% humidity is approximately double the attenuation with a water-vapor density of 7.5 g/m^3. Water-vapor attenuation is much higher than the oxygen attenuation at 35 or 95 GHz from Fig. 3.30, and clear air attenuation is approximately proportional to the water-vapor density. Hence, with 1 mm/h of rain, the total attenuation is about 0.4 dB/km at 35 GHz and about 2.0 dB/km at 95 GHz.

The clutter-to-noise ratio of the SLAR must be high enough for good ground painting; consequently, the reflectivity σ_0 of the terrain is important. As discussed in Sec. 3.3, Fig. 3.40 plots the average σ_0 values computed with Eq.

(3.110) for a grass or crop background, which is assumed for this analysis. From Eqs. (3.27) and (3.30) in Sec. 3.2, the clutter cross section σ_c is

$$\sigma_c = \sigma_0 \beta_a R \rho_r \sec\psi \ . \tag{3.147}$$

For example, at 95 GHz at a ground range of 10 km, $\psi = 5.7$ deg or 0.0995 rad. Substituting 0.0995 rad into Eq. (3.110) with the constants in Fig. 3.40, $\sigma_0 = -15.7$ dB. Because $\beta_a = 1.2 \times 0.00316/3 = 0.001264$ rad, $\sigma_c = 2.36$ dBsm.

We assume that the transmitter tube has a peak power of 10 kW at both 35 and 95 GHz. Magnetrons with a peak power of over 100 kW are available in the United States at 35 GHz, and magnetrons with more than 10 kW of peak power have been built in the former USSR at 95 GHz, but may not be available from domestic sources. With a PRF of 5000 Hz, the duty cycle is $33 \times 10^{-9} \times 5000 = 0.017\%$, and the first range ambiguity is at 30 km.

3.5.2.2 Clutter-to-Noise Computations. The SLAR must have an adequate clutter-to-noise ratio to make good strip maps. In the clutter-to-noise computations, the average clutter RCS in a resolution cell σ_c becomes the signal, and σ_c replaces the RCS σ of a point target in the radar equation. Thus, the single-pulse clutter-to-noise ratio (CNR) is

$$\mathrm{CNR}_1 = \frac{P_t G(\psi)^2 \lambda^2 \sigma_c(\psi)}{(4\pi)^3 R^4 k T_0 \mathrm{B F L_{rf}} L_{at}} \ . \tag{3.148}$$

Tables 3.11 and 3.12 list the radar parameters assumed at 35 and 95 GHz and show the clutter-to-noise computations for clear air at a 10-km range. The parameters are similar at the two frequencies, but the noise figure, rf losses, and atmospheric losses are higher at 95 GHz. Figure 3.81 plots the single-pulse CNR at the two frequencies for the clutter models assumed.

The SLAR output data are displayed on a CRT or recorded on film or digital tape, and noncoherent integration increases the CNR. The noncoherent integration gain depends on the number of pulses integrated and on the loss resulting from envelope detection rather than coherent detection. The number of pulses n noncoherently summed from a clutter cell equals

$$n = \frac{\beta_a R \, \mathrm{PRF}}{V} \ , \tag{3.149}$$

where β_a is the azimuth -3-dB beamwidth, R is the slant range, PRF is the pulse repetition frequency, and V is the velocity of the airborne platform.

The return from a clutter cell becomes decorrelated with time or with changes in aspect angle.[40,47] From Fig. 3.44, for example, the normalized frequency spectrum for deciduous trees is 3 dB down at about 7 Hz for wind speeds of 0 to 5 mph at 35 or 95 GHz. It can be shown that the decorrelation time is approximately 31 ms for a Gaussian spectrum; consequently, tree clutter from a resolution cell is constant for only about 150 pulses with a 5000-Hz PRF. Strom[47] developed an approximate equation for the decorrelation angle β_d of a clutter patch for a SLAR,

Table 3.11 Clutter-to-Noise Ratio Computations at 35 GHz

$\beta_a = 3.43 \times 10^{-3}$ rad, $V = 250$ m/s				
	Parameter	Value	dB+	dB−
P_t	Peak power	10,000 W	40.0	
$G(\psi)^2$	Antenna gain squared	2×41.4 dB	82.8	
λ^2	Wavelength squared	$(0.00857)^2$		41.3
σ_c	Clutter cell RCS	−0.20 dBsm		0.2
$(4\pi)^3$				33.0
R^4	Range to fourth power	$(10{,}000)^4$ m^4		160.0
kT_0	Boltzmann's const. × 290 deg	−204 dB (J)	204.0	
F	Noise figure	7 dB		7.0
B	Bandwidth	30 MHz		74.8
L_M	Microwave losses	4 dB		4.0
L_{at}	Atm. losses	0.1 dB/km × 20 km	____	2.0
		Totals	326.8	322.2
CNR_1 = single-pulse CNR = 326.8 − 322.2 = 4.6 dB				
Number of Pulses Integrated				
$n = \dfrac{\beta_a \text{RPRF}}{V} = \dfrac{3.43 \times 10^{-3} \times 10^4 \times 5000}{250} = 686$				
Integration Gain				
$G_i(n) = \dfrac{n(\text{CNR}_1)}{2(\text{CNR}_1 + 2.3)}\left\{1 + \left[1 + \dfrac{9.2(\text{CNR}_1 + 2.3)}{n(\text{CNR}_1)^2}\right]^{1/2}\right\}$				
$G_i(n) = \dfrac{686 \times 2.9}{2(2.9 + 2.3)}\left\{1 + \left[1 + \dfrac{9.2(2.9 + 2.3)}{686 \times 2.9^2}\right]^{1/2}\right\} = 25.8$ dB				
Integrated Clutter-to-Noise Ratio CNR_{int}				
$(\text{CNR})_{\text{int}} = (\text{CNR})_1 + G_i(n) - L_{\text{BS}} = 4.6 + 25.8 = 30.4$ dB				

$$\beta_d \cong \frac{2.5\lambda}{R\beta_a} \cong \frac{2.5L}{R} \quad . \tag{3.150}$$

Equation (3.150) assumes that the range resolution is much finer than the azimuth resolution and that the clutter patch becomes decorrelated when more than 20% of the scatterers undergo a phase shift of π radians. With the 3-m aperture, the decorrelation angle is 0.75 mrad at a 10-km range; consequently, there are about 4.6 and 1.7 independent clutter samples at 35 and 95 GHz when the radar flies past the clutter patch at a 10-km range.

The noncoherent integration gain is approximated by the gain for a nonfluctuating target because the fluctuation in the clutter cross section is expected to be small, and because it would be difficult to model and account for the fluctuation. Urkowitz[123] has developed an expression for the integration gain

Table 3.12 Clutter-to-Noise Ratio Computations at 95 GHz

$\beta_a = 1.264 \times 10^{-3}$ rad, $V = 250$ m/s

	Parameter	Value	dB+	dB−
P_t	Peak power	10,000 W	40.0	
$G(\psi)^2$	Antenna gain squared	2 × 45.7 dB	91.4	
λ^2	Wavelength squared	$(0.00316)^2$		50.0
σ_c	Clutter cell RCS	2.3 dBsm	2.3	
$(4\pi)^3$				33.0
R^4	Range to fourth power	$(10{,}000)^4$ m^4		160.0
kT_0	Boltzmann's const. × 290 deg	−204 dB (J)	204.0	
F	Noise figure	8 dB		8.0
B	Bandwidth	30 MHz		74.8
L_M	Microwave losses	7 dB		7.0
L_{at}	Atm. losses	0.5 dB/km × 20 km	____	10.0
		Totals	337.7	342.8

CNR_1 = single-pulse CNR = 337.7 − 342.8 = −5.1 dB

Number of Pulses Integrated

$$n = \frac{\beta_a \text{RPRF}}{V} = \frac{1.264 \times 10^{-3} \times 10^4 \times 5000}{250} = 253$$

Integration Gain

$$G_i(n) = \frac{n(\text{CNR}_1)}{2(\text{CNR}_1 + 2.3)}\left\{1 + \left[1 + \frac{9.2(\text{CNR}_1 + 2.3)}{n(\text{CNR}_1)^2}\right]^{1/2}\right\}$$

$$G_i(n) = \frac{253 \times 0.31}{2(0.31 + 2.3)}\left\{1 + \left[1 + \frac{9.2(0.31 + 2.3)}{253 \times 0.31^2}\right]^{1/2}\right\} = 15.6 \text{ dB}$$

Integrated Clutter-to-Noise Ratio CNR_{int}

$(\text{CNR})_{\text{int}} = (\text{CNR})_1 + G_i(n) - L_{\text{BS}} = -5.1 + 15.6 = 10.5$ dB

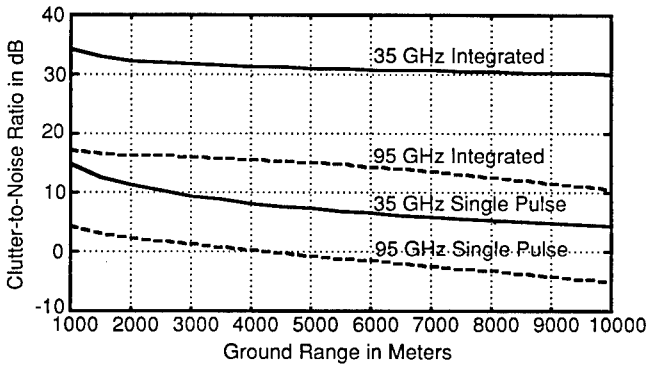


Fig. 3.81 Clutter-to-noise ratio versus ground range with clear air (grass and crop clutter model from Fig. 3.40).

G_i for the noncoherent integration of n pulses when the single-pulse SNR $(\text{SNR})_1$ is known; this equation is based on Barton's empirical expression for envelope detection loss discussed in Sec. 3.2.

$$G_i(n) = \frac{n(\text{SNR})_1}{2[(\text{SNR})_1 + 2.3]}\left(1 + \left\{1 + \frac{9.2\,[(\text{SNR})_1 + 2.3]}{n(\text{SNR})_1^2}\right\}^{1/2}\right) . \tag{3.151}$$

Because the average clutter in a resolution cell is the signal, $(\text{CNR})_1$ replaces $(\text{SNR})_1$ in Eq. (3.151). The integrated CNR in decibels then becomes

$$(\text{CNR})_{\text{int}} = (\text{CNR})_1 + G_i(n) \quad [\text{dB}] . \tag{3.152}$$

Tables 3.11 and 3.12 also show the integration-gain computations, and Fig. 3.81 shows plots of the integrated CNR at 35 and 95 GHz. The integration gain is higher at 35 GHz than at 95 GHz, because many more pulses are integrated with the wider antenna beamwidth at 35 GHz.

3.5.2.3 Signal-to-Interference Computations. Consider a tank which has an RCS of about 15 dBsm (31.6 m^2); Fig. 3.82 shows plots of the SCR for a 15-dBsm target on a grass or crop background. The SCR is higher at 95 GHz than at 35 GHz at ranges less than 4700 m because of the smaller resolution cell, but the 35-GHz CNR is higher at longer ranges because the clutter model predictions for σ_0 are considerably lower at 35 GHz at small grazing angles. The signal-to-interference (signal-to-background) ratio determines the probability of detection of a point target in a radar ground map. From Sec. 3.2.4, the signal-to-interference ratio is

$$\text{SIR} = \frac{1}{1/(\text{SNR}) + 1/(\text{SCR})} , \tag{3.153}$$

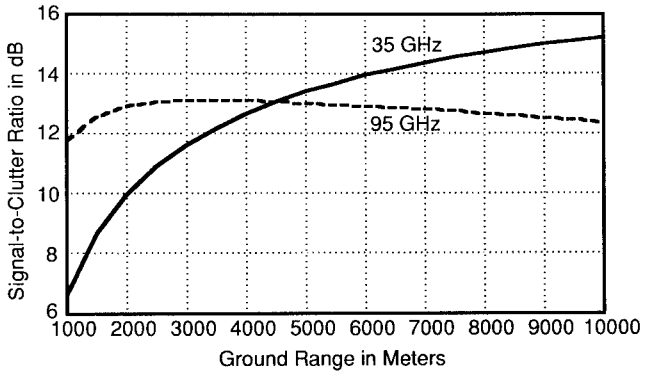


Fig. 3.82 Signal-to-clutter ratio versus ground range (target RCS = 15 dBsm, grass and crop model from Fig. 3.40).

where SCR is the signal-to-clutter ratio; i.e., the ratio of the target RCS to the clutter cell RCS given by Eq. (3.147). As Fig. 3.81 indicates, the integrated CNR is over 10 dB at 95 GHz and over 30 dB at 35 GHz. The SNR will be higher than the CNR shown in Fig. 3.81 by the SCR in Fig. 3.82. Thus, for the design example with clear air operation,

$$\text{SIR} \cong \text{SCR} \ , \tag{3.154}$$

and the background clutter limits target detection. As shown in the probability of detection curves in Sec. 3.2, a signal-to-background ratio on the order of 12 to 14 dB is needed for the reliable detection of a nonfluctuating target in a noise background. The probability of detection in clutter backgrounds is often less than that for a noise background because of the higher tails in the clutter probability density functions; consequently, the signal-to-ground clutter ratio of 12 to 14 dB may be marginal for target detection in clutter.

3.5.2.4 Performance in Rain. The presence of rain results in a significant decrease in performance. Figure 3.83 shows plots of the integrated and single-pulse CNRs at 35 and 95 GHz with a rainfall rate of 1 mm/h (light rain) covering the entire propagation path. At 95 GHz, the increased attenuation along the propagation path has lowered the CNR so much that ground mapping is not possible.

As discussed in Sec. 3.2.4, the backscattered power σ_B from rain in a resolution cell is the product of the reflectivity η and volume of the resolution cell

$$\sigma_B = \eta \frac{\pi}{4}(0.75)^2 \theta_a \theta_e \rho_r R^2 \ . \tag{3.155}$$

From Eq. (3.104), the reflectivity with a rainfall rate of 1 mm/h is 1.2×10^{-6} m^2/m^3 at 35 GHz and 1.5×10^{-5} m^2/m^3 at 95 GHz. Values of σ_B were computed,

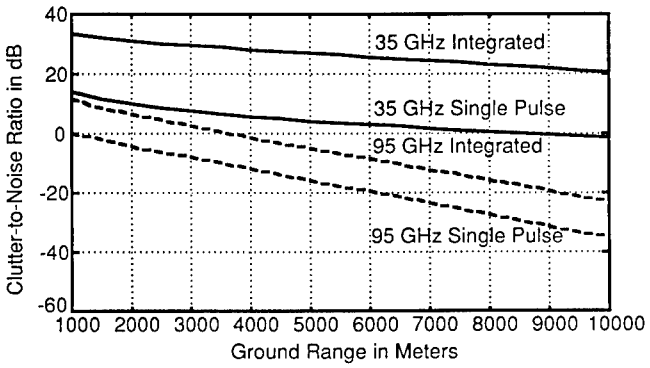


Fig. 3.83 Clutter-to-noise ratio versus ground range with 1 mm/h of rain (grass and crop clutter model from Fig. 3.40).

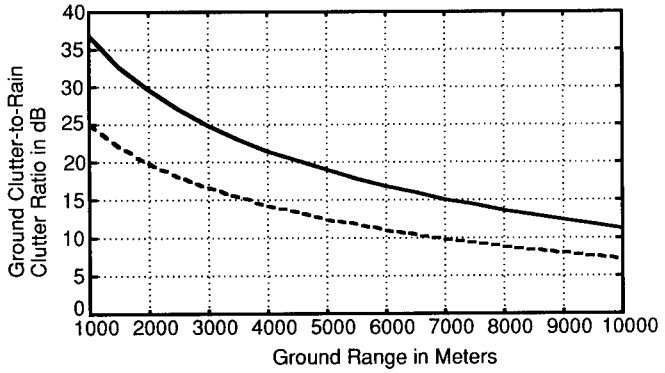


Fig. 3.84 Ground clutter-to-rain clutter ratio with 1 mm/h of rain (grass and crop clutter model from Fig. 3.40).

and Fig. 3.84 shows plots of the ratio of the ground clutter cross section σ_c to the rain backscatter cross section σ_B. Because σ_B is proportional to R^2, and σ_c is proportional to R, the ratio decreases significantly with range. The 95-GHz system cannot make useful ground maps with 1 mm/h of rain because the CNR is too low (see Fig. 3.83); consequently, the low ratio of ground clutter to rain clutter is of no importance at that frequency. With higher rain rates, however, the rain backscatter reduces map contrast at the longer ranges at 35 GHz.

3.5.2.5 Trade-Off Summary. The CNR of the 35-GHz system was better than that of the 95-GHz system with both clear air and 1 mm/h of rain; the CNR of the 95-GHz system was too low for ground mapping in the 1 mm/h rain. The improved azimuth resolution of the 95-GHz system resulted in a higher SCR only at ground ranges less than 4700 m with the grass and crop clutter models. The SCRs were marginal for tank detection at both frequencies. Higher rain rates significantly reduce map contrast at 35 GHz.

3.5.3 Seekers and Munition Guidance

3.5.3.1 Background. One millimeter-wave radar application that has great potential is the radar seeker for smart weapons. Smart weapons include sensor-fused munitions that detect the target and fire a warhead and terminally guided munitions, where the radar sensor searches an area, detects a target, and guides the missile until it hits the target, or is close enough to fire a warhead. There are no civilian applications for missile seekers, and complete descriptions of current systems are not available in the open literature. Missile applications include air to air, air to surface, surface to air, and surface to surface; there are beam-rider, semiactive, and autonomous guidance systems.[124] With a beam-rider system, the missile senses its position in a narrow beam aimed at the target and follows the beam until impact. With semiactive guidance, a separate radar illuminates the target, and the missile homes on the reflected energy. An autonomous system can be active or passive: active systems radiate signals and home on the reflected energy, whereas passive

systems home on the radiation emitted from the target. Although millimeter-wave radars have been employed in many different missile systems, a major use has been in air-to-ground weapons for attacking tanks, armored personnel carriers, and other high-value ground targets. The advantages of millimeter-wave operation include small component size and weight and a narrow antenna beamwidth with the limited aperture (15 to 20 cm) available on a missile.

In the last 15 or 20 yr of millimeter-wave seeker development, two major problems have surfaced. The first is the detection and identification of targets in the presence of heavy ground clutter.[125,126] The earlier seekers relied on the target's signal-to-background ratio for ground target detection and identification, but terrain backgrounds frequently have radar returns as high as the target's return so that the false-alarm rate per square kilometer of terrain is too high. As a consequence, other discriminants for target selection are necessary, and work on target discrimination is continuing. The use of target polarization properties and fine resolution appears promising; the most powerful polarization discrimination techniques require the measurement of the entire scattering matrix to provide all the polarization information available.[127]

The second major seeker problem is the high cost.[125,128] The assembly of millimeter-wave systems using discrete components is so expensive that the Defense Advanced Research Projects Agency (DARPA) has sponsored the monolithic microwave integrated circuits (MMIC) program to develop techniques for manufacturing microwave and millimeter-wave integrated circuits more cheaply. Millimeter-wave seekers are one obvious application for millimeter-wave integrated circuits.

3.5.3.2 Terminal Guidance. Millimeter-wave seekers find application in sensor-fused munitions and in missiles for autonomous terminal guidance for attacking ground targets.[43,129,130] Figure 3.85 shows block diagrams of millimeter-wave sensors for sensor-fused (smart) and terminally guided munitions. Sensor fusing is used for free fall and parachute-suspended munitions. With terminal guidance and lock-on-after-launch, the missile flies along a prescribed path while the antenna scans back and forth across the ground track, as shown in Fig. 3.86. After detection, the seeker guides the missile to impact. For terminal guidance, conical scan has the advantage of simplicity compared with the monopulse system, but it is not as good for tracking, particularly in a countermeasures environment. Millimeter-wave seekers normally operate at 35 or 95 GHz: 95 GHz has the advantages of a narrower antenna beam, higher antenna gain, greater Doppler sensitivity, and smaller components, whereas the advantages of 35 GHz include lower clutter reflectivity, better operation in adverse weather, and higher transmitter power.

Seekers usually use a short pulse or an FM-cw waveform. With a short pulse, the source is an IMPATT diode with a pulse width of the order of tens of nanoseconds to provide a range resolution of 5 to 10 m; the peak output power is 5 to 10 W. The short pulse receiver is a superheterodyne receiver with a local oscillator to downconvert the returns to an intermediate frequency.

With the FM-cw waveform, Fig. 3.87, a long pulse whose frequency varies linearly with time is transmitted, and the return from a point target mixes with a portion of the transmitted signal, resulting in a difference frequency

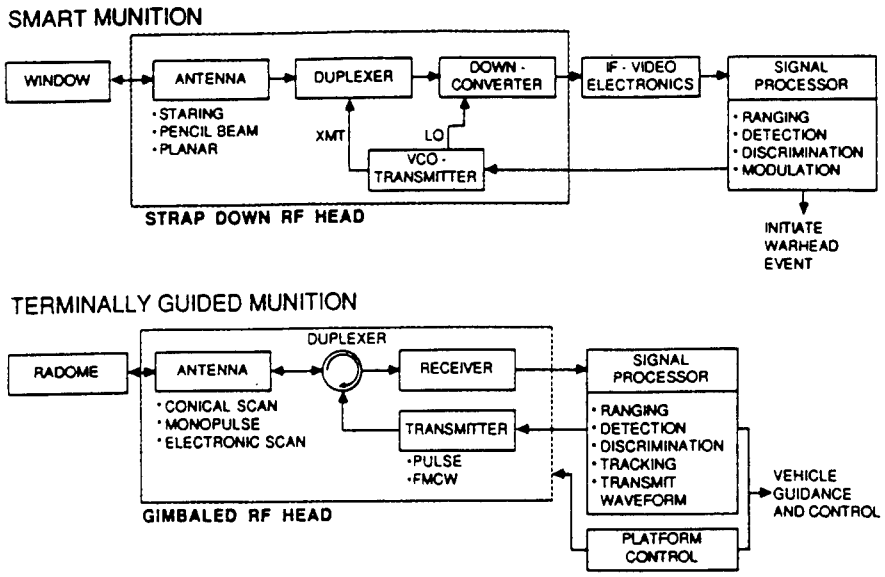


Fig. 3.85 MMW radar sensor block diagram.[131] (© 1988 SPIE, reprinted with permission)

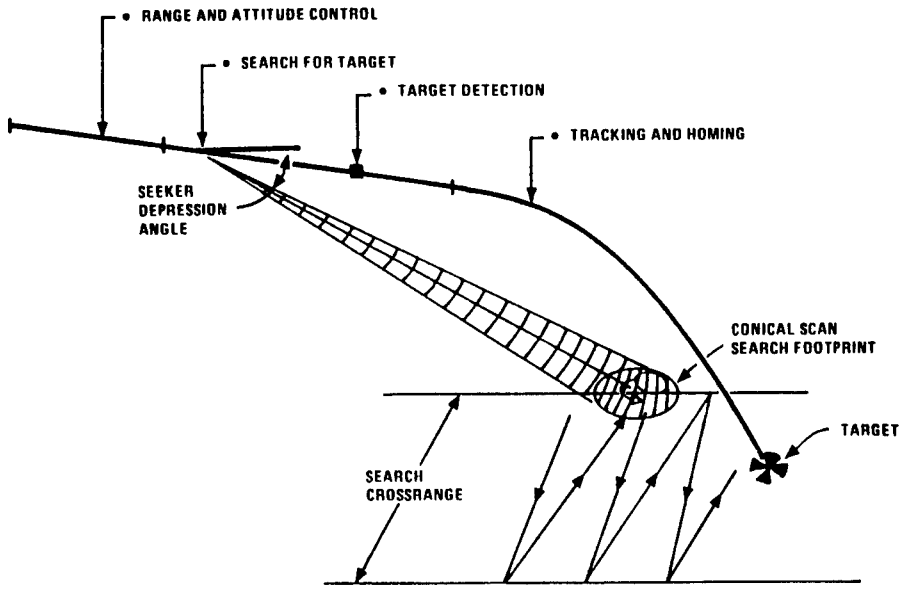


Fig. 3.86 Functions for a terminal guidance seeker.[43] (© 1981 Academic Press, reprinted with permission)

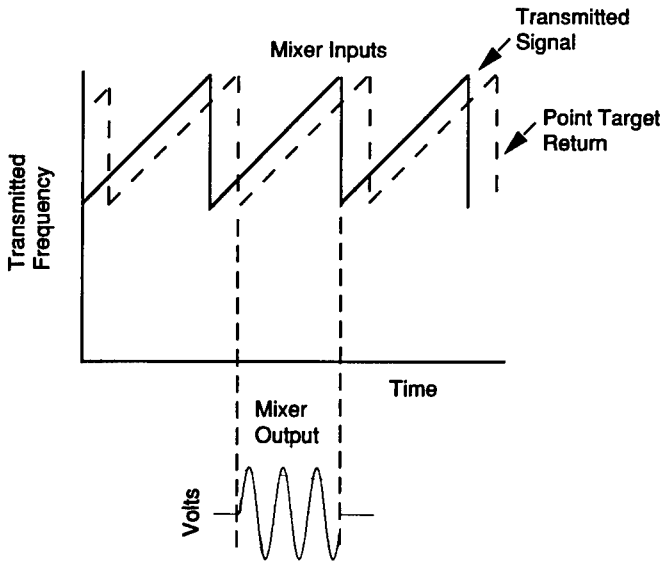


Fig. 3.87 FM-cw waveform.

Δf, which is proportional to the FM rate γ of the transmitted waveform and the range R to the target:

$$\Delta f = \frac{2R\gamma}{c} \quad . \tag{3.156}$$

Here c is the velocity of light. The FM-cw waveform has the advantages of high average power, receiver simplification (no separate local oscillator), and good short-range performance; FM-cw is a good waveform for sensor-fused munitions, which operate at short range. With this waveform, the bandwidth is held constant, and the sweep time or the beat frequency is changed to account for range changes. The source for the FM-cw waveform is usually a Gunn oscillator; the waveform requires good sweep frequency linearity and transmitter-receiver isolation.

3.5.3.3 Dual-Mode Seekers. Most seekers have two or more operating modes, because a single mode does not provide adequate performance.[43,131,132] In the 1970s and early 1980s millimeter-wave seekers combined active and passive modes. In the passive or radiometric mode, the seeker senses the temperature difference between a target and the background. With a metallic target such as a tank, the reflection of the cold sky from the metallic surface results in a lower temperature than the clutter background. The dual-mode millimeter-wave seeker operates in the active mode for search and initial tracking, and it switches to the passive mode for terminal homing. In the active mode, target glint causes an angular tracking error that is not present in the passive mode; hence, angular tracking in the passive mode is more accurate. The operating range of the passive mode is, however, shorter because the target must fill most of the antenna beam to provide sufficient temperature contrast.

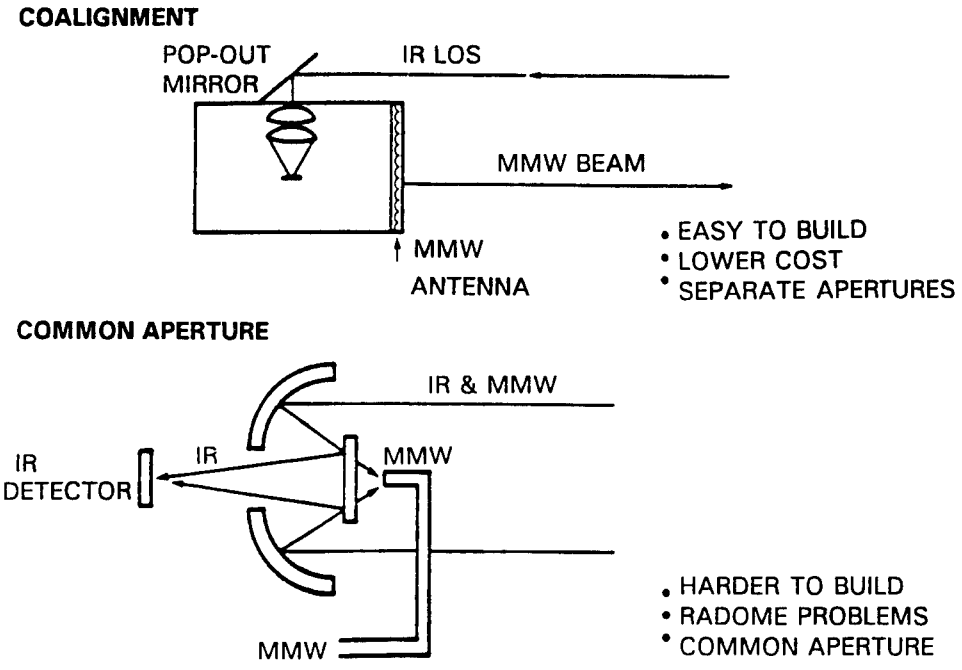


Fig. 3.88 Sensor fusion methods.[131] (© 1988 SPIE, reprinted with permission)

Seekers combining millimeter-wave and active or passive optical sensors are now under development; the fusion of an active millimeter-wave and a passive IR sensor is the predominant dual-mode combination.[131,132] The strengths of the active millimeter-wave sensor include providing range information, operation in adverse weather, and the lack of susceptibility to IR countermeasures. The passive IR sensor, which detects target hot spots, provides more accurate angular information than the active millimeter-wave sensor, and it is not susceptible to radar countermeasures. With the dual-mode sensor, the signal processor combines information from both sensors for target detection and discrimination. As with any seeker, the key to successful operation is a low false alarm rate.

As shown in Fig. 3.88, there are two methods of combining an active millimeter-wave and a passive IR sensor in the same aperture. With coalignment, both sensors have separate lines of sight but a common boresight. With the common-aperture approach, the two sensors use the same aperture, and a secondary reflector, which is transparent at one frequency and reflective at the other, separates the two signals.

3.5.4 Airborne Millimeter-Wave Radars

3.5.4.1 Background. A big advantage of using a millimeter wavelength for an airborne ground-mapping radar is a narrower antenna beamwidth for the limited aperture size available; other advantages include the availability of

wider bandwidths and smaller and lighter components.[121] Atmospheric losses in clear air and in rain are much higher than at microwave frequencies, and the transmitter power available is much less; consequently, the maximum operating range of a millimeter-wave airborne radar system may be 10 to 20 km, compared with hundreds of kilometers at lower frequencies.

There are two basic types of ground-mapping radars: real beam radars and synthetic aperture radars. The physical beamwidth of the antenna determines the azimuth resolution of a real beam radar, which can have a scanning or a side-looking antenna. As discussed in Sec. 3.2.7, synthetic aperture radars (SARs) generate the effect of an antenna that is much longer than the physical aperture of the antenna by means of signal processing.

3.5.4.2 Real Beam Radars. Most airborne scanning radars operate at K_u band (12 to 18 GHz), but a number of scanning radars have been built at 35 GHz, where the azimuth resolution is reduced by a factor of 2 for the same aperture size. Such radars provide ground maps on a plan position indicator (PPI), which displays a map in slant range-azimuth coordinates. Scanning radars may also have moving target and terrain avoidance capabilities, and the antenna normally has a shaped elevation pattern to provide ground returns that are relatively independent of range.[121]

The AN/APQ-137, or MOTARDES radar, is an example of an operational 35-GHz radar that is mounted in the nose of a helicopter.[121] This radar has a shaped elevation beam for scanning and a pencil beam for target tracking; the system uses noncoherent moving target indication (MTI) for detecting moving targets. Another radar with a K_a-band-mapping capability is the dual-frequency AN/APQ-122(V). This radar provides ground mapping out to 200 nmiles at X band; the K_a-band mode is used at short range for fine resolution ground mapping for the delivery of aerial cargo.[121] Experimental scanning systems have also been built[133] at 95 GHz.

The side-looking airborne radar, or SLAR, is the other type of real beam millimeter-wave radar in use; for good azimuth resolution, the SLAR has a long antenna that is approximately parallel to the fore-aft axis of the airborne platform. Such radars frequently have separate maps for fixed and moving targets, and data can be stored digitally or on film. The AN/APQ-97 built by Westinghouse[121] is an example of a millimeter-wave SLAR. The center frequency of this radar is 34.85 GHz, the peak power is 100 kW, and effective resolution of the antenna is 0.1 deg.

3.5.4.3 Synthetic Aperture Radars (SARs). Few millimeter-wave synthetic aperture radars have been built, because it is possible to build SARs with fine enough resolution at K_u or X band, where higher transmitter power and better components are available. The maximum operating range at these frequencies is greater because of the higher power and lower atmospheric losses, and system construction is easier at the lower frequencies. Simpler millimeter-wave radars with a real beam can sometimes replace SARs operating at lower frequencies, when the azimuth resolution is acceptable with a millimeter-wave antenna of reasonable size. Nevertheless, future development of millimeter-wave SAR systems can be expected for millimeter-wave measurements and other applications requiring finer resolution than that possible with real beam systems.

Table 3.13 ADTS Parameters (from Ref. 134)

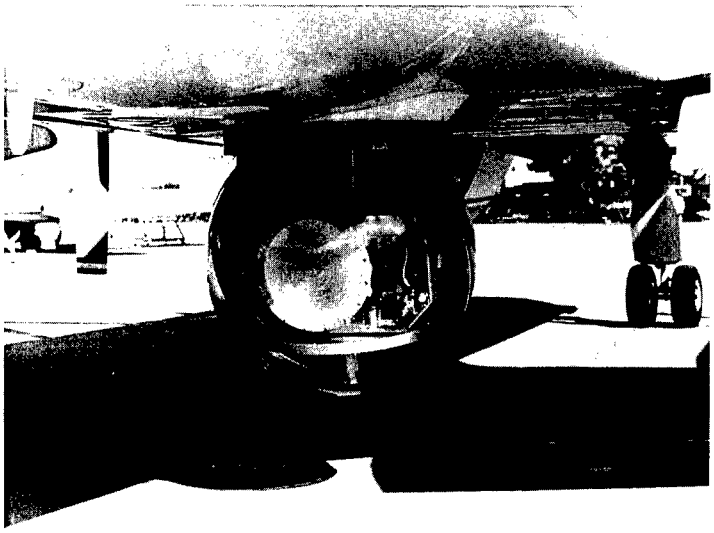
Frequency	33.56 GHz
rf bandwidth	600 MHz
SAR resolution	1 ft × 1 ft
Transmit polarization	Horizontal/vertical or left/right circular
Receiver polarization	Dual channels
Transmitter power	84 W peak
Receiver noise figure	6.8 dB
Antenna beamwidth	2.1 deg
Antenna peak sidelobes	−25 dB
Platform	Gulfstream G1
Recording media	28-channel high-density tape video recorder
Waveform	Linear FM

An example of a millimeter-wave SAR is the Advanced Detection Technology Sensor (ADTS) constructed by the Loral Corporation under a DARPA-funded program established at Lincoln Laboratory.[134,135] The ADTS radar, which is installed in a Gulfstream G-1 aircraft, collects ultra high resolution, fully polarimetric images of targets and clutter in both real- and synthetic-aperture modes to provide data for studies of the detection of stationary targets in clutter.

Table 3.13 lists key parameters for the radar that operates at a center frequency of 33.6 GHz and provides 1-ft resolution in both range and azimuth in the two SAR modes, strip map and spotlight. The antenna beamwidth, 2.1 deg, determines the cross-range resolution in the real-aperture mode, which is about 70 m at the normal operating range of 7250 m. The ADTS transmits orthogonal polarizations, either circular or linear, on alternate pulses and receives parallel and cross polarizations simultaneously in a two-channel receiver. Figure 3.89(a) shows a photograph of the ADTS radome and antenna, and Fig. 3.89(b) is an example of ADTS imagery.

3.5.5 Ground-Based Millimeter-Wave Radars

3.5.5.1 Surveillance and Target Acquisition. One ground-based radar application for which millimeter-waves are attractive is ground surveillance and target acquisition at moderate ranges: millimeter-wave operation can provide good resolution in both range and azimuth. The first millimeter-wave radar to receive an AN nomenclature, the AN/MPS-29 built at Georgia Tech 30 years ago, was designed for combat surveillance of ground targets.[136] Fine resolution was obtained with a 50-ns pulse width and a 0.02-deg azimuth antenna beamwidth. The radar scanned a 30-deg sector 20 times per second using a folded geodesic Luneberg lens antenna. The maximum detection range was 10 to 15 km for light vehicles and 18 km for a 2.5-ton truck. The output tube, a magnetron, had a peak power of 15 kW at the operating frequency of 70 GHz. Georgia Tech later built a rapid scan radar that also used a folded geodesic Luneberg lens antenna.[137]



(a)

(b)

Fig. 3.89 (a) ADTS radome and antenna (built by Loral Corporation) and (b) ADTS image of trees (after polarimetric whitening filter for speckle reduction). (Courtesy of MIT Lincoln Laboratory)

Table 3.14 STARTLE Parameters (data from Refs. 1 and 6)

Center frequency	94 GHz
Beamwidth	11 mrad
Search sectors	40 deg or 15 deg azimuth by 7.5 deg elevation 5 deg azimuth by 2.5 deg elevation
Scan rate	150 deg/s
Peak power	4 W
Average power	0.1 to 0.5 W
Target detection	3 km with 100-m visibility
Target tracking	0.5-mrad accuracy at 2 km with 100-m visibility

The STARTLE radar is another example of a ground-based millimeter-wave radar; STARTLE is the acronym for *s*urveillance and *t*arget *a*cquisition *r*adar for *t*ank *l*ocation and *e*ngagement.[6,138,139] As the name implies, the purpose of this radar is target acquisition and fire control for tank engagements. The radar complements a tank's IR thermal sight in adverse weather. Table 3.14 lists the STARTLE parameters. The radar uses IMPATT diodes as the transmitter power source, and it displays detected targets on an angle-angle display. The radar has a monopulse antenna for angle tracking, a coherent MTI mode for targets moving radially, and a noncoherent area MTI mode with a spread spectrum for smoothing clutter and glint with creeping and tangential targets.

3.5.5.2 Airborne Target Tracking. The tracking of airborne targets at low grazing angles has long been a problem for ground-based radars because of multipath and clutter. When a target is at a low grazing angle over a specular surface, the radar receives returns reflected from the surface along with the direct returns, causing large amplitude and angular fluctuations in the total received signal. The largest angular fluctuation errors are usually in elevation where the radar is apt to track between the target and the multipath reflection, but azimuth multipath errors are also possible. The most effective way of reducing multipath effects is to use a narrow antenna beamwidth to keep multipath reflections and ground clutter out of the antenna mainlobe.[140]

The use of millimeter wavelengths for low angle tracking is advantageous because of the narrow antenna beamwidths available compared with microwave frequencies; however, this advantage must be balanced against the increased propagation losses at millimeter lengths. A dual frequency radar, TRAKX (*tr*acking *r*adar at K and X bands), was designed to take advantage of the strong points of K- and X-band operation. The TRAKX radar used a single 8-ft dish with dual feeds to transmit and receive at K and X bands simultaneously; the radar could track a target in range, azimuth or elevation independently at either frequency.[141] Table 3.15 lists system parameters, and Fig. 3.90 is a system block diagram.

SEATRACKS (*s*mall *e*levation *a*ngle *tr*ack and *s*urveillance) is another K_a-band radar designed to take advantage of the narrow antenna beamwidth available at millimeter wavelengths to reduce multipath tracking errors.[142] This radar combines low-angle surveillance with monopulse tracking for short-range air

Table 3.15 TRAKX Parameters (data from Ref. 141)

Antenna	8-ft dish with combined feed
Beamwidth	1 deg X band, 0.25 deg K_a band
Polarization	Linear
Frequencies	9 GHz and 35 GHz
Pedestal	Converted Nike-Hercules
Peak power	300 kW X band, 130 kW K_a band
Pulse width	0.25 μs
PRF	160 to 1600
Noise figure	6.25 dB X band, 13 dB K_a band
Tracking	Three-channel monopulse, use of each independent coordinate
Displays	TV, A-scope, alphanumeric
System operation	Computer operation, manual backup
Calculated ranges on 1-m^2 target	Tracking to 28 nmiles at K_a band, normal day; acquisition out to 75 nmiles at X band

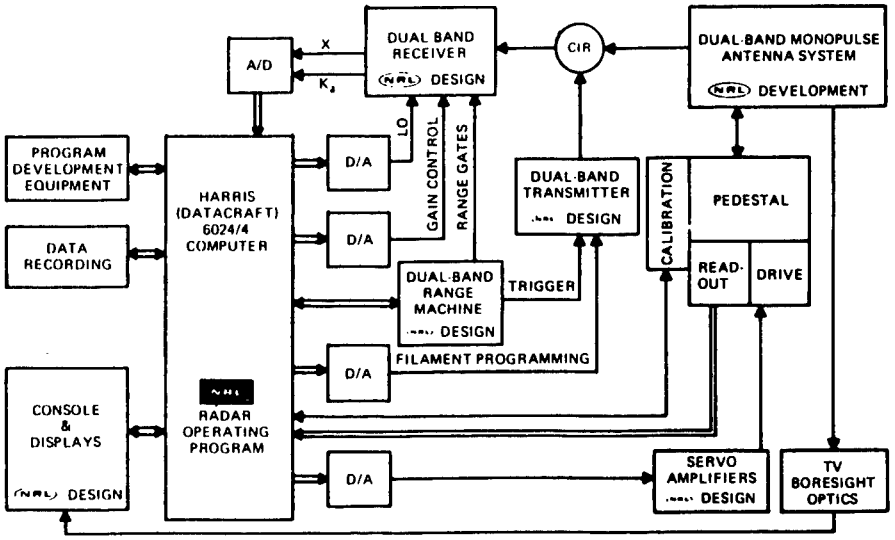


Fig. 3.90 TRAKX block diagram.[141] (© 1976 *Microwave Journal,* reprinted with permission)

Table 3.16 SEATRACKS Parameters (data from Ref. 142)

Antenna	2-ft Cassegrain with four-horn monopulse feed
Beamwidth	1 deg
Frequency	34.67 GHz
Peak power	100 kW (controllable)
Transmitter source	Coaxial magnetron
Pulse width	0.1, 0.5, or 1.0 μs
PRF	500–5500 PPS
Polarization	Vertical, horizontal, or circular
Displays	A scope, B scope
Receiver	Three-channel monopulse
Receiver noise figure	5.1 dB
Predicted detection range on 1-m^2 nonfluctuating target	29 km clear air, 10 km in 4 mm/h rain

defense. Table 3.16 lists the SEATRACKS parameters, and Fig. 3.91 is a simplified block diagram.

3.5.5.3 Space Object Applications. Millimeter-wave radars have also been proposed for space-object identification. The wide bandwidth available at millimeter wavelengths has the potential of providing better range resolution and thus more structural detail than at lower frequencies. The short wavelength also permits more precise measurement of Doppler frequencies resulting from satellite spinning and tumbling. Inverse synthetic aperture (ISAR) techniques can provide images with much smaller apertures than at lower frequencies.

Reference 143 discusses one of the earliest 95-GHz radars reported in the literature; the goal of this radar was obtaining echoes to identify objects in orbit. The radar used a 1-ms pulse width for high average power and a 1-GHz

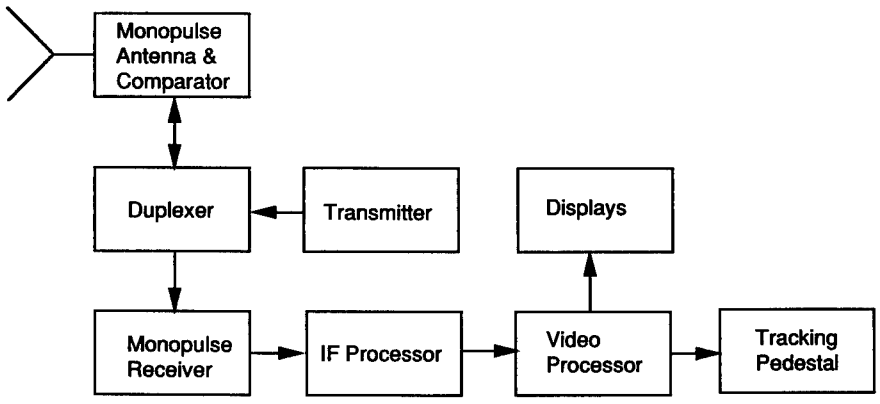


Fig. 3.91 SEATRACKS simplified block diagram.

bandwidth to obtain the desired 6-in. range resolution. Reference 144 updates requirements for space-object identification, and Ref. 145 briefly discusses millimeter-wave radar applications for ballistic missile defense.

3.5.6 Ground-Based Instrumentation Radars

3.5.6.1 Millimeter-Wave Measurements. Millimeter-wave instrumentation radars are needed for a variety of data. Extrapolation of lower frequency measurements to millimeter wavelengths leads to inaccuracies because the measurement results are wavelength dependent. Data of importance to millimeter-wave systems include atmospheric attenuation, attenuation and backscatter from hydrometeors, backscatter from land and sea, and the RCS of military and civilian targets. Indoor installations include anechoic chambers and compact ranges, whereas outdoor instrumentation may be at ground level, mounted on a tower, or on a truck or trailer for portable measurements.

As at lower frequencies, a number of different types of millimeter RCS measurements are possible. For the simplest type of measurement, the radar transmits a pulse long enough to contain the entire target, the target is rotated, and the RCS is recorded continuously as a function of azimuth angle. If the radar has a wide transmitter bandwidth for fine range resolution, the radar can resolve individual scattering centers on the target and record RCS versus slant range. In recent years, ISAR techniques have become very popular for making fine resolution, two-dimensional (range, cross-range) images of a rotating target; such images can provide diagnostic information about the locations and RCS values of individual scattering centers on the target. Because of the intense interest in polarization for target discrimination, many of the millimeter-wave instrumentation radars now being built are polarimetric; i.e., they have the capability of measuring the polarization scattering matrix at a single frequency, with range profiles, or with ISAR images.

At millimeter wavelengths, data are frequently collected on scale models as well as on full-sized targets.[31–35] In many cases, a full-sized target is not available, or it is too large and heavy for the measurement facility. Scale models are used most frequently with metallic targets where there are no problems associated with dielectric scaling. It is often assumed that both the full-scale target and its scale model have infinite conductivity, although this assumption is only an approximation. Measurements made on a $1/n$'th-scale model at $1/n$'th of the full-scale wavelength provide full-scale RCS values when the scale-model RCS values are multiplied by n^2. For example, the RCS values measured at 95 GHz on a 1/10 scale target should be multiplied by 100 to provide full-scale data at 9.5 GHz. Problems associated with scale-model measurements involve the geometric fidelity of the model and the frequency scaling of the target's electromagnetic properties.

3.5.6.2 Instrumentation Radar Examples. A number of millimeter-wave instrumentation radars have been built for target and background measurements; see, for example, Refs. 1, 85, 100, and 146 to 149. Measurement instrumentation is commercially available for single-frequency, range-profile, and ISAR RCS measurements.

Figure 3.92 is a simplified block diagram of a coherent 95-GHz front end developed for RCS measurement programs, and Table 3.17 lists the key op-

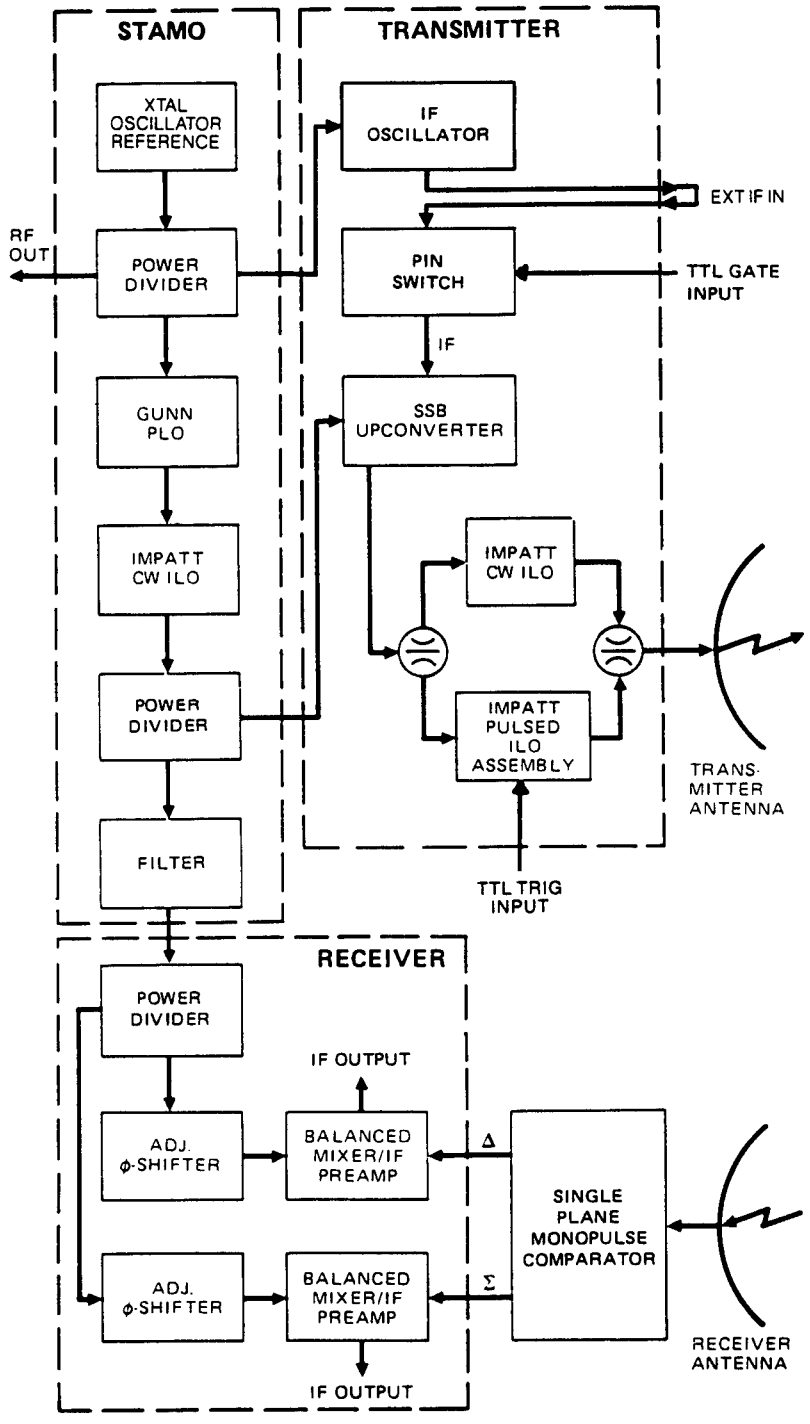


Fig. 3.92 Functional block diagram of model 42266H-2000 coherent radar front end. (Courtesy of Hughes Aircraft Company)

Table 3.17 Parameters of Model Number 42266H-2000 Radar Front-End Instrumentation (Hughes Aircraft Company)

Upconverter Inputs	
cw mode	
IF input power	17 dBm min.
Frequency	3 GHz nominal
Bandwidth	200 MHz max.
FM waveform	Ramp, sawtooth, or sinusoidal
Pulsed mode	
Pulse width	20–200 ns
PRF	60 kHz max.
Frequency	3 GHz nominal
Bandwidth	200 MHz max.
Transmitter	
cw mode	
Output power	20 dBm min.
Output frequency	95 GHz nominal
Bandwidth	200 MHz min.
Pulsed mode	
Output power	1 W min.
Output frequency	95 GHz nominal
p-i-n switch on/off ratio	30 dB min.
Pulse width	10–70 ns (5-ns steps)
PRF	60 kHz max.
Receiver	
SSB noise figure	10 dB max.
RF/IF gain	20 dB min.
IF	3 GHz ± 200 MHz
Antennas	
Gain	47 dB nominal
Polarization	Linear

erating parameters for this multifunction equipment.[100] The 100-MHz crystal oscillator serves as a reference frequency for the phase-locked Gunn oscillator (PLO) whose output is amplified in an injection-locked cw IMPATT oscillator (ILO); the PLO and the ILO form the stable master oscillator (STAMO) for the system.

The output of the IMPATT cw ILO goes to the single-sideband (SSB) upconverter whose other input is an IF frequency of 3 GHz ± 100 MHz, which is locked to a harmonic of the 100-MHz master oscillator. The output of the SSB upconverter passes through a waveguide switch to either an IMPATT cw ILO or an IMPATT pulsed ILO. Another waveguide switch connects the transmitter antenna to the desired ILO.

The receiver has a single-plane monopulse comparator connected to two receiver channels; the sum (Σ) and difference (Δ) comparator outputs feed balanced mixer/IF preamplifiers. The STAMO serves as the local oscillator for the mixers whose IF bandwidth is 3 GHz. Two manual phase shifters provide separate phase adjustments for the two channels.

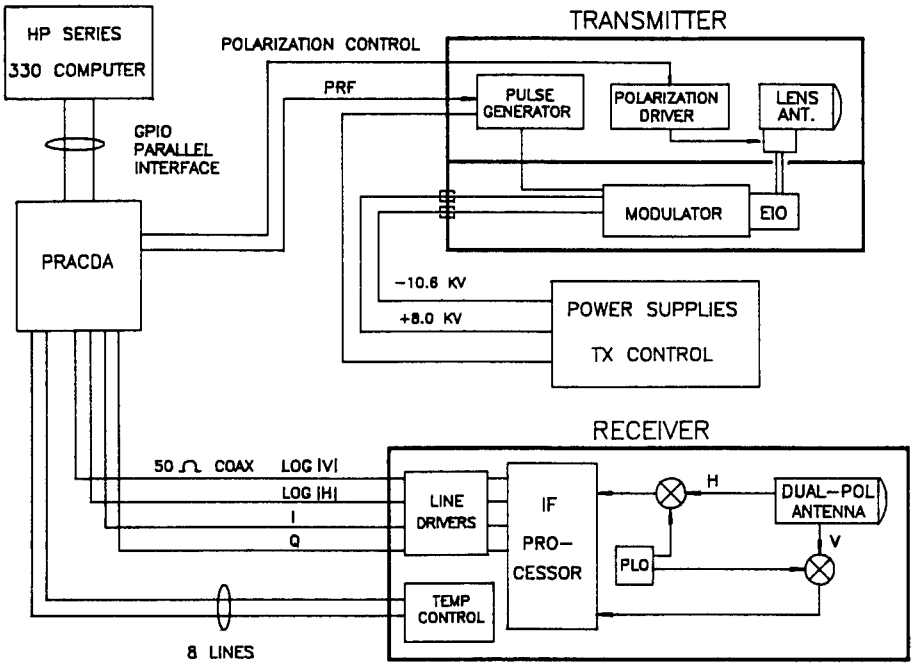


Fig. 3.93 A 225-GHz polarimetric radar block diagram.[151] (© 1990 IEEE, reprinted with permission)

The system serves as a multifunction, coherent front end with either cw or pulsed operation. In the cw mode, the IF input to the SSB upconverter controls the transmitter frequency, which can be single frequency or a swept frequency such as a sawtooth, a ramp, or a sinusoid. In the pulsed mode, the radar operates at a single frequency, with pulse-to-pulse frequency agility, with coherent chirp, or with a combination of both. See Ref. 100 for more details.

Polarimetric radars, which measure the entire scattering matrix, provide more information on targets and backgrounds than radars that measure only a single polarization, and there is currently great interest in the collection of polarimetric millimeter-wave data. References 150 and 151 describe an incoherent 225-GHz radar for measuring the polarimetric properties of targets and backgrounds. Figure 3.93 is a simplified block diagram of the system, which contains a transmitter with multiple polarization capability, a dual-polarized receiver, a *p*olarimetric *ra*dar *c*ontrol and *d*ata *a*cquisition (PRACDA) system, and a data-logging computer; Table 3.18 lists the system parameters.

The transmitter has a multiple-polarization lens antenna, Fig. 3.94, that can produce any linear polarization as well as right- and left-hand circular polarizations. The radiation from the scalar horn passes through two quarter-wave plates and illuminates the 6-in. dielectric lens after reflection from the flat mirror; the nominal beamwidth is 0.61 deg. The rotation of the quarter-wave plates provides the desired polarization. The transmitter uses a model VKY-2429R extended interaction oscillator (EIO) manufactured by Varian

Table 3.18 Parameters for 225-GHz Polarimetric Radar (from Ref. 151; © 1990, IEEE)

Transmitter	
Center frequency	225.63 GHz
Peak output power	60 W
Maximum duty cycle	0.005
Maximum PRF	20 kHz
Pulse width	50–500 ns
Receiver	
SSB noise figure	15 dB
IF center frequency	194.8 MHz
IF bandwidth	Up to 50 MHz
Outputs	$\log\|V\|$, $\log\|H\|$, I_{rel}, Q_{rel}
Dynamic range	70 dB
Antennas	
Receiver	Dual polarization, H and V channels, 0.61-deg beamwidth
Transmitter	Multiple polarization, manually switchable, 0.61-deg beamwidth

Canada Microwave Products that produces a 60-W output pulse, 50 to 600 ns long. The EIO limits the maximum system duty cycle to 0.005.

Figure 3.95 is a simplified block diagram of the receiver. The incoming signal passes through a 6-in. dielectric lens to a linear array of 0.001-in.-diam wires spaced 300 wires per inch; this wire array separates the incoming field into vertical and horizontal components that are received by separate scalar feeds. Second harmonic mixers are connected to the scalar feeds; the local oscillator for the mixers is the 112.717-GHz output of an InP Gunn oscillator,

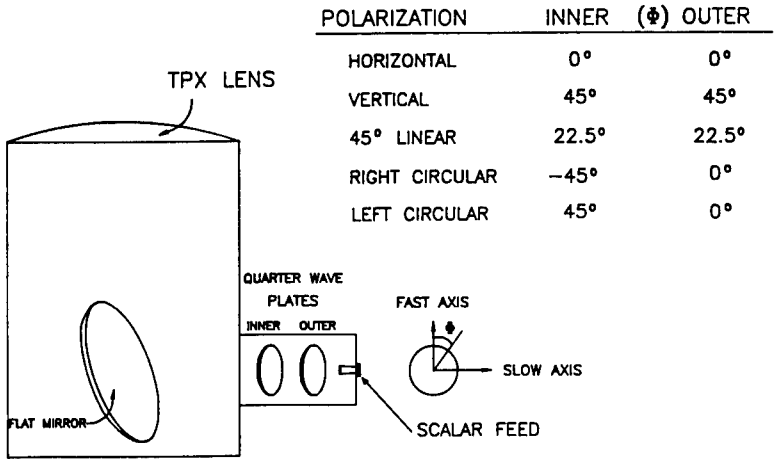


Fig. 3.94 Multiple polarization lens antenna.[151] (© 1990 IEEE, reprinted with permission)

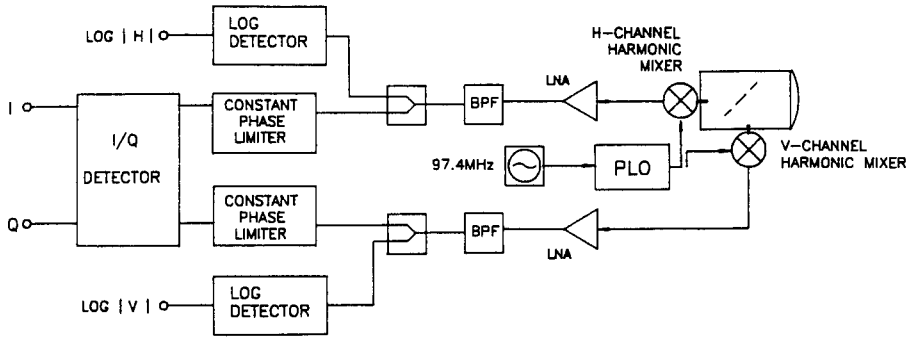


Fig. 3.95 A 225-GHz polarimetric radar receiver.[151] (© 1990 IEEE, reprinted with permission)

which is phase locked to a 97.4-MHz crystal oscillator. The IF passband at the mixer output is 50 to 500 MHz. The IF receiver components measure the magnitude of the returns in the vertical and horizontal channels and the relative phase between the channels, and this information is sufficient to completely characterize the polarization state of the measured field. Reference 151 discusses calibration and the principles of incoherent polarimetry used by the system.

3.5.7 Automotive Radars

Automotive radar systems provide a warning or a warning and braking when there is no longer a safe distance between vehicles in the same lane or when there are obstacles in the lane. References 152 to 156 describe automotive collision avoidance radar systems. Work on automotive radar for collision avoidance began about 1960 and is still continuing, but problems with achieving adequate reliability and developing the required technology have precluded widespread use. A number of organizations that have had automotive radar programs in the past are currently inactive.

As shown in Fig. 3.96, the collision avoidance radar mounts at the front of the vehicle, and a narrow-beam antenna illuminates the lane down which the vehicle is traveling. Studies have shown that the optimum horizontal beamwidth, about 2.5 deg, provides a 4-m spot at a 100-m range; overpasses and other factors limit the vertical beamwidth to about 3.5 deg. Although some automotive radars have operated at 10 or 17 GHz, most systems have a higher center frequency, usually 35 GHz or above, to reduce the size of the antenna and the space required at the front of the vehicle.

Automotive radars use a short-pulse or an FM-cw waveform; the short-pulse systems measure the range and determine relative velocity by range changes. The FM-cw radar transmits a cw signal whose frequency is a sawtooth similar to Fig. 3.87, and the transmitted signal mixes with the return echoes, generating a difference frequency proportional to the FM rate and the range to the target. Because an FM-cw system can determine both the range and relative velocity from this signal, it is less complex than a short-pulse system. Depending on the system, the maximum range is of the order of 50 to 120 m.

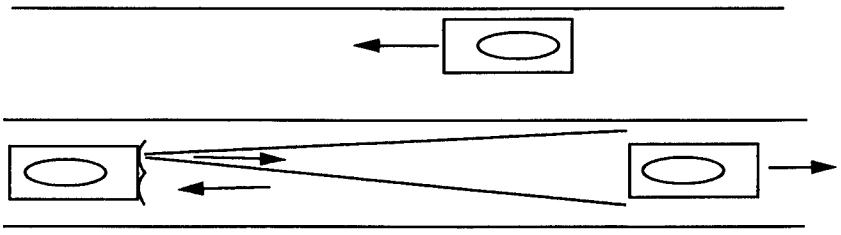


Fig. 3.96 Collision avoidance radar.

There are usually separate transmitting and receiving antennas to provide good isolation between the transmitter and the receiver. The transmitter uses a low-power Gunn oscillator as a source, and its other components are solid state. The radar output passes to a signal processor whose other input is the vehicle velocity. The signal processor determines a safe spacing for the relative velocities of the two vehicles and the road conditions, and it provides a warning or initiates braking when conditions warrant. Obstacles such as road signs, guard rails, bridges, and traffic in adjacent lanes can cause false alarms, particularly on a curve, and the signal processor must discriminate against the false alarms based on the echo behavior. The maximum operating range is usually reduced on turns when the radar beam can illuminate objects outside the lane of the vehicle.

References

1. N. C. Currie and C. E. Brown, Eds., *Principles and Applications of Millimeter-Wave Radar,* Artech House, Norwood, MA (1987).
2. E. K. Reedy and J. C. Wiltse, "Fundamentals of millimeter-wave radar systems," in *Aspects of Modern Radar,* Eli Brookner, Ed., Artech House, Boston (1988).
3. S. L. Johnston, "Millimeter radar," *Microwave Journal* **20**(11), 16–28 (1977).
4. J. C. Wiltse, "History of millimeter and submillimeter waves," *IEEE Transactions on Microwave Theory and Techniques* **32**(9), 1118–1127 (1984).
5. E. K. Reedy and G. W. Ewell, "Millimeter radar," Chap. 2 in *Infrared and Millimeter Waves, Vol. 4: Millimeter Wave Systems,* K. J. Button and J. C. Wiltse, Eds., Academic Press, New York (1981).
6. P. Bhartia and I. J. Bahl, *Millimeter Wave Engineering and Applications,* John Wiley & Sons, New York (1984).
7. S. L. Johnston, Ed., *Millimeter Wave Radar,* Artech House, Dedham, MA (1980).
8. D. K. Barton, *Modern Radar Systems Analysis,* Artech House, Norwood, MA (1988).
9. M. I. Skolnik, *Introduction to Radar Systems,* 2nd ed., McGraw-Hill, New York (1980).
10. D. R. Wehner, *High Resolution Radar,* Artech House, Norwood, MA (1987).
11. N. Levanon, *Radar Principles,* Chap. 8, John Wiley & Sons, New York (1988).
12. A. W. Rihaczek, *Principles of High-Resolution Radar,* McGraw-Hill, New York (1969).
13. L. V. Blake, *Radar Range-Performance Analysis,* Artech House, Norwood, MA (1986).
14. L. V. Blake, "A guide to basic pulse-radar maximum range calculation, Part 1," Naval Research Laboratory Report 6930 (Dec. 1972).
15. J. V. Difranco and W. L. Rubin, *Radar Detection,* Prentice-Hall, Englewood Cliffs, NJ (1968).
16. J. I. Marcum, "A statistical theory of target detection," (Rand Memo RM-754, Dec. 1947), *IRE Transactions on Information Theory* **6**(2), 59–267 (1960).

17. D. K. Barton, "Simple procedures for radar detection calculations," *IEEE Transactions on Aerospace Electronic Systems* **5**(5), 837–846 (1969).
18. P. Swerling, "Probability of detection of fluctuating targets," (Rand Memo RM-1217, March 17, 1954), *IRE Transactions on Information Theory* **6**(2), 269–308 (1960).
19. F. E. Nathanson, *Radar Design Principles,* 2nd ed., McGraw-Hill, New York (1991).
20. H. A. Mayer and D. P. Meyer, *Radar Target Detection,* Academic Press, New York (1973).
21. G. R. Heidbreder and R. L. Mitchell, "Detection probabilities for log-normally distributed signals," *IEEE Transactions on Aerospace Electronic Systems* **3**(1), 5–13 (1967).
22. D. K. Barton and H. R. Ward, *Handbook of Radar Measurements,* Prentice-Hall, Englewood Cliffs, NJ (1969).
23. J. H. Dunn, D. D. Howard, and K. B. Pendleton, "Tracking radar," Chap. 21 in *Radar Handbook,* M. I. Skolnik, Ed., McGraw-Hill (1970).
24. L. J. Cutrona, "Synthetic aperture radar," Chap. 21 in *Radar Handbook,* 2nd ed., M. Skolnik, Ed., McGraw-Hill, New York (1990).
25. J. L. Walker, "Range-Doppler imaging of rotating objects," *IEEE Transactions on Aerospace Electronic Systems* **16**(1), 23–52 (1980).
26. D. A. Ausherman et al., "Developments in radar imaging," *IEEE Transactions on Aerospace Electronic Systems* **20**(4), 363–400 (1984).
27. D. L. Mensa, *High Resolution Radar Imaging,* Artech House, Dedham, MA (1981).
28. W. M. Brown and C. J. Palermo, *Random Processes, Communications and Radar,* Chap. 8, McGraw-Hill, New York (1969).
29. *Synthetic Aperture Radar: Technology and Applications,* University of Michigan College of Engineering Short Course Notes (July 16–20, 1990).
30. R. O. Harger, *Synthetic Aperture Radar,* Academic Press, New York (1970).
31. E. F. Knott, J. F. Shaeffer, and M. T. Tuley, *Radar Cross Section,* Artech House, Norwood, MA (1985).
32. J. W. Crispin, Jr. and K. M. Siegel, *Methods of Radar Cross-Section Analysis,* Academic Press, New York (1968).
33. G. T. Ruck et al., *Radar Cross Section Handbook,* Plenum Press, New York (1970).
34. A. L. Maffett, *Topics for a Statistical Description of Radar Cross Section,* John Wiley & Sons, New York (1989).
35. M. M. Horst and B. Perry, "MMW modeling techniques," Chap. 7 in *Principles and Applications of Millimeter-Wave Radar,* N. C. Currie and C. E. Brown, Eds., Artech House, Norwood, MA (1987).
36. E. L. Johansen, "Top-hat reflectors cap radar calibration," *Microwaves* **20**(13), 65–66 (1981).
37. W. A. Holm, "Polarimetric fundamentals and techniques," Chap. 20 in *Principles of Modern Radar,* J. L. Eaves and E. K. Reedy, Eds., Van Nostrand Reinhold, New York (1987).
38. D. Giuli, "Polarization diversity in radars," *Proceedings of the IEEE* **74**(2), 245–269 (1986).
39. F. T. Ulaby and C. Elachi, Eds., *Radar Polarimetry for Geoscience Applications,* Artech House, Norwood, MA (1990).
40. M. W. Long, *Radar Reflectivity of Land and Sea,* 2nd ed., Artech House, Dedham, MA (1983).
41. A. L. Maffett, "Scattering matrices," Chap. 3 in *Methods of Radar Cross-Section Analysis,* J. W. Crispin and K. M. Siegel, Eds., Academic Press, New York (1968).
42. M. H. Dawson and F. F. Rechlin, Dare Technology, Inc., "100 GHz RCS measurement and diagnostic imaging facility," U.S. Army Contract DAAK02-72-C-0441 Final Report (1973).
43. C. R. Seashore, "Missile guidance," Chap. 3 in *Infrared and Millimeter Waves Vol. 4: Millimeter Wave Systems,* K. J. Button and J. C. Wiltse, Eds., Academic Press, New York (1981).
44. M. E. Beebe et al., Hughes Aircarft Co., "94 GHz sensor tower test program," Naval Weapons Center Contract N60530-75-C-0221 Final Report (1976).
45. J. Preissner, "The influence of the atmosphere on passive radiometric measurements," *Millimeter and Submillimeter Wave Propagation and Circuits,* AGARD Conference Reprint No. 245 (1978).
46. A. J. Bogush, Jr., *Radar and the Atmosphere,* Artech House, Norwood, MA (1989).
47. L. D. Strom, "Applications for millimeter radars," Report No. 108, System Planning Corporation (1973).

48. H. J. Liebe and D. H. Layton, "Millimeter-wave properties of the atmosphere: laboratory studies and propagation modeling," NTIA Report No. 87-224 (1987).

49. B. R. Bean and E. J. Dutton, "Radio Meteorology," National Bureau of Standards (U.S.), Monograph 92 (1966).

50. E. K. Smith, Jr. and S. Weintraub, "The constants in the equation for atmospheric refractive index at radio frequencies," *Proceedings of the IRE* **41**(7), 1035–1037 (1953).

51. L. J. Ippolito, *Radiowave Propagation in Satellite Communications,* Van Nostrand Reinhold, New York (1986).

52. H. J. Liebe, "An updated model for millimeter wave propagation in moist air," *Radio Science* **20**(5), 1069–1089 (1985).

53. M. L. Meeks, *Radar Propagation at Low Altitudes,* Chap. 4, Artech House, Dedham, MA (1982).

54. R. W. McMillan, J. C. Wiltse, and D. E. Snider, "Atmospheric turbulence effects on millimeter wave propagation," *IEEE EASCON-79 Conference Record* **1,** 42–47 (1979).

55. F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive, Volume 1: Microwave Remote Sensing Fundamentals and Radiometry,* Chap. 5, Addison-Wesley, Reading, MA (1981).

56. C. J. Gibbins, "Improved algorithms for the determination of specific attenuation at sea level of dry air and water vapor in the frequency range 1-350 GHz," *Radio Science* **21**(6), 949–954 (1986); also correction, *Radio Science* **24**(2), 261 (1989).

57. E. E. Altshuler, "A simple expression for estimating attenuation by fog at millimeter wavelengths," *IEEE Transactions on Antennas and Propagation* **32**(7), 757–758 (1984); revision in **37**(11), 1477 (1989).

58. H. J. Liebe, T. Manabe, and J. P. Stricklen, "Millimeter-wave attenuation and delay for a fog event," *IEEE Digest: 12th International Conference Infrared and Millimeter Waves,* Orlando, FL (1987).

59. R. N. Trebits, "MMW propagation phenomena," Chap. 4 in *Principles and Applications of Millimeter-Wave Radar,* N. C. Currie and C. E. Brown, Eds., Artech House, Norwood, MA (1987).

60. J. E. Knox, "Millimeter-wave propagation in smoke," *IEEE EASCON-79 Conference Record* **2,** 357–361 (1979).

61. F. C. Petito and E. W. Wentworth, "Measurements of millimeter-wave radar transmission and backscatter during dusty infrared test II (DIRT II)," Report DELNV-TR-0011, AD A086189 (1980).

62. J. D. Lindberg et al., "Measured effects of battlefield dust and smoke on visible, infrared, and millimeter wavelength propagation: a preliminary report on dusty infrared test (DIRT-1)," Atmospheric Sciences Laboratory, White Sands Missile Range, AD B035420L (1979).

63. S. A. A. Abdulla, H. M. Al-Rizzo, and M. M. Cyril, "Particle-size distribution of Iraqi sand and dust storms and their influence on microwave communication systems," *IEEE Transactions on Antennas and Propagation* **36**(1), 114–126 (1988).

64. H. A. Corriher, Jr., "Multipath phenomena and effects," Chap. 4 in *Principles of Modern Radar,* J. L. Eaves and E. K. Reedy, Eds., Van Nostrand Reinhold, New York (1987).

65. S. Shibuya, *A Basic Atlas of Radio-Wave Propagation,* Chap. 2, John Wiley & Sons, New York (1987).

66. N. C. Currie, "Clutter characteristics and effects," Chap. 10 in *Principles of Modern Radar,* J. L. Eaves and E. K. Reedy, Eds., Van Nostrand Reinhold, New York (1987).

67. N. C. Currie, "MMW clutter characteristics," Chap. 5 in *Principles and Applications of Millimeter-Wave Radar,"* N. C. Currie and C. E. Brown, Eds., Artech House, Norwood, MA (1987).

68. N. C. Currie, S. P. Zehner, and F. B. Dyer, "MMW clutter model update," *IEE Radar* **87,** 217–221 (Oct. 1987).

69. R. K. Moore, "Ground echo," Chap. 12 in *Radar Handbook,* 2nd ed., M. Skolnik, Ed., McGraw-Hill, New York (1990).

70. F. T. Ulaby and M. C. Dobson, *Handbook of Radar Scattering Statistics for Terrain,* Artech House, Norwood, MA (1989).

71. R. L. Cosgriff, W. H. Peake, and R. C. Taylor, *Terrain Scattering Properties for Sensor System Design (Terrain Handbook II),* Engineering Experiment Station Bulletin, Vol. 29 (3), the Ohio State University (1960); the values listed here are 4 to 5 dB low; see remarks in *IEEE Transactions on Antennas and Propagation* **24,** 896–899 (1976).

72. T. F. Haddock and F. T. Ulaby, "140-GHz scatterometer system and measurements of terrain," *IEEE Transactions on Geoscience and Remote Sensing* **28**(4), 492–499 (1990).

73. R. M. Narayanan and R. E. McIntosh, "Millimeter-wave backscatter characteristics of multilayered snow surfaces," *IEEE Transactions on Antennas and Propagation* **38**(5), 693–703 (1990).

74. R. M. Narayanan, C. C. Borel, and R. E. McIntosh, "Radar backscatter characteristics of trees at 215 GHz," *IEEE Transactions on Geoscience and Remote Sensing* **26**(3), 217–228 (1988).

75. J. D. Echard, "Target detection in noise and clutter," Chap. 3 in *Principles and Applications of Millimeter-Wave Radar,* N.C. Currie and C. E. Brown, Eds., Artech House, Norwood, MA (1987).

76. J. M. Miller, "Characterization of clutter and its use in maintaining CFAR operation," *Electromagnetics* **4**(2–3), 185–203 (1984).

77. L. B. Wetzel, "Sea clutter," Chap. 13 in *Radar Handbook,* 2nd ed., M. Skolnik, Ed., McGraw-Hill, New York (1990).

78. F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing, Active and Passive,* Vol. 3, Chap. 20, Artech House, Norwood, MA (1986).

79. H. Masuko et al., "Measurement of microwave backscattering signatures of the ocean surface using X-band and K_a-band scatterometers," *Journal of Geophysical Research* **91**(C11), 13,055–13,083 (1986).

80. G. V. Morris, *Airborne Pulsed Radar,* p. 351, Artech House, Norwood, MA (1988).

81. D. G. Bodnar, "MMW antennas," Chap. 11 in *Principles and Applications of Millimeter-Wave Radar,* N. C. Currie and C. E. Brown, Eds., Artech House, Norwood, MA (1987).

82. D. G. Bodnar, "Radar antennas," Chap. 6 in *Principles of Modern Radar,* J. L. Eaves and E. K. Reedy, Eds., Van Nostrand Reinhold, New York (1987).

83. R. C. Johnson, "Introduction to antennas," Chap. 1 in *Antenna Engineering Handbook,* 2nd ed., R. C. Johnson and H. Jasik, Eds., McGraw-Hill, New York (1984).

84. A. W. Love, "Horn antennas," Chap. 15 in *Antenna Engineering Handbook,* 2nd ed., R. C. Johnson and H. Jasik, Eds., McGraw-Hill, New York (1984).

85. TRG Division of Alpha Industries, "Millimeter microwave antenna catalog and handbook," Woburn, MA (n.d.).

86. L. E. Rhoades and J. Eisenberger, "Large millimeter-wave antenna systems," *Microwave Systems News* **18**(12), 19–28 (1988).

87. J. J. James and A. Henderson, "Planar millimeter-wave antenna arrays," Chap. 3 in *Infrared and Millimeter Waves, Vol. 14; Millimeter Components and Techniques, Part V,* K. J. Button, Ed., Academic Press, Orlando (1985).

88. R. E. Munson, "Microstrip antennas," Chap. 7 in *Antenna Engineering Handbook,* R. C. Johnson and H. Jasik, Eds., McGraw-Hill, New York (1984).

89. F. Lalezeri, C. D. Massey, and D. F. Hunter, "Millimeter-wave tactical/small-aperture antennas," *Microwave Systems News* **18**(12), 41–50 (1988).

90. E. B. Joy, "MMW radomes," Chap. 12 in *Principles and Applications of Millimeter-Wave Radar,* N. C. Currie and C. E. Brown, Eds., Artech House, Norwood, MA (1987).

91. G. K. Huddleston and H. L. Basset, "Radomes," Chap. 44 in *Antenna Engineering Handbook,* R. C. Johnson and H. Jasik, Eds., McGraw-Hill, New York (1984).

92. G. W. Ewell, "Radar transmitters," Chap. 5 in *Principles of Modern Radar,* J. L. Eaves and E. K. Reedy, Eds., Van Nostrand Reinhold, New York (1987).

93. V. L. Granastein and I. Alexeff, Eds., *High-Power Microwave Sources,* Chaps. 4 and 5, Artech House, Norwood, MA (1987).

94. J. C. Butterworth and T. V. Wallace, "High-power MMW transmitters," Chap. 9 in *Principles and Applications of Millimeter-Wave Radar,* N. C. Currie and C. E. Brown, Eds., Artech House, Norwood MA (1987).

95. R. Espinosa, "Tubes still vital to microwave systems," *The Microwave System Designer's Handbook. Microwave Systems News* **13**(12), 126–150 (1983).

96. Varian Canada, Inc., "Introduction to millimeter extended interaction klystrons" (1986).

97. J. W. Hansen, "US TWTs from 1 to 100 GHz," *Microwave Journal 1989 State of the Art Reference,* pp. 179–193 (1989).

98. R. W. McMillan, "MMW solid-state sources," Chap. 8 in *Principles and Applications of*

Millimeter-Wave Radar, N. C. Currie and C. E. Brown, Eds., Artech House, Norwood, MA (1987).

99. Y. C. Shih and H. J. Kuno, "Solid-state sources from 1 to 100 GHz," *Microwave Journal 1989 State of the Art Reference,* pp. 145–161 (1989).
100. Hughes Aircraft Company, Microwave Products Division, "Hughes millimeter-wave products catalog" (n.d.).
101. Millitech Corporation, "Millimeter and submillimeter-wave components catalog" (n.d.).
102. F. B. Fank, "InP emerges as near-ideal material for prototype millimeter wave devices," *Microwave Systems News* **12**(2), 56 (1982).
103. J. Browne, "Component family brings coax coverage to millimeter waves," *Microwaves and RF* **27**(10), 123–131 (1988).
104. M. M. Maury, Jr., "Microwave coaxial connector technology: a continuing evolution," *Microwave Journal 1990 State of the Art Reference,* pp. 39–59 (1990).
105. R. E. Collin, *Field Theory of Guided Waves,* Chap. 5, McGraw-Hill, New York (1960).
106. T. Itoh, "Waveguides and resonators," Chap. 30 in *Reference Data for Engineers: Radio, Electronics, Computer, and Communications,* 7th ed., E. C. Jordan, Ed., H. W. Sams, Indianapolis, IN (1985).
107. F. J. Tischer, "Experimental attenuation of rectangular waveguides at millimeter wavelengths," *IEEE Transactions on Microwave Theory and Techniques* **27**(1), 31–37 (1979).
108. T. Itoh, "Transmission lines," Chap. 29 in *Reference Data for Engineers: Radio, Electronics, Computer, and Communications,* 7th ed., E. C. Jordan, Ed., H. W. Sams, Indianapolis, IN (1985).
109. K. C. Gupta, R. Garg, and I. J. Bahl, *Microstrip Lines and Slot Lines,* Artech House, Norwood, MA (1979).
110. A. Cardiasmenos, "Design considerations examined for millimeter-wave defense-electronics systems," *Microwave Systems News: The Microwave System Designer's Handbook,* 4th ed. **16**(7), 280–296 (1986).
111. C. H. Currie and N. C. Currie, "MMW reflectivity measurement techniques," Chap. 17 in *Principles and Applications of Millimeter-Wave Radar,* N. C. Currie and C. E. Brown, Eds., Artech House, Norwood, MA (1987).
112. G. P. Rodrique, "Circulators from 1 to 100 GHz," *Microwave Journal 1989 State of the Art Reference,* pp. 115–132 (1989).
113. W. E. Hord, "Microwave and millimeter-wave ferrite phase shifters," *Microwave Journal 1989 State of the Art Reference,* pp. 81–94 (1989).
114. A. K. Sharma, "Solid-state control devices: state of the art," *Microwave Journal 1989 State of the Art Reference,* pp. 95–112 (1989).
115. P. F. Goldsmith, "Quasioptical techniques offer advantages at millimeter frequencies," *Microwave Systems News* **13**(12), 65–84 (1983).
116. P. F. Goldsmith and E. L. Moore, "Gaussian optics lens antennas," *Microwave Journal* **27**(7), 153–157 (1984).
117. P. F. Goldsmith, "Designing quasioptical systems," *The Microwave System Designer's Handbook, Microwave System News* **17**(8), 182–192 (1987).
118. R. K. Hoffmann, *Handbook of Microwave Integrated Circuits,* Chap. 1, Artech House, Norwood, MA (1987).
119. R. Douville and M. G. Stubbs, "MIC technology for phased arrays," *Microwave Journal* **31**(3), 143–163 (1988).
120. G. L. Lan et al., "A wideband mm-wave monolithic receiver," *Microwave Journal* **33**(2), 93–106 (1990).
121. G. V. Morris, "MMW airborne mapping radar," Chap. 15 in *Principles and Applications of Millimeter-Wave Radar,* N. C. Currie and C. E. Brown, Eds., Artech House, Norwood, MA (1987).
122. N. C. Currie, S. W. Parker, and R. B. Efurd, "MMW system trade-offs," in *Proceedings of the 1988 IEEE National Radar Conference,* Ann Arbor, MI (Apr. 20–21, 1988).
123. H. Urkowitz, "Closed form expressions for noncoherent radar integration gain and collapsing loss," *IEEE Transaction Aerospace Electronic Systems* **9**(5), 781–783 (1973).
124. A. Ivanov, "Radar guidance of missiles," Chap. 19 in *Radar Handbook,* 2nd ed., M. Skolnik, Ed., McGraw-Hill, New York (1990).

125. J. O. Gobien, "Introductory address," in *Proceedings of the Tenth DARPA/Tri-Service Millimeter Wave Symposium,* pp. 5–12, Report DARPA-TIO-86-1 (1986).

126. T. W. Glynn, "Millimeter-wave radar seekers: an historical perspective," *Microwave System News* **18**(12), 63–66 (1988).

127. W. A. Holm, "MMW radar signal processing techniques," Chap. 6 in *Principles and Applications of Millimeter-Wave Radar,* N. C. Currie and C. E. Brown, Eds., Artech House, Norwood, MA (1987).

128. T. Corrado, "Smart munitions—are they affordable?" *Microwave Journal* **18**(12), 32–36 (1988).

129. C. R. Seashore, "mm-wave sensors for missile guidance," *Microwave Journal* **26**(9), 133–144 (1983).

130. S. O. Piper, "MMW seekers," Chap. 14 in *Principles and Applications of Millimeter-Wave Radar,* N. C. Currie and C. E. Brown, Eds., Artech House, Norwood, MA (1987).

131. J. A. Hoschette and C. R. Seashore, "IR and MMW sensor fusion for precision guided munitions," *Proceedings of the SPIE* **931,** 124–130 (1988).

132. J. E. Malpass, B. M. Sundstrom, and D. H. Getts, "Joint USAF/Army dual mode IR/MMW seeker development program," in *Proceedings of the Tenth DARPA/Tri-Service Millimeter Wave Symposium,* pp. 161–170, Report DARPA-TIO-86-1 (1986).

133. J. A. Scheer and P. P. Britt, "Solid state tracking radar," *Microwave Journal* **25**(10), 59–65 (1982).

134. J. C. Henry and P. W. Goetz, "35 GHz images created by the advanced detection technology sensor," in *The 34th Tri-Service Radar Symposium Record,* pp. 157–163, AD C044 065L, (21–23 June 1988).

135. L. M. Novak et al., "Optimal processing of polarimetric synthetic-aperture radar imagery," *The Lincoln Laboratory Journal* **3**(2), 273–289 (1990).

136. M. W. Long, W. K. Rivers, and J. C. Butterworth, "Combat surveillance radar AN/MPS-29 (XE-1)," in *The 6th Tri-Service Radar Symposium Record,* pp. 230–244 (1960) (reprinted in Ref. 7).

137. F. B. Dyer and R. M. Goodman, Jr., "Vehicle mounted millimeter radar," in *The 18th Tri-Service Radar Symposium Record,* pp. 473–495 (1972) (reprinted in Ref. 7).

138. C. R. Barrett, Jr., and D. A. Ryan, "Surveillance and target acquisition radar for tank location and engagement (STARTLE)," in *Proceedings of the Military Microwave Conference,* pp. 1–8, Microwave Exhibitors & Publishers Ltd., Kent, United Kingdom (1980).

139. J. Fawcette, "Army to test STARTLE radar while navy readies gyrotron," *Microwave System News* **9**(7), 23–24 (1987).

140. J. Bruder, "MMW low angle tracking radars," Chap. 16 in *Principles and Applications of Millimeter-Wave Radar,* N. C. Currie and C. E. Brown, Eds., Artech House, Norwood, MA (1987).

141. D. Cross et al., "TRAKX: a dual-frequency tracking radar," *Microwave Journal* **19**(9), 39–41 (1976) (reprinted in Ref. 7).

142. G. E. Layman, "SEATRACKS—a millimeter wave radar fire control system," *IEEE EASCON—78 Conference Record,* pp. 211–216 (1978) (reprinted in Ref. 7).

143. L. A. Hoffman et al., "A 94-GHz radar for space object identification," *IEEE Transactions on Microwave Theory and Techniques* **17**(12), 1145–1149 (1969) (reprinted in Ref. 7).

144. R. B. Dybdal, "Millimeter radar application to SOI," *IEEE EASCON-77 Conference Record,* pp. 16-4A–1 (1977) (reprinted in Ref. 7).

145. G. B. Jones, "Potential applications of millimeter wave radar to ballistic missile defense," *IEEE EASCON-77 Conference Record,* pp. 16-3A–C (1977) (reprinted in Ref. 7).

146. E. L. Johansen and A. L. Maffett, *35-GHz Active Target Signature Measurements,* Report No. AFATL-TR-79-3, ERIM, Box 134001, Ann Arbor, MI (1979).

147. *WAAM Target/Background Signature Measurement Program Final Report, Vol. II: System Description,* ERIM Report No. 140400-24-T, ERIM, Box 134001, Ann Arbor, MI (1981).

148. W. C. Parnell, "Multiband design boosts resolution of imaging radar," *Microwaves and RF* **26**(9), 85–94 (1988).

149. J. A. Scheer et al., "MMW radar cross section range characterizes targets," in *Proceedings of the 1988 IEEE Radar Conference,* pp. 209–213 (1988).

150. R. E. McIntosh and J. B. Mead, "Polarimetric radar scans terrains for 225-GHz images," *Microwaves and RF* **28**(10), 91–102 (1989).

151. J. B. Mead and R. E. McIntosh, "A 225 GHz polarimetric radar," *IEEE Transactions on Microwave Theory and Techniques* **38**(9), 1252–1258 (1990).

152. G. Neininger, "An FM/CW radar with high resolution in range and Doppler; anti-collision radar for vehicles," in *Radar 77, IEE Conference Publication No. 155,* pp. 526–530, London (1977).

153. Y. K. Wu and C. P. Tresselt, "mm radar for highway collision avoidance," *Microwave Journal* **20**(11), 39ff. (1977).

155. E. H. Dull and H. J. Peters, "Collision avoidance system for automobiles," SAE 780263, Detroit (1978).

156. T. Takehana, Y. Isogai, and T. Sakamoto, "Automative radar using mm wave," pp. 123–128, IMechE C178/81 (1981).

CHAPTER 4

Fiber Optic Systems

Norris E. Lewis
Michael B. Miller
Litton Systems, Inc.
Blacksburg, Virginia

CONTENTS

4.1 INTRODUCTION

This chapter treats fiber optic systems, beginning with descriptions of the components. Features of each component necessary for system design and application are detailed. Examples of several different types of systems are given, along with the basic rules that govern fiber optic system configurations.

Table 4.1 lists the symbols, nomenclature, and units used in this chapter.

4.2 FIBER AND CABLE

Figure 4.1 shows a diagram of a typical optical cable. Although many variants of the basic cable have been developed for different applications, the figure identifies the generic features common to most cable designs. An optical fiber that has been coated with a buffer material resides at the center of the cable. Around this fiber is an inner tube of a plastic material. This tube may be tightly or loosely bound to the fiber. Next is a layer of fibrous or stranded material, such as Kevlar, which runs the length of the cable and provides axial strength. Finally, an outer jacket of plastic material covers the cable assembly. The entire cable is contained within a diameter of 1 to 3 mm. Larger cables containing from 4 to 32 or more fibers are common, but all of them include the basic features shown here.

4.2.1 Fiber Basics

Optical fibers can be categorized in two basic ways. First, a fiber is either single mode or multimode, depending on the number of optical modes it can support. These two types of fiber have very different properties and applications. Second,

Table 4.1 Symbols, Nomenclature, and Units

Symbol	Nomenclature	Units
a	Core diameter	m
A_s	Source emitting area	m^2
B	Bandwidth	Hz
C	Speed of light in vacuum	m/s
e	Electron charge, 1.6×10^{-19}	C
E(n)	Excess loss as a function of number of ports	dB
g_m	Transconductance	mhos
G_{opt}	Optimum APD gain	---
k	Boltzmann constant	---
L	Optical loss	dB
λ	Wavelength	m
N.A.	Numerical aperture	---
n(r)	Refractive index profile	---
n_o	Refractive index of center of fiber core	---
P	Power	W
R_f	Feedback resistance	ohms
S	Fiber–fiber end separation	m
T	Temperature	K
T_b	Bit period	S
T_r	Risetime, 10–90%	S
τ	Pulse broadening	S
θ	Ray angle	radian or degree
V	Number of modes	---

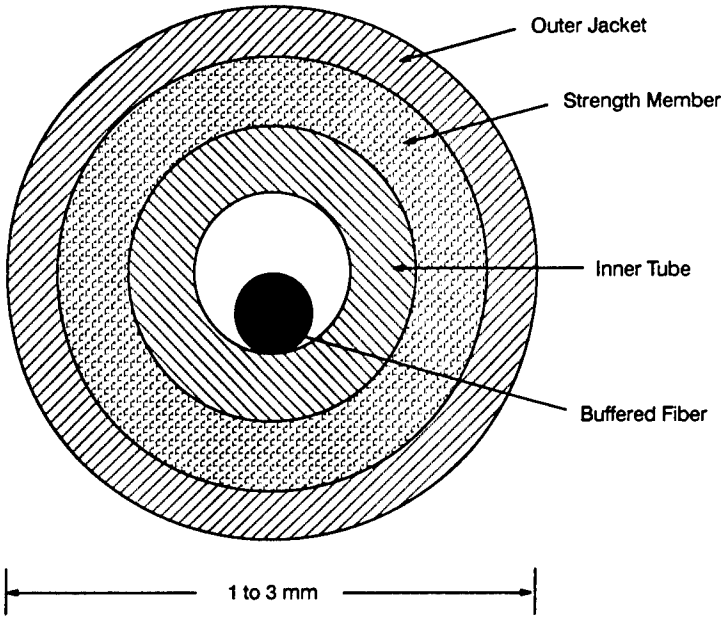


Fig. 4.1 Typical optical cable construction.

the fibers can be further categorized by the shape of their refractive index profile. The two most common profiles are step index and graded index, however, variations exist within these two categories. Index profiles are customized to achieve particular performance characteristics for the fibers.

To completely describe the propagation characteristics of optical fibers, they are treated as cylindrical dielectric waveguides. The application and solution of Maxwell's equations under the appropriate conditions provides the information necessary to describe the electromagnetic field within the fiber. This can be a very involved problem, and does not lie within the scope of this chapter. The reader is referred to any of a number of textbooks for a detailed description of the techniques (Refs. 1 to 3, for example). We use the results of this process and concentrate on the geometric optics models of propagation in the fiber for clarity.

4.2.1.1 Step-Index Multimode Fiber. The refractive index profile of an optical fiber can be described by the following relation (Ref. 4, p. 50):

$$n(r) = n_0\left[1 - \left(\frac{r}{a}\right)^i\right]^{1/2}, \tag{4.1}$$

where r is the radial distance from the center of the core, n_0 is the refractive index of the center of the core, and i is a parameter that defines the shape of the profile. For step-index fiber, $i = \infty$ for $r < a$, and $i = 0$ for $r > a$.

Figure 4.2 shows the structure of a step-index multimode fiber. The refractive index profile shows two different indices for core and cladding, with the

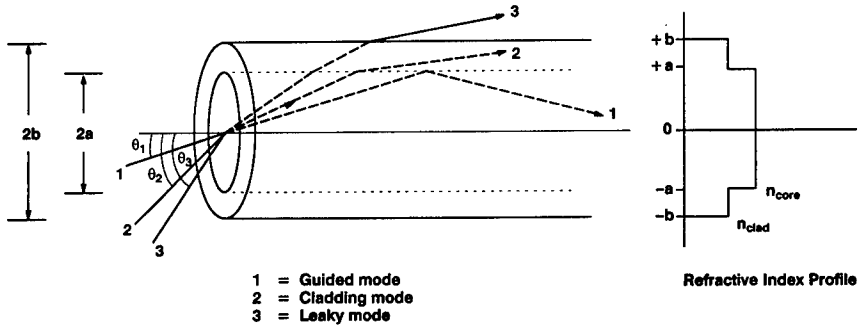


Fig. 4.2 Step-index fiber.

core index being the larger of the two. Several typical light rays are shown in the figure. Ray 1 enters the core at an angle θ_1 with respect to the axis of the fiber. This ray refracts on entering the core glass and propagates until it encounters the boundary between the core and cladding. At this point, the ray is reflected via total internal reflection (TIR) if the original angle $\theta_1 < (n^2_{\text{core}} - n^2_{\text{clad}})^{1/2}$. This limiting angle, viewed from outside the fiber, is called the *numerical aperture* (NA) and it is an important factor in determining the amount of optical power that can be coupled into the fiber from a light source or another fiber. The NA is typically measured with a short fiber (a few meters). When the NA is measured for a long fiber, it is typically found to be smaller. This results from the relatively greater attenuation of the high-order (high-angle) modes or rays in the fiber.[5]

Rays entering the fiber at angles higher than θ are not guided within the fiber core. As shown in Fig. 4.2, a ray, such as ray 2, that enters the fiber core at an angle $\theta_2 > \theta_1$, when it encounters the core-cladding boundary is not reflected. This ray refracts into the cladding and may be trapped there. This ray is called a *cladding mode.* A ray, such as ray 3, that enters the fiber at an even greater angle $\theta_3 > \theta_2$, is not contained within the cladding but leaks out of the fiber altogether.

Dimensions of the step-index fibers vary considerably, depending on the intended applications. Core diameters range from 35 μm to more than 1000 μm, but are most commonly found to be 100 to 200 μm. Fiber outer diameters are typically 1.2 to 2.5 times the core diameter, giving core diameter to outside diameter ratios of 50/125, 100/140, 200/280 μm, etc. Various fiber size standards have been developed by different manufacturers.

The number of modes supported by a step-index fiber (Ref. 6, p. 12) is given by

$$N_m = \frac{V^2}{2} , \tag{4.2}$$

where

$$V = 2\pi\left(\frac{a}{\lambda}\right)(n^2_{\text{core}} - n^2_{\text{clad}})^{1/2} = 2\pi \frac{a}{\lambda}(\text{NA}) ,$$

and V is called the *normalized frequency* of the fiber. Another useful definition is a measure of the difference between the refractive indices of the core and cladding:

$$\Delta = \frac{n_{\text{core}} - n_{\text{clad}}}{n_{\text{core}}} . \tag{4.3}$$

Because the effective path length inside the fiber is not the same for all modes, it is possible for light rays emanating from the source at the same instant to arrive at the far end of the fiber at different times. This causes a broadening of the width of pulses of light used to communicate in the fiber. The overall effect limits the effective bandwidth of the fiber. The time of flight, or transit time, for an axial ray (which is the fastest) in a fiber of length L (Ref. 4, p. 58) is given by

$$t_{\text{axial}} = L\frac{n_{\text{core}}}{c} , \tag{4.4}$$

where c = speed of light in vacuum. Then the difference between the fastest and slowest rays in the fiber is

$$\tau = t_{\text{axial}}\Delta . \tag{4.5}$$

This is the maximum possible pulse broadening and is independent of core diameter and wavelength. Of course, not all modes carry equal amounts of optical power, and so the effective pulse broadening is usually smaller than this.

Another type of dispersion is caused by the difference in refractive index within the fiber for different wavelengths of light. This is a property of the glass, and is called *chromatic* or *material dispersion*. Typically, the material dispersion is 0 at about 1300 nm.

As an example, the maximum pulse broadening in a step-index fiber can be calculated as follows. The transit time of an axial ray is given as

$$t_{\text{axial}} = L\frac{n_{\text{core}}}{c} = (1\text{ m})\left(\frac{1.5}{3 \times 10^8\text{ m/s}}\right) = 5 \times 10^{-9}\text{ s} , \tag{4.6}$$

or 5 μs/km. Then, assuming $\Delta = 0.01$, or 1%, the pulse broadening is

$$\tau = t_{\text{axial}}\Delta = (5\ \mu\text{s/km})(0.01) = 50\text{ ns/km} . \tag{4.7}$$

The resulting maximum bandwidth (Ref. 4, p. 73) is approximately

$$B = \frac{1}{2\tau} = 0.01\text{ ns}^{-1} = 10\text{ MHz} . \tag{4.8}$$

To make it easier for the system designer, fiber manufacturers usually state the bandwidth of their fiber rather than specifying the refractive indices. A typical value for 100/140 fiber with NA = 0.22 is 100 MHz km.

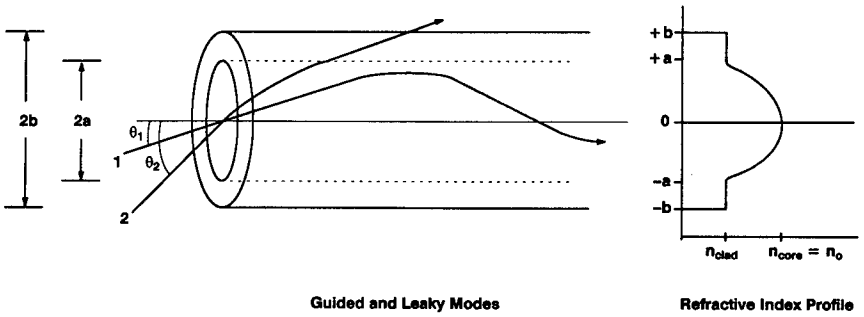


Fig. 4.3 Graded-index fiber.

4.2.1.2 Graded-Index Multimode Fiber. To increase the bandwidth of the fiber, the various modes must be made to arrive at the end of the fiber at the same time. Because the modes traverse different effective optical path lengths, they must be made to travel at different speeds to achieve the same rate down the fiber. This can be accomplished by causing the refractive index of the glass to vary from the center of the core to the edge. Figure 4.3 shows the resulting *graded-index fiber.* The refractive index can again be found from the general expression

$$n(r) = n_0\left[1 - \left(\frac{r}{a}\right)^i\right]^{1/2}, \tag{4.9}$$

where $2.0 \leq i \leq 2.25$, typically. Light rays entering the fiber at or below the critical angle, such as θ_1, are guided in curved paths through the fiber core. High-angle rays travel faster than low-angle rays because of the different value of refractive index at the center and edge of the core. The result is that the various modes arrive at nearly the same time. The NA of the fiber is dependent on where on the core the ray begins. The number of modes (Ref. 4, p. 58) is given by

$$N_m = \Delta\left(\frac{2\pi a n_{\text{core}}}{\lambda}\right)^2\left(\frac{i}{i+2}\right), \tag{4.10}$$

where i is defined above. The pulse broadening is

$$\tau = t_{\text{axial}}\frac{\Delta^2}{2}. \tag{4.11}$$

Using our previous example with $t_{\text{axial}} = 5$ μs/km, the pulse broadening is given by

$$\tau = 5\ \mu\text{s/km}\left(\frac{0.01^2}{2}\right) = 2.5 \times 10^{-10}\ \text{s}, \tag{4.12}$$

and the bandwidth is

$$B = \frac{1}{2\tau} = \frac{1}{5 \times 10^{-10}\ \text{s}} = 2 \times 10^9\ \text{Hz} = 2000\ \text{MHz} \ . \tag{4.13}$$

Core diameters of graded-index fibers are typically between 50 and 100 μm.

4.2.1.3 Single-mode Fiber. The maximum possible bandwidth is realized when only one mode (the *axial ray*) is allowed to propagate in the fiber core. This can happen if the core is made very small and Δ is limited so that the NA is small. Figure 4.4 shows a typical single-mode fiber. The normalized frequency of the fiber (Ref. 6, p. 13) is

$$V = \left(\frac{2^{3/2}\pi a}{\lambda}\right)\Delta^{1/2} \ . \tag{4.14}$$

The fiber operates in the single-mode region for $V \leq 2.405$ ($N_m = 1$). Thus, it is seen that this condition can be met by varying either a or Δ. Because fiber core size and Δ are fixed by the manufacturer, it is instructive to calculate the wavelength at which the fiber becomes single mode. This wavelength, called the *cutoff wavelength,* is given by[7]

$$\lambda_c = \frac{2\pi a n_{\text{core}}\sqrt{2\Delta}}{2.405} \ . \tag{4.15}$$

As an example, let us assume that the core diameter $a = 9$ μm, $n_{\text{core}} = 1.485$, and $\Delta = 0.0007$. We then calculate that the cutoff wavelength is

$$\lambda_c = \frac{2\pi(9\ \mu\text{m})1.485(2 \times 0.0007)^{1/2}}{2.405}$$

$$= 1.306\ \mu\text{m} \ . \tag{4.16}$$

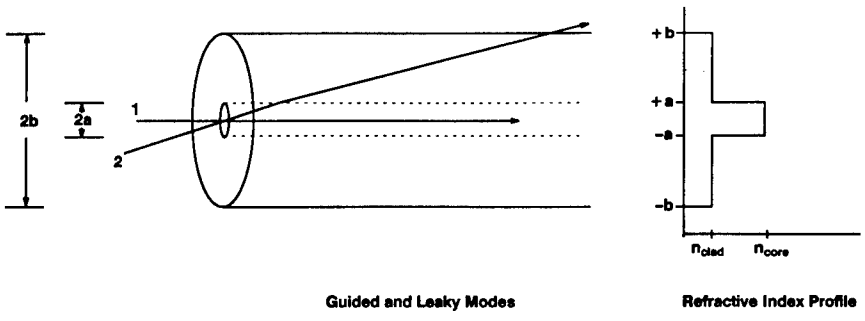


Fig. 4.4 Single-mode fiber.

Thus, this fiber will support only one mode if the wavelength of the light is greater than 1.306 μm. At shorter wavelengths, more modes appear. As wavelengths become much longer than 1.306 μm, the mode is more weakly guided. The energy is contained less and less within the core as the wavelength gets longer than λ_c and is therefore more susceptible to leaking out of the fiber at bends. As shown in Fig. 4.4, light rays entering the fiber at the correct angle are guided, and any other rays are not guided.

Because there is only one mode, no modal dispersion exists, but chromatic and waveguide dispersion do exist. Chromatic dispersion is usually zero at 1300 nm, but the index profile can be modified to shift the zero to 1550 nm, where the minimum attenuation occurs. In this way, it is possible to produce single-mode fiber transmission links at hundreds of megabits per second without repeaters, for distances of tens or hundreds of kilometers.

The shape of the core of single-mode fiber can be varied in many ways to produce various fiber effects, such as control of polarization. Discussion of these fibers can be found in several of the references and is not included here.

4.2.2 Fiber Attenuation

In addition to bandwidth, the other major performance-limiting characteristic of optical fibers is their attenuation. The attenuation is a measure of how much light is lost in a section of fiber per unit length. There are two types of causes of attenuation, intrinsic and extrinsic. Intrinsic causes are a result of various physical or material properties of the fiber. They are (1) material absorption, (2) scattering, (3) waveguide attenuation, and (4) leaky modes. Material absorption is caused by mechanical resonances in the crystalline structure of the waveguide material (SiO_2) and by absorption peaks resulting from impurities in the glass, such as metal ions and water in the form of OH ions. Some of the energy of the light signal is absorbed and changed into mechanical vibrations of the molecules (heat), resulting in less signal being left for communication. Both linear and nonlinear scattering effects may be present in the fiber. Linear scattering effects are minimized by proper waveguide design. Nonlinear effects have threshold power levels of a few tens of milliwatts or more. Waveguide attenuations are caused by bending of the fiber and joints within a length of fiber. Extrinsic losses are absorption increases caused by exposure to radiation.

The fiber attenuation is expressed as a ratio of output to input power per unit length of fiber, and is normally given in logarithmic units:

$$\text{total attenuation} = 10 \log_{10}\left(\frac{P_{\text{out}}}{P_{\text{in}}}\right) \text{ per } L \ . \tag{4.17}$$

The units of attenuation are decibels, and the attenuation of the fiber is normally given as decibels per kilometer.

Three *windows,* or low-loss wavelength ranges, occur in normal optical fibers. These are areas of the spectrum where the attenuation is low and optical sources and detectors are available. These windows nominally occur at 850, 1300, and 1550 nm. Figure 4.5 shows a graph of fiber attenuations versus

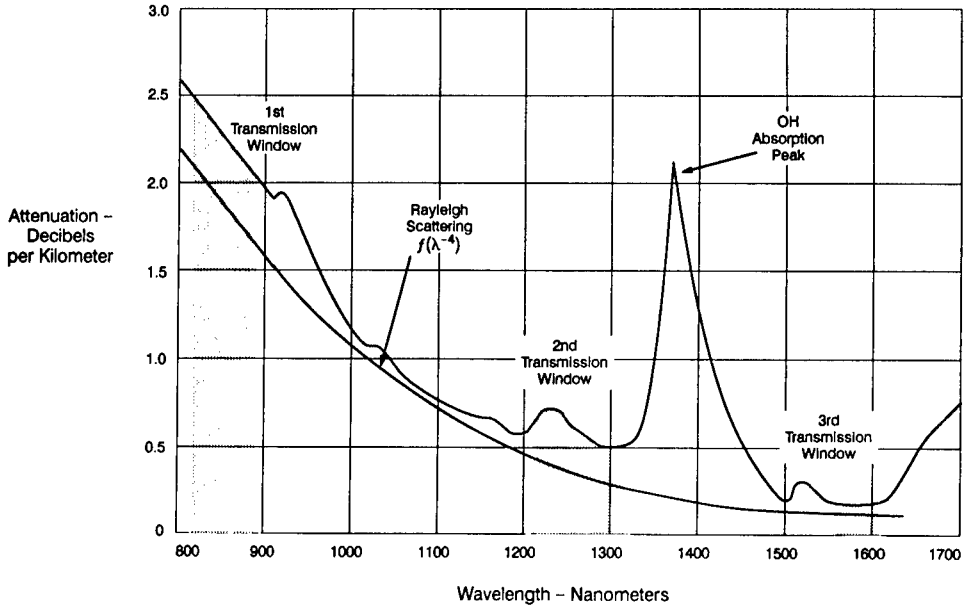


Fig. 4.5 Optical fiber attenuation as a function of wavelength. (After Ref. 4, p. 20)

wavelengths (Ref. 4, p. 20). Table 4.2 compares the various operating parameters for different fiber types.

4.3 OPTICAL SOURCES AND TRANSMITTERS

In the analysis of a fiber optic system, one must consider the transmission rate, transmission distance, the available optical power budget and the fiber dispersion and loss characteristics. Therefore, any topology analysis must concentrate on addressing the power characteristics and related parameters of optical transmitters as well as the achievable sensitivity of optical receivers.

Table 4.2 Fiber Comparisons

Parameter	Multimode		Singlemode
	Step Index	Graded Index	
Core / Clad Dia. (μm)	50/125 85/125 100/140 200/240 400/480	50/125 62.5/125 100/140	(5–10) / (80–125) ~5μm for 850nm ~9μm for 1300nm
Numerical Aperture (NA)	0.2–0.3 ~0.4 (plastic clad)	0.2–0.3	0.1
Bandwidth (MHz x km)	20–100	400–1500	up to 10^3GHz x km
Attenuation (dB/km)	3–10	0.5–3.0	0.15

4.3.1 Optical Sources

There are two types of sources that are used for fiber optic communication systems and fiber optic sensor systems: the light-emitting diode (LED) and the laser diode. These devices are made as semiconducting *p-n* junction diodes that operate in the 750- to 1550-nm region of the optical spectrum. The sources that operate in the 750- to 860-nm region are categorized as short-wavelength devices and those that operate in the region of 1300 to 1550 nm are categorized as long-wavelength devices. The materials that are utilized for LED and laser devices are group III–V compound semiconductors. A III–V compound is formed when more than one group III element is randomly distributed on group III lattice sites or more than one group V element is randomly distributed in group V lattice sites (Ref. 8, p. 285). Short-wavelength devices utilize gallium aluminum arsenide (GaAlAs) and long-wavelength devices are based on alloys of indium gallium arsenide phosphide. These materials are direct band-gap semiconductors. Direct band-gap semiconductors are characterized by radiative lifetimes of excited states that are shorter and more efficient than indirect band-gap semiconductors.

A *p-n* junction is formed when *n*-type and *p*-type regions form a contiguous interface in a single crystal. When energy in the form of a current source is applied to a *p-n* junction, light is emitted from the regions where electron-hole recombination takes place. There are three processes for interaction between a photon and an electron in a solid. These processes are absorption, spontaneous emission, and stimulated emission (Ref. 8, p. 253). The processes that are of interest here are spontaneous emission and stimulated emission. Light-emitting diodes are *p-n* junctions that emit spontaneous radiation in the ultraviolet, visible, or infrared regions of the optical spectrum. Each spontaneous emission event is independent, and for this reason, light sources of this type are affected by mechanisms that can broaden linewidths. When an atom is in an excited state, it can be stimulated to make a radiative transition to a lower energy state (e.g., the ground state) by an incident photon. The phase of the radiated photon and the incident process photon are the same. All photons given off in the stimulated emission process prove to have the same energy and are in phase and, for this reason, produce coherent monochromatic radiation.

4.3.2 Light-Emitting Diodes

The light-emitting diodes that are in general use today fall into two categories: the surface-emitting diode (Burrus LED) and the edge-emitting diode. These LEDs differ in the manner in which the light is radiated from the active region. The surface emitter radiates in an isotropic manner. Efficient coupling into a fiber is achieved by locating the fiber in close proximity to the radiating surface (see Fig. 4.6).

The edge-emitting diode constrains the light to a waveguided region. The recombination region is determined by the current flow through the active area, which in turn is determined by the electrical contact geometry. The refractive index of surrounding regions is greater than the recombination region and, therefore, guides the light parallel to the junction toward the fiber to be coupled. Figure 4.7 illustrates the double heterostructure of the edge emitter.

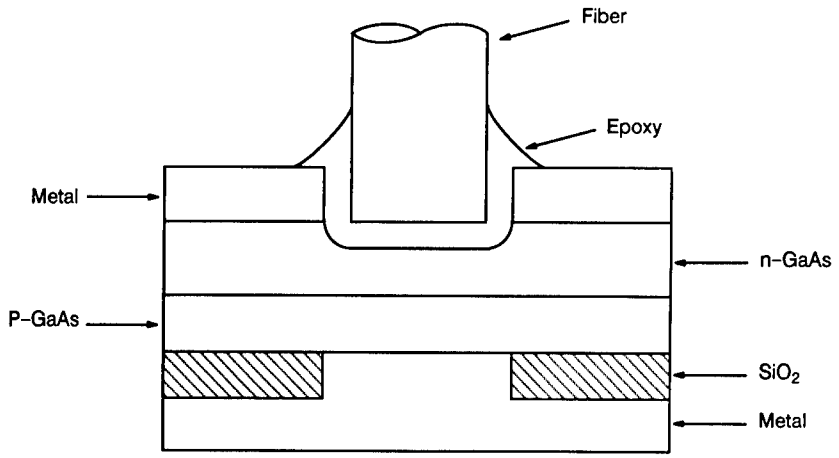


Fig. 4.6 Surface-emitting LED.

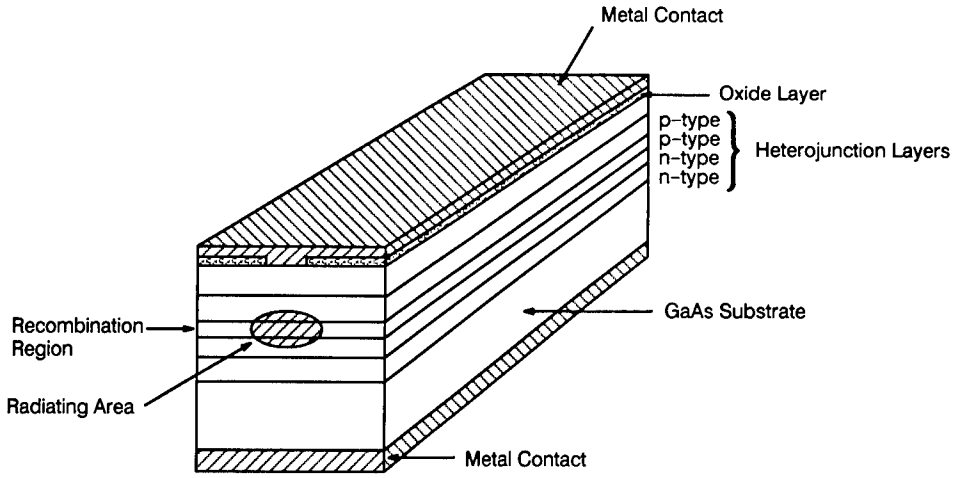


Fig. 4.7 Typical construction of an edge-emitting LED.

4.3.3 Semiconducting Lasers

One criterion that must be established before any material can lase is that a population inversion of the emitting species in the upper versus the lower laser energy states must exist. When the Fermi levels on the p and n sides of a p-n junction are adjusted appropriately by applying a forward bias, a population inversion of holes in the valence band and electrons in the conduction band occurs (Ref. 8, pp. 270–271). If the light is confined to a waveguiding region, the level for stimulated emission may be exceeded. Containment is achieved by adjusting the refractive index of the surrounding materials to a lower value than that used for the laser. Most semiconducting laser materials have a refractive index of about 3 (Ref. 9, p. 49).

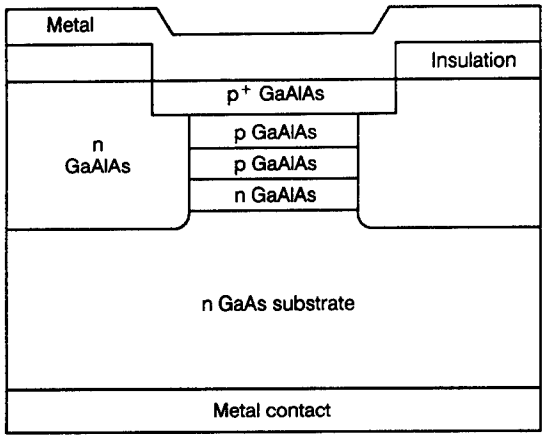


Fig. 4.8 Buried-heterostructure laser.

A diagram of a heterostructure is shown in Fig. 4.8. The heterostructure is used to guide the light through the active region. The oxide layer and the electrical contact geometry determine the current-conducting area of the lasing material. Although the energy level structure of the laser material determines the appropriate wavelength of oscillation, the distance L between the cleaved faces of the laser determines the precise wavelength(s) at which it will emit light.

Because an integral number of half-wavelengths will satisfy the constructive interference criterion, the semiconducting laser can operate over a 30- to 50-Å range (Ref. 8, p. 277).

$$L \times n = m\left(\frac{\lambda}{2}\right) , \tag{4.18}$$

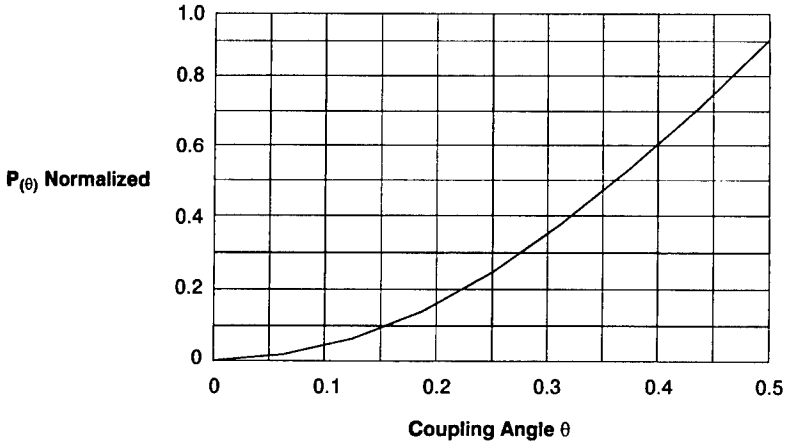
where

n = refractive index of lasing medium
L = length of cavity
λ = wavelength
m = integer.

When it is desirable to limit the laser output to a single mode, it is necessary to change the cavity. One approach is to construct a Bragg reflector internal to the laser structure. This periodicity of this structure determines the mode that is output. Lasers that utilize this type of structure are classified as *distributed-feedback* (DFB) *lasers.*

4.3.4 Source-Fiber Coupling

The parameters controlling source-fiber coupling are the source surface area, the emitting angle, the field of view, and the numerical apertures of the emitter



$$P_{(\theta)} = 2\pi A_s \int_0^{\theta} B_o \cos\theta \sin\theta \, d\theta$$

$A_s = 1, B_o = 1$

θ = 0 to 0.45 (radians)
or = 0 to .30 (NA)

$P_{(\theta)}$ are the values of the integral for each θ

Fig. 4.9 Source-fiber coupling without lenses.

and the receiver (Ref. 10, pp. 97–99). The power received by a multimode fiber through direct coupling can be calculated by using an equation of the form

$$P = 2\pi A_s \int B_0 \cos\theta \sin\theta \, d\theta \ , \tag{4.19}$$

where

B_0 = radiance of the source
θ = angle measured from the normal of the element ds
A_s = area of source.

To obtain values from this expression, θ must be integrated from 0 to θ_{NA}, where θ_{NA} is the numerical aperture of the fiber (see Fig. 4.9 for a graph of this expression versus numerical aperture).

A variety of other techniques exist for coupling a source to an optical fiber (Ref. 9, p. 54). Spherical lenses between the source and the fiber can result in some improved matching of source and fiber characteristics. A hemispherical end can be formed directly on the fiber after a tapering operation has been completed. A third lens technique uses a graded-index lens. In all of these cases, a limited amount of matching between source and fiber characteristics is achieved.

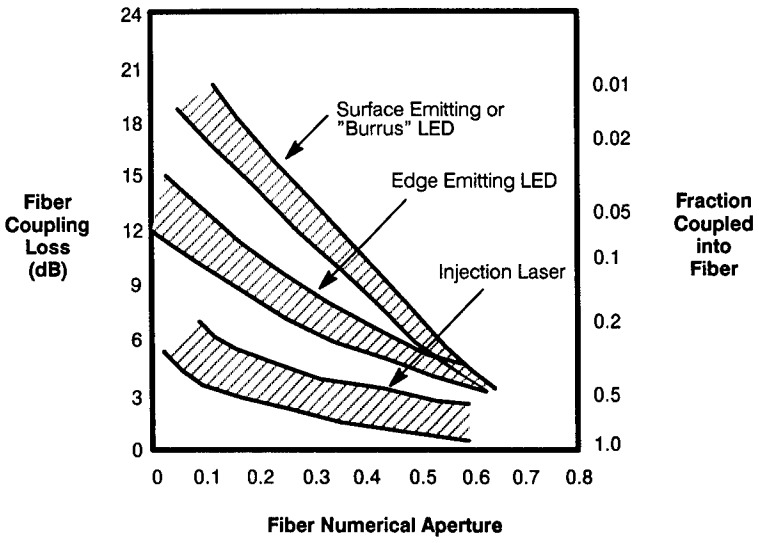


Fig. 4.10 Coupling efficiency into step-index fibers of varying numerical aperture.

4.3.5 Source Trade-Offs

The transmitters used in networks utilize either an LED, for high reliability, low drive circuit complexity, high optical extinction ratio, and low cost, or a laser. Lasers are typically used where high speed and high optical power requirements are present. In loss-limited systems, the laser offers an important advantage—more coupled optical power[11] (see Fig. 4.10). One alternative to the laser is the superluminescent LED. It can inject about 6 dB more power into an optical fiber than a conventional LED. In the case of some fiber optic sensors, a coherent source is required. Also in the case of certain distributed fiber optic sensors systems, a laser operated in a pulsed mode can couple hundreds of milliwatts of power into the system. If a coherent source is not required for the sensor, an LED is generally used because of simplicity and size. The use of lasers involves greater circuit complexity and cost.

4.3.5.1 Power Consumption. The cooling circuitry on any type of laser system has its greatest impact on power supply demand. Six watts of power can be consumed at 50°C to maintain a constant laser threshold point. Because laser wavelength is temperature dependent, temperature stability is of concern in networks involving wavelength-division multiplexing.

The output wavelength of a laser varies because the index of refraction of the waveguide material, the bandgap of the material, and the cavity length are all temperature dependent. Sudden changes in wavelength occur because of mode hopping associated with cavity length changes. A change of 40 Å in output wavelength can occur for a short-wavelength laser over a 20°C temperature range (Ref. 9, p. 57).

There is a wavelength shift with temperature for LEDs because of the temperature dependence of the bandgap of the LED material. Figures 4.11 and 4.12 illustrate typical shifts in wavelength and changes in output power,

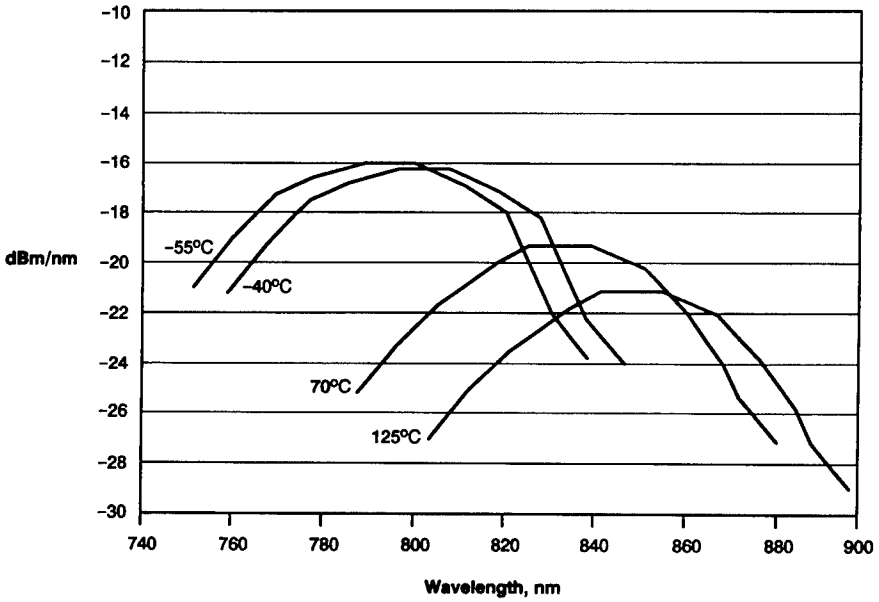


Fig. 4.11 LED output wavelength versus temperature for 830-nm LED.

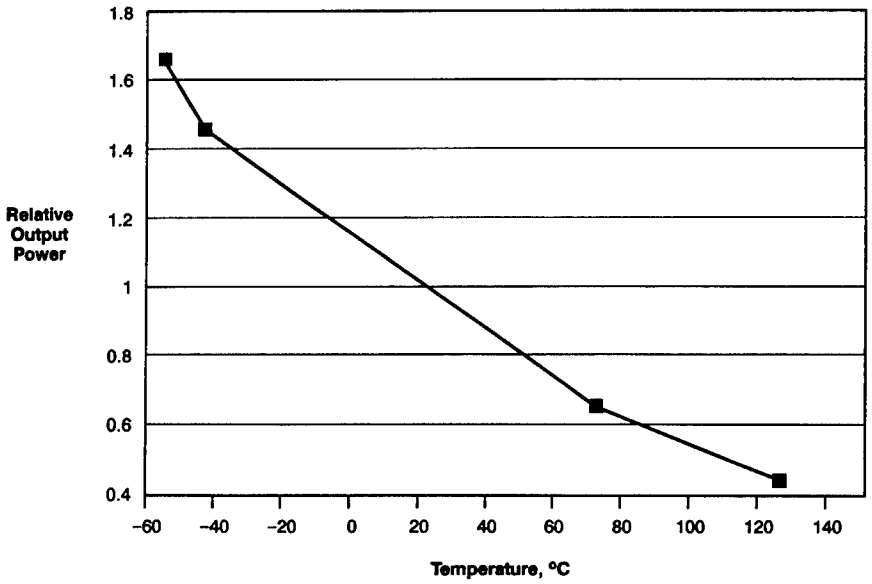


Fig. 4.12 LED output power versus temperature for 830-nm LED.

respectively. It is important to be able to manage these shifts when utilizing broadband sources that have been constructed from multiple LEDs for wavelength-division-multiplexed-based sensors. When wavelength slicing is done, the shift of the intensity versus wavelength profile is important to ensure that adequate power remains in all wavelengths required by the sensor.

4.3.5.2 Extinction Ratio. One of the important operational differences between LED sources and laser sources is their on-to-off ratio, or *extinction ratio.* Because the current through an LED can be turned completely off, there is no optical power incident on the detector for a logical 0 state. Therefore, the optical difference between logical 0 and logical 1 at the detector is maximized, resulting in the highest possible signal level. A laser, on the other hand, is generally current biased just below the lasing threshold so that the signal drive current serves to pull the laser from the LED operating region (off state) to the lasing region (on state). Thus, even in the off state the laser emits optical power into the system. This means that there is always optical power incident on the detector, even in the logical 0 state. This optical power effectively reduces the maximum signal swing at the receiver, thereby degrading system performance.

Another operational difference between the LED and laser is the laser control circuitry. Lasers are generally equipped with photodiodes to allow monitoring of the laser output from the back facet. The signal from the photodiode is used in a feedback circuit to maintain laser output power at some average level, usually 0 to -3 dBm. When no data are present, the stabilization circuit continues to drive the laser so that the average output level is maintained. Thus, even with no data, the receiver sees an average signal level equivalent to the signal level when data are present. In data systems where the data are continuous, this situation does not cause a problem. However, in many data bus systems, the data occur only in bursts. Therefore, the receivers must be able to sense when the bus is quiet and also must be able to sense data collisions. In such cases, the use of laser transmitters is often impractical or impossible.

4.3.5.3 Reliability. Although laser reliability is constantly being improved on, it is still not competitive with the LED. Mean times between failures (MTBF) are greater than 10^6 h at 70°C for LEDs, whereas at 22°C, the laser MTBF lies in the vicinity of 100,000 h. Figure 4.13 gives time to failure for AlGaAs DH lasers tested under a variety of conditions. These failures are caused by several factors, some of which are defects in the active-region geometry, nonradiative recombination in the active region, and facet damage. Circuitry can help to extend laser life through cooling and bias current control consisting of feedback loops that monitor laser heat sink temperature and the back facet optical output. LEDs do not require those elements. Tables 4.3 and 4.4 summarize the trade-offs for different types of light sources used in fiber optic systems.

4.4 OPTICAL DETECTORS AND RECEIVERS

4.4.1 Detectors

The situation with detectors is somewhat analogous to that of light sources (i.e., there are two primary choices):

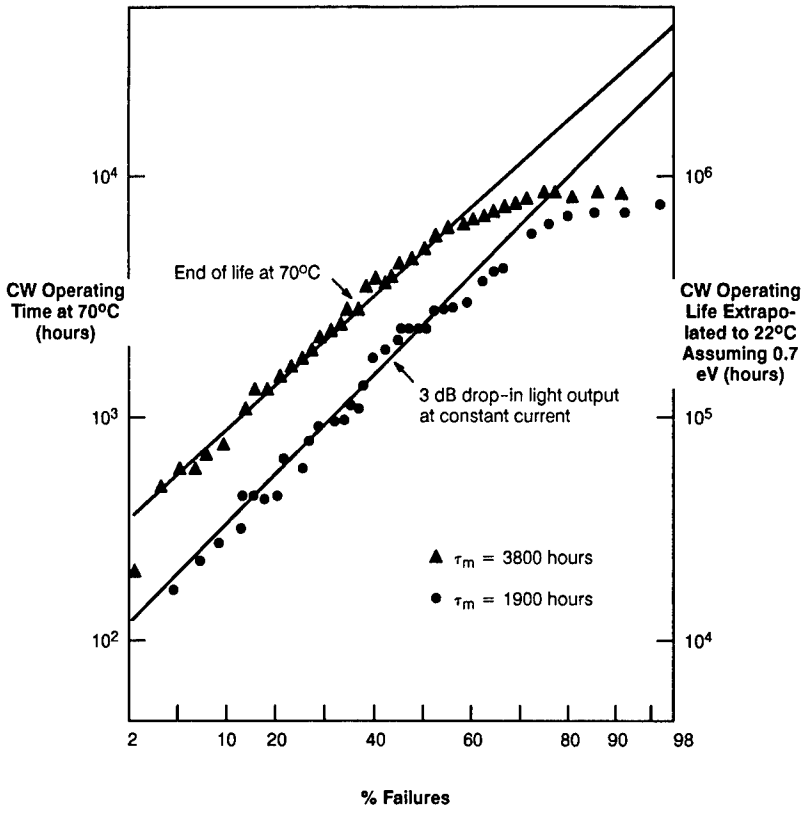


Fig. 4.13 Time to failure at 70°C heat sink temperature on log-normal coordinates for 40 low-threshold (50-mA) oxide-defined stripe AlGaAs DH lasers. Two failure points are assumed, the first when the laser drops to half its initial output at constant current (3-dB life) and the second where the laser can no longer emit 1.25 mW at the 70°C heat sink temperature (end of life). The τ_m is the time needed for 50% of the lasers to fail. The estimated lifetime at 22°F is shown assuming 0.7-eV activation energy (Ref. 10, p. 247).

Table 4.3 Source Summary Short Wavelength

Device	*Launch Power (100/140 0.29 NA)*	*Risetime*	*Temp. Range*	*MTBF (25 C)*
Laser Diode	>1 mW (0 dBm)	0.5 ns	-10 to +60	10^6 hrs.
LED (Burrus)	.160 mW (-8 dBm)	2.5 ns	-55 to +125	10^6 hrs.
LED (Edge-emitter)	.445 mW (-3.5 dBm)	2.5 ns	-55 to +125	10^7 hrs.
LED (Superluminescent)	.750 mW (-1.2 dBm)	3.0 ns	-55 to +125	10^6 hrs.

Table 4.4 Light Sources for Fiber Optic Sensors

	Typical Source Width	*Power*	*Average Power / nm*	*Coherence Length*
Laser: Single Mode	0.1 nm	1 mw	10 mw/nm	20 m
Laser: Multimode	1.5 nm	1.5 mw	1 mw/nm	10 mm
Superluminescent LED	15–30nm	4 mw	0.3 mw/nm	50 um
LED	80 nm	100 uw	1 uw/nm	-
LED Bank	150 nm	40 uw	0.3 uw/nm	-
Broadband LED	120 nm	40 uw	0.3 uw/nm	-
Broadband Light	1000+ nm	200 uw	0.2 uw/nm	-

1. A *p-i-n* photodiode
2. An avalanche photodiode (APD).

The *p-i-n* photodiode is in more widespread use than the APD. Incident photons create electron-hole pairs in the intrinsic region. The electric field established by a reverse bias moves the electrons and holes to the contact regions. Figure 4.14 illustrates the cross section of a *p-i-n* photodiode. The thickness of the intrinsic region can be adjusted to optimize frequency response and quantum efficiency at a given wavelength. As the thickness of the intrinsic layer is increased, the quantum efficiency is improved at the expense of frequency response. Typical thickness for the intrinsic layer varies from 10 to 20 μm. Quantum efficiencies of *p-i-n* diodes are in the range of 0.5 to 0.9. Bias voltages can be as low as 5 V, but may vary to 50 V, depending on the application.

Avalanche photodiodes have a similar structure to *p-i-n* photodiodes. A higher reverse bias potential is used to achieve gain. If the reverse bias voltage

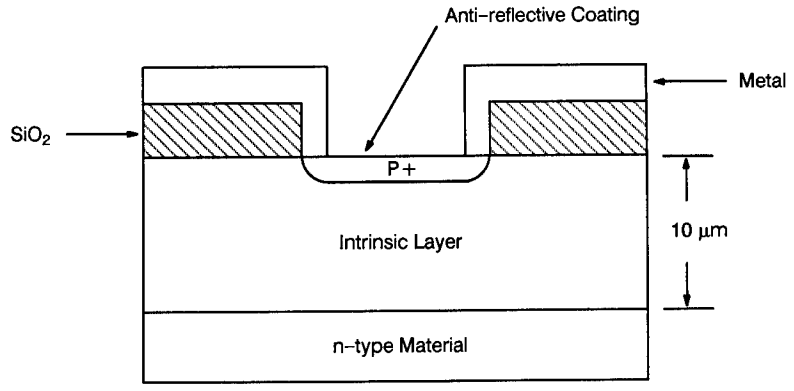


Fig. 4.14 Cross-sectional view of a *p-i-n* diode.

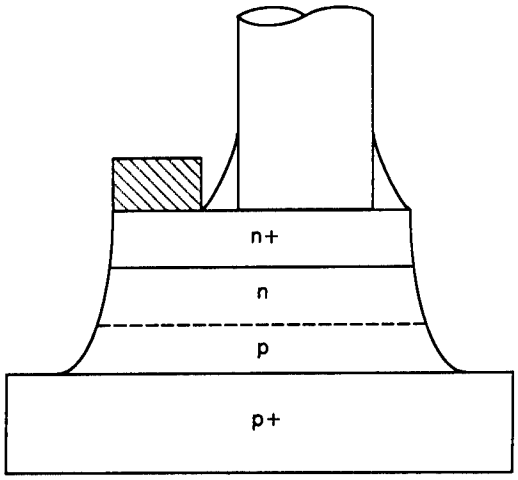


Fig. 4.15 Pigtailed avalanche photodiode. Cross section is for a mesa structure.

is such that an electric field of $>10^5$ V/cm is achieved, electrons and holes can acquire enough kinetic energy to produce electron-hole pairs through inelastic collisions (Ref. 10, p. 247). Bias levels are selected such that a gain of 10 to 100 can be achieved. The ionization coefficients, which determine the gain characteristics, are temperature dependent.[13] It is for this reason the APDs must be implemented with a temperature compensation circuit. A typical APD structure is illustrated in Fig. 4.15.

4.4.2 Receiver Characteristics

The receiver utilized in fiber optic applications entails much more design complexity than does a typical transmitter design. Of most importance to the optical receiver are the photodetector and the low-noise preamplifier. Together, these two elements determine many of the receiver characteristics and ultimate system performance.

The receiver can be expected to be sensitive to low optical power levels as well as being able to handle very high optical power levels. The difference between these levels is the dynamic range. The saturation level of the preamplifier sets the upper limit, whereas the noise level at the input of the receiver sets the lower limit of the dynamic range. The task is to design a receiver of the lowest noise possible for greatest sensitivity but with a high saturation point to achieve a wide dynamic range. A wide dynamic range is important because it allows flexibility and convenience in system configuration. The ability to receive a wide range of optical power levels means that the receiver can be used for both short and long links. A wide dynamic range requirement is also critical in local area network applications, where the transmitted optical power may have to go through a different number of access couplers and splitters before reaching each terminal.

In point-to-point data links and local area networks, the ability of the optical receiver to accommodate a wide range of bit rates and unrestricted data pat-

terns is often desirable. Bit-rate transparency refers to the ability of the optical receiver to operate over a range of bit rates. Bit pattern independency refers to the ability of the optical receiver to operate with an unrestricted coding format (e.g., no restriction on the number of bits per transition, dc average level, etc.).

The acquisition time is the time required for the receiver to stabilize after receiving a burst of input data from its idle state (e.g., the state of no input optical power). In both point-to-point data links and local area network applications, the receiver should respond immediately to the input data burst, otherwise a long preamble sequence has to be used and the data transmission efficiency is reduced.

The receiver sensitivity is a measure of the minimum optical power level required for the receiver to operate with the required bit error rate. It is often given as the average optical power P required for a bit error rate of 10^{-9}. Average power is based on the assumption that the transmitter is producing a bit pattern with a 50% duty cycle. In some cases, receiver sensitivities are expressed as a peak number, which means the transmitter is in a constant on condition. Because there is 50% less power in an average signal, average power numbers are 3 dB lower than peak numbers. When comparing receiver performance, it is important to be aware of this difference in the way sensitivities are quoted.

4.4.2.1 Noise Analysis. The receiver requirements of high sensitivity and wide dynamic range can be met by utilizing a field effect transistor (FET) with a transimpedance amplifier when a *p-i-n* detector is employed, and either an FET or a bipolar transistor and a transimpedance amplifier when an ADP is employed. The APD receiver is the most optically sensitive choice. The selection of a detector involves certain trade-offs and considerations that necessitate a system constraint analysis and a receiver analysis. The receiver analysis centers mostly on calculating noise currents, because that is what sets one extreme of the dynamic range. Much of this analysis relies on the work done by Personick[14] who pioneered fiber optic receiver design. The noise analysis presented in a more recent paper[15] is described here.

For an FET front-end amplifier, the input equivalent noise current power is given by

$$\langle i_{\text{na}}^2 \rangle = \frac{4kT}{R_f} I_2 B + 2eI_L I_2 B + \frac{4kT\Gamma}{g_m}(2\pi C_T)^2 f_c I_f B^2 + \frac{4kT\Gamma}{g_m} \tag{4.20}$$

where

B = operating bit rate
R_f = feedback resistance
I_L = total leakage current (FET gate current and unmultiplied dark-current component of the photodiode)
g_m = FET transconductance

C_T = total input capacitance (including photodiode and stray capacitance)
f_c = the $1/f$ noise corner frequency of the FET
Γ = numerical constant
k = Boltzmann constant
T = absolute temperature.

The weighting functions I_2, I_3, and I_f are dependent only on the input optical pulse shape to the receiver and the equalized output pulse shape [for NRZ (Sec. 4.4.3) coding format and equalized output pulse with a full raised cosine spectrum, we have $I_2 = 0.562$, $I_3 = 0.0868$, and $I_f = 0.184$, whereas for RZ (Sec. 4.4.3) code with 50% duty cycle, these parameter values are $I_2 = 0.403$, $I_3 = 0.0361$, and $I_f = 0.0984$].

The parameter Γ is a noise factor associated with channel thermal noise and gate-induced noise in the FET. It is normally taken to be 0.7 for Si FETs and 1.1 for GaAs MESFETs. However, it has been recently argued that when induced gate noise and its correlation with channel thermal noise are taken into account, Γ for short-channel Si MESFETs and GaAs MESFETs is higher, typically, 1.03 and 1.75, respectively.

The total capacitance C_T is given by

$$C_T = C_{ds} + C_{gs} + C_{gd} + C_f \ , \tag{4.21}$$

where

C_{ds} = photodetector and stray capacitance at the input
C_{gs} = gate-to-source capacitance of the FET
C_{gd} = gate-to-drain capacitance of the FET
C_f = stray capacitance of the feedback resistor.

The first noise term results from the feedback or load resistance, which is negligible if the high-impedance design is used. The second and third noise terms are contributions from the leakage current and $1/f$ noise. Finally, the fourth noise term results from the channel thermal noise and induced gate noise.

For the case of a p-i-n photodiode, the average optical power required at the receiver input for a desired error rate can be written

$$nP = Q\frac{h\gamma}{e}(\langle i_{na}^2 \rangle)^{1/2} \ , \tag{4.22}$$

where

$$n = \text{quantum efficiency} \ , \tag{4.23}$$
$$h\gamma = \frac{hc}{\lambda} \ \text{photon energy} \ ,$$

and Q is related to the desired bit error rate by the expression

$$P(E) = \frac{1}{\sqrt{2\pi}} \int_Q^\infty \exp(-x^2/2)\, dx \ . \tag{4.24}$$

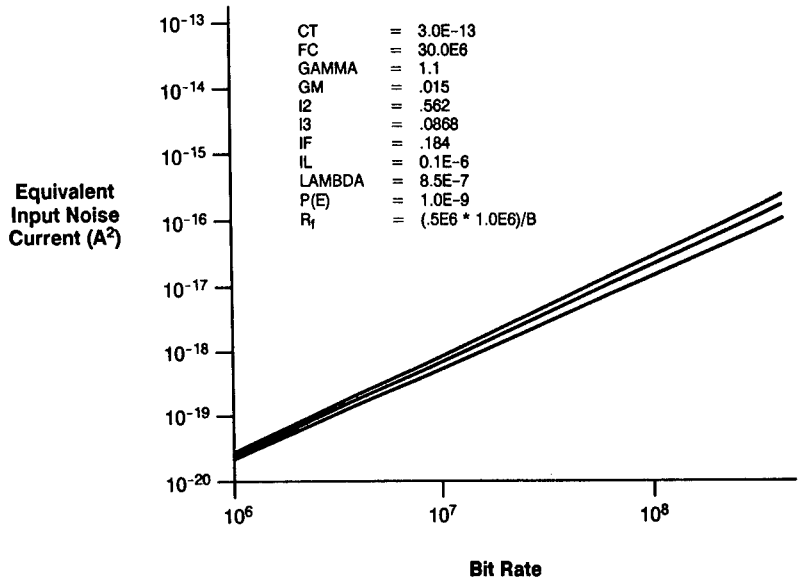


Fig. 4.16 Noise equivalent power versus bit rate.

One example of calculated equivalent noise power is shown in Fig. 4.16 as a function of bit rate. Figures 4.17 through 4.20 are presented to demonstrate the effect of the varying component values and preamp parameters such as g_m, R_f, C_T, and I_L. In all cases, the equations have been evaluated at three temperatures: 215 K (−58°C), 310 K (37°C), and 400 K (127°C). These results show that in the 1- to 100-Mbits/s region, the R_f and I_L dependences dominate.

4.4.2.2 Dynamic Range. The preamp saturation level, or the APD saturation level, determines the upper limit of the dynamic range. As an example, consider a preamp saturation level of 1.8 V. The ultimate dynamic range can be calculated if actual data for premium transistors and detectors are used. Using a *p-i-n* receiver sensitivity of 51 dBm (8×10^{-9} W), the dynamic range (DR) for a 2-MHz receiver, assuming a *p-i-n* responsivity of 0.5 A/W and peak optical power is used, is written

$$\mathrm{DR} = 10 \log \frac{V_{\mathrm{sat}}/R_f}{P_{\min} \times 0.5\ \mathrm{A/W}}\ , \tag{4.25}$$

where V_{sat} is the preamp saturation voltage and R_f is the feedback resistance. Using a value for R_f of $1 \times 10^6\ \Omega$, the dynamic range is 25 dB average, or 22 dB peak. The APD receiver offers about 10 dB more sensitivity but about 8 to 10 dB less dynamic range. This is the result of the gain mechanism in the device. Again, this assumes that burst data are being received and, therefore, that the bias voltage that controls the APD gain cannot be lowered quickly enough to preserve the data pulse shape or will induce false data into the preamp.

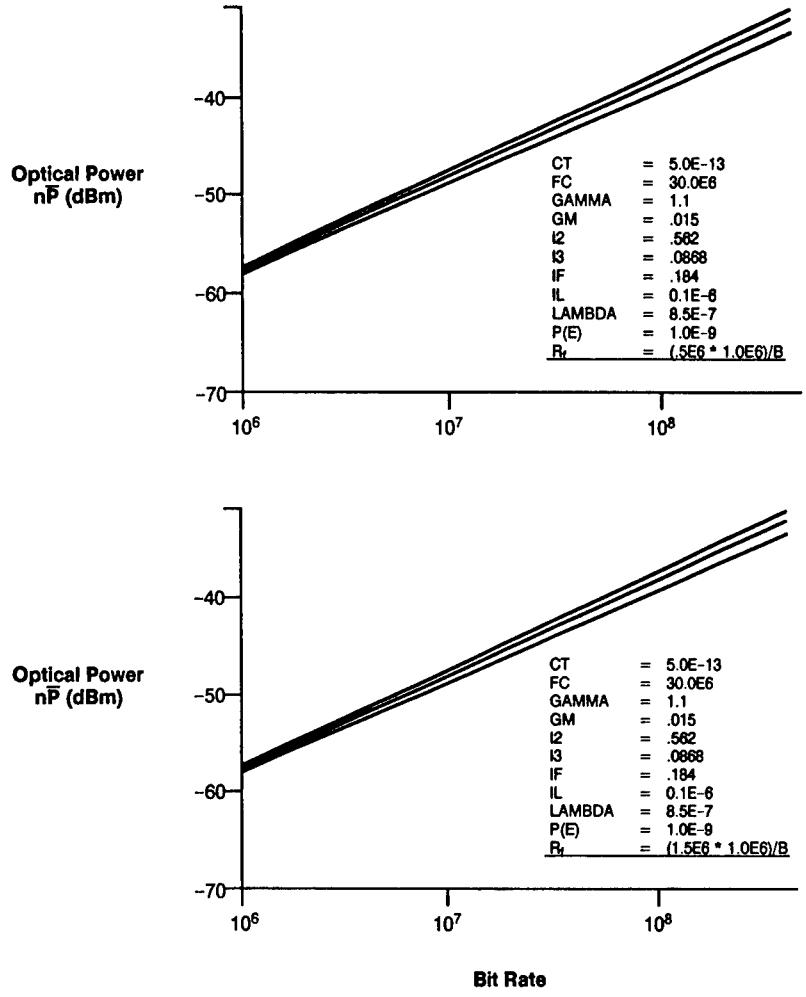


Fig. 4.17 Effect of feedback resistor value R_f on receiver sensitivity.

4.4.2.3 APD Gain. If an APD is utilized, an optimum gain exists that will maximize the SNR:

$$G_{\text{opt}} \sim \left(\left\langle \frac{i_{\text{na}}^2}{ekQBI_1} \right\rangle^{1/2} \right)^{1/2}, \tag{4.26}$$

where k is the excess noise factor of the APD and differs from device to device, Q is the required factor for a 10^{-9} bit error rate (which is 12 if peak power is considered), and B is the bandwidth. Typical optimum gains lie between 50 and 150.

4.4.2.4 Post-Amplifier Circuitry. It has been shown in many texts that the remainder of the fiber optic receiver, although important, contributes negli-

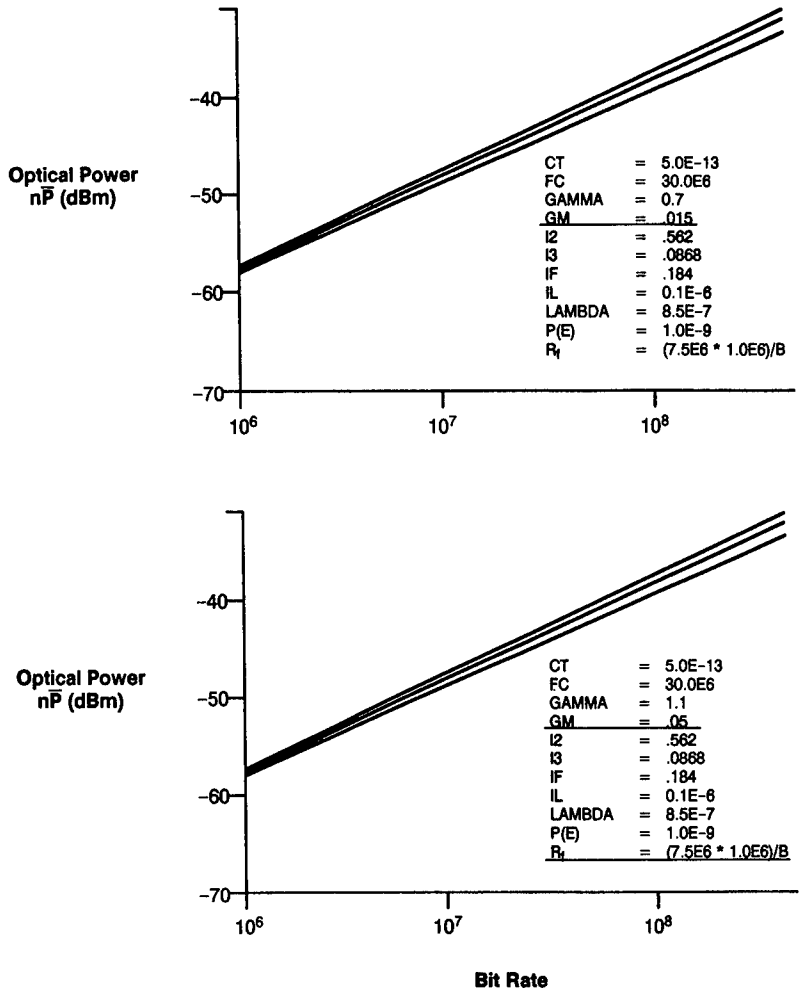


Fig. 4.18 Receiver sensitivity at two values of g_m.

gibly to the noise baseline and therefore to the ultimate sensitivity of the receiver. Some products are designed so that the bandwidth of the preamp is tailored to meet the required bandwidth of the incoming signal, avoiding intersymbol interference problems. This convenience simplifies the receiver such that adding equalization and filtering would not significantly improve receiver performance. A block diagram for a receiver is shown in Fig. 4.21. The time equalizer allows the adaptive thresholding to take place before the first bit arrives at the comparator. Other types of postamp circuits can be designed for different applications. Tables 4.5 and 4.6 summarize the trade-offs for various detectors used in fiber optic systems.

4.4.3 Encoding Techniques

The following is a description of different encoding techniques available for use in a digital communication system. The encoding schemes are divided into

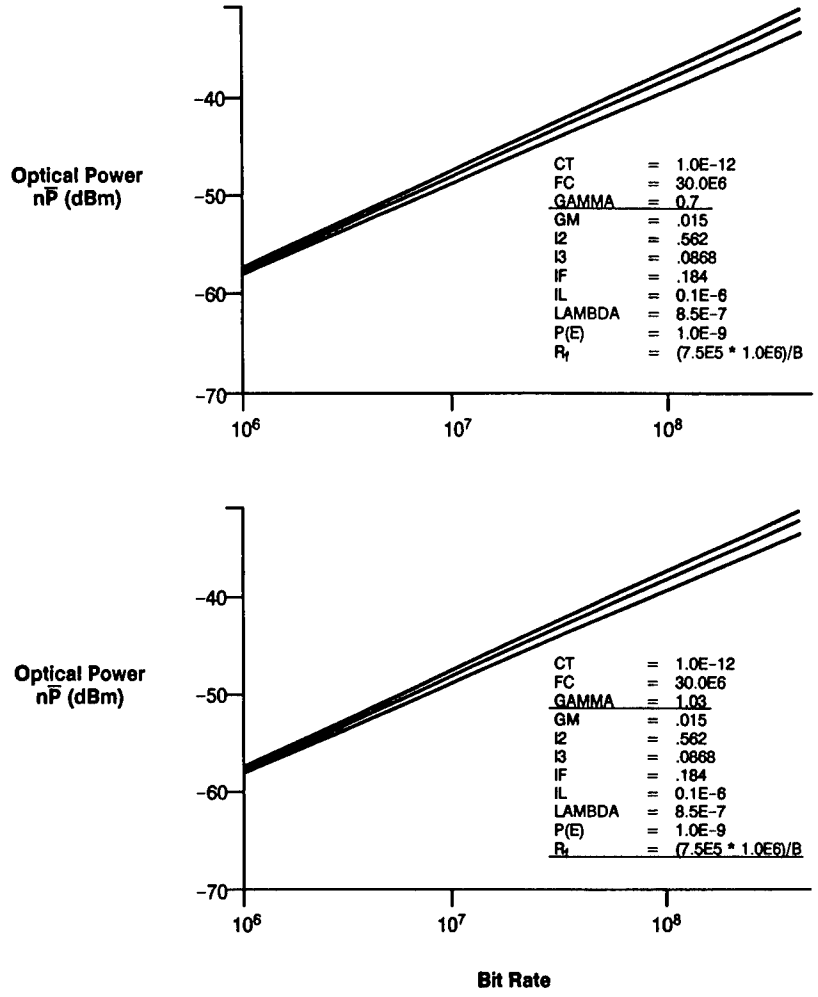


Fig. 4.19 Receiver sensitivity at different values of gamma.

two groups: base-band techniques and bandpass techniques. A base-band encoding scheme does not use a carrier frequency and has a power spectral density (PSD) greatest near dc levels (frequency = 0). The base-band techniques described here include the unipolar nonreturn to zero (NRZ), the polar NRZ, the unipolar return to zero (RZ), the bipolar, the Manchester, and the 4B/5B scheme. A bandpass waveform includes a carrier frequency at which the PSD is centered. The bandpass techniques include on-off keying (OOK), binary phase-shift keying (BPSK), and frequency shift keying (FSK). More complex forms of bandpass signaling, such as quadrature phase-shift keying (QPSK) and M-ary phase-shift keying (MPSK) are neglected because their complexity is useful only for extremely high data rates (of the order of gigabits per second). The different encoding schemes are described below along with a short theoretical analysis of the bandwidth requirements.

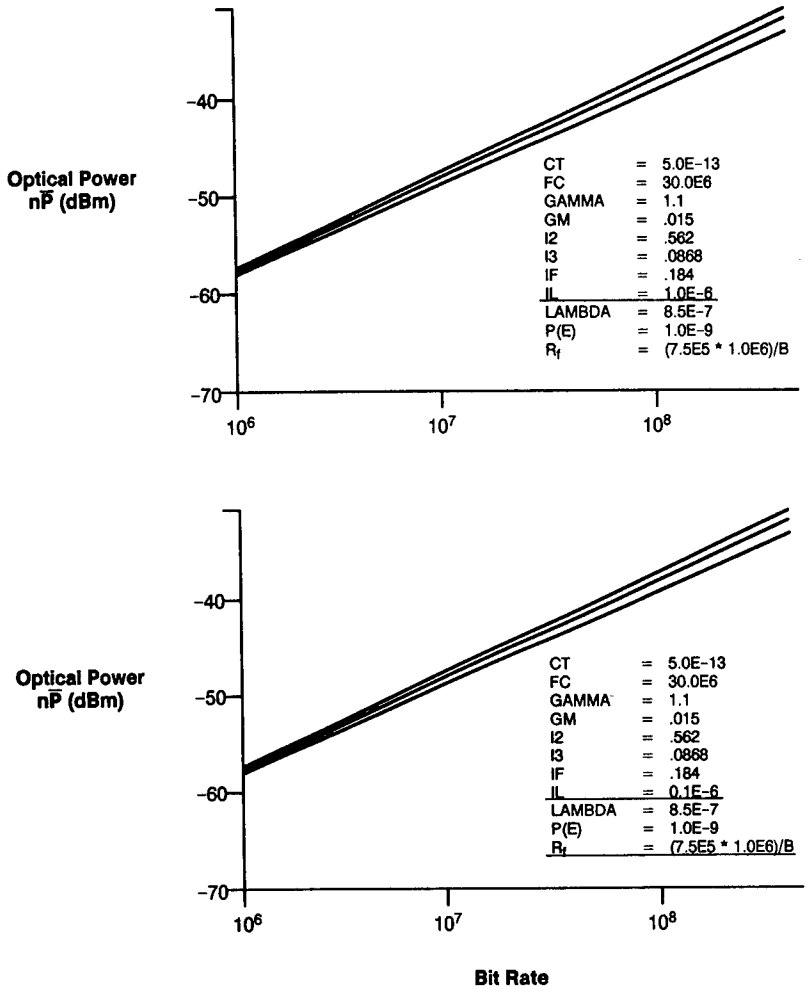


Fig. 4.20 Receiver sensitivity for different values of FET leakage current I_L.

4.4.3.1 Base-band Signaling Techniques. A diagram showing the different base-band signaling techniques is given as Fig. 4.22. A description of each follows. To calculate the frequency requirements for each of these techniques, the power spectral density must be obtained. The PSD is obtained by taking the Fourier transform of the bit shape and multiplying it by the autocorrelation of the bits (which takes into account the bit pattern variations). The frequency requirements of the system are then calculated from the PSD. Figure 4.23 plots the power spectral densities for each of the encoding techniques.

Unipolar NRZ. The unipolar NRZ pulse code represents 1's as a positive voltage (+A) and 0's as 0 voltage. The voltages for both 1's and 0's last for

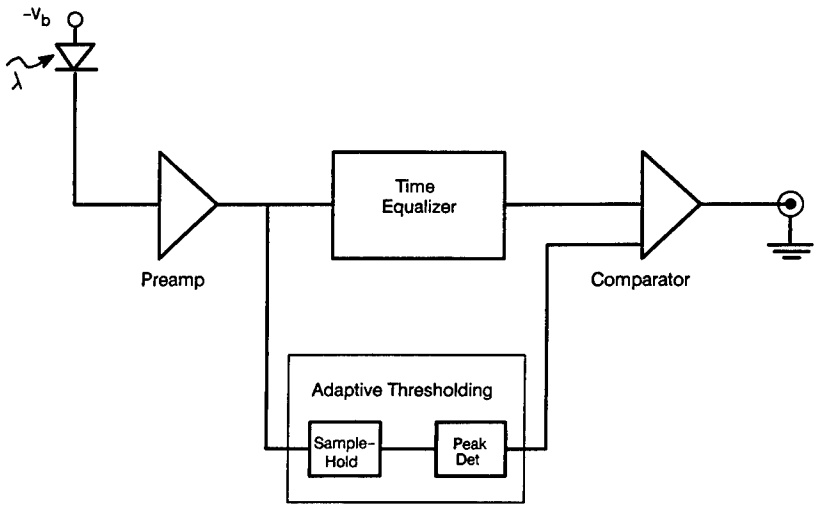


Fig. 4.21 Receiver block diagram.

the duration of the bit time. The PSD for unipolar NRZ with rectangular pulses is given by

$$P_{\text{unipolar}}(f) = \frac{1}{2}T_b\left[\frac{\sin(\pi f T_b)}{\pi f T_b}\right]^2 + \frac{1}{2}\delta(f) \ , \tag{4.27}$$

where T_b is the bit duration and is equal to $1/R$, where R is the bit rate. The delta function is present because of the average dc level of the system. By using the full width at half maximum (FWHM) of the PSD, the frequency range required for the unipolar NRZ encoding format is found to be $0.5R$.

Table 4.5 Detector Summary

Device	*Risetime*	*Sensitivity*	*Dynamic Range*	*Temp. Range*	*MTBF (25 C)*
PIN	< 1 ns	-51 dBm at 1 Mbps -34 dBm at 50 Mbps	≥ 25 dB	-55 to +125	10^9 hrs.
APD	< 1 ns	-61 dBm at 1 Mbps -44 dBm at 50 Mbps	≤ 20 dB	-55 to +125	10^9 hrs.

Table 4.6 Detector Arrays for Fiber Optic Sensors

Arrays of Photodiodes:	*Sensor Size*	*Sensor Spacing*	*Circuitry Required*	*Spectral Sensitivity Range*	*Sensitivity*	*Response Time*
Silicon	.010"	.025"	Separate preamp for each element	0.4 - 1 um	-70 dBm / element @ 100 Hz	1 ns
Germanium	.010"	.025"	Separate preamp for each element	0.9 - 17 um	-50 dBm / element @ 100 Hz	1 ns
InGaAs	.010"	.025"	Separate preamp for each element	0.6 - 1.8 um	-70 dBm / element @ 100 Hz	0.5 ns
CCD/Linear	.001"	.001"	Scanning circuit	0.4 - 1 um	-110 dBm / element	@ 16 ms
CCD/Area	.0005"	.0005"	Scanning circuit	0.4 - 1 um	-70 dBm / element	@ 1 ms

Polar NRZ. The polar NRZ encoding format represents a 1 as a positive voltage (+A) and a 0 as a negative voltage (−A). Again, the voltage level is maintained for the duration of the bit. The average value of this signal is therefore 0. The PSD for a polar NRZ with rectangular pulses is given by

$$P_{\text{polar}}(f) = T_b\left[\frac{\sin(\pi f T_b)}{\pi f T_b}\right]^2 . \tag{4.28}$$

Again using the FWHM of the PSD, a polar NRZ encoding format requires a frequency range of $0.5R$.

Unipolar RZ. The unipolar RZ format codes 1's as a positive voltage (+A) for half of the bit period and then a return to 0 for the second half. Zeros are coded as 0 V for the entire bit period. The PSD of the unipolar RZ format is given by

$$P_{\text{RZ}}(f) = \frac{T_b}{4}\left[\frac{\sin(\pi f T_b/2)}{\pi f T_b/2}\right]^2\left[1 + \frac{1}{T_b}\sum_{n=-\infty}^{\infty}\delta\left(f - \frac{n}{T_b}\right)\right] . \tag{4.29}$$

Because a 1 is represented by a pulse half as wide as the NRZ case, it follows that the frequency range is twice as wide as the NRZ case, and the PSD shows this. The frequency range required for a unipolar RZ format is R.

Bipolar. For the bipolar encoding format, the 1's are represented alternately by positive and negative RZ pulses. The 0's are given by 0 V for the duration of the bit period. This means that the detector must be able to detect three levels instead of two. The PSD for this format is given by

$$P_{\text{bipolar}}(f) = \frac{T_b}{4}\left[\frac{\sin(\pi f T_b/2)}{\pi f T_b/2}\right]^2 [2\cos(2\pi f T_b)] . \tag{4.30}$$

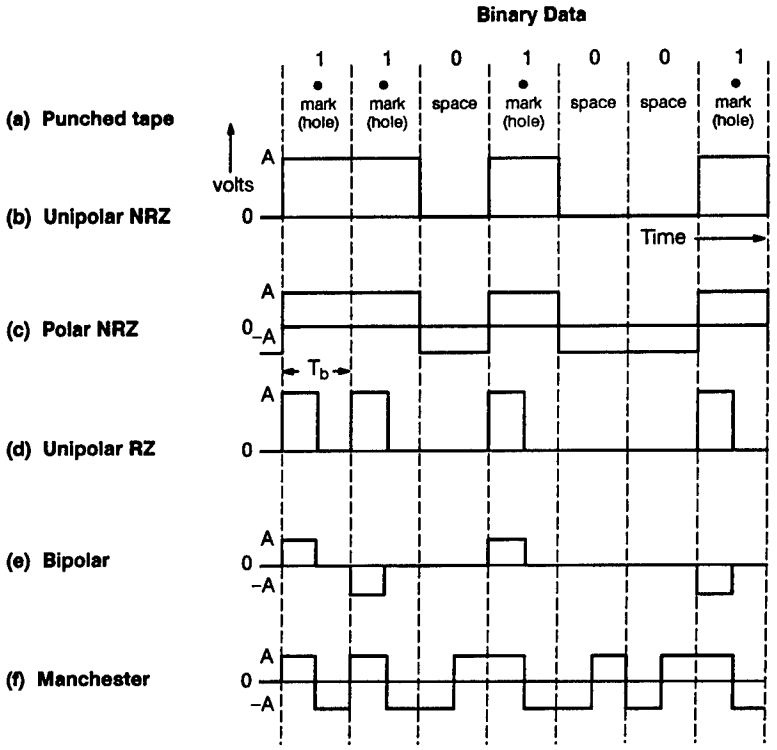


Fig. 4.22 Base-band binary signaling formats.

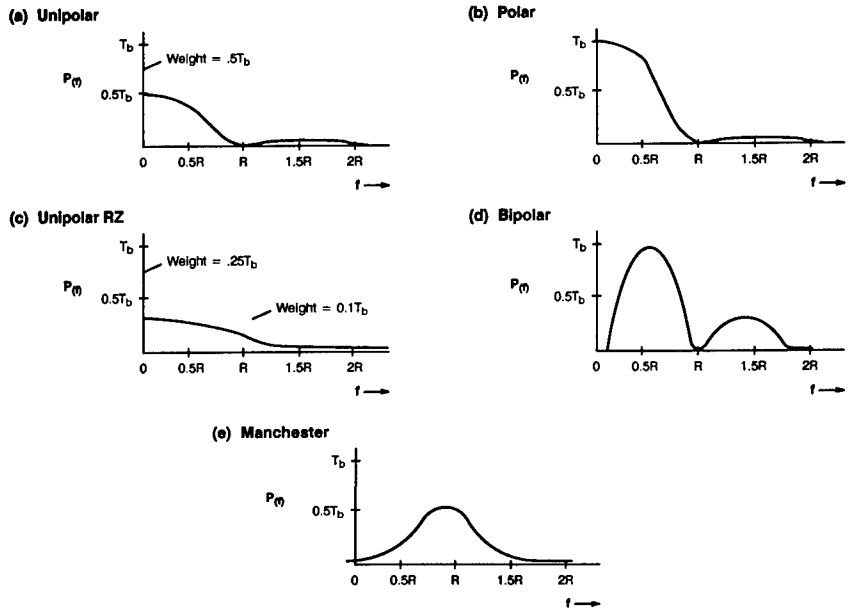


Fig. 4.23 Power spectral densities for the base-band signaling formats.

The bandwidth is taken from the FWHM of the first lobe of the PSD, which may mean that some of the power is lost in the second lobe. The frequency range for bipolar is $0.5R$.

Manchester. For the Manchester code, 1's are represented by a positive voltage (+A) for the first half of the bit period and a negative voltage (−A) for the second half. The 0's are represented by the reverse, i.e., negative first half and positive second half. This allows for exact clock recovery independent of the data stream. The PSD of the Manchester code is given by

$$P_{\text{Manchester}}(f) = T_b\left[\frac{\sin(\pi f T_b/2)}{\pi f T_b/2}\right]^2 \sin^2(\pi f T_b/2) \ . \tag{4.31}$$

This gives a FWHM frequency range of approximately $0.85R$. Note that this is a polar Manchester format, i.e., both positive and negative voltages are used (this includes the PSD in Fig. 4.23). For a fiber optic system, only positive intensities are available and a unipolar Manchester is used. This change reduces the frequency range to $0.5R$ and changes the high and low values for frequency to R and $0.5R$, respectively.

4B/5B. For more advanced encoding schemes, variations of the unipolar and polar NRZ formats are used. In the case of the 4B/5B format, every 4 bits must have at least one transition. Therefore, 5 bits are used for every four information bits to allow for the clock recovery, error detection, and control data bits. This gives a bandwidth around $0.7R$. These schemes require much more elaborate electronics and are not necessary at lower bit rates.

4.4.3.2 Bandpass Signaling Techniques. Figure 4.24 is an illustration of the following bandpass signaling techniques. A comparison to the unipolar and polar NRZ formats is also given. Again, the PSD techniques for calculating the frequency requirements are used. However, because the signals include a carrier frequency, the PSDs given are centered around this carrier frequency. The technique used to transform the base-band PSDs to the bandpass PSDs involves the equation

$$P_s(f) = \frac{1}{4}[P_g(f - f_c) + P_g(-f - f_c)] \ , \tag{4.32}$$

where $P_g(f)$ represents the PSD of the envelope (the corresponding base-band signal) and f_c is the carrier frequency.

On-Off Keying (OOK). OOK involves turning on and off an oscillator with 1's represented by the on state and 0's represented by the off state. This is essentially a unipolar signal with a carrier frequency. The PSD of this signal is identical to the unipolar NRZ, except that it is positioned around the center frequency and is twice the width. This also indicates that the frequency range of the OOK signal is twice that of the unipolar NRZ signal, or the bandwidth equals R. Figure 4.25 shows the PSD for OOK and BPSK.

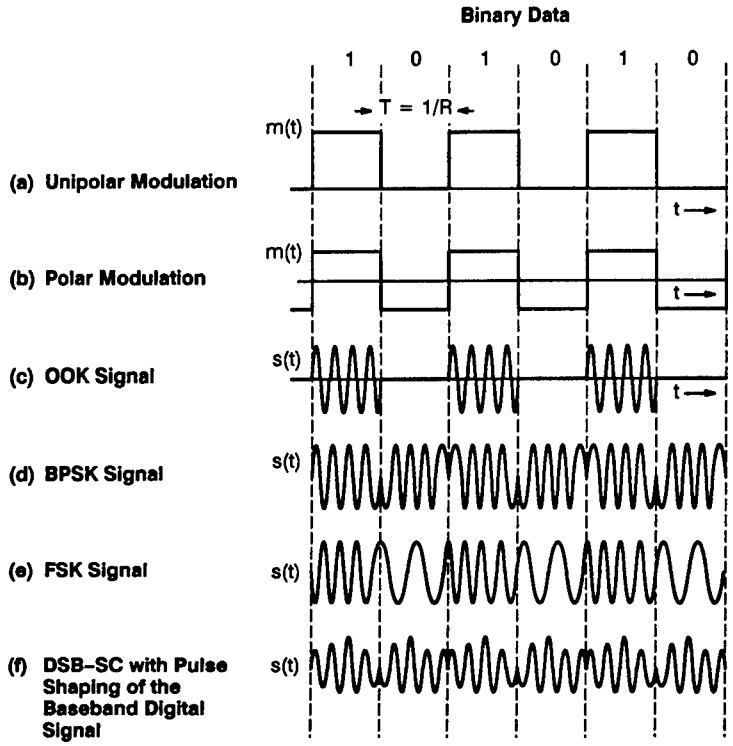


Fig. 4.24 Bandpass signaling formats.

Binary Phase-Shift Keying (BPSK). BPSK consists of shifting the phase of the carrier signal between 0 and 180 deg with the 0 deg representing a 1 and 180 deg representing a 0. This can also be thought of as a polar NRZ signal format multiplied by a carrier frequency. Using the formula for calculating the PSD of the signal, it can be seen that the BPSK signal is identical to the polar NRZ signal, except that it is centered around the carrier frequency and is twice as wide. Therefore, the bandwidth of this signal is twice that of the polar NRZ format, or R.

Frequency Shift Keying (FSK). FSK uses two different frequencies to represent the 1's and 0's. The carrier frequency is then considered to be the average of these two frequencies. Calculating the bandwidth for the FSK signal is much more difficult than for the other types of signals, but it can be approximated by

$$B_T = 2\Delta F + 2B \ , \tag{4.33}$$

where $2\Delta F = |f_2 - f_1|$ and B is the bandwidth of the square-wave modulation signal and is equal to R for this case. So, substituting, we get

$$B_T = 2\Delta F + 2R \ . \tag{4.34}$$

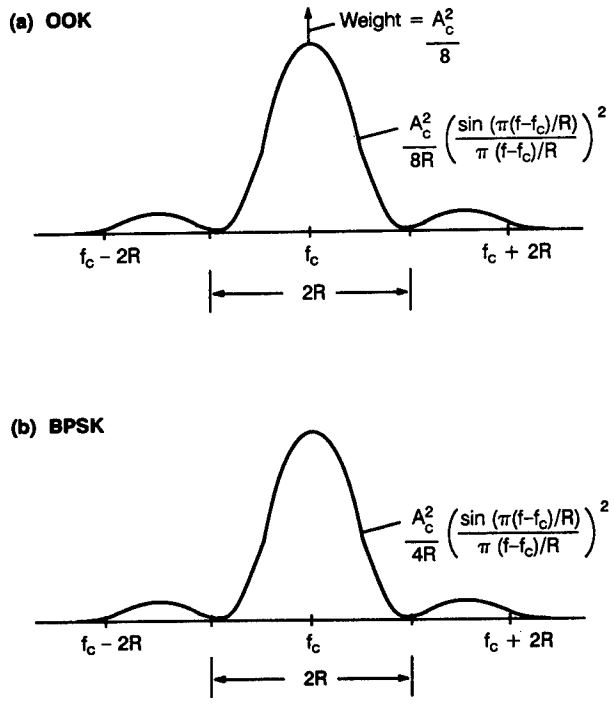


Fig. 4.25 Power spectral densities for the bandpass signaling formats.

4.4.3.3 Amplifier Range Requirements. The amplifier range is an important factor when considering the overall cost of the receiver. The main problem encountered when considering the range of frequencies to be amplified is the need to amplify from dc out to some large frequency. This is a factor close to infinity in terms of increase in frequency from the low value to the high value. For example, a signal format with a range of frequencies from 1 to 2 MHz is only a factor of 2. An amplifier for this type of signal is much easier to build than one that must amplify dc up to 1 MHz, even though they have the same numerical range. Therefore, signal formats with dc power levels are less attractive from an overall system standpoint even though their bandwidths are very small.

4.4.3.4 Rise Time Requirements. For base-band-type signals, a square-pulse train is used. In the frequency domain, the PSD actually includes a mainlobe plus many sidelobes. This results from the harmonics that are involved in creating a perfectly square wave. For the above analysis, the bandwidth was calculated using the first null bandwidth approach, i.e., only the first lobe is considered for the bandwidth. This technique indicates that the square wave is actually represented by a sinusoidal wave. To make the sinusoidal wave more closely resemble the square wave, the harmonics must be included in the bandwidth. This obviously increases the amount of bandwidth required for

Table 4.7 Bandwidth Requirements Versus Time for Pulses

t_r, Rise Time, ns	f,* Bandwidth, MHz
1.00	350
1.50	233
2.00	175
2.50	140
3.00	116
3.50	100
4.00	87.5
5.00	70.0
6.00	58.4
7.00	50.0

* f is the upper-bandpass frequency at the 3-dB point for a low-pass filter. The frequency f can also be considered the upper (highest) frequency component (with a significant amount of energy) in the waveform generated when a logic device switches states.

the system. A representative progression of the bandwidth requirement versus the rise time of the leading edge of the pulse is given in Table 4.7. An important fact to consider here is that the signal in the fiber does not have to resemble a square wave to work well. Using good level detectors, a sine wave produces the same output at the receiver as a square wave.

4.4.3.5 Comparison Chart for Different Signaling Schemes.

Table 4.8 compares the attributes of the different signaling schemes. The frequency range, including minimum and maximum required frequencies, is given along with some advantages and disadvantages for each of the formats. The chart is divided into two sections because the bandpass formats have true bandwidths associated with them and the base-band formats have a frequency range. This frequency range is similar to a bandwidth (BW) and can be used as such for comparison purposes. Therefore, for the bandpass formats, the frequency range is actually the bandwidth.

4.5 MULTIPLEXING TECHNIQUES

One of the main reasons that fiber optics technology is so useful for communications is its tremendous bandwidth capability. Because light waves have higher frequencies than radio waves, their information-carrying capacities are correspondingly greater. A typical 850-nm source has a frequency of 10^{14} Hz. Of course, the only way to realize this potential capacity is by using carrier modulation techniques, such as heterodyne, etc. Such systems are currently still under development, and so will not be treated here. There are two other methods of increasing the effective bandwidths of current fiber optic systems, however. These are optical and electronic multiplexing of signals.

4.5.1 Spatial Multiplexing

The simplest type of multiplexing is spatial multiplexing. In spatial multiplexing, the extremely small size of the fibers (compared to twisted shielded

Table 4.8 Comparison of Attributes for Various Signaling Formats

Format		Frequency Range	Min. Freq.	Max. Freq.	Advantages	Disadvantages
BASEBAND	Unipolar NRZ	1 MHz	0 MHz	1 MHz	small BW	dc power level rise time harmonics
	Polar NRZ	1 MHz	0 MHz	1 MHz	small BW	dc power level rise time harmonics
	Unipolar RZ	2 MHz	0 MHz	2 MHz	clocking possible on ones	dc power level rise time harmonics
	Bipolar	1 MHz	.5 MHz	1.5 MHz	error detection	rise time harmonics three level det.
	Manchester (polar)	1.68 MHz	.64 MHz	2.32 MHz	clock recovery always possible no dc power	rise time harmonics
	Manchester (unipolar)	1 MHz	1 MHz	2 MHz		
BANDPASS	OOK	2 MHz	f_c - 1 MHZ	f_c - 1 MHZ	no dc power	must detect carrier reference signal
	BPSK	2 MHz	f_c - 1 MHz	f_c - 1 MHz	no dc power	synchronous detection needed
	FSK	2 MHz + $\lvert f_2 - f_1 \rvert$	$f_c - \Delta F$ - 1 MHz	$f_c - \Delta F$ - 1 MHz	no dc power	need frequency detector or coherent det.

pairs or coaxial cable, for example) is used to place several fibers in a space smaller than a single wire. Each fiber carries one data transmission in the simplest case, and so the number of multiplexed signals depends on the number of fibers placed together. In most cases, spatial multiplexing is used in conjunction with one or more other multiplexing techniques to further increase system capacity, or to provide growth or redundancy capability. For some telephony applications, 12 ribbon cables, each containing 12 fibers, are placed together in an assembly that is less than 0.5 in. square in cross section.

4.5.2 Wavelength Multiplexing

As the name implies, wavelength multiplexing is used to combine many optical channels on a single fiber by bandlimiting the optical spectrum of each channel. For example, the optical spectrum of a typical 850-nm LED, approximately 45 nm, could be divided into four channels of 10 nm each with a 1-nm guard band between channels.

The parameters that control and determine the degree of multiplexing are the available spectral width and the bandlimiting properties of the optical multiplexing/demultiplexing components. Figure 4.26 shows the spectrum available using typical LEDs. In most cases, visible LEDs are unsuitable for communication purposes and are therefore not treated here. The remaining LEDs can all be used for wavelength multiplexing.

The most straightforward way to optically multiplex is to use sources from two or more of these regions to transmit signals over the same fiber. For example, because 850-nm detectors have essentially no responsivity to 1300-nm wavelengths, the detector serves as its own low-pass optical filter. Two optical signals, one at 850 nm and one at 1300 nm, traveling in the same fiber simultaneously, can be divided and directed to two separate detectors, one responsive at 0.85 μm and one at 1.3 μm. The 850-nm detector, as shown in Fig. 4.27, is not able to respond to the 1300-nm optical signal, so no interference between the two signals occurs at 850 nm. Although the 1300-nm detector can respond to 850-nm wavelengths, its responsivity is typically at least 8 to 10 dB lower than at 1300 nm, so any 850-nm optical signals reaching the detector

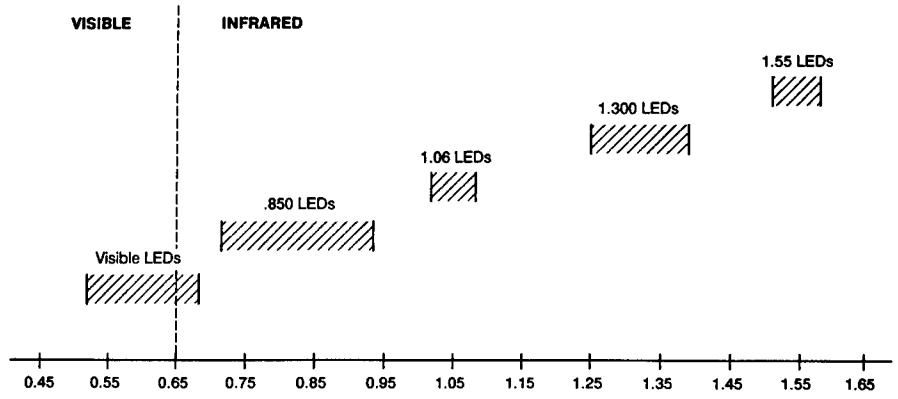


Fig. 4.26 Source spectra available from LEDs.

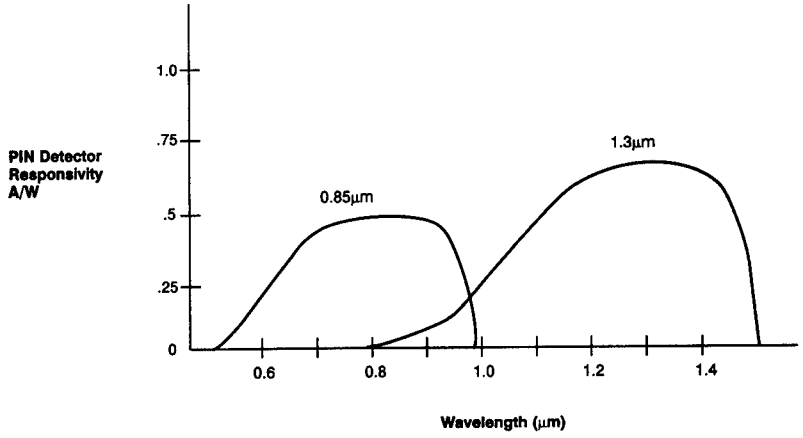


Fig. 4.27 Detector responsivity.

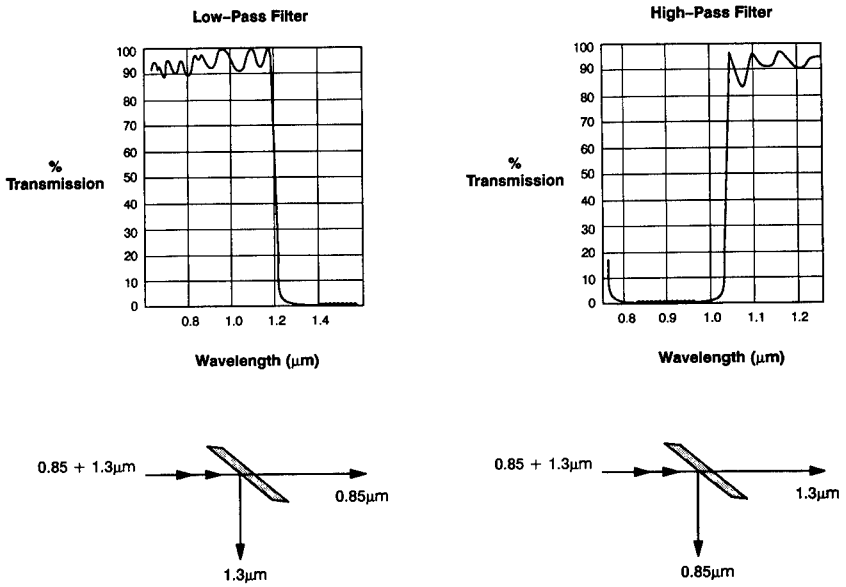


Fig. 4.28 Dichroic beamsplitters.

are attenuated about 10 dB relative to the 1300-nm signals. This is often sufficient attenuation, given typical system power budgets, to produce effective two-channel multiplexing.

In many cases, 30 dB or more of relative attenuation is necessary. Filters like those shown in Fig. 4.28 can be used to accomplish this. The filters are multilayer dielectrics designed to selectively reflect and transmit different

wavelengths. The quality of the filter is judged by the maximum percentage transmission or reflection and the steepness of the cutoff edge. The percentage transmission affects the attenuation of the signal, and the steepness of the edge determines how closely in wavelength the two channels can be placed.

Figure 4.29 shows a dichroic two-color multiplexing device constructed from such filters. Two light signals having different wavelengths (850 nm and 1300 nm, for example) enter the device from the fiber as shown in the figure. The light is collimated and the collimated beam is incident on the filter where one wavelength is reflected and the other is transmitted. The resulting two collimated beams are refocused into separate fibers. This device works in reverse as well, so bidirection transmission paths are possible over a single fiber. The quality of the two channels can be measured by how much of the wrong wavelength remains in the channel. Typically, a single filter can produce about 30 dB of optical isolation (one part in 1000). Higher quality filters are possible, but it is usually more cost effective to place one or more "blocking" filters in front of the output fibers. In an 850/1300-nm dichroic system, a 1300-nm blocking or reflecting filter would be placed in front of the 850-nm fiber to produce up to 30 dB more isolation.

Figure 4.30 shows the wavelength versus percentage transmission of another type of filter, which can be made from dielectric filters, called the *bandpass filter*. This filter is designed to reflect all light except a narrow bandwidth, which is transmitted. A number of such filters can be stacked to select out multiple optical channels. This can also be accomplished by stacking high- or low-pass filters having slightly different cutoff wavelengths, as shown in Fig. 4.31. The number of channels possible is limited by the filter characteristics as already described.

One drawback of dichroic filters used in multiplexing optical signals is that the percentage transmission decreases as the optical bandpass of the filter decreases. Therefore, to multiplex several channels within the spectral region

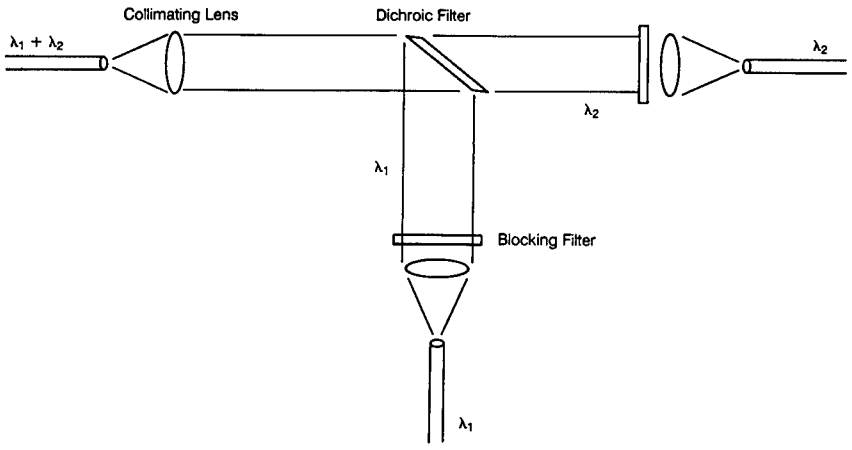


Fig. 4.29 Dichroic coupler configurations.

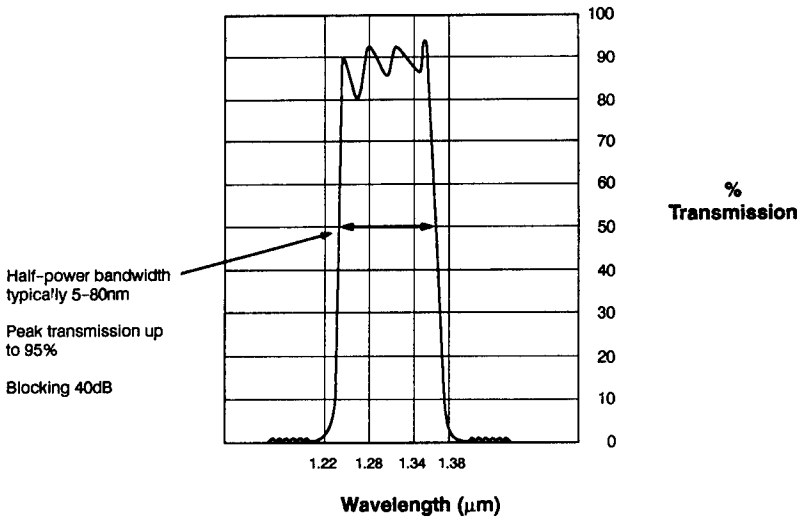


Fig. 4.30 Bandpass filters.

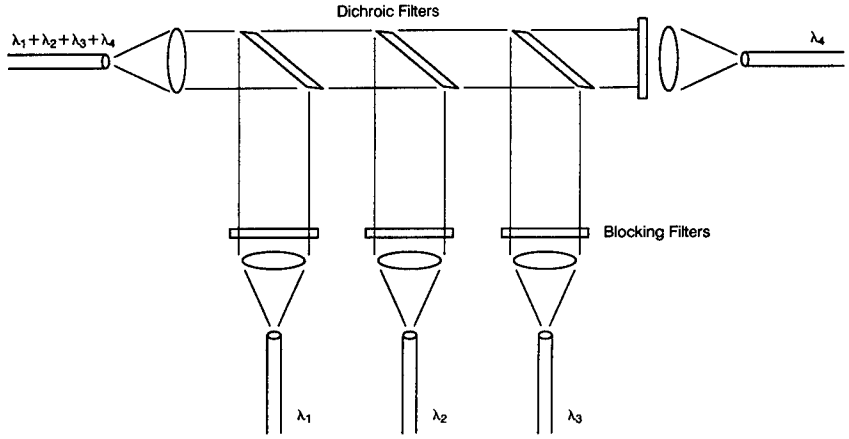


Fig. 4.31 Multiple-channel dichroic multiplex system.

of a single type of LED (850 nm, for example), the exceptionally narrow filter bandpass required increases the wavelength division multiplexing (WDM) insertion loss considerably.

To avoid this problem, WDM that uses a dispersive element can be employed. A schematic of such a device is shown in Fig. 4.32. In the demultiplexing mode, light of many different wavelengths ($\lambda_1 \ldots \lambda_n$) emerges from the input fiber and is collimated before it strikes a reflective diffraction grating. The grating diffracts the light, sending a collimated beam for each λ back to different fibers. For a discrete set of input λs, a discrete set of output positions is gen-

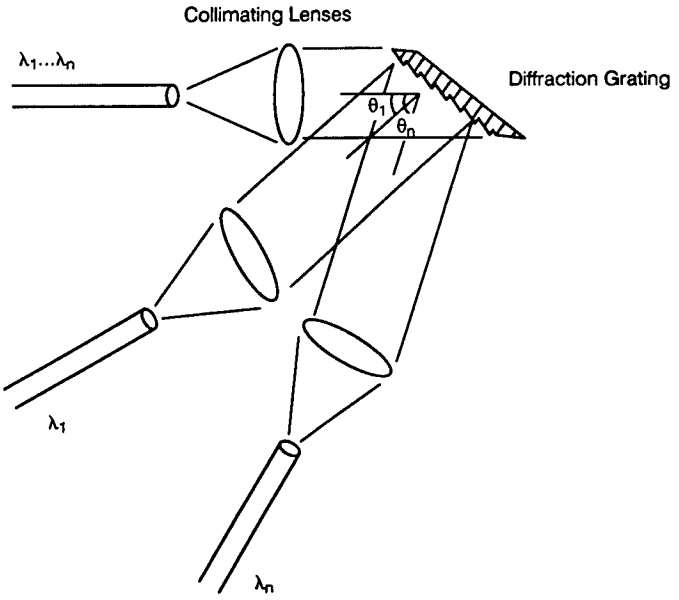


Fig. 4.32 Dispersive WDM schematic diagram.

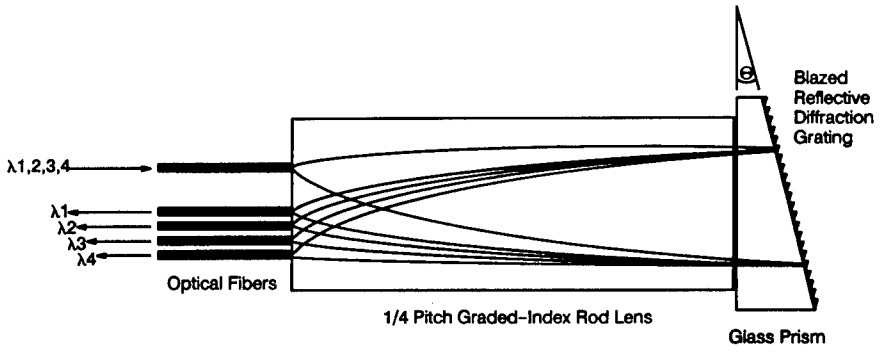


Fig. 4.33 Dispersive element WDM.

erated. If the input spectrum is continuous, then any number of output channels can be obtained with proper placement of the output fibers. In that case, the centers of the output channels are determined by the fiber positions, whereas the bandpasses of each channel are determined by the size of the fiber core.

A practical device such as the one just described is shown in Fig. 4.33. In this device, one lens serves for both output and input fibers. For convenience, a graded-refractive-index (GRIN) rod is used to focus the light. The grating is replicated on the back of a prism whose angle closely matches the blaze angle, so that the grating is used in the Littrow configuration. Figure 4.34 shows the bandpass characteristics of this device when used to multiplex or demultiplex four channels. Figure 4.35 shows a four-channel system constructed from a

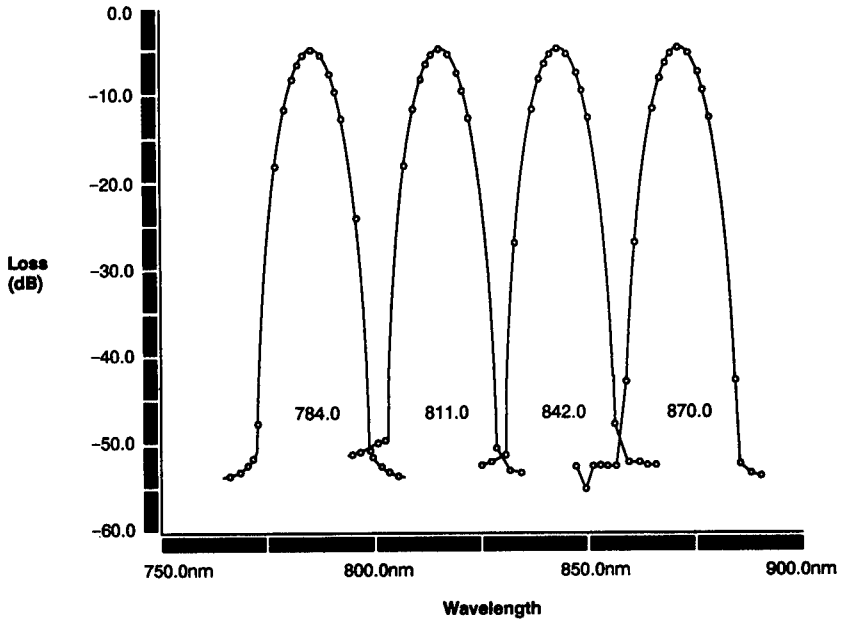


Fig. 4.34 WDM bandpass.

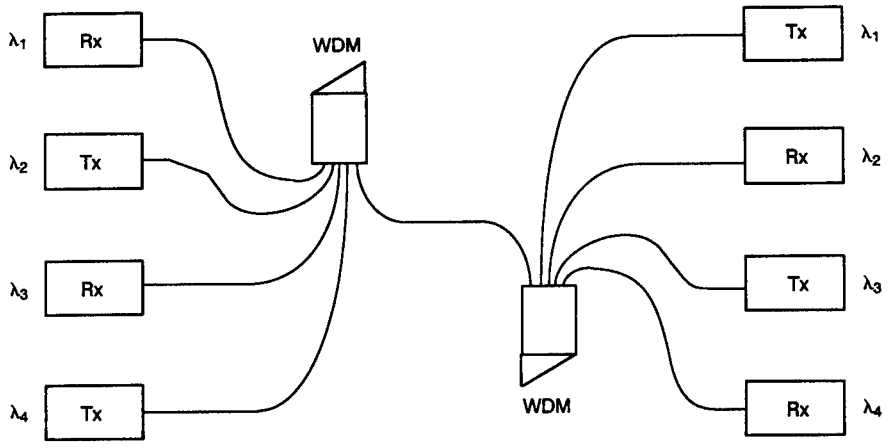


Fig. 4.35 Four-channel WDM system.

pair of these bidirectional devices. Typical performance of this type of device is given in Table 4.9.

Note that the effective attenuation of the narrow bandpass channels is greater than the center-band attenuation. For example, if the channel width is 12 nm and the LED source spectral width is 50 nm, the effective channel attenuation is the sum of the WDM loss and the filtering effect of the WDM

Table 4.9 Dispersive WDM Performance

Grating Constant	1200 lines/mm
Blaze Angle	17.45 degrees
Input Fiber	100/140
Output Fiber	100/140
Minimum Channel Spacing	24 nm
3 dB Channel Width	12 nm
Crosstalk Isolation	>30 dB
Centerband Attenuation	3 dB

channel on the LED spectrum. So, if an LED unfiltered would launch −10 dBm into 100/140 fiber (P_0), then the power that would exit the output of the WDM would be

$$P_0 - (L_{\text{WDM}} + L_{\text{bandpass}}) \ , \tag{4.35}$$

where P_0 = LED full spectral power, L_{WDM} = WDM center channel attenuation, and L_{bandpass} = "filtering" attenuation in decibels, which is defined as

$$-10 \log\left(\frac{\text{channel FWHM}}{\text{LED FWHM}}\right) \ . \tag{4.36}$$

Using $P_0 = -10$ dBm, $L_{\text{WDM}} = 3$ dB, channel FWHM = 12 nm, and LED FWHM = 50 nm, we obtain

$$\begin{aligned} P &= -10 \text{ dBm} - \left[3 \text{ dB} - 10 \log\frac{12}{50}\right] \\ &= -10 \text{ dBm} - (3 \text{ dB} + 6 \text{ dB}) \\ &= -19 \text{ dBm} \ . \end{aligned} \tag{4.37}$$

This is the total effective launch power in the WDM channel. A full consideration of system effects is given in Sec. 4.7.5

4.5.3 Electronic Multiplexing

Because optical fiber links can support high data rate transmission over longer distances than wire or coax links, it is often advantageous to combine standard electronic multiplexing techniques with fiber optics to produce multiplexed data paths. A diagram of a digitally multiplexed generic link is shown in Fig. 4.36.

Digital Multiplexing. Both analog and digital data can be multiplexed using a digital technique called *time division multiplexing* (TDM). If the data to be combined are already in digital form, it is a simple matter to multiplex the data using a parallel-to-serial converter, such as a shift register, as shown in Fig. 4.37. In this technique, each data input, which is itself a serial data stream, is sampled in turn, 1 bit from each input. The shift register is then clocked out at a rate that is at least eight times the highest data rate of the individual

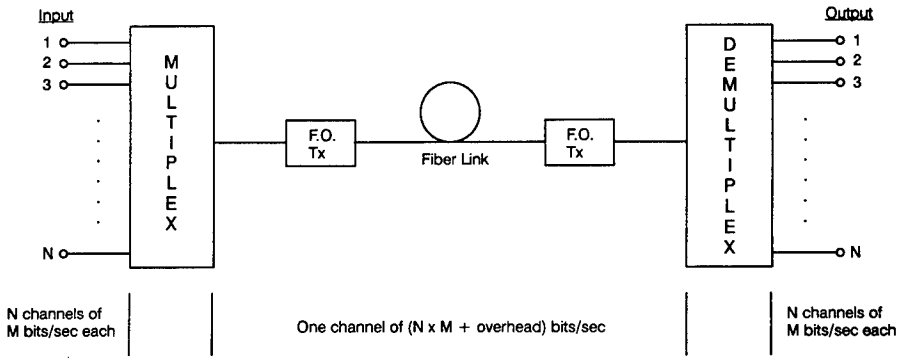


Fig. 4.36 Electronic multiplexing block diagram.

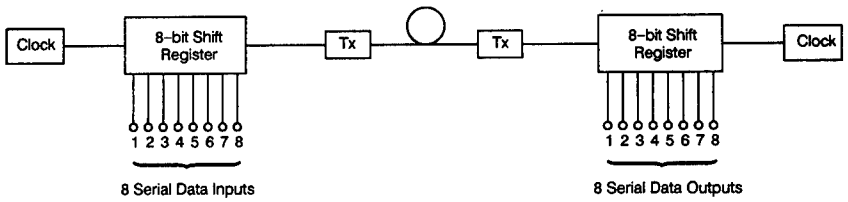


Fig. 4.37 TDM using shift registers.

inputs. These serial data are then transmitted over the optical fiber, after which they are fed into another shift register used in the serial-to-parallel mode.

Conceptually, this is simple. However, it cannot practically be used in this form, normally. The main problem is synchronization between the two circuits. To correctly rebuild the eight individual data streams, it is necessary to make sure that each sampled channel from the multiplexer (MUX) comes out of the demultiplexer (DEMUX) on the same channel. In addition, the demultiplexer circuit clock must be timed accurately to the multiplexer clock. These problems can be overcome by transmitting separate clock and synchronization channels over separate fibers, or via wavelength multiplexing. This is a somewhat inelegant and risky approach, however, given the timing problems involved and the original reason for multiplexing, which was to reduce the number of fibers needed to transmit the data.

Figure 4.38 shows the typical solution to these problems. The data inputs are sampled as before, but extra "framing bits" are added to the data stream to provide a means for the receiver to resynchronize the channels. A data encoder is also used to insert the clock into the final serial stream. This encoding can be any technique, such as Manchester or m-B/n-B, which will guarantee an adequate number of transitions for the clock to be recovered at the receiver side of the link. This also has the added benefit of increasing the lowest frequency that is transmitted through the optical link, simplifying the receiver design. The serial output of the optical receiver passes through a clock recovery circuit. The recovered clock is now used to achieve correct timing in

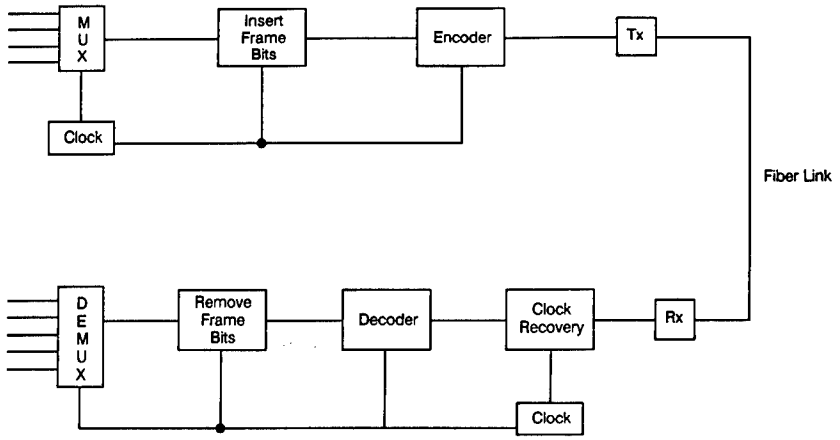


Fig. 4.38 TDM with full synchronization.

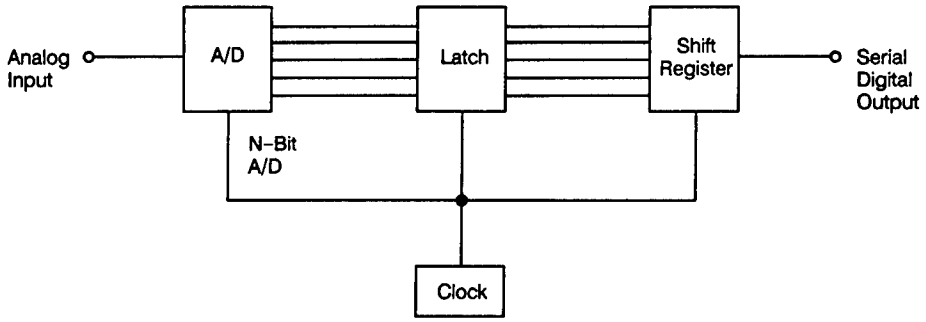


Fig. 4.39 TDM with analog signals.

the remaining circuitry. Next, the data are decoded, leaving only the actual data and frame bits. In some cases, the clock recovery and decoder functions are provided by the same circuit. Finally, the framing bits are removed and the serial data are restored to their original parallel-sampled form. In most cases, latches are required on both the input to the MUX and the output of the DEMUX. Several silicon and GaAs IC manufacturers now offer circuits that provide the entire MUX and DEMUX functions on a chip. Some of these products can be used to multiplex up to 40 channels of data at a resulting combined data rate of 500 Mbits/s. Other circuits can be obtained, in a less integrated form, to multiplex eight channels at a combined rate of 1.6 Gbits/s.

To multiplex analog data using this technique, the use of analog-to-digital (A/D) converters is required. Figure 4.39 shows how this is accomplished. The analog data are converted to n-bit parallel digital output, which is latched and converted to serial format by the shift register. Often, the latches are included in the output of the A/D converter. To avoid aliasing effects, the sample rate of the analog signal must be at least twice the highest frequency contained

within it. To multiplex several analog signals, the number of channels in the MUX must be at least $n \times m$, where n is the number of bits of resolution of the A/D converter and m is the number of signals to be multiplexed.

4.6 SYSTEM COMPONENTS

In the foregoing discussions, the key elements of a fiber optic link have been described (transmitters, receivers, fiber, and multiplexers). In this section, the remaining kinds of components most often used in fiber optic systems are described and their important performance features explained, especially as they relate to fiber optic system performance.

4.6.1 Connectors and Splices

The most basic component to be considered in a fiber optic system after the transmitter, receiver, and fiber is the fiber connector or splice. There are many different means used to make a fiber-fiber optical connection, but they fall into two main categories: splices and connectors. Although factors such as mode mixing and modal noise can be important in some types of systems, the fundamental performance parameter is attenuation. Connection attenuation is composed of the following components.

4.6.1.1 Fresnel Loss. Figure 4.40 shows an exaggerated view of the interface between two fiber cores. A light ray is shown propagating from a fiber core with refractive index n_1 through an intervening medium of index n_2 and into a second fiber core of index n_3. In general, there is a reflection at each interface where the indices on either side of the interface are different. The magnitude of the reflected power is given by

$$R = (n_a - n_b)^2/(n_a + n_b)^2 \ , \tag{4.38}$$

where n_a and n_b are the refractive indices of the media on each side of the interface. The Fresnel loss is given in decibels by

$$L_F = 10 \log_{10}(1 - R) \ . \tag{4.39}$$

In most cases, we consider an air-glass interface where $n_1 = 1.5$ and $n_2 = 1.0$, resulting in a Fresnel reflection of 0.04 (4%) and a Fresnel loss of 10 $\log_{10}$ (0.96) = 0.18 dB for each interface. In the general case shown in Fig. 4.40, there are two such interfaces with $n_1 = n_3 = 1.5$ and $n_2 = 1.0$, which give

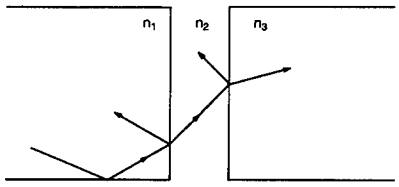


Fig. 4.40 Fiber-fiber interface.

two reflections of 4% each. The resulting Fresnel loss is $(0.96)^2$ or 0.35 dB. This is the case for a typical fiber optic connector. To reduce this loss, n_2 must be made more nearly equal to n_1. This can be accomplished either by using an index-matching fluid or by eliminating the interface altogether by fusing the two fiber cores together. Because this Fresnel loss is relatively small, it is normally ignored. In some cases, elimination of the reflected light is more important than the Fresnel loss, because the reflected light can cause a degradation of performance in some types of systems. It is also important to note that if the two interfaces are sufficiently flat and parallel, the gap is of the proper dimension, and the refractive index difference is correct, a Fabry-Pérot cavity is made between the two fiber cores. The interference effects of this cavity can either reduce or amplify the reflection, depending on the wavelength of the light being transmitted. This can produce unpredictable loss in the typical single-wavelength system, and can have disastrous effects in a wavelength-multiplexed system.

4.6.1.2 Mismatch Loss. Figure 4.41 shows the situation where two fibers having different core sizes are coupled together. Light rays propagating from the smaller core to the larger core suffer no additional attenuation. However, assuming fully filled conditions, the light coupled from the larger core to the smaller sees a mismatch loss of

$$L_{\text{diameter}} = 10 \log(D_1^2/D_2^2) \quad [\text{dB}] \ , \tag{4.40}$$

where D_1 and D_2 are the diameters of the two cores, as shown. Figure 4.42 shows a similar effect related to numerical aperture mismatch. Again, the loss is given by the ratios of the squares of the two NAs.

$$L_{\text{NA}} = 10 \log(\text{NA}_1^2/\text{NA}_2^2) \quad [\text{dB}] \ . \tag{4.41}$$

Figure 4.43 is a graph from which the loss can be found if the diameter or NA ratios are known.

4.6.1.3 Fiber End Separations. Figure 4.44 shows two fibers of the same diameter D and NA, separated by a distance S. To find the loss under these conditions, the end face area of the fiber from which the light rays emanate is projected at an angle corresponding to the NA, at a distance S from the face. Then the diameter mismatch relation is used. The resulting formula for the loss is

$$L_{\text{gap}} = 20 \log\left[1 + 2\frac{S}{D}(\text{NA})\right] \quad [\text{dB}] \ . \tag{4.42}$$

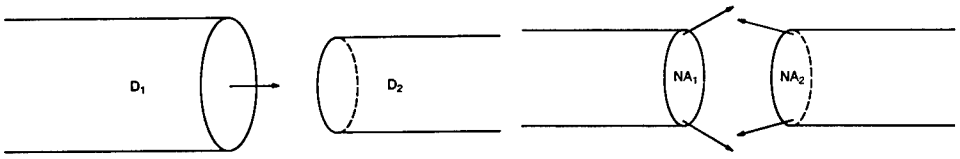


Fig. 4.41 Fiber core mismatch.

Fig. 4.42 Numerical aperture mismatch.

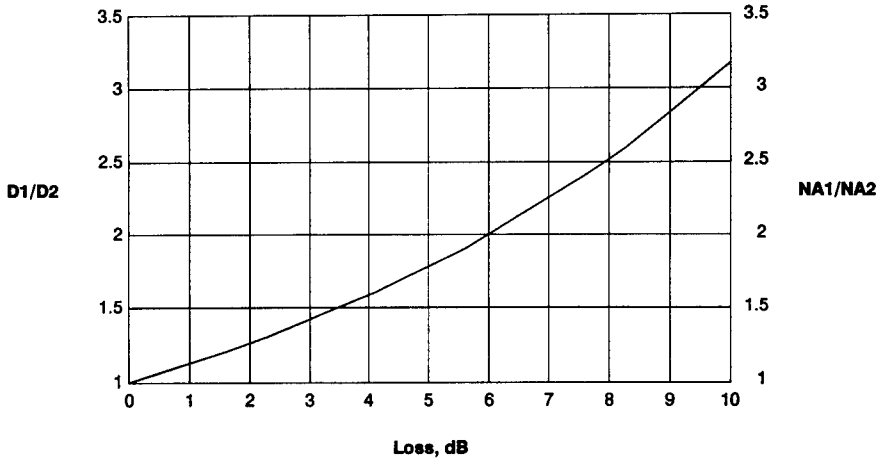


Fig. 4.43 Mismatch losses.

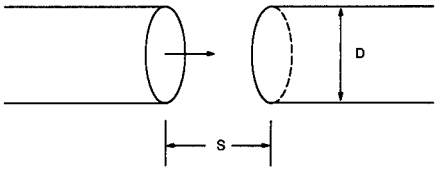


Fig. 4.44 End separation loss.

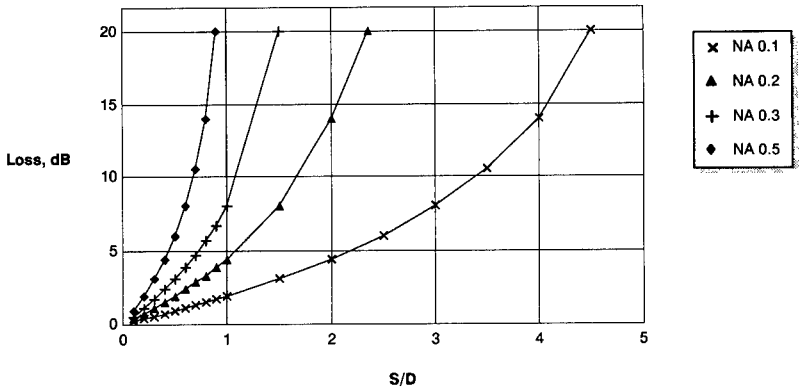


Fig. 4.45 End separation loss graph.

This result is valid only for fibers with matching diameters and NA, which is the usual case. The loss for other situations may be derived from the foregoing discussions. A graph of the gap loss is given in Fig. 4.45.

4.6.1.4 Splice Types. There are two basic methods of making splices. Fusion splices, which have been discussed, are made by removing the protective

buffer layers from the fiber, cleaning the fiber to remove contaminants, and then cleaving the ends to produce a clean, flat face. The two ends are aligned and held under slight compression while being heated. When the glass melting point is reached, the fibers flow together and the heat is removed. A protective buffer layer is added to the bare glass, and the splice is placed in a protective tube. The splice loss is normally less than 0.1 dB and the spliced fiber is mechanically as strong as an unspliced fiber. Manual or automatic machines are available that can perform the fusion process.

The second type of splice is the permanent or semipermanent mechanical splice. Fibers are prepared in the same manner, but instead of being fused together, the fiber ends are butt coupled using an index-matching medium inside a precision tube or groove mechanism. In some cases, the index-matching medium is an epoxy, especially one that cures when exposed to ultraviolet light. Insertion loss of a mechanical splice is rarely less than 0.1 dB, and can be several tenths of a decibel or more, depending on the quality of the mechanical components. Even though the mechanical splice exhibits higher loss than the fusion splice, its low cost and ease of use make it popular among many system designers.

4.6.1.5 Connector Types. In most cases, a fiber splice is intended to be a relatively permanent fiber-fiber coupling. The connector, on the other hand, is intended to be mated and demated a number of times during its service life. There are two basic types of optical connectors—butt coupled and expanded beam.

Butt-Coupled Connectors. Figure 4.46 shows the alignment technique used in butt-coupled connectors. Fibers are epoxied into metal or ceramic ferrules, whose inner and outer diameters have been carefully machined or otherwise made very accurate and concentric. Two ferrules are aligned inside a precision alignment tube, which accurately positions the two ferrules with respect to each other. The accuracy of the alignment of the two fiber cores depends on the combined tolerances of the individual components. Connection losses are typically 0.5 to 1.5 dB. Standard connectors of this type are known by their types: SMA, ST, and FC. Several manufacturers produce each type. The ends of the ferrules and fibers can be polished convex so that reflections from the fiber ends are minimized when the ends touch. This is the physical contact (PC) technique, and is usually found in ST or FC connector styles. In another

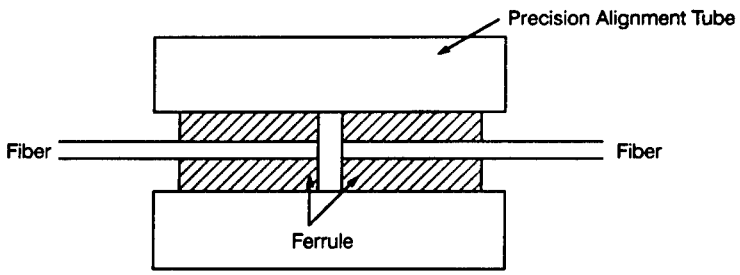


Fig. 4.46 Butt-coupled connector.

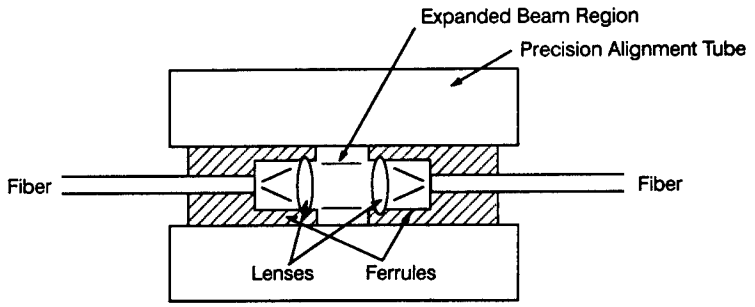


Fig. 4.47 Expanded-beam connector.

variation, called the *biconic connector,* the ferrules and alignment tube have matching cone shapes with the apex at the fiber end face (although the fibers are either flat or PC polished). The conical shapes help maintain the alignment of the fibers.

Expanded-Beam Connectors. Figure 4.47 also shows the basic components of an expanded-beam connector. The ferrules are made such that a lens is placed one focal length away from the end of the fiber. Then the light exiting the fiber is collimated into a parallel beam, which is refocused into the mating fiber by a lens in the second ferrule. The major advantages of the expanded-beam connector are its relative insensitivity to lateral misalignment and contamination. In a butt-coupled connector, a single particle of contamination that is only 50 μm in diameter can greatly affect the performance of the connector, even to the point of causing the fiber ends to shatter. A typical expanded-beam connector, on the other hand, contains a lens of 1 to 3 mm diameter. This expanded beam is much less sensitive to a small particle, and the lens can be wiped clean in the field without significant damage. Lateral alignment tolerances are also greatly relaxed because of the expanded beam. However, there is a corresponding increase in the requirement for accuracy in angular alignment. In addition, losses for the expanded-beam connector are usually greater, typically 1.5 to 2.0 dB or more.

4.6.1.6 Connector Systems. In the foregoing discussions we have actually been discussing the terminus portion of a connector, rather than the connector itself. Over the years, a great deal of effort has gone into the development of connector systems—i.e., the mechanical fasteners in which the termini are housed. In general, any of the standard connection systems (subminiature, D-style, and circular, for example) can be modified for optical terminus systems. In fact, the SMA optical connector is an optical version of an electrical connector (subminiature-A). Several connector manufacturers have already adapted their standard connectors for optical fibers. It is currently possible to obtain MIL-T-29504 optical contacts for MIL-C-38999 circular connectors, many times configured for both optical and electrical connections. This area of fiber optic technology changes rapidly in response to market requirements.

4.6.2 Couplers

In many fiber optic systems, it is necessary to distribute or combine optical paths between several transmitters and receivers. To accomplish this, an optical coupler is used.

4.6.2.1 Tap Coupler. The simplest type of coupler is the tap coupler, also known as a 2 × 2, 1 × 2, T-coupler, or Y-coupler. Figure 4.48 shows the various configurations.

This simple coupler allows the light from one fiber core to leak into another fiber core, so that light from port 1 comes out of ports 2 and 4. The following parameters define the performance of the 2 × 2 coupler, with port 1 being the input port.

split ratio = port 2/port 4

directivity = port 3/port 1

excess loss = port 1 − (port 2 + port 4)

splitting loss = port2/port 1 or port 4/port 1 .

The split ratio is defined as the ratio of the light in the two output ports. For example, 50/50 would be an equal amount of light (one half of the total input light power) in each output. Because 50% of the light is 3 dB, this type of coupler is often called a *3-dB coupler*. Directivity is a measure of the amount of light that scatters backward into port 3. In a perfect coupler, this would be zero, but it typically measures −30 to −40 dB or less, when the output ports are index matched. Excess loss is the difference between the input power and output power. It is a measure of the efficiency of the coupler. It can range from less than 0.1 dB for a single-mode coupler to 1.5 dB for large-core multimode couplers. The splitting loss is the fraction of light appearing at each output port, neglecting the excess loss. For example, a 50/50 coupler has a splitting loss of 10 log(0.5) = 3 dB at each port, whereas a 90/10 coupler has a splitting loss of 10 log(0.9) = 0.46 dB on one port and 10 log(0.1) = 10 dB on the other port. The actual loss of the coupler is found by adding the excess loss to the splitting loss. For example, a 50/50 coupler with a 1.0-dB excess loss would have a total loss of 3 dB + 1 dB = 4 dB. This is the loss value that would be used in a system power budget.

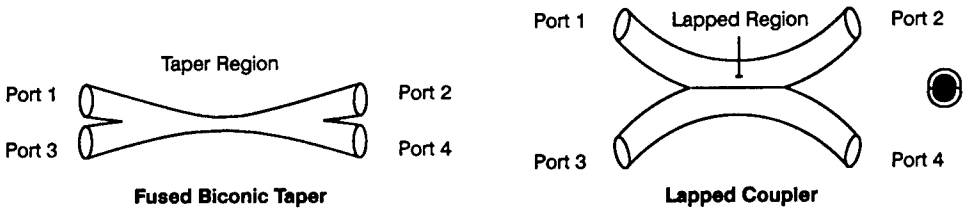


Fig. 4.48 Configurations of 2 × 2 tap couplers.

The 2 × 2 coupler can be fabricated in a number of ways, but the most common are shown in Fig. 4.48. In the fused biconic taper (FBT) coupler, the two fibers are held together while being slowly heated under tension. As the glass melting point is reached, the two fibers fuse together and are drawn down into a taper that is smaller than either of the individual fibers. The tapering process produces two cone-shaped regions in the coupler. The change in angle of the core-cladding interface is enough to convert many core modes into cladding modes in the taper. At the exit end of the taper, the reverse situation allows the cladding modes back into the expanding core. Because light in the taper region is contained in the cladding, any contamination such as finger oil or any index-matching packaging materials that come in contact with the glass surface allow light to leak out of the coupler. For this reason, the FBT couplers are normally suspended in air within a housing.

A different technique is used to produce a lapped coupler. A fiber is mounted on a radius tool and then a flat is polished or "lapped" on one side, exposing an oval-shaped region of the core (for multimode couplers). A second fiber is prepared in a similar manner, and the two flat sides are pressed together, making a coupler. To make this coupler permanent, the two fibers may be bonded together by a laser welding process. The resulting coupler is very strong, and, because no cladding mode conversion takes place, it can be encapsulated in epoxies to withstand a more rugged environment than the FBT coupler. The split ratio is determined by the polish depths of the two fibers, and the excess loss is determined by the quality of the polished surfaces, the relative alignments of the two flats, and the quality of the weld. By varying the fiber sizes and polish depths of the two fibers it is possible to make a coupler that has a different split ratio, depending on whether port 1 or port 3 is the input. This type of coupler is called *asymmetric* and can be used effectively in linear data bus systems.

Another method of making couplers is through the use of planar waveguide technology, as shown in Fig. 4.49. Two basic structures are possible. Photoresist and microlithographic techniques can be used to etch thin layers of glass that have been deposited on silica substrates, leaving raised waveguides, as shown. Similar techniques can be used to diffuse impurities into glass or other materials, creating areas of altered refractive index resulting in buried wave-

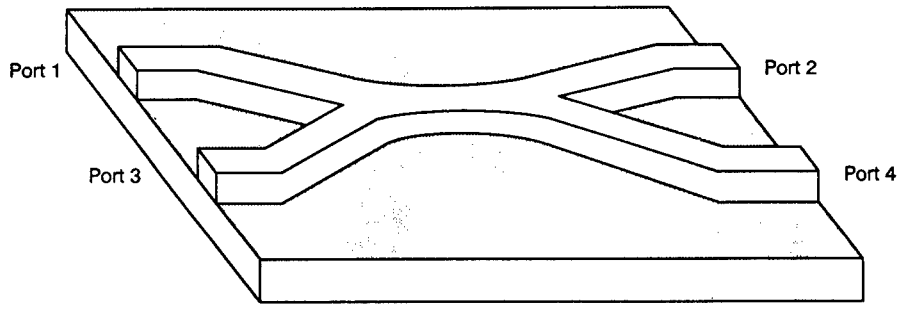


Fig. 4.49 Waveguide coupler.

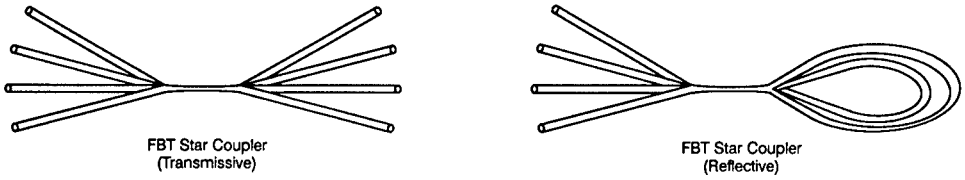


Fig. 4.50 Star couplers.

guides. Losses for waveguide-type couplers are comparable to those of FBT couplers if the waveguide material is low loss.

4.6.2.2 Star Couplers. Using these techniques, it is possible to make couplers having a larger number of input and output ports. Couplers larger than 2×2 are generally called *star couplers*. Figure 4.50 shows how FBT techniques can be used to construct 4×4 star couplers, both transmissive and reflective. In transmissive stars, input and output ports are on opposite ends of the coupler. In reflective stars, input and output ports are on the same side of the coupler. Star coupler loss is given by

$$\text{loss} = 10 \log \frac{1}{n} + E(n) \quad [\text{dB}] \ , \tag{4.43}$$

where n is the number of ports and $E(n)$ is the excess loss, which may be related to the number of ports. In most cases, the value of $E(n)$ includes port-to-port variations, and for multimode couplers is typically 1 to 1.5 dB. For single-mode couplers, it can be as low as 0.1 dB. Excess loss for reflective star couplers is typically larger than for transmissive stars, because the fuse-taper region is traversed twice.

4.7 SYSTEM ANALYSIS

Given the component data and specifications from the preceding sections it is now possible to assemble complete fiber optic systems and to calculate the performance of those systems. Two basic tasks must be performed in the system analysis: power budget analysis and bandwidth budget analysis. In this section, the analyses are explained and several examples of typical fiber optic systems are given.

4.7.1 Power Budget Analysis

Power budget analysis is the calculation of the amount of light power that arrives at the detector in comparison to the amount of light power required for reliable system operation. If the power budget is positive, then the system will have excess light, or margin. Systems having too little optical power to operate reliably (i.e., zero or negative margin) are called *loss-limited systems,* compared to systems that have adequate optical power but insufficient bandwidth, which are *bandwidth-limited systems.* The following steps should be performed in determining the power budget.

1. Determine the launch power. In most cases, the manufacturer of the transmitter or light source will specify the launch power into various fiber core sizes. If the total emitted power is specified, the power launched into the fiber core can be found using the graphs in Sec. 4.3. Linear units (microwatts, for example) should be converted into log units (dBm) for easier calculation of margins. For example, a launch power of 100 μW should be converted to −10 dBm. For modulated launch power, either peak or average power may be specified, as defined in Fig. 4.51. For most digital systems, the difference between peak and average power is 3 dB. In analog systems, such as video intensity modulated links, this may not be the case.
2. Total the system losses. Add the attenuations of all the passive optical components.
3. Calculate the light power incident on the detector. This is done by subtracting the system losses from the launch power. Again, this is easier to do when the launch power is in dBm units.
4. Find the receiver sensitivity. The receiver manufacturer will specify the minimum light power that produces adequate SNR for the appropriate analog signal reception or bit error rate for digital signal reception. This should be converted into dBm and should be in peak or average levels to match the launch power.
5. Calculate the power margin. Subtract the receiver sensitivity from the light power incident on the receiver. The difference is the power margin. A positive number indicates extra light power is available.

A power margin of zero indicates that there is just sufficient optical power at the receiver to allow the system to operate. Obviously, it is desirable to have somewhat more power than this. Launch power, system losses, and receiver sensitivity may all change as a function of environmental conditions and age. A typical system allows a 3- to 5-dB margin to allow for these factors. Figure 4.52 is a nomograph that can be used to perform power budget analyses. By knowing any two of launch power, system loss, and receiver sensitivity,

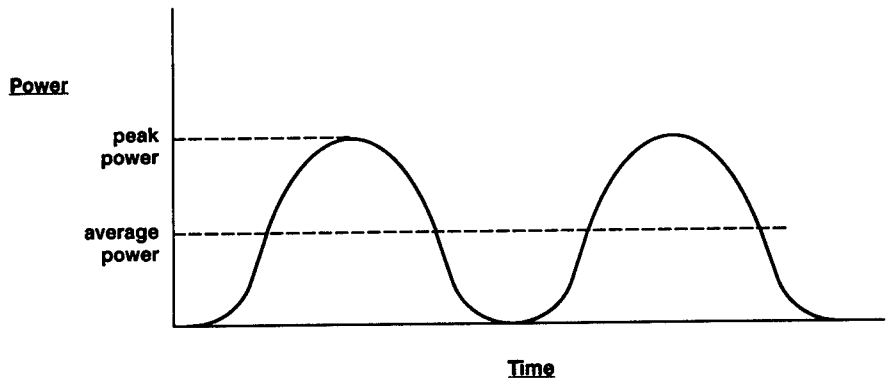


Fig. 4.51 Peak and average power.

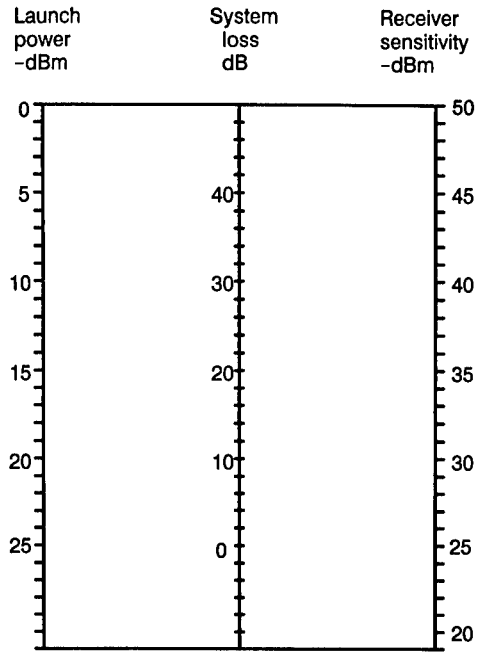


Fig. 4.52 System power budget chart.

the third can be found on the chart by drawing a straight line between the two known values and noting where this line intersects the third scale.

In most cases, the tolerances of the various components are assumed to combine in such a way as to produce worst-case operation, and the margin is calculated under this condition. It is equally likely, however, that the tolerances will combine in a way to produce a best-case system. It is necessary that the receiver dynamic range be capable of operating under both of these conditions. Therefore, once the power margin is found to be sufficient, it is a good idea to check the probable dynamic range requirements.

4.7.2 Bandwidth Budget Analysis

In addition to the system optical power requirements, it is necessary that the system bandwidth be adequate to allow transmission of the data. Bandwidth requirements for different data types are discussed in Sec. 4.4. Bandwidths of fibers are discussed in Sec. 4.2, and other passive components generally have no effect on bandwidth. The following relations are used to find the system bandwidth:

$$\mathrm{BW} = \frac{0.35}{\mathrm{Tr}} , \tag{4.44}$$

where BW is the bandwidth and Tr is the 10% to 90% risetime.

$$\mathrm{Tr}_{\mathrm{sys}}^2 = \mathrm{Tr}_{\mathrm{Tx}}^2 + \mathrm{Tr}_{\mathrm{Rx}}^2 + \mathrm{Tr}_{\mathrm{fiber}}^2 , \tag{4.45}$$

where

Tr_{sys} = system risetime
Tr_{Tx} = transmitter risetime
Tr_{Rx} = receiver risetime
Tr_{fiber} = fiber risetime.

If the system risetime equation is satisfied, then the system will operate at the required data rate with the correct SNR or bit error rate (BER). If the system bandwidth is inadequate, then a power penalty is necessary to re-establish SNR (i.e., additional optical power is required). Of course, there is a limit to the amount of power penalty that can be used to make up for bandwidth.

4.7.3 Simple Data Link

Figure 4.53 shows a simple data link consisting of a transmitter, a receiver, and a length of fiber. The object of this example is to calculate the maximum possible length of fiber, given the known performance limits of the transmitter and receiver. The following information is known:

1. data rate = 100 Mbits/s, NRZ format
2. fiber is 100/140, 0.29 NA, bandwidth = 200 MHz km
3. transmitter launch power = −10 dBm at 850 nm
4. receiver sensitivity = −37 dBm at 100 Mbits/s, NRZ
5. fiber attenuation = 4 dB/km at 850 nm.

First, calculate the power budget:

launch power Rx sensitivity:	−10 dBm
	− −37 dBm
allowed loss:	27 dB
safety margin:	−4 dB
remaining margin:	23 dB
length of fiber:	23 dB / 4 dB/km = 5.75 km.

Next, calculate the bandwidth budget:

system risetime for NRZ data:
$$T_{R_{sys}} = \frac{0.7}{\text{data rate}} = \frac{0.7}{100 \text{ Mbits/s}} = 7.0 \times 10^{-9} \text{ s}$$

system bandwidth:
$$BW_{sys} = \frac{0.35}{T_{R_{sys}}} = \frac{0.35}{7 \times 10^{-9}} = 50 \text{ MHz}$$

fiber bandwidth:
$$BW_{fiber} = 200 \text{ MHz km} / 5.75 \text{ km} = 40 \text{ MHz}.$$

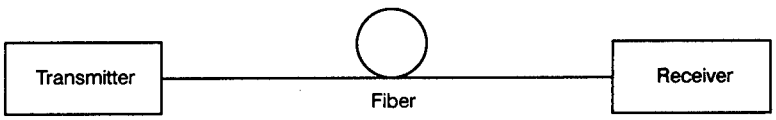


Fig. 4.53 Simple data link.

Because the fiber bandwidth at 5.75 km is less than the required system bandwidth, the system is bandwidth limited. Now the new length of fiber must be calculated to satisfy the bandwidth limitations.

Calculate bandwidth-limited fiber length and new power margin:

200 MHz km / 50 MHz = 4 km	
Allowed loss	27 dB
−4-km fiber	−16 dB
margin	11 dB .

Now the receiver must be able to operate at −26 dBm.

4.7.4 Full-Duplex Link with Dichroic Couplers

Figure 4.54 shows a full-duplex (bidirectional) link in which data are traveling in opposite directions simultaneously in the fiber. This is accomplished by using two different wavelengths (850 and 1300 nm) and dichroic couplers. The following information is known:

1. data rate = 10 Mbits/s, NRZ
2. fiber is 50/125, 0.2 NA, 1000 MHz km, 2-km long, with attenuation of 4 dB/km at 850 nm and 2 dB/km at 1300 nm
3. 850-nm Tx launch power = −13 dBm
4. 1300-nm Tx launch power = −16 dBm
5. 850-nm Rx sensitivity = −30 dBm
6. 1300-nm Rx sensitivity = −32 dBm
7. dichroic coupler loss = 1.5 dB
8. dichroic coupler crosstalk = 30 dB.

First, calculate the bandwidth budget:

fiber bandwidth = 1000 MHz/km / 2 km = 500 MHz .

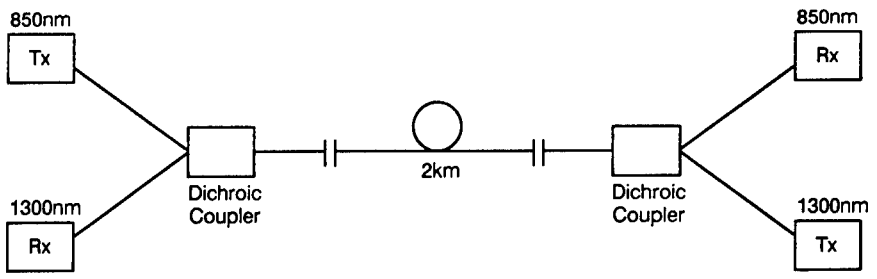


Fig. 4.54 Full-duplex link.

System bandwidth required is 5 MHz (see Sec. 4.7.3). Since the other system components do not contribute to bandwidth, the system is not bandwidth-limited.

Second, calculate power budget:

	850 nm	1300 nm
Launch power	−13.0 dBm	−16.0 dBm
Dichroic coupler	1.5 dB	1.5 dB
Connector	1.0 dB	1.0 dB
Fiber (2 km)	8.0 dB	4.0 dB
Connector	1.0 dB	1.0 dB
Dichroic coupler	1.5 dB	1.5 dB
Received power	−26.0 dBm	−25.0 dBm
Rx sensitivity	−30.0 dBm	−32.0 dBm
Margin	4.0 dB	7.0 dB

Next, check crosstalk levels. Because this is a full-duplex system, any signal leaking from one channel to the other will appear as noise to the receiver, and will increase the BER if it is greater than 15 dB below the actual signal level.[16]

For the 1300-nm Rx:

launch (850 nm)	−13.0 dBm
crosstalk isolation	30.0 dB
crosstalk level	−43.0 dBm
1300-nm received power	−25.0 dBm
difference	18.0 dB .

Therefore, the 1300-nm link is okay.

For the 850-nm Rx:

launch (1300 nm)	−16.0 dBm
crosstalk isolation	30.0dBm
crosstalk level	−46.0 dBm
850-nm received power	−26.0 dBm
difference	20.0 dB .

Therefore, the 850-nm link is acceptable.

4.7.5 Four-Channel Wavelength Multiplexed Link

In this example, shown in Fig. 4.55, a four-channel WDM is used to combine four data channels onto a single fiber. The following information is known:

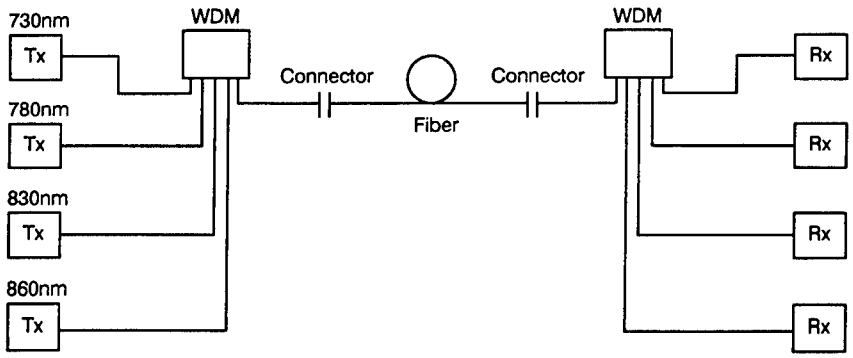


Fig. 4.55 Four-channel WDM link.

1. Fiber is 100/140, 0.29 NA, 200 MHz km, 1-km long with attenuation of 4 dB/km.
2. WDM loss is 4 dB, channel width is 15 nm.
3. Wavelength-related performance summary:

	730 nm	780 nm	830 nm	860 nm
Launch power	−7.0 dBm	−12.0 dBm	−10.0 dBm	−3.0 dBm
LED spectral width	30 nm	40 nm	45 nm	50 nm
Receiver λ penalty	2 dB	1 dB	0 dB	0 dB

The receiver wavelength penalty results from reduced photodiode responsivity at the shorter wavelengths. The actual WDM attenuation is a combination of the basic loss of the device and the filtering effect of the WDM channels on the LED spectra. The filter factor is calculated as described in Sec. 4.5.2. An additional loss of about 1 dB is normally budgeted for a mismatch between the multiplexer and demultiplexer. The object of this example is to calculate the optical power that will arrive at the receivers.

	730 nm	780 nm	830 nm	810 nm
Launch	−7.0	−12.0	−10.0	−3.0
WDM	4.0	4.0	4.0	4.0
Filter factor	3.0	4.3	4.8	5.2
Connector	1.0	1.0	1.0	1.0
Fiber	4.0	4.0	4.0	4.0
Connector	1.0	1.0	1.0	1.0
WDM	4.0	4.0	4.0	4.0
Mismatch	1.0	1.0	1.0	1.0
Rx penalty	2.0	1.0	0	0
Received power	−27.0 dBm	−32.3 dBm	−29.8 dBm	−23.2 dBm

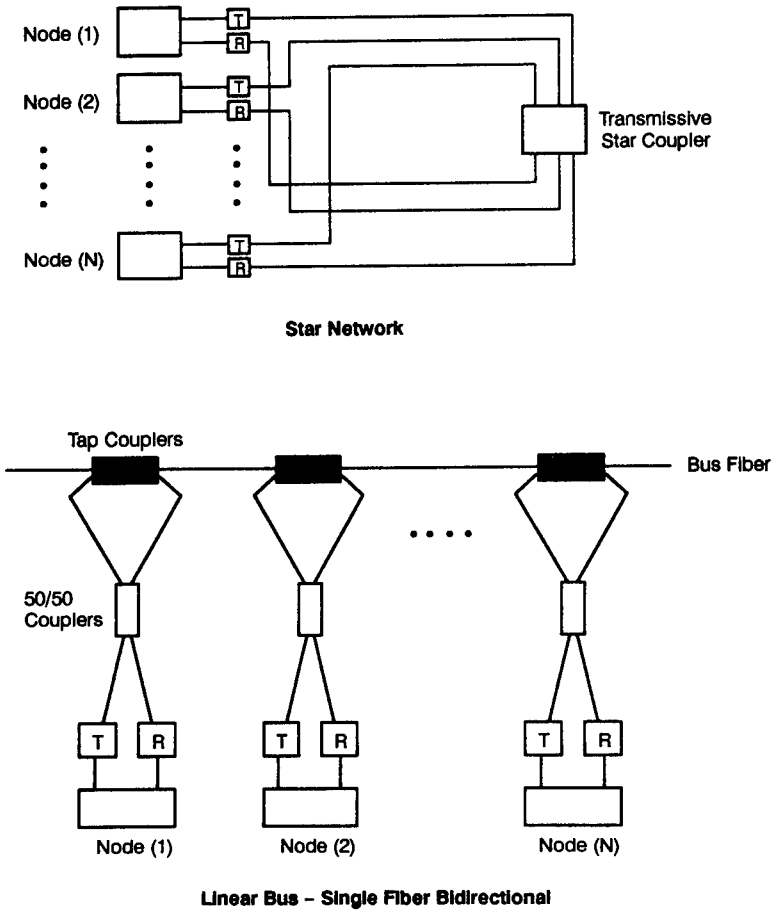


Fig. 4.56 Basic network topologies.

4.7.6 Fiber Optic Networks

There are an almost endless number of ways to configure fiber optic components into networks. However, the various topologies can be broken down basically into stars or linear busses. Figure 4.56 shows these two topologies.

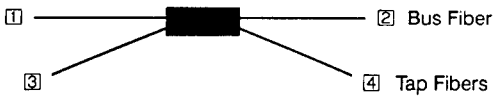
4.7.6.1 Star Network. In the star network, each node or terminal is furnished with a transmitter and a receiver. All transmitters are connected to the input side of an $n \times n$ star coupler. All of the receivers are connected to the output ports of the star. The power budget is inversely proportional to the number of nodes, exactly following the star coupler loss. The advantages of a star network are that it exhibits the minimum loss of any passive topology, and that it requires the minimum dynamic range, because each node sees the same loss. The main disadvantages are that the star is located remotely from the nodes, requiring a large number of fibers to provide all of the necessary connections, and that a failure of the star can defeat the entire network.

4.7.6.2 Linear Bus. The linear bus is characterized by a bus fiber and some number of nodes where light is tapped and inserted. The linear bus is a serial

construct rather than a parallel one like the star network. The loss from one node to another is additive and is, therefore, different for each node to node combination. The figure shows one way to obtain bidirectionality at each node on a single fiber by using a 50/50 coupler. The coupler increases the node-node loss by about 8 dB. A bidirectional bus without 50/50 couplers would rely on two bus fibers, one to carry information in each direction, and twice the number of transmitters and receivers. The advantages of the linear bus are that it requires only one or two fibers to connect all of the nodes, and that there is no single point failure such as with a star coupler (although, depending on the network protocols, a broken bus fiber might produce the same result). In many installations, a linear topology is the most sensible physical layout for a network. The main disadvantage is the higher loss of the linear bus.

Linear busses can be configured from either symmetric or asymmetric tap couplers. Figure 4.57 shows the difference in operation between the two. The basic difference is that the asymmetric couplers allow more light to be injected onto the bus fiber, effectively stretching out the number of nodes that can be accessed with limited launch power and receiver sensitivity.

Figure 4.58 is a graph showing relative losses for different numbers of nodes for four different basic topologies. The reflective star coupler topology includes



90/10 Symmetric Coupler Losses

In \ Out	1	2	3	4
1	--	0.95	--	10.8
2	0.95	--	10.8	--
3	--	10	--	0.95
4	10.8	--	0.95	--

200/50 Asymmetric Coupler Losses

In \ Out	1	2	3	4
1	--	0.3	--	17.5
2	0.3	--	17.5	--
3	--	3.0	--	19.0
4	3.0	--	19.0	--

Fig. 4.57 Tap couplers for linear bus.

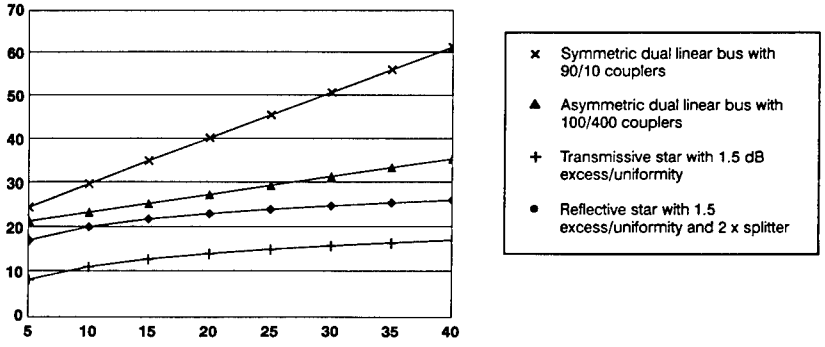


Fig. 4.58 Network optical losses.

the loss of a 2 × 2 coupler at each node. The linear topologies use losses shown in Fig. 4.57, with splices at 0.1 dB each used to connect the couplers.

References

1. A. Yariv, *Introduction to Optical Electronics,* 2nd ed., pp.1–17, Holt, Rinehart and Winston, New York (1980).
2. A. W. Snyder and John D. Love, *Optical Waveguide Theory,* Chapman and Hall, New York (1983).
3. M. Born and E. Wolf, *Principles of Optics,* 6th ed., pp. 1–70, Pergamon Press, New York (1980).
4. Helmut F. Wolf, Ed., *Handbook of Fiber Optics: Theory and Applications,* Garland STPM Press, New York (1979).
5. Technical Staff of CSELT, *Optical Fiber Communication,* p. 267, McGraw-Hill, 1981.
6. E. E. Basch, Ed., *Optical Fiber Transmission,* 1st ed., p. 12, Howard W. Sams (Macmillan) (1987).
7. Donald B. Keck, "Fundamentals of optical waveguide fibers," *IEEE Communications Magazine* **23**(5) (1985).
8. S. M. Sze, *Semiconductor Devices,* John Wiley & Sons, New York (1985).
9. Eric Udd, *Fiber Optic Sensors,* John Wiley & Sons, New York (1991).
10. P. K. Cheo, *Fiber Optics Devices and Systems,* Prentice-Hall, Englewood Cliffs, NJ (1985).
11. M. Ettenberg, unpublished.
12. M. Ettenberg and H. Kressel, *IEEE Journal of Quantum Electronics* **QE-16,** 170 (1980).
13. B. Culshaw, *Optical Fibre Sensing and Signal Processing,* p. 55, Peter Peregrinus Ltd., London, 1984.
14. S. D. Personick, "Receiver design for digital fiber optic communication systems, Part I and Part II," *Bell System Technical Journal* **52,** 843–886 (July–Aug. 1973).
15. T. V. Muoi, "Receiver design for high speed optical fiber systems," *IEEE Journal of Lightwave Technology* **LT-2**(3), 243–267 (1984).
16. B. S. Vidula, W. M. Muska, and T. W. Leonard, "Crosstalk penalty in digital fiber optic systems employing wavelength division multiplexing couplers/decouplers," in *Proceedings of FOC-LAN 83 Information Gatekeepers,* Boston, MA, p. 139, Information Gatekeepers (1983).

Index